

Modeling Fee Shifting With Computational Game Theory

Michael Abramowicz
abramowicz@law.gwu.edu
George Washington University Law School

October 16, 2022

Abstract

While modern mathematical models of settlement bargaining in litigation generally seek to identify perfect Bayesian Nash equilibria, previous computational models have lacked game theoretic foundations. This article illustrates how algorithmic game theory can complement analytical models. It identifies equilibria by applying linear programming techniques developed in von Stengel et al. (2002) to a discretized version of a cutting-edge model of settlement bargaining, Dari-Mattiacci and Saraceno (2020). This approach makes it straightforward to alter some assumptions in the model, including that the evidence about which the parties receive signals is irrelevant to the merits and that the party with a stronger case on the merits also has better information. The computational model can also toggle easily to explore cases involving liability rather than damages and can incorporate risk aversion. A drawback of the computational model is that bargaining games may have many equilibria, complicating assessments of whether changes in equilibria associated with parameter variations are causal.

1 Introduction

The literature analyzing the effects of fee shifting confronts a daunting analytic challenge. Settlement bargaining is a two-player asymmetric information game. The gold standard solution to such a game is a pair of common knowledge strategies that form a perfect Bayesian equilibrium. The perfection requirement, as defined by Fudenberg and Tirole (1991) [10], specializes the general Nash [29] equilibrium criterion in an imperfect information game, by insisting that at no point in the game may a player have any incentive to change the player's strategy. Each player applies Bayesian reasoning to incorporate new information, such as signals of case quality, into estimates of trial outcomes, and the player's probabilistic beliefs are required to be correct given such information. Complicating the challenge of crafting such equilibria, each party might decide not to

contest the litigation, one or both parties may be risk averse, players may or may not have asymmetric information, a case may concern liability or damages, and the loser may or may not be required to pay the winner’s fees.

Incorporating anywhere near all of these considerations into a single model of settlement bargaining has proven elusive. The settlement bargaining modeler stands before a smorgasbord of potentially critical game features, but faces the admonition to choose no more than a few. The result, Daughety and Reinganum (1993) [6] observed, is a literature that “has grown in a disorganized fashion, resulting in a multitude of models involving different informational endowments and timing structures.” This statement remains true nearly three decades later, with the settlement-bargaining-modeling art making unmistakable but limited progress. The earliest models of Landes (1971) [25], Posner (1973) [32], and Gould (1973) [12] had ignored the challenges of Bayesian inference. Later came models, such as Bebchuk (1984) [2] and Polinsky and Rubinfeld (1998) [30], in which one party knows the probability of liability or the amount of damages while the other party knows only the distribution, and Daughety and Reinganum (1994) [8], in which one party has information on liability and the other party, on damages. The latest generation of scholarship, including Friedman and Wittman (2006) [9], Klerman, Lee, and Liu (2018) [23], and Dari-Mattiacci and Saraceno (2020) [5], models two-sided asymmetric information, in which each of the plaintiff and defendant has independent private information about the same issue, for example about the level of damages. The last of these even succeeds at the Herculean task of incorporating fee shifting, but we will see that even it does not escape the curse of dimensionality, as it must adopt a number of restrictive assumptions that make it difficult to assess the generality of its conclusions.

The literature is extraordinarily clever, in both the positive and negative senses of the word. It takes advantage of mathematical assumptions to make otherwise intractable problems tractable. We can expect further progress from relaxing different assumptions, but the goal of developing a single mathematical model that allows exploration of different values of a large number of variables may be unattainable. The literature develops critical intuitions about settlement bargaining, including how changing fee shifting rules may augment or diminish the effectiveness of the litigation system, and review articles, like Katz and Sanchirico (2012) [22], informally integrate various models’ conclusions about how fee shifting might affect trial rates or settlement rates with different structures to the litigation game. Even with such reviews, however, it is difficult to generalize about the wisdom of fee shifting, because of the interactivity between trial and settlement rates. If, for example, increased ease of settlement leads to plaintiffs’ bringing and defendants’ defending more cases, total litigation expenditures in principle could rise.

Scholars have studied settlement bargaining with other methodologies, but these have their own limitations. Empirical analyses are limited to studying the rare cases in which a change in fee-shifting rules occurs, as in the examination by Hughes and Snyder (1995) [18] of a briefly-lived policy experiment in Florida. Laboratory experiments offer another approach, with contributions

by Coursey and Stanley (1988) [4], Inglis et al. (2005) [20], Rowe and Vidmar (1988) [35], and Main and Park (2000) [27]. It is not clear, however, whether the artificial stakes and quick decisions in a lab produce results similar to those of real litigation. A final methodology in the literature is computer-based simulation. Priest and Klein (1984) [33], Katz (1987) [21], Hause (1989) [15], and Hylton (1993) [19] were pioneers in using computation, either independently or as complements to formal models. While these articles all include innovations building on the Landes-Gould-Posner model, they share a significant limitation: Unlike the math models, the simulations do not seek perfect Nash equilibria.

It is, however, possible to harness computational power in the quest for perfect Bayesian equilibria, by turning to computational game theory. The settlement bargaining literature has acknowledged the importance of game theoretic concepts, but articles build at most relatively small game trees. Acknowledging that litigation can be viewed as “a particular extensive-form bargaining game,” Spier (1994, pp. 202-03) [38] sensibly worries that the results would be sensitive to issues such as “the structure of the asymmetric information.” This concern suggests that solvable game theoretic models are insufficiently rich to encapsulate critical aspects of the litigation game. Computational game theory, however, allows for the identification of equilibria in games that could not practically be solved by hand. A subliteration focuses directly on the solution of two-player (and sometimes n -player) general sum games. A litigation game between plaintiff and defendant is general sum, which can be more difficult than a zero-sum game to solve, because the players may transfer wealth not only from defendant to plaintiff, but also from both litigants to lawyers. A useful review of algorithms that can help solve such games is von Stengel (2002) [39].

A publicly available open source software package known as Gambit by McKelvey et al. (2016) [28] features a number of these algorithms. This article, however, applies an algorithm not included in Gambit, specifically an algorithm described in an article in *Econometrica*: von Stengel, van den Elzen, and Talmann (2002) [40]. This algorithm, described further below, is guaranteed to produce perfect Bayesian Nash equilibria in a finite game in which the players have perfect recall. Sometimes, these equilibria are pure, with players acting deterministically conditional on the information that they possess, but at other times, they are mixed, with the players randomly choosing at certain moments of the game between or among equally good strategies, each with some nonzero probability. Mixed strategies need not reflect explicit randomization by players; they may be understood as describing balanced populations of litigants who take different approaches, none better than others given opponents’ strategies, for reasons exogenous to the model. The authors test their algorithm on games of up to 1,023 nodes. This article pushes the computational limits of the algorithm, applying it to each of a large number of games with up to 16,111 nodes. By separately identifying equilibria corresponding to different information and game structures, we can assess the conditions in which changing fee-shifting rules may increase or decrease the effectiveness of the litigation system, as manifested in accuracy or trial rates, assuming the litigation game is played in equilibrium by rational actors. This approach thus enables modeling of a richer and more

diverse litigation environment than any single prior approach.

To illustrate how this approach can complement analytic models, this Article focuses on the model of Dari-Mattiacci and Saraceno (2020), the first article to integrate both two-sided asymmetric information and fee-shifting. In this model, both players know the true quality of the litigation, but the judgment depends not only on this value, but also on the sum of signals independently received by the parties. Part 2 describes the Dari-Mattiacci and Saraceno model, and it identifies a number of assumptions that may be necessary for mathematical tractability but are unrealistic or greatly narrow the model’s applicability. Part 3 then presents a discretized version of the Dari-Mattiacci and Saraceno model that can be solved computationally. It also addresses two potential concerns: first, that discretization might itself significantly affect equilibrium strategies, and second, that the game might have multiple equilibria.

Part 4 provides the central results. It first reports accuracy and trial rates for values of litigation quality within the range permitted by the Dari-Mattiacci and Saraceno model, and it obtains results consistent with their conclusions, at least when an implicit assumption of their model that trial costs are relatively small is met. It then relaxes a number of assumptions of their model, extending the range of permissible litigation quality and litigation quality parameters, disentangling strength of information from litigation quality, allowing the evidence that parties receive to be relevant to the merits, adopting other fee-shifting rules, granting the parties an outside option not to litigate, and featuring risk-averse litigants. The results generally illustrate the importance of Dari-Mattiacci and Saraceno’s assumption that the parties’ evidentiary signals do not reveal information about the merits. Under this assumption, the English rule tends to improve accuracy, but when the information is about the merits, the English rule results in greater inaccuracy. When liability is at issue, however, the English rule appears slightly to reduce inaccuracy. All of these conclusions, however, are sensitive to the definition of accuracy, and an alternative definition that focuses on per-case accuracy instead of average accuracy produces quite different results. Overall, the results are sufficiently nuanced and sensitive to model specification that the exercise highlights the difficulty of reaching conclusions about the effects of fee-shifting rules that are broadly applicable.

2 Analytical Models of Two-Sided Asymmetric Information

Because this article’s goal is to illustrate how computational models can build on limitations of an analytical model, and vice-versa, we will focus on the structure and methodological choices in Dari-Mattiacci and Saraceno (2020). We begin by reviewing the Friedman and Wittman (2006) [9] upon which Dari-Mattiacci and Saraceno build, before summarizing the Dari-Mattiacci and Saraceno model and exploring some of the assumptions inherent in the article.

2.1 Friedman and Wittman’s Averaged Signals Model

In the one-sided information models, the structure of bargaining often affects which party receives most of the surplus. Friedman and Wittman avoid this problem by adopting the bargaining protocol of Chatterjee and Samuelson (1983) [3]. In Chatterjee-Samuelson bargaining, the plaintiff and defendant simultaneously submit offers. If the plaintiff’s exceeds the defendant’s, the case definitively settles at the midpoint; otherwise, bargaining has failed. Costs of trial are borne only in the event of bargaining failure. Friedman and Wittman justify this bargaining structure not on the ground that the protocol is commonly used (it is not), but on the ground that it provides a useful reduced form of a more complicated bargaining process. In contrast to divergent expectations models, with Chatterjee-Samuelson bargaining, a case may go to trial even though there is a social surplus from settlement given the parties’ expectations. The reason is that the parties may shade their offers in the hope of deriving a larger portion of the settlement surplus, even at the risk of bargaining failure.

The informational structure is arrestingly simple. The plaintiff observes a signal θ_p drawn from a known distribution, and the defendant independently observes a signal θ_d drawn from the same distribution. The principal results of the paper apply to a “basic litigation model” in which the distribution is the uniform distribution; this extends without loss of generality to any uniform distribution between a lower bound of L and an upper bound of U . In the event that settlement fails, a judgment is entered in the amount of the average $(\theta_p + \theta_d)/2$. Perhaps one can imagine situations in which this might be realistic. For example, the parties might have information about different components of damages in a case in which liability is uncontested, and should trial ensue, the information will be revealed and the judgment will be the sum. But a more intuitively appealing model in most situations would reverse the causality. Signals would depend on the underlying truth to be revealed at judgment, rather than the judgment depending on signals. Friedman and Wittman cleverly recognize, however, that modeling litigation in this way makes the model tractable.

Friedman and Wittman derive a Nash equilibrium in the basic litigation game. In this equilibrium, the plaintiff will ordinarily offer $\frac{2}{3}\theta_p - 2c + \frac{1}{2}$, and the defendant will ordinarily offer $\frac{2}{3}\theta_d + 2c - \frac{1}{6}$, where c represents each party’s trial cost. The word “ordinarily” signals what may seem a mild caveat: Neither party will ever make an offer beyond the range of the other party’s possible offers. Thus, the plaintiff’s offers are truncated above at $\min(1, 2c + \frac{1}{2})$ and below at $\max(0, 2c - \frac{1}{6})$, while the defendant’s offers are truncated above at $\min(1, \frac{7}{6} - 2c)$ and below at $\max(0, -2c + \frac{1}{2})$. Friedman and Wittman do not eliminate the possibility that there might be some nonlinear Nash equilibrium, but they prove that the equilibrium they derive is the unique nontrivial piecewise linear equilibrium. There are also infinitely many trivial equilibria, in which the plaintiff’s settlement demands always exceed the defendant’s.

Friedman and Wittman’s model permits them to focus on the trial rate. They derive a piecewise quadratic formula for the trial rate, and they also examine, in the tradition of Priest-Klein, how the trial rate varies near the midpoint

of the decision spectrum. They show that when trial costs are sufficiently low ($c < \frac{1}{6}$), the probability of a trial is higher, the farther the judgment would be from $\frac{1}{2}$, and when trial costs exceed this threshold, the probability of a trial is highest at the $\frac{1}{2}$ point. The intuition is that when costs are high, the parties become more generous, and so the plaintiff's range of offers will fall below the defendant's. The truncations then ensure that cases at the extremes, where either both parties receive a low signal or both parties receive a high signal, are more likely to settle. When trial costs are low, the parties are less generous, and the plaintiff's range of offers will be above the defendant's. Cases at the extremes are then less likely to settle. With the basic litigation game, the $\frac{1}{6}$ cost threshold occurs where the parties' range of offers are equal. Friedman and Wittman also offer a graphical argument that extends to other continuous distributions, though they do not expressly consider the case where liability rather than damages is uncertain.

2.2 Dari-Mattiacci and Saraceno's Evidentiary Signals Model

Dari-Mattiacci and Saraceno (2020) illustrate the challenge of building on Friedman and Wittman by successfully extending the model to fee shifting. The article includes an online appendix with 60 pages of proofs. The difficulty stems from the need to address four principal cases, depending on relative values of parameters, and within these principal cases, to make various calculations that depend on the relative values of other parameters, including in many instances five different formulas for five different ranges of a variable. The resulting product is testimony both to human ingenuity and to endurance, and it makes breakthroughs in our understanding of the effects of fee-shifting with two-sided asymmetric information.

As in Friedman and Wittman, plaintiff and defendant receive signals, now denoted θ_Π and θ_Δ , respectively, and the judgment is an average of the signals. Now, however, both parties have common knowledge of the true merits of the litigation, denoted by q . The signals thus do not serve the function of informing the parties of the true merits, but rather of providing the parties with evidence that they may use to convince the court. The plaintiff's signal θ_Π is drawn from a uniform distribution on the interval $(0, q)$, and the defendant's signal, on the interval $(q, 1)$. Because the defendant's signal can be no less than q , the plaintiff's best possible evidence, where $\theta_\Pi = q$, would convince the court that the judgment must be at least q . Similarly, the defendant's best possible evidence, where $\theta_\Delta = q$, would convince the court that the judgment must be no more than q . But the litigants do not always draw the best possible evidence.

The fee shifting rule that Dari-Mattiacci and Saraceno primarily analyze is triggered based on (1) whether the final judgment is above or below $\frac{1}{2}$ (i.e., which party "wins" in the sense of being awarded more than half of the contested damages), and (2) whether the evidence of the winning party is sufficiently strong. If the judgment is less than $\frac{1}{2}$, then the defendant might be able to shift its costs to the plaintiff, but only if the defendant's signal falls below some

threshold, i.e. $\theta_\Delta < t$, where $0 \leq t \leq 1$. Likewise, if the judgment is greater than $\frac{1}{2}$, then the plaintiff might be able to shift its costs to the defendant, but only if the plaintiff's signal exceeds a threshold, i.e. $\theta_\Pi > 1 - t$. An intuition is that if a party wins a case merely because its opponent has produced little evidence, a court will not order fee-shifting; another is that a court will only order shifting of fees when those fees were spent on producing strong evidence. Note that when $t = 0$, fees will never be shifted, so this extreme is the American rule of no fee shifting, and when $t = 1$, fees will always be shifted to a winning party (i.e., to the plaintiff if the final judgment exceeds $\frac{1}{2}$ and to the defendant if the final judgment is less than $\frac{1}{2}$), so that extreme is the English rule of universal fee shifting. The analysis thus effectively allows for a continuum of fee shifting rules.

This information structure enables Dari-Mattiacci and Saraceno to derive the offers that the parties will make. They prove that each party's offer function is a best response to its opponent's offer function and thus that a Bayesian Nash equilibrium exists. They also derive formulas for settlement amounts, along with identification of the ranges of parameters values where such settlements occur, and accordingly of the litigation rate. They prove that the litigation rate depends only on c (now representing the combined trial cost of the two parties) and is thus independent of both case quality q and the fee-shifting rule t . This produces the surprising conclusion that the litigation rate is the same under both the American and the English rule. Finally, they offer a calculation of litigation accuracy, and they prove that when costs are below a certain threshold, the English rule produces more accuracy than the American rule, while the reverse is true when costs are above a certain threshold. The stylized fact that litigation is cheaper in England may thus help explain the choice of rule in each country.

2.3 Assumptions in Dari-Mattiacci and Saraceno's Model

The Dari-Mattiacci and Saraceno model adopts a number of assumptions. Many of these assumptions appear to be driven, quite reasonably, by the demands of mathematical tractability, and it is difficult to develop strong intuitions for whether they matter. In identifying these assumptions, we create a series of challenges for a computational model that aspires to assess the robustness of the analytical model.

Parameter values

Balanced asymmetric information Meanwhile, the true merits variable is constrained so that $\frac{1}{3} \leq q \leq \frac{2}{3}$. The reason for this constraint is that with more extreme values of q , the increasingly one-sided nature of asymmetric information leads the pure strategy equilibria derived by the authors to break down. This highlights once again the problematic nature of asymmetric information quality equivalence, because it means that the authors not only cannot model situations with relatively high information asymmetry, but also that they

cannot model situations in which the true merits of a case are near the extremes of the probability distribution. Perhaps a computational model might be able to find an equilibrium with relatively extreme quality values and/or with relatively extreme information asymmetry, and this could help extend the understandings provided by the model.

Low or moderate cost Dari-Mattiacci and Saraceno follow Friedman and Wittman in implicitly assuming that the cost variable is not so high that the plaintiff's untruncated offer range is entirely below the defendant's untruncated offer range. The truncation functions defined by Friedman and Wittman are undefined, because when their c is sufficiently high, they instruct that the plaintiff's offers should be truncated above at 1 and below at a number greater than 1, and similarly the defendant's offers are truncated below at 0 and above at a number less than 1. With sufficiently high costs, there will be many Nash equilibria; the parties will be determined not to go to trial, but neither party would deviate from any positive allocation of the surplus from settlement. Literal application of the Dari-Mattiacci and Saraceno formula, however, would lead to both players truncating their bids and would not choose any of these equilibria.

In a subtle way, the Dari-Mattiacci and Saraceno cost assumption is more restrictive than Friedman and Wittman's. As we will see shortly, a computational model can be used to assess whether parties' strategies form an equilibrium, and with sufficiently high costs, the bid functions identified by Dari-Mattiacci and Saraceno in some cases do not form equilibria. This can be traced, at least in part, to a complication in what Dari-Mattiacci and Saraceno call Case 4B. They implicitly assume that the bid functions that they derive would each contain a discontinuity, but if $6c(1 - q) > 1$, the plaintiff's bid function consists only of a single line segment. It can be shown, for example, that for the parameters $t = 0.8, q = 0.4, c = 0.3$, the plaintiff's strategy cannot be a best response, because, given the defendant's presumed strategy, the plaintiff would be very slightly better off with a bid function in which it always bids one-third of its normalized signal. In correspondence, Dari-Mattiacci and Saraceno have acknowledged this complication and that their model implicitly assumes that c is not too high. This is a reasonable assumption, but a challenge for the computational model is to overcome it.

Structural constraints

Piecewise linearity Dari-Mattiacci and Saraceno explicitly assume a linear relationship between the parties' signals and their offers. They allow, however, for discontinuities in the linear relationship. In this sense, the strategies they model are similar to those of Friedman and Wittman, and indeed Dari-Mattiacci and Saraceno similarly truncate the parties' strategies. The assumption is somewhat stronger, however, in that Friedman and Wittman demonstrated that the piecewise linear strategies they derived would be a Nash

equilibrium even when nonlinear strategies are possible. On the other hand, Dari-Mattiacci and Saraceno allow for additional discontinuities at points where fee-shifting would change. This is central to the design of their model and the thrust of their analysis. Because fee-shifting depends partly on the quality of the evidence possessed by the winning party, a litigant will know whether it will be entitled to fee-shifting if it wins, and the signal values at which this fact changes are points at which Dari-Mattiacci and Saraceno are able to break the problem down into smaller pieces. Piecewise linearity thus allows for explicit modeling of the effects of changes in a fee-shifting rule, but because it is unclear how restrictive this assumption is, it is a prime candidate for relaxation in a computational model.

Asymmetric information quality equivalence Recall that the plaintiff receives a signal in the range $(0, q)$ and the defendant, in $(q, 1)$. As a consequence, when $q > \frac{1}{2}$, the plaintiff’s signal has a greater potential effect than the defendant’s, and when $q < \frac{1}{2}$, the reverse is true. The single variable q thus serves two, independent functions in the model: one is to represent the “true merits” of the case, while the other is to represent the degree of information asymmetry. This greatly increases the tractability of the model, and plausibly it allows for consideration of both issues related to accuracy and issues related to information asymmetry. The problem, though, is that the issues are necessarily conflated; where a case is at an extreme of the probability distribution, there is always high information asymmetry. There is no obvious reason to believe that true merits should generally track information asymmetry in this way. The question thus arises whether the results would be the same if the model allowed independent variation of true merits and information asymmetry.

Signals independent of true merits Dari-Mattiacci and Saraceno refer to the signals that the parties receive as “evidence” of the true merits of the case, but there is a paradox: The parties are assumed to know the true merits of the case (q) and indeed use this information in constructing their offer functions. Thus the variance in the signals that each party may receive has nothing to do with the merits. Given the fixed value of q , whether the plaintiff receives a signal slightly above 0 or slightly below q tells the plaintiff nothing about the true merits. What receipt of the signal accomplishes is to inform the plaintiff about the plaintiff’s likely ability to persuade the judge about the true merits. The judge does not know the true merits, but is trying to guess the true merits. The higher q , the higher the parties’ signals will tend to be, so the judge’s strategy is reasonable, even if non-Bayesian. But the result is that from the perspective of the parties, for whom q is fixed, the randomness in case outcomes has to do only with who is lucky in finding promising evidence.

This point can be more clearly seen in a transformation of the model that Dari-Mattiacci and Saraceno offer. They note that the signals θ_Π and θ_Δ can be mapped one-to-one onto signals from 0 to 1, which they label z_Π and z_Δ . These signals are thus independent signals from a unit uniform distribution, and the

θ signals can be derived from them according to the formulas $\theta_{\Pi} = qz_{\Pi}$ and $\theta_{\Delta} = q + (1 - q)z_{\Delta}$. This highlights that the θ signals result from commingling the true merits of the case and the random uniform distribution draws. With these transformations, the judgment depends on the following formula:

$$J(z_{\Pi}, z_{\Delta}) = \frac{1}{2}q + \frac{1}{2}(qz_{\Pi} + (1 - q)z_{\Delta}) \quad (1)$$

As this presentation makes clear, the decision is half based on the true merits of the case, independent of any evidence presented by the parties. Meanwhile, the decision is half based on a weighted average of the parties' uniform distribution draws, with the weights equal to q . Recall that q represents the degree of information asymmetry. Thus, in effect, half of the judge's decision is based on the true merits and half of the judge's decision depends on a weighted average of signals that are entirely independent of the true merits. The only reason that this makes sense from the perspective of the judge is that the judge does not observe z_{Π} and z_{Δ} directly. The Dari-Mattiaci and Saraceno model could be realistic in some contexts, in which the parties have asymmetric information about their persuasive abilities, independent of the merits. In any event, their approach may be necessary in a mathematical model that seeks, as theirs does, to measure accuracy. It is considerably easier (though still extraordinarily difficult) to measure outcomes relative to the constant q than it would be relative to a function of q and the parties signals.

A challenge for the computational model is to assess whether results about accuracy continue to obtain when the true merits are defined to be inclusive of the parties' normalized signals. On this formulation, q would represent knowledge that the parties share about the true merits, and the z signals represent private information about the true merits. When the judge adds these together according to the above formula, the judge obtains not only the judgment, but also the true merits. This is thus a conceptual reformulation with no implications for which cases settle. It requires only an alteration of the definition of accuracy.

Game structure

Fee-shifting structure Fee shifting in Dari-Mattiaci and Saraceno's model depends not only on which party wins more than half of the judgment at trial, but also on the quality of the evidence produced by the winning party. This is mathematically convenient, because each party knows the quality of its own evidence and thus whether fee-shifting will occur for any given value of the opponent's signal and any value of t . An alternative approach, however, would be for fee-shifting to depend on both parties' evidence. Indeed, Dari-Mattiaci and Saraceno explicitly consider fee-shifting based on the margin of victory. If we redefine t to represent the margin-of-victory parameter, then if $\theta_{\Pi} + \theta_{\Delta} < t$, the plaintiff must pay the defendant's fees, and if $\theta_{\Pi} + \theta_{\Delta} > 2 - t$, then the defendant must pay the plaintiff's fees. In this regime, if $t = 0$, no fee shifting

occurs (the American rule), and if $t = 1$, fee shifting always occurs absent an evenly split judgment (the English rule); thus, the margin-of-victory approach converges with the other approaches at the extremes. Dari-Mattiacci and Saraceno explicitly calculate the parties' offers under this approach, but they do not prove their results related to accuracy. This raises the question whether their accuracy results are robust to the alternative specification. One might also imagine other fee-shifting rules, such as a simple rule in which a party always occurs when the party wins half of the judgment but the proportion of fees shifted may vary from 0 to 1.

Damages issue Dari-Mattiacci and Saraceno explicitly describe their model as one in which the parties are arguing about how to divide a disputed asset, such as in a case of divorce, and they point out that without loss of generality, this can be extended to a judicial determination of damages between some minimum and maximum value. An extension would be to consider cases where liability is at issue, i.e. where the plaintiff will receive 1 if $\theta_{\Pi} + \theta_{\Delta} > 1$ and 0 otherwise. For example, they might generalize the model to an arbitrary cumulative distribution function mapping $\theta_{\Pi} + \theta_{\Delta}$ onto the judgment, but this would add considerable challenge. Once again, this should be trivial in a computational model, which need only transform the judgment values in particular cases, either to 0 or 1 or based on some other distribution.

Contestation Dari-Mattiacci and Saraceno's model assumes that the plaintiff always files the lawsuit and that the defendant always answers. [ADD REFERENCE TO SHAVELL (1982) AND ON OTHER SIDE] In their Online Appendix, they offer some informal considerations of how fee shifting might affect filing. They reason, but do not explicitly prove, that the greater the fee-shifting parameter t , the more often the plaintiff will not file or the defendant will not contest litigation. This could affect the settlement bids that the parties make. Thus, we cannot be sure that an equilibrium in the original model will remain an equilibrium when the parties are allowed not to contest the litigation.

Risk neutrality The plaintiff and defendant are assumed to be risk neutral. Incorporating risk aversion into the model would likely add considerable challenge, though the argument could still proceed in case-by-case fashion. Incorporating risk aversion is virtually costless to a computational model, requiring only the transformation of the parties' utilities in any game outcome.

Accuracy definition Dari-Mattiacci and Saraceno define inaccuracy in their appendix as "the square distance between the expected outcome E_t and the merits q ." We have already explored how we might reconceive the definition of the merits so that all evidence is counted as part of the true merits, so let us continue moving from right to left in this definition.

The definition of E_t is complex, involving double integrals over both costs and the parties' signals. The essence is that it is a measure of the expected

outcome of a dispute, taking into account both the settlements and the trials. The outcome in the event of trial that they calculate is represented by G , which “captures both the decision on the merits and fee shifting.” For example, if the judgment is for 0.45 and the plaintiff pays costs of 0.10 to the defendant, then $G = 0.35$. The inclusion of fee shifting costs reflects that imposition of fee shifting not only affects settlement negotiations, but also affects the amount that the plaintiff must pay to the defendant at trial. It is reasonable to view the difference between the expected value of G and the value of q as a measure of accuracy, but the question remains whether conclusions about accuracy would be robust to alternative specifications.

Outcome expectation The specification chosen focuses on the expectation of settlement or trial results, rather than on the actual result in particular cases. It is a comparison of the expectation of the result with the true merits, not a measure of the error. If, for example, there are two scenarios in which the correct result based on the true merits would be for the defendant to pay the plaintiff 0.50, and in one scenario the defendant pays 0 and in the other scenario the defendant pays 1, then this measure would count the legal system as perfectly accurate. Because the parties are risk-neutral, they would be indifferent between receiving perfectly accurate results and results that are correct on average.

An alternative measure, which arguably would be more appropriate at least for risk-averse parties, would aggregate the distance between the actual outcome and the ideal outcome in each case. In more technical terms, instead of calculating a measure of inaccuracy that is a function of E_t , the authors might have calculated a measure of expected inaccuracy in which the inaccuracy is calculated within each case rather than based on an average across cases. Easier said than done, of course. In the analytical model, this would require moving a minus q term and a squared term within the double integrals in the current E_t definition.

Accounting for costs The accuracy measure also ignores the pre-fee shifting costs that the parties pay. Dari-Mattiacci and Saraceno note “that the plaintiff receives $G - \frac{c}{2}$ and the defendant pays $G + \frac{c}{2}$.” Imagine a case with very high costs and no fee shifting, where each party spends a million dollars and the court arrives at precisely the correct conclusion that the defendant owes the plaintiff 50 cents. From this definition’s perspective, this outcome counts as a perfectly accurate result. That is a plausible definition of accuracy, but one that offers no comfort to the parties. An alternative definition of accuracy would consider any amounts actually spent at trial, for example counting the outcome from the plaintiff’s perspective as $G - \frac{c}{2}$. A similar definition could measure accuracy from the defendant’s perspective. Either of these two approaches captures three distinct aspects of costs: (1) the costs impact settlement negotiations; (2) when trial occurs, the costs are deadweight losses to society at large; and (3) costs may reduce (or perhaps in some cases increase) the accuracy of adjudication viewed

as a black box from the perspective of each individual litigant. Although it is reasonable for Dari-Mattiacci and Saraceno to define accuracy entirely independently of cost, an interesting question is whether any conclusions based on this definition will extend to definitions that incorporate costs.

Squared accuracy Finally, one might quibble about the use of a squared term rather than an absolute value. Of course, it is conventional to measure (in)accuracy using the ℓ_2 norm rather than the ℓ_1 norm. The convention reflects the dominance of ordinary least squares regression over least absolute deviation regression, but that dominance stems as least in part from the greater tractability of the former. Portnoy and Koenker (1997) [[portnoykoenker](#)] note that computational power mitigates this advantage, and that an advantage of the ℓ_1 norm is that it is more robust to outliers. Because the Dari-Mattiacci and Saraceno definition compares the outcome expectation with q , any results on accuracy necessarily extend to the ℓ_1 norm. The computational results here will be reported using the ℓ_1 norm, because interpretation is more intuitive, because this will make it more straightforward to compare different accuracy measures, and because this will not change any results about when accuracy rises or falls.

3 Litigation as an Extensive Form Game

This section describes an algorithm that can be used to find equilibria in any extensive form game. It illustrates that this approach can come close to replicating the simpler Dari-Mattiacci and Saraceno results by applying the algorithm to a model in which each party selects a slope for its bid function and a truncation. The more flexible approach, however, relaxes the assumption of piecewise linearity by discretizing each player’s signals and offers.

3.1 The von Stengel, van den Elzen, and Talman Algorithm

In principle, any two-player extensive form game with a finite number of strategies can be presented in strategic form, with the parties’ payoffs embedded in a matrix. The row player may use a pure strategy and select a single row from the matrix or a mixed strategy, a probability distribution over the matrix rows. The column player analogously chooses a column or a probability distribution over matrix columns. Nash (1951) famously proved that every strategic game involving a finite number of players has at least one equilibrium, now called a Nash equilibrium, in which neither player could increase the player’s payoff by switching to a different strategy given the opponent’s strategy. Moreover, given a game in strategic form, any equilibrium in pure strategies can be identified relatively easily by the algorithm of iterative elimination of dominated strategies. If a cell of the matrix exists where the column player’s utility is greater than in any other cell in its row and the row player’s utility is greater than in any other cell in its column, then that cell represents a Nash equilibrium.

This approach is insufficient to find mixed strategy equilibria. The problem of finding Nash equilibria, however, can be converted into a linear programming problem. The reader with a strong interest in the algorithm applied here should read von Stengel (2002) for a comprehensive overview, including proofs that it produces equilibria. The reader interested solely in legal ramifications may skip this section altogether. A virtue of using computational game theory to identify equilibria is that one may choose to treat algorithms as black boxes, particularly because all equilibria found in this article were computationally verified.

Still, we will provide an introduction to the algorithm for the reader with an intermediate level of interest. Let the payoffs for players 1 and 2, respectively, be represented by the $M \times N$ matrices A and B , whose entries are positive; any game with negative payoffs can be scaled to an equivalent such game. Player 1's strategy can thus be represented by a vector $x \in \mathbb{R}^M$ whose components sum to 1, and Player 2's strategy can be represented by a vector $y \in \mathbb{R}^N$ whose components have the same property. Using matrix notation, if $E = [1, \dots, 1] \in \mathbb{R}^{1 \times M}$ and $F = [1, \dots, 1] \in \mathbb{R}^{1 \times N}$, then $Ex = 1$, $Fy = 1$, and $x, y > 0$. Player 1's strategy is called a "best response" to a strategy y if it maximizes the expected payoff $x^T Ay$ subject to the $Ex = 1$ constraint, and similarly a strategy y maximizing $x^T By$ subject to the $Fy = 1$ constraint is a best response to x . The strategy pair (x, y) forms a Nash equilibrium if each is a best response to the other. To find a Nash equilibrium, one can take advantage of a linear programming principle known as duality. Under strong duality, if there is an optimal solution to a "primal LP," then a dual LP has the same optimal solution. In the dual LP, the variables and constraints are reversed and the objective (maximization or minimization) is inverted, relative to the primal LP. Thus, the primal problem of maximizing $x^T Ay$ subject to $Ex = e$ (where $e = 1$) has a dual problem of finding u that minimizes $e^T u$ subject to $E^T u - Ay \geq 0$. A similar dual minimization problem can be used to represent player 2's constrained maximization of $x^T By$.

In von Stengel (2002), Theorem 2.4 shows that (x, y) form a Nash equilibrium if and only if, for some u and v , all of the following hold:

$$\begin{aligned} x^T (E^T u - Ay) &= 0 \\ y^T (F^T v - B^T x) &= 0 \\ Ex &= e \\ Fy &= f \\ E^T u - Ay &\geq 0 \\ F^T v - B^T x &\geq 0 \\ x, y &\geq 0 \end{aligned}$$

These conditions collectively define a mixed linear complementarity problem (LCP). The word "complementarity" refers to the fact that because x , y , A , and B are nonnegative, the first line implies that x and $E^T u - Ay$ cannot have a positive component in the same position. The same holds for y and $F^T v - B^T x$.

The best known algorithm for finding a solution to a linear complementarity problem is that of Lemke and Howson (1964) [26].

Lemke and Howson is inadequate to the settlement bargaining problem, however, because the bimatrix game will be too large. When an extensive form game is converted to a matrix, the matrix must contain a separate row for all permutations of the choices that the row player may make at each information set, and similarly a column for each permutation of the column player’s information set choices. For example, if each player has 10 different information sets (perhaps corresponding to 10 different possible signals), and may make 10 different moves at each information set, then $10(10)$ rows or columns would be needed for that player. Several different scholars, the earliest being Romanovskii (1962) [34], offer a way around this dilemma. The trick, as explained by Koller, Megiddo, and von Stengel (1996) [24], is to craft a matrix in which the rows or columns represent not a player’s pure strategies, but instead the sequences that the player may play at any stage of the game. In the above example, there would be 101 sequences per player (10 for each information set plus an empty sequence). If we inserted an opportunity by each player to quit after receiving its signal, then there would be 121 sequences per player (the above, plus decisions to quit or not at each of the 10 signals). Some sequences may be incompatible, because it may be impossible for a player to play a certain information set given a move by another; thus, the relevant matrix includes a zero in such cells.

The von Stengel, van den Elzen, and Talman (2002) algorithm adapts the “sequence form” approach by combining techniques used in von Stengel (1996) with developments in van den Elzen and Talman (1999), which adapts the Lemke-Howson algorithm to ensure perfect equilibria. The result is an algorithm that produces perfect, Nash equilibria. The requirement of perfection would not be needed if players committed to strategies at the outset of the game. But in litigation, such commitments do not occur. Thus, even if an equilibrium meets the Nash criterion, meaning that neither player will have an incentive at the outset of the game to change its strategy choice given the opponent’s strategy choice, it may be imperfect, if a player might have an incentive to deviate in its strategy choice after learning new information. The solution concept of perfect Bayesian equilibrium, a term in use since at least the 1960s, has various formal definitions in the literature, including that of Fudenberg and Tirole (1991) [10], and is the imperfect information analogue of the subgame perfection equilibrium concept proposed by Selten (1965) [36].

The algorithm is initialized by setting every information set to a fully mixed behavior strategy, that is a strategy in which each action has some positive probability of being played. A data structure called a tableau is initialized in turn based on these values. This tableau encodes not only action probabilities, but also the relationships among information sets, specifically which information sets may follow other information sets in sequence. The initialization is designed so that the equations reflected by the tableau reflect a “basic feasible solution” to some of the equations in the linear complementarity problem, albeit ordinarily not a solution that corresponds to a perfect Nash equilibrium. The algorithm then proceeds through a series of pivoting steps, which represent linear algebraic

manipulations in which variables are added to or removed from the list of non-zero variables known as the basis. When the variable initially added to the basis eventually leaves the basis, a Nash equilibrium is guaranteed.

The algorithm is not without its limitations. First, the algorithm is not guaranteed to produce all Nash equilibria. One may attempt to obtain multiple equilibria by choosing different initializations of the information sets. Second, to ensure that an equilibrium is reached, the algorithm must be performed using exact arithmetic with rational numbers, rather than floating point arithmetic. Even given the generous amount of storage allowed by modern 64-bit computer architectures, floating point operations involve rounding, and that may prevent the algorithm from converging. The use of exact arithmetic is cumbersome, however, given the frequent need to calculate greatest common factors of integers with many digits. Third, the algorithm is roughly linear in the size of the game tree, not in the size of the set of player sequences. The game tree grows exponentially as more moves are added to the game.

3.2 A Near Replication

To replicate Dari-Mattiacci and Saraceno as closely as possible, we must restrict each player’s bid function to consist of piecewise linear segments. Yet, there are an infinite number of possible combinations of piecewise linear segments, and so we must adopt some restrictions. In theory, we could, for example, require each player to choose one of 100 starting values and one of 100 ending values for each of 10 signal ranges; for example, plaintiff might bid 0.47 at $z = 0.5$ and 0.92 at $z = 0.6$, representing one such segment. But, for each player, this would produce $(100 \times 100)^{10}$ possible strategies, and the game tree would be the square of this in size.

For this approach to be feasible, the strategies must be severely constrained. For this preliminary exercise, we thus allow each of the plaintiff and defendant (denoted P and D) to choose an offer \mathcal{B} line with a slope m in $\{\frac{1}{3}, \frac{1}{2}, \frac{2}{3}, 1\}$ and a minimum value of ϑ in $\{0.02, 0.06, 0.10, \dots, 0.98\}$. Because truncations are critical to the Dari-Mattiacci and Saraceno equilibria, each player also chooses a portion t of this line to truncate in $\{0, \frac{1}{9}, \frac{2}{9}, \dots, 1\}$. Thus, $\mathcal{B}_P = \max(m_P \cdot z + \vartheta_P, t_P)$ and $\mathcal{B}_D = \min(m_D \cdot z + \vartheta_D, t_D)$. Each player may in effect choose from $25 \times 4 \times 10 = 1,000$ strategies, so the game tree consists of 1,000,000 final nodes. At each of these nodes, for each of 10,000 combinations of the players’ signals, the players’ utilities are calculated by determining based on the corresponding strategies, whether the case settles and the resulting outcome of the game.

Figure 1 illustrates a result of the algorithm, in this case for parameters $q = 0.4$, $c = 0.2$, and $t = 0.4$. The panels on the left represent the plaintiff’s strategies; on the right, the defendant’s. The results of the computational model, on the bottom, are reasonably similar to the equilibrium for these parameters identified in the analytical model of Dari-Mattiacci and Saraceno, on the top, though hardly an exact match. This result is an exact equilibrium in the game as defined with these restrictions, but it is only an approximate equilibrium in the original Dari-Mattiacci and Saraceno game, which does not impose these

restrictions.

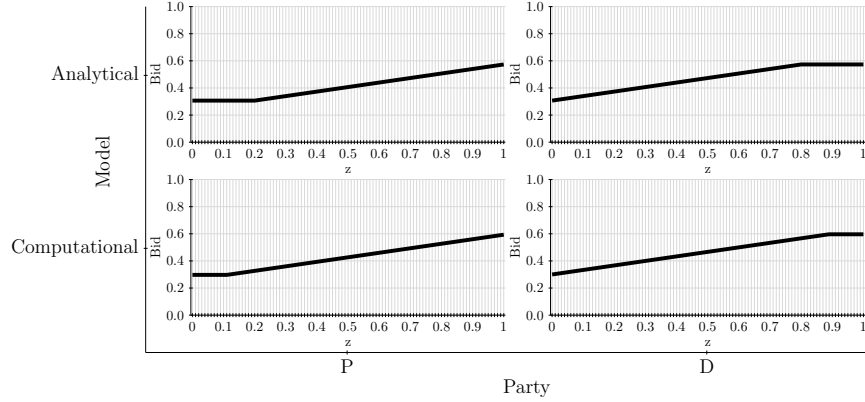


Figure 1: Attempted replication with $q = 0.4$, $c = 0.2$, and $t = 0.4$

Still, by combining a number of such parameter sets, one can perform a task similar to that of Dari-Mattiacci and Saraceno, observing how changes in parameters affect the equilibrium. The results of a number of such observations are available in the Online Supplement to this article in the “DMS Replication” folder, corresponding to parameter values that are multiples of 0.1 that meet Dari-Mattiacci and Saraceno’s assumptions regarding q and c and for which in their analytical model neither party’s strategy consists of more than a single line segment. The results in Figure 1 are typical, though there are some surprises, like Figure 2, in which the algorithm calculates a mixed strategy equilibrium, with the relative darkness of each of the parties’ strategies corresponding to the probability that a party will play them. Though it seems unlikely that any individual would play such a strategy, a mixed strategy can be interpreted as a set of pure strategies that might be played by different litigants, where neither litigant knows the other litigant’s type. Given the significance of mixed strategies in games like chicken that bear some resemblance to settlement bargaining, the ability of the algorithm to find mixed strategies is a strength of the computational approach, but it complicates any attempt at replication.

Although these results illustrate that the algorithm can approximately replicate some analytical results, they may be disappointing. The most interesting results of Dari-Mattiacci and Saraceno occur in cases in which the strategies are discontinuous, and we have not even attempted to replicate those cases. The reason that the algorithm is so constrained in this preliminary analysis is that it does not fully take advantage of the sequence form. Each party has a limited number of information sets, in many of which it considers a large number of possibilities (such as which of 25 minimum values to select) that in turn dictate the party’s strategy over a range of signals. The result is not much better than could be achieved by applying iterated elimination of dominated strategies to the full bimatrix game. Still, the exercise is sufficiently simple that one might

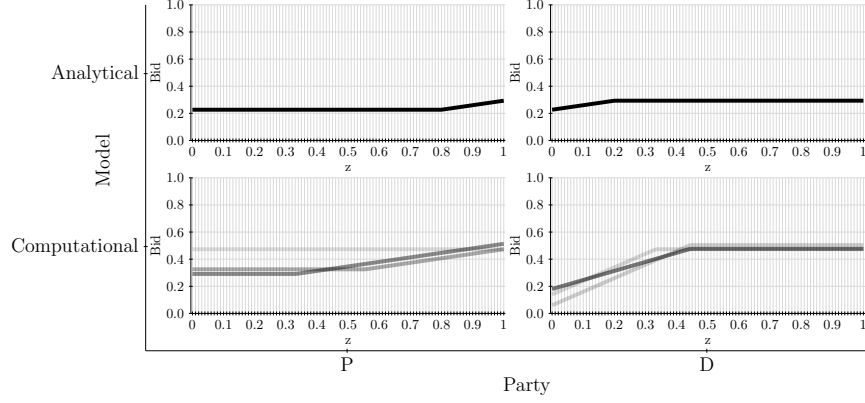


Figure 2: Attempted replication with $q = 0.4$, $c = 0.3$, and $t = 0.6$

wonder why the settlement bargaining literature has eschewed even that well known technique. This approach is more flexible than a simple linear model, and it can be applied to a wide range of game specifications that are analytically intractable. Still, a more fruitful strategy is to discretize the model, so that each player has separate information sets based on the signals that it receives.

3.3 A Discretized Evidentiary Signals Game

In the discretized model, each party receives a discrete signal from a set of n_S possible signals and then must make a discrete offer from n_D available offers. For legibility, Figure 3 illustrates a highly simplified version of this game for $n_S = 2$ and $n_D = 2$, but we will generally use $n_S = 10$ and $n_D = 10$, which produces a game tree consisting of 11,111 nodes. The circles identify the players (Chance, Plaintiff, or Defendant), as well as information set numbers. The plaintiff does not observe the signal received by the defendant and vice-versa. Thus, for example, at each of the four points in the tree labeled as “D1,” the defendant has the same information set, in which it has received the signal 1 (corresponding to $z_\Delta = 0.25$) instead of the signal 2 (corresponding to $z_\Delta = 0.75$). Thus, the defendant must assign the same move probabilities to its two alternative offers (corresponding to 0.25 and 0.75) at each of these points. The diagram illustrates an equilibrium identified by the algorithm for this simple game, in which each player is always aggressive, with the plaintiff demanding 0.75 and the defendant offering only 0.25. Under the Chatterjee and Samuelson bargaining protocol, this does not result in a settlement.

The strategies permitted with this approach are in some ways more constrained and in some ways less constrained than the strategies in Dari-Mattiacci and Saraceno. The analytical model restricts the litigants from playing nonlinear strategies between the points that produce discontinuities in their preferred equilibria. The computational model may allow for nonlinearities between the critical points, but only at the 10 discrete signal values. A further restriction of

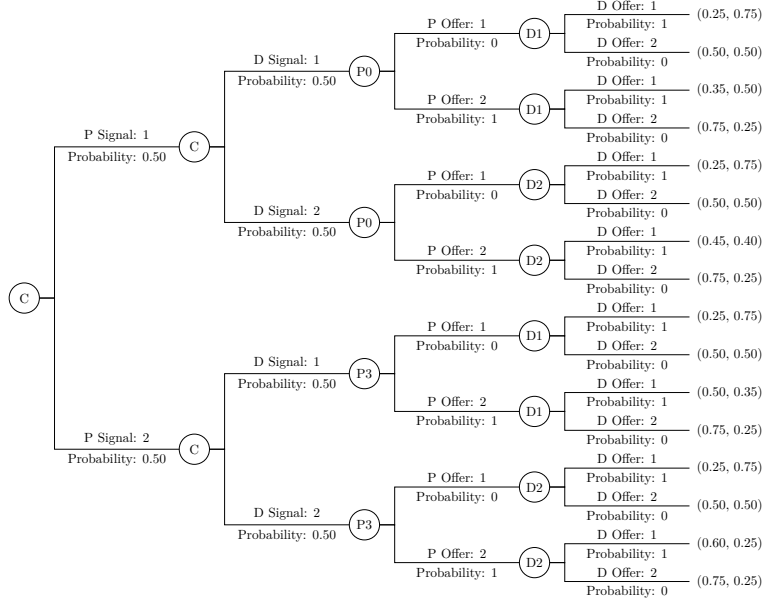


Figure 3: The game tree with $n_S = 2$ and $n_B = 2$

the computational model is that offers can be made only at 10 discrete points. Thus, analytical and computational approaches both impose admittedly arbitrary constraints on what strategies are permissible. This is, of course, not unusual, and many articles in the settlement bargaining literature impose tighter constraints, such as in models where a party receives one of two signals instead of one of ten.

An important question is whether these differing restrictions significantly affect the equilibria. One can reason that these assumptions do not seem critical to game dynamics, but it is difficult to be sure. There is no way to test whether the piecewise linearity assumption is driving the analytical model's key conclusions without some other model that relaxes those assumptions. Fortunately, the computational model can be such a model. Later, we will use the computational model to calculate equilibria employing the discretization assumption. But we can also use computation to measure the extent to which, for any given parameters, a discretized version of the Dari-Mattiacci and Saraceno model deviates from equilibrium.

This assessment requires two steps. First, we calculate each player's offer value at the discrete signal values, and then identify the discrete offer value nearest that value. This produces discretized versions of the bid functions derived by Dari-Mattiacci and Saraceno. Second, we determine each player's best response to the other player's strategy using an algorithm detailed in [CITE] and use these to calculate the exploitability of those strategies. That is, let σ_p

represent the strategy of player p at the discretized equilibrium, let σ_{-p} represent the strategy of that player's opponent, and let $\mathcal{U}_p(\sigma_p, \sigma_{-p})$ represent the expected return of player p given these strategies. Now, let $\sigma_{\bar{p}}$ represent the best response of player p to σ_{-p} . We then define exploitability of player $-p$'s strategy $\mathcal{E}_{-p} = \mathcal{U}_p(\sigma_{\bar{p}}, \sigma_{-p}) - \mathcal{U}_p(\sigma_p, \sigma_{-p})$. Allowing P and D to represent, respectively, the plaintiff and defendant in the discretization of the Dari-Mattiacci and Saraceno equilibrium, we define overall exploitability $\mathcal{E} = \frac{\mathcal{E}_P + \mathcal{E}_D}{2}$, i.e. the average of the amounts that each player can improve its score by changing its strategy holding the other's strategy constant. By definition, at a Nash equilibrium, overall exploitability is zero, and so \mathcal{E} provides a measure of how much discretization moves the strategies from Nash equilibrium.

Figure 4 illustrates the results. The outer horizontal c axis represents trial cost. The outer vertical p axis represents litigation quality. All of the later simulations will be executed with $q \in \{\frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}\}$, but we omit the row for $q = \frac{2}{3}$ where it is identical or symmetric to the row for $q = \frac{1}{3}$, and we omit the extreme values of q because they are outside the bounds assumed by Dari-Mattiacci and Saraceno. (Diagrams including these rows, both for this and for other truncated diagrams produced later, are available in the Online Repository.) Each mini-graph represents the measurement of exploitability \mathcal{E} for each value of the fee-shifting threshold t in $\{0, 0.01, \dots, 1\}$. We can see that for $c \leq \frac{1}{8}$, exploitability is very close to zero; overall all of these cases, the average \mathcal{E} value is 0.00047. This indicates that the discretization changes the strategic dynamics of the game very little. With higher trial costs, considerably higher exploitability values (up to 0.067) can be seen in Figure 4. This is likely due to Dari-Mattiacci and Saraceno's implicit assumption, noted earlier, that c is not too high.

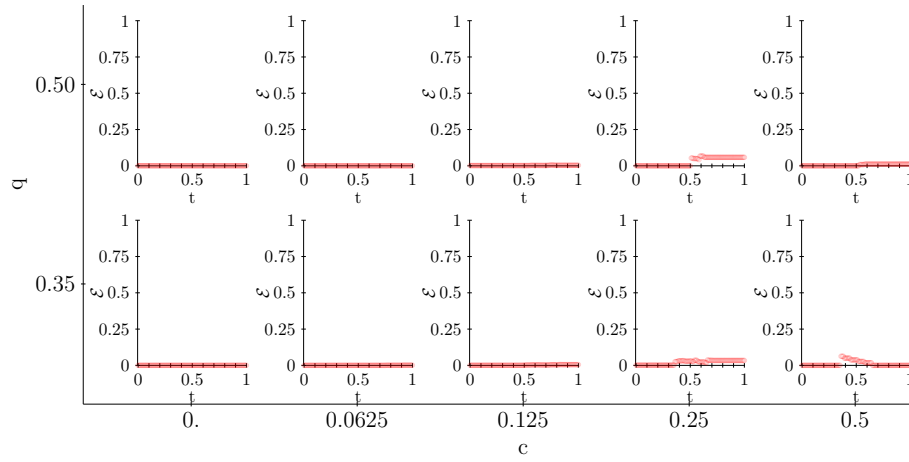


Figure 4: Exploitability of discretized Dari-Mattiacci and Saraceno strategies

Using the discretization of the Dari-Mattiacci and Saraceno strategies, we can calculate outcome variables. Results on accuracy are illustrated in Figure

5. The accuracy A is defined analogously to Dari-Mattiacci and Saraceno, i.e. by comparing each party's expected outcome (taking into account settlements and judgments, including fee shifting, and not taking into account each party's own legal costs) to the true merits q , though for easier visual interpretability, the absolute value instead of the square of the difference is shown. This reflects, as Dari-Mattiacci and Saraceno prove, that for relatively low costs, the English rule produces at least as good accuracy (relative low A) in comparison to the American rule. This does not reflect that the reverse is true with relatively high costs, again presumably due to their implicit assumption governing costs. Note also that accuracy is very high when $q = \frac{1}{2}$. This does not indicate that each case produces the correct result, but that the Dari-Mattiacci and Saraceno accuracy measure is based on averages across cases, and with $q = \frac{1}{2}$, the cases in which the plaintiff receives too much exactly balance those in which the plaintiff receives too little.

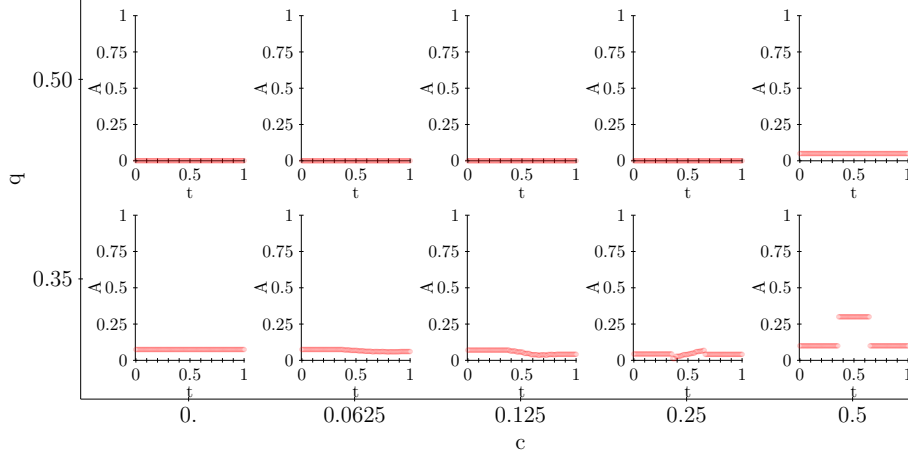


Figure 5: Accuracy with discretized Dari-Mattiacci and Saraceno strategies

Figure 6, meanwhile, illustrates the analytical model's conclusions regarding the trial rate L . As Dari-Mattiacci and Saraceno show, and as is reflected here at least with $c \leq \frac{1}{8}$, trial rates are invariant to the degree of fee shifting and the quality of the true merits, though small wobbles are observable here as a result of discretization. Also as expected, trial rates fall as the cost of trial rises. Note that trial rates are relatively high, perhaps in part because of the assumption of risk neutrality.

The accuracy and trial results, it is worth emphasizing, are simply calculations based on discretizations of Dari-Mattiacci and Saraceno's constructed equilibria, not equilibria calculated by application of the von Stengel et al. algorithm. These measurements, however, provide a baseline that we can compare to results from equilibria calculated by that algorithm. We will thus be able to assess whether exact, perfect equilibria calculated by the algorithm produce sim-

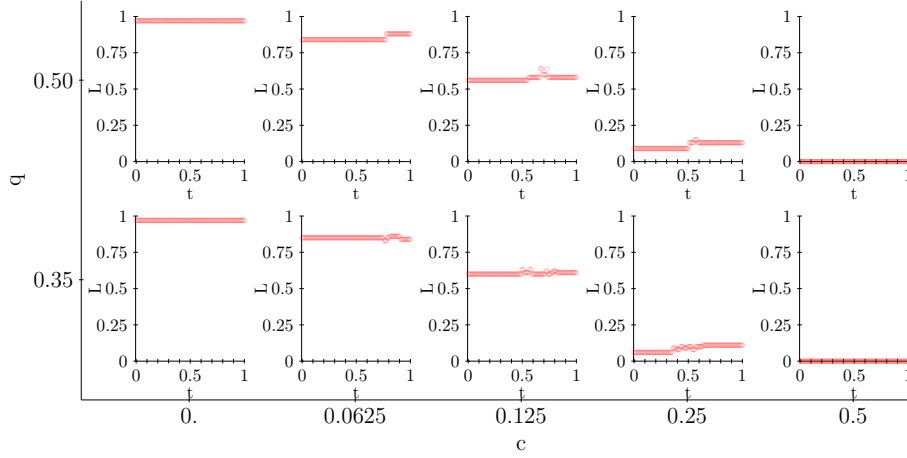


Figure 6: Trial with discretized Dari-Mattiaci and Saraceno strategies

ilar results in the discretized version of the original additive evidence game and in variations generated by relaxing various assumptions of the Dari-Mattiaci and Saraceno model.

3.4 Equilibria for a single set of parameters

Which equilibrium of a game the von Stengel et al. algorithm identifies may depend on the initialization of the information sets. We ran the algorithm 5,000 times for a single set of parameter values representing the middle of the values we are exploring each, i.e. $t = \frac{1}{2}, q = \frac{1}{2}, c = \frac{1}{8}$, and identified 194 different equilibria. Figure 7 illustrates for each player an average strategy in which each equilibrium strategy is played with equal probability, without regard to whether the player's opponent chooses the corresponding equilibrium strategy. For example, when receiving a signal corresponding to $z_{\Pi} = 0.05$, the plaintiff makes an offer $\mathcal{B}_P = 0.35$ approximately 8% of the time.

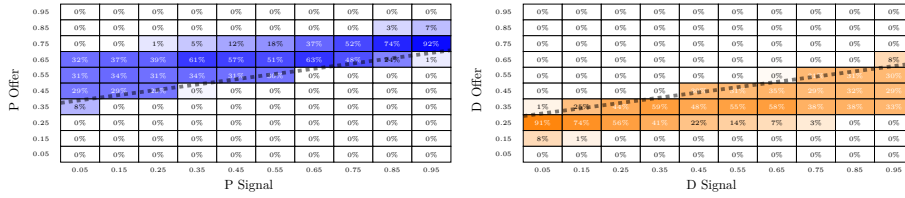


Figure 7: Average strategies across 194 equilibria for $t = \frac{1}{2}, q = \frac{1}{2}, c = \frac{1}{8}$

Although there are a large number of equilibria, they cluster near the equilibrium derived by Dari-Mattiaci and Saraceno, represented by dotted lines, though some of the equilibria reflect somewhat more aggressive game play. Some

of the equilibria differ from others in ways that have no bearing on the game outcome. For example, if plaintiff chooses $\mathcal{B}_P \geq 0.75$ for some signal, settlement will never occur, and so an offer of 0.75 is functionally equivalent to an offer of 0.85. Of the 194 equilibria identified, 106 are pure strategy equilibria while 88 reflect mixed strategies, such as one in which given a particular signal, a party made one offer with exact probability $\frac{4,164}{15,277}$ and another with probability $\frac{11,113}{15,227}$. Correlated strategies over the 194 equilibria would by definition also be exact Nash equilibria. The average strategy in 7 is not an exact equilibria, but is close, with exploitability (E) = 0.0027.

It will not be practical to seek out 5,000 equilibria for each parameter value for each variation of the model in the remainder of the paper, as this exercise took nearly three and a half hours of computer time. But our later results will continue to reflect initialization of each information set to a random determination of a fully mixed strategy for each player, and we will run the algorithm for every value of $t \in 0, 0.01, \dots, 1$. It will thus be possible to identify different families of equilibria that arise for very close values of t .

4 Results

This section first identifies multiple equilibria with a single set of parameters. It then aggregates results from the additive evidence game as defined by Dari-Mattiacci and Saraceno over many different sets of parameters. Finally, it progressively relaxes assumptions of their model that are identified above.

4.1 Aggregating results over parameter values

Figures 5 and 6 calculated accuracy and trial results using the strategies constructed by Dari-Mattiacci and Saraceno. Figures 8 and 9 show the analogous results calculating discretized equilibria with the von Stengel et al. algorithm. This is a snippet of the aggregated results. For each of the 1,515 permutations of parameter values executed under these assumptions, the Online Repository includes files with the exact equilibria calculated, heatmaps illustrating the parties' strategies, reports elaborating various outcome variables, and logs of program execution. Such files are also available for each permutation for each of the other assumptions to be explored later in this article.

These results are generally close to the results corresponding to Dari-Mattiacci and Saraceno's constructed equilibria, at least for $c \leq \frac{1}{8}$. The shapes of the accuracy curves are very similar, reflecting that their results concerning the better accuracy of the English rule with low trial costs extend to the computational model. Trial rates in the computational model are a bit higher, and multiple equilibria are apparent. The most notable difference is that with sufficiently high costs, the English rule appears to increase trial rates. For example, with $q = \frac{1}{3}$ and $c = \frac{1}{4}$, some equilibria with around half of cases being litigated exist at all fee shifting levels, but equilibria with much lower trial rates are apparent only with relatively low levels of fee shifting. We must, however, acknowledge

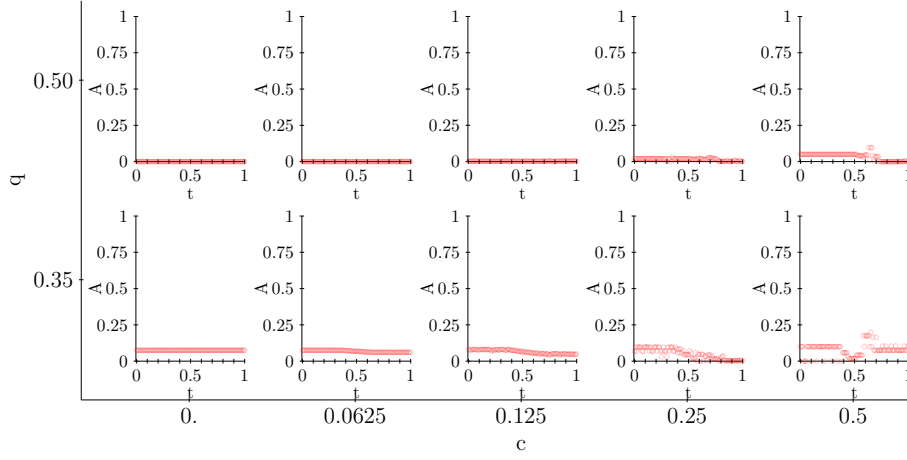


Figure 8: Accuracy with calculated equilibria

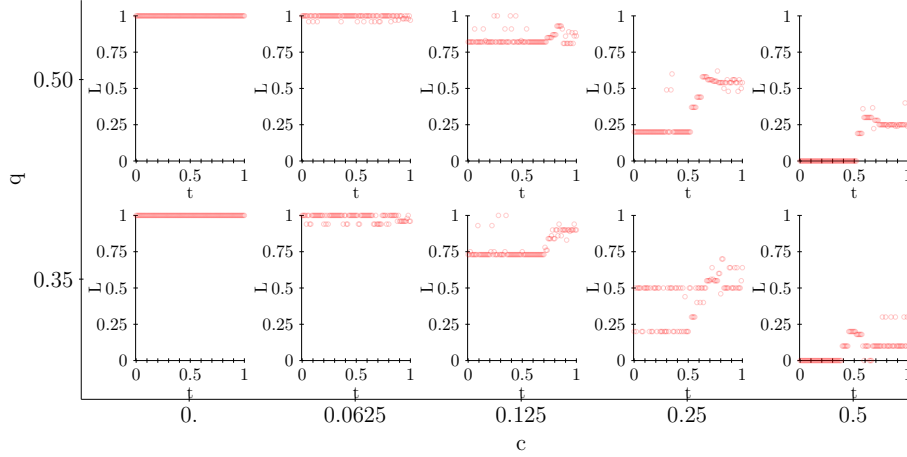


Figure 9: Trial rates with calculated equilibria

a limitation of the computational model here. The absence of evidence of equilibria with lower trial rates does not prove that no such exact equilibria exist, only that the algorithm did not identify them over a range of fee shifting values with different randomized initial conditions.

4.2 Relaxing assumptions

The analysis so far has endeavored to track Dari-Mattiacci and Saraceno as closely as possible. Use of the computational model, which discretizes signals and allows for different bid for each signal, reinforces their central results while clarifying the importance of their implicit assumption regarding trial costs. The

computational model’s potential lies not so much in replication, but in changing critical assumptions. In Section 2.3, we identified a number of restrictive assumptions in Dari-Mattiacci and Saraceno’s model. Though entirely justifiable by needs of tractability, these assumptions either seem unrealistic or, as, for example, with the focus on damages rather than liability disputes, apply to some but not all cases. In this subsection, we will identify how various assumptions may be relaxed.

The model in Figures 8 and 9 already relaxes three assumptions: balanced asymmetric information, low or moderate cost, and piecewise linearity. Although only q values permitted by the Dari-Mattiacci and Saraceno model are shown, a broader range of such values is included in the Online Repository. Figures 8 and ?? illustrate a broader range of costs than Dari-Mattiacci and Saraceno’s implicit costs assumption allows, and discretizing signals and bids inherently relaxes piecewise linearity.

With the computational model, we can also replace the assumption of asymmetric information quality equivalence with an assumption of equal information strength. (The Online Repository also includes a model in which the plaintiff has one-fourth of the available private information.) This assumption avoids the conflation of the true merits with the degree of information asymmetry. The judgment is thus computed as follows:

$$J(z_{\Pi}, z_{\Delta}) = \frac{1}{2}q + \frac{1}{4}(z_{\Pi} + z_{\Delta}) \quad (2)$$

As in Equation 1, half of the judgment depends on q , but in Equation 2, the remainder of the judgment no longer weights the parties’ private signals by q and $1 - q$.

With the change embodied in Equation `redefinition:equalInfo`, it also becomes straightforward to remove the assumption that signals are independent of the true merits with an assumption that the signals form part of the merits. That is, q now represents shared information about half of the merits, and J in Equation 2 now represents the merits. In other words, the normalized information signal that each party has about how the court will rule is now evidence of the merits rather than just information that may sway the court even though it has nothing to do with the merits. Because this is a change only in interpretation, it does not require running of any additional models. It merely requires a different baseline for calculating accuracy. (We will consider other changes to the accuracy definition below.) A possible limitation of this assumption is that it ignores that trial judgments themselves may reflect noise or bias, but it seems more plausible that parties would have asymmetric information about noise rather than asymmetric information about the merits. We leave the task of modeling judicial error and shared information about judicial bias to future work.

The computational model also makes it straightforward to change the fee-shifting rule. First, we consider the margin-of-victory fee shifting for which Dari-Mattiacci and Saraceno calculate equilibria but do not offer results regarding accuracy. Second, we also consider what we will call “ordinary fee shifting.”

With this approach, the party that wins more than half of the judgment receives the benefit of fee-shifting, but the proportion of the winning party’s legal fees paid by the other party is multiplied by t . Note that t serves a different function in the original Dari-Mattiacci and Saraceno definition and these two alternatives.

The computational model also enables easy toggling for a damages (or asset allocation) issue, as in Dari-Mattiacci and Saraceno, and a liability issue. This requires only changing the pre-fee-shifting payoffs so that in cases in which a plaintiff wins a judgment for greater than $\frac{1}{2}$, the plaintiff receives a payoff of 1, and in all other cases, the plaintiff wins 0. (This reflects an assumption that the plaintiff bears the burden of proof, an issue that is irrelevant in Dari-Mattiacci and Saraceno’s model, where the probability that $J = \frac{1}{2}$ is zero.)

A more challenging task for the computational model is to allow the parties to decide sequentially whether to file and answer. If the plaintiff does not file, the plaintiff receives 0, and if the plaintiff files but the defendant does not answer, the plaintiff receives 1, but either avoids the payment of trial costs. This redefinition changes the game tree. Figure 10 shows a portion of the game tree in the simplified game tree in which $n_S = 2$ and $n_B = 2$. Note that something like this branch appears in the game tree for each combination of plaintiff and defendant signal. We also make a further change, allowing $n_B = 12$, with the highest and lowest bids defined, respectively, to correspond to situations in which the plaintiff or the defendant refuses to negotiate; without this change, a party would never refuse to file or answer, because it could always guarantee a settlement at a slightly more favorable value. With these changes, the game tree contains 16,111 nodes instead of 11,111.

The assumption of risk neutrality can be replaced with an assumption of mild risk aversion by modifying the parties’ payoffs in the game, so that utility is a nonlinear function of wealth. Specifically, let $U = -e^{-\alpha W}$, where the parameter $\alpha = 1$ and W , representing each party’s initial wealth, is set at 10. Of course, these parameters can be changed easily to simulate higher degrees of risk aversion.

Table 1 identifies the various models that we will execute and for which the next section will calculate accuracy results. Model A represents the original model, while Model B adds the assumption of equal information strength. Model C reflects the same equilibrium as Model B but calculating accuracy relative to a baseline that includes the parties’ private information. Models D and E build on these assumptions to test margin-of-victory and ordinary fee shifting, respectively. Models F, G and H return to the same fee-shifting rule as Dari-Mattiacci and Saraceno, incorporating all of the same assumptions as Model C, but with liability at issue instead of damages. In Models G and H, the parties may choose whether to contest litigation, and in Model H only, the parties are subject to mild risk aversion.

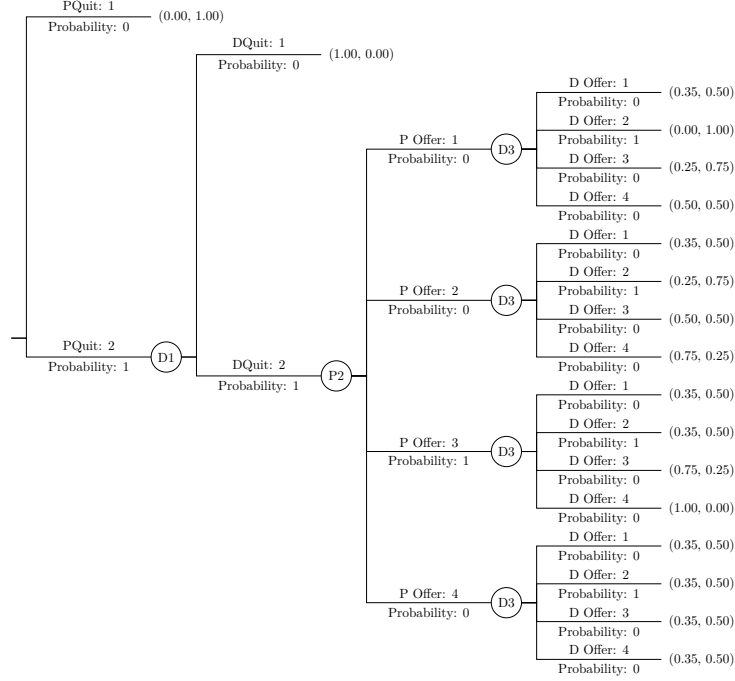


Figure 10: A portion of the simplified game tree with file and answer decisions

4.3 Results

Figure 11 aggregates the results. Each row represents one of the models described above. The (in)accuracy values are averages of the values calculated for each $q \in \{\frac{1}{6}, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}, \frac{5}{6}\}$. Because Models C through H count the parties' signals within the merits, with $c = 0$, because all cases go to trial, there is by definition no inaccuracy. As c rises, however, the English rule *increases* tends to increase inaccuracy in these models instead of decreasing it in these models, revealing the importance of this assumption. With higher costs, the parties' offers will track their shared information on q more closely, but this effect is much more pronounced with the English rule. (This can be observed in diagrams contained in the Online Repository showing average plaintiff and defendant offers.) Results correspondingly are less influenced by the parties' private information. When q represents the true merits, the English rule therefore tends to improve accuracy, but when the parties' private information is about the merits rather than simply representing noise, the English rule tends to increase inaccuracy.

Changing the definition of fee shifting, as represented by Models D and E, does not alter the fundamental pattern. The most significant difference is that inaccuracy increases gradually over the range of t values. With the Dari-Mattiaci and Saraceno definition, given low values of t , fee shifting will be rare,

Assumption	A	B	C	D	E	F	G	H
Broad costs range	✓	✓	✓	✓	✓	✓	✓	✓
Broad quality range	✓	✓	✓	✓	✓	✓	✓	✓
Discretized signals	✓	✓	✓	✓	✓	✓	✓	✓
Equal information strength		✓	✓	✓	✓	✓	✓	✓
Signals form part of merits			✓	✓	✓	✓	✓	✓
Margin-of-victory fee shifting				✓				
Ordinary fee shifting					✓			
Liability issue						✓	✓	✓
Contestation choice							✓	✓
Risk aversion								✓

Table 1: Assumptions used in subsequent models

occurring only where a party wins and its own evidence is very strong. With these other models of fee-shifting, it will be relatively rare for a party to be able to rule out the possibility of fee-shifting.

The increase in inaccuracy with the English holds does not generally manifest when liability is at issue. Indeed, with sufficiently high costs and no risk aversion (in Models F and G), there appears to be a slight reduction in inaccuracy with the British rule. The explanation can be traced in part to a greater incidence of trial in the winner-take-all liability context. With relatively high trial costs, trial rates are higher in Model F than in Model C. Because all trials by definition produce no inaccuracy in these models, the increase in trial rates tends to improve accuracy, all else equal. Risk aversion, meanwhile, drives trial rates back down in Model H, and this tends to increase inaccuracy.

Another important phenomenon in Models G and H is that the parties may decide not to file or not to answer. Unsurprisingly, parties are more likely to give up (a) when q indicates that they have a bad case; (b) when t is relatively high; and (c) when the parties are risk averse. (This is evident in diagrams of quit rates available in the Online Repository.) Decisions not to contest litigation will tend to improve accuracy. Usually, the party that decides not to contest will be the party that would lose, and so, by avoiding a settlement, the accurate result is achieved.

There are many nuances above that the summaries above admittedly do not capture. Interestingly, for example, with $c = \frac{1}{16}$ and $q = \frac{1}{6}$, the plaintiff in Model G will decide not to file with a probability of approximately 0.7, yet the plaintiff will always file with $c = \frac{1}{8}$ and $q = \frac{1}{6}$ if $t < \frac{1}{2}$. When the plaintiff does file with $c = \frac{1}{16}$, the case always goes to trial, but with $c = \frac{1}{8}$, the case will sometimes be settled. One can see here the intricate dance between the file decision and the parties decisions whether to contest litigation. Because higher costs make settlement more likely, they may lead a party to be more willing to contest litigation in the first place. Once t is sufficiently large, the plaintiff will quit at greater rates, because it recognizes that the defendant will have greater resolve in settlement negotiations. Granted, even this is a simplified story. In an

article of this length, we cannot account for every strategic nuance revealed by the computational results, and any textual explanation will not fully describe the dynamics of changes in the Nash equilibrium with changes in parameters.

We will conclude by using a different definition of accuracy for all of Models A-H. This definition considers a plaintiff's *net* recovery, taking into account costs and fee-shifting. Suppose, for example, that the court would award a judgment of 0.65 but the plaintiff receives only 0.55, either as a settlement or on net at trial after each party pays its fees and fee-shifting is resolved. Under this definition of inaccuracy, which focuses on the danger of undercompensation, this counts as an error of 0.10. If the plaintiff receives more than the court would award at trial, that will count as 0 inaccuracy, but it would lead to an error in a similar measure of accuracy focusing on excessive payments by defendants and thus on concerns of overdeterrence. Figure 12 shows the plaintiff-centered measure, but the defendant-centered measure would show a mirror of these effects, given that the game is entirely symmetrical.

A comparison of Figures 11 and 12 illustrates how different definitions of accuracy can produce considerably different results. First, inaccuracy levels generally appear higher with the undercompensation inaccuracy measure, highlighting that very high levels of inaccuracy on our earlier measures may correspond to situations in which some plaintiffs receive way too much while others receive way too little. Even with risk neutrality, which measure is more relevant to social welfare may depend on factors such as whether parties might be able to predict the errors before they engage in actions with the potential to lead to litigation. Second, Model A, the computational version of Dari-Mattiacci and Saraceno's original model, demonstrates improvements in accuracy with greater fee shifting in Figure 11 but not in 12. Third, Models F, G and H, corresponding to a winner-take-all liability issue, illustrate that the danger of undercompensation inaccuracy becomes much greater with high costs. The English rule appears to reduce undercompensation inaccuracy in these models, while raising undercompensation inaccuracy in other models, but whatever the merits of the fee-shifting rule, Figure 12 illustrates a far more malign view of the implications of high trial costs than Figure 11.

5 Conclusion

As a broad generalization, the results here reinforce that many of the conclusions of Dari-Mattiacci and Saraceno's elegant model are robust to alternative specifications. Greater trial costs increase settlement rates, but have relatively limited effects on the values at which settlements are struck. In addition, at least with relatively low trial costs, the English rule tends to improve accuracy, measured as an average across all cases. These conclusions do not appear to be a result of assumptions such as balanced asymmetric information, asymmetric information quality equivalence, or the independence of signals from the true merits. These conclusions suggest that continued development of these models may be fruitful. At the same time, the computational model indicates

References

- [1] Lucian Arye Bebchuk and Howard F. Chang. “An Analysis of Fee Shifting Based on the Margin of Victory: On Frivolous Suits, Meritorious Suits, and the Role of Rule 11”. In: *Journal of Legal Studies* 25 (1996), pp. 371–403.
- [2] Lucian Ayre Bebchuk. “Litigation and Settlement Under Imperfect Information”. In: *RAND Journal of Economics* 15 (1984), pp. 404–415.
- [3] Kalyan Chatterjee and William Samuelson. “Comparison of Arbitration Procedures: Models with Complete and Incomplete Information”. In: *IEEE Transactions on Systems, Man and Cybernetics* 31 (1983), pp. 835–851.
- [4] Don L. Coursey and L.R. Stanley. “Pretrial Bargaining Behavior with the Shadow of the Law: Theory and Experimental Evidence”. In: *International Review of Law and Economics* 8 (1988), pp. 161–179.
- [5] Giuseppe Dari-Mattiaci and Margherita Saraceno. “Fee Shifting and Accuracy in Adjudication”. In: *International Review of Law and Economics* 63 (2020), pp. 1–15.
- [6] Andrew F. Daughety and Jennifer F. Reinganum. “Endogenous Sequencing in Models of Settlement”. In: *Journal of Law, Economics and Organization* 9 (1993), pp. 314–348.
- [7] Andrew F. Daughety and Jennifer F. Reinganum. “Settlement”. In: *Procedural Law and Economics*. Ed. by Chris William Sanchirico. 2012, pp. 386–471.
- [8] Andrew F. Daughety and Jennifer F. Reinganum. “Settlement Negotiations with Two-Sided Asymmetric Information: Model Duality, Information Distribution, and Efficiency”. In: *International Review of Law and Economics* 14 (1994), pp. 283–298.
- [9] Daniel Friedman and Donald Wittman. “Litigation with Symmetric Bargaining and Two-Sided Incomplete Information”. In: *Journal of Law, Economics, and Organization* 23 (2006), pp. 98–126.
- [10] Drew Fudenberg and Jean Tirole. “Perfect Bayesian Equilibrium and Sequential Equilibrium”. In: *Journal of Economic Theory* 53.2 (1991), pp. 236–260.
- [11] Jonah P. Gelbach. “The Reduced Form of Litigation Models and the Plaintiff’s Win Rate”. In: *Journal of Law and Economics* 61 (2018), pp. 125–157.
- [12] John P. Gould. “The Economics of Legal Conflicts”. In: *Journal of Legal Studies* 2 (1973), pp. 279–300.
- [13] Chris Guthrie. “Better Settle Than Sorry: The Regret Aversion Theory of Litigation Behavior”. In: *University of Illinois Law Review* (1999), pp. 43–90.

- [14] John C Harsanyi. “Games with incomplete information played by ‘Bayesian’ players, I–III Part I. The basic model”. In: *Management science* 14.3 (1967), pp. 159–182.
- [15] John C. Hause. “Indemnity, Settlement, and Litigation, or I’ll Be Suing You”. In: *Journal of Legal Studies* 18 (1989), pp. 157–179.
- [16] Peter H. Huang. “Lawsuit Abandonment Options in Possibly Frivolous Litigation Games”. In: *Review of Litigation* 23 (2004), pp. 47–113.
- [17] William H.J. Hubbard. “Sinking Costs to Force or Deter Settlement”. In: *Journal of Law, Economics, and Organization* 32 (2016), pp. 545–585.
- [18] James W. Hughes and Edward A. Snyder. “Litigation and Settlement under the English and American Rules: Theory and Evidence”. In: *Journal of Law and Economics* 38 (1995), pp. 225–250.
- [19] Keith N. Hylton. “An Asymmetric-Information Model of Litigation”. In: *International Review of Law and Economics* 22 (2002), pp. 153–175.
- [20] Laura Inglis et al. “Experiments on the Effects of Cost Shifting, Court Costs, and Discovery on the Efficient Settlement of Tort Claims”. In: *Florida State University Law Review* 33 (2005), pp. 89–117.
- [21] Avery Katz. “Measuring the Demand for Litigation: Is the English Rule Really Cheaper?” In: *Journal of Law, Economics, and Organization* 3 (1987), pp. 143–176.
- [22] Avery Wiener Katz and Chris William Sanchirico. “Fee Shifting”. In: *Procedural Law and Economics*. Ed. by Chris William Sanchirico. 2012, pp. 271–307.
- [23] Daniel Klerman, Yoon-Ho Alex Lee, and Lawrence Liu. “Litigation and Selection with Correlated Two-Sided Incomplete Information”. In: *American Law and Economics Review* 20 (2018), pp. 382–459.
- [24] Daphne Koller, Nimrod Megiddo, and Bernhard von Stengel. “Efficient Computation of Equilibria for Extensive Two-Person Games”. In: *Games and Economic Behavior* 14 (1996), pp. 247–259.
- [25] William M. Landes. “An Economic Analysis of the Courts”. In: *Journal of Law and Economics* 14 (1971), pp. 61–107.
- [26] C. E. Lemke and J. T. Howson. “Equilibrium Points of Bimatrix Games”. In: *Journal of the Society for Industrial and Applied Mathematics* 12 (1964), pp. 413–423.
- [27] Brian G.M. Main and Andrew Park. “The Impact of Defendant Offers into Court on Negotiation in the Shadow of the Law: Experimental Evidence”. In: *International Review of Law and Economics* 22 (2002), pp. 177–192.
- [28] Richard D. McKelvey, Andrew M. McLennan, and Theodore L. Turocy. *Gambit: Software Tools for Game Theory, Version 16.0.1*. 2016. URL: <http://www.gambit-project.org>.

- [29] John Nash. “Non-Cooperative Games”. In: *Annals of Mathematics* 54 (1951), pp. 286–295.
- [30] A. Mitchell Polinsky and Daniel Rubinfeld. “Does the English Rule Discourage Low Probability of Prevailing Plaintiffs?” In: *Journal of Legal Studies* 27 (1998), pp. 519–535.
- [31] A. Mitchell Polinsky and Daniel Rubinfeld. “Optimal Awards and Penalties When the Probability of Prevailing Varies Among Plaintiffs”. In: *RAND Journal of Economics* 27 (1996), pp. 269–280.
- [32] Richard A. Posner. “An Economic Approach to Legal Procedure and Judicial Administration”. In: *Journal of Legal Studies* 2 (1973), pp. 399–458.
- [33] George L. Priest and Benjamin Klein. “The Selection of Disputes for Litigation”. In: *Journal of Legal Studies* 13 (1984), pp. 1–56.
- [34] I.V. Romanovskii. “Reduction of a Game with Complete Memory to a Matrix Game”. In: *Soviet Mathematics* 3 (1962), pp. 678–681.
- [35] Thomas D. Rowe and Neil Vidmar. “Empirical Research on Offers of Settlement: A Preliminary Report”. In: *Law and Contemporary Problems* 51 (1988), pp. 13–39.
- [36] Reinhard Selten. “Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit”. In: *Zeitschrift für die Gesamte Staatswissenschaft* 121 (1965), pp. 667–689.
- [37] Steven Shavell. “Suit, Settlement, and Trial: A Theoretical Analysis Under Alternative Methods for the Allocation of Legal Costs”. In: *Journal of Legal Studies* 11 (1982), pp. 55–81.
- [38] Kathryn E. Spier. “Pretrial Bargaining and the Design of Fee-Shifting Rules”. In: *RAND Journal of Economics* 25 (1994), pp. 197–214.
- [39] Bernhard von Stengel. “Computing Equilibria for Two-Person Games”. In: *Handbook of Game Theory with Economic Applications*. Ed. by R.J. Aumann and S. Hart. Vol. 3. 2002, pp. 1723–59.
- [40] Bernhard von Stengel, Antoon van den Elzen, and Dolf Talman. “Computing Normal Form Perfect Equilibria for Extensive Two-Person Games”. In: *Econometrica* 70 (2002), pp. 693–715.

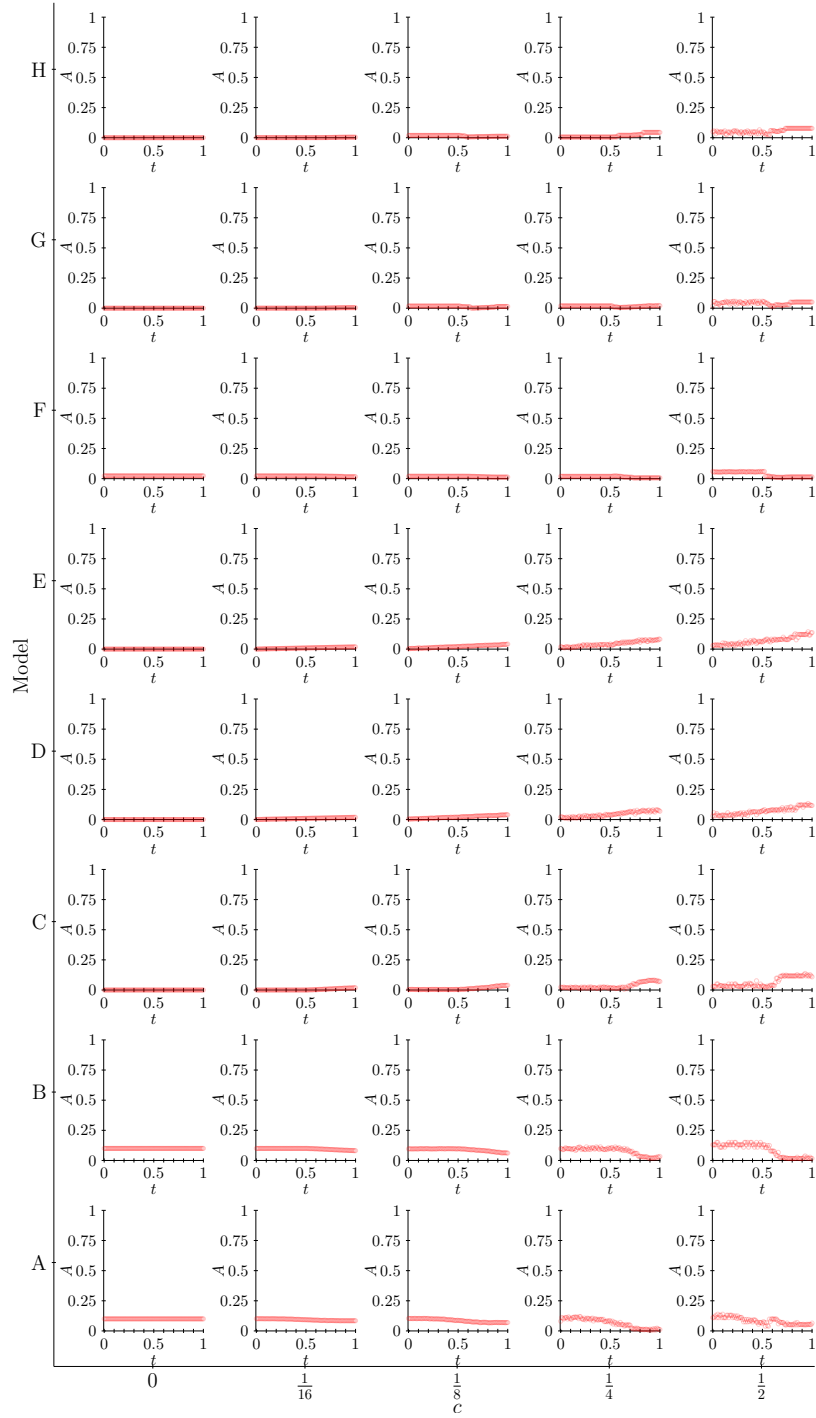


Figure 11: Overall accuracy averaged across values of q

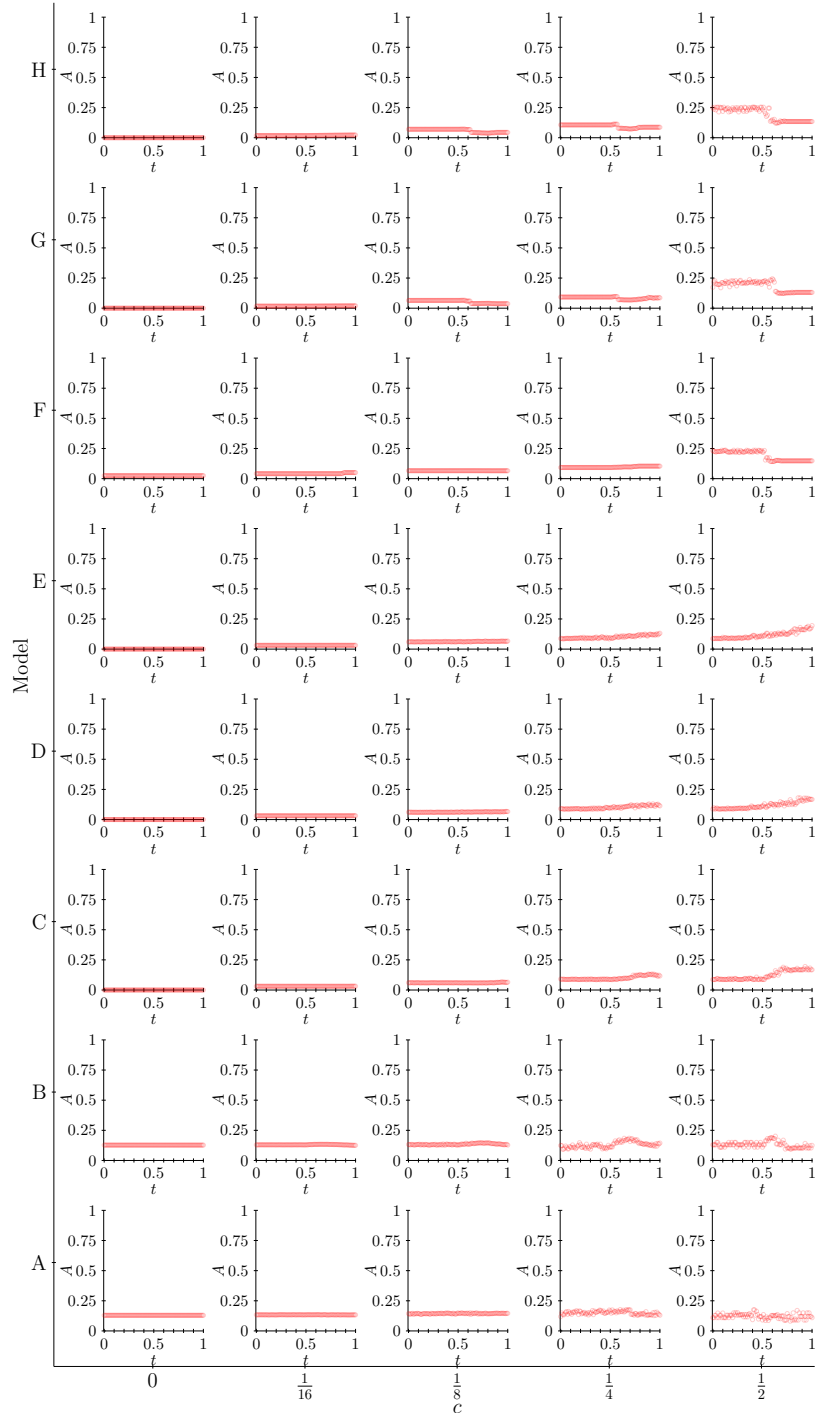


Figure 12: Undercompensation inaccuracy averaged across values of q