2018

Does the Priest and Klein Model Travel? Testing Litigation Selection Hypotheses with Foreign Court Data

Yun-chien Chang

William Hubbard

Does the Priest and Klein Model Travel?

Testing Litigation Selection Hypotheses with Foreign Court Data

Yun-chien Chang & William Hubbard†

Abstract

Law and economics theorists have long advanced theories of litigation and settlement, including the canonical Priest and Klein (1984) model that predicts that, as settlement rates rise, plaintiff win rates approach 50%. Empiricists have tested this and other predictions from the theoretical literature, finding qualified support for the Priest and Klein model. So far, though, this literature has relied primarily on data from the U.S. Yet, because this model rests on simple assumptions about human behavior in the context of a legal dispute, not on any distinctive features of American litigation, its basic insights should apply everywhere. In this paper, we put the universality of the Priest and Klein model to the test, using a unique administrative data set from Taiwan, comprising about a half-million observations. We find strongly suggestive evidence in favor of the Priest and Klein model. We consider variants of the model incorporating risk aversion and compare results testing potential implications of models of settlement under asymmetric information.

Keywords

Priest-Klein model, 50 percent hypothesis, divergent expectations, asymmetric information, risk aversion, trial rate, settlement rate, plaintiff win rate, attorney representation

---

† Chang: Research Professor & Director of Center for Empirical Legal Studies, Institutum Iurisprudentiae, Academia Sinica, Taiwan. Email: kleiber@sinica.edu.tw. Hubbard: Professor of Law and Ronald H. Coase Teaching Scholar, the University of Chicago Law School. Email: whubbard@uchicago.edu. We thank Dawn Chutkow, Hanjo Hamann, Michael Heise, Dan Klerman, John Liu, JJ Prescott, Giovanni Ramello, Hiroharu Saito, Simone Sepe, Jeanette Shao, and participants at the First CELS in Asia, 12th Conference on Empirical Legal Studies, Georgetown University Law Center Law and Economics Workshop, George Mason University Law and Economics Colloquium, Faculty Workshop at Academia Sinica, 2017 Asian Law and Society Association Annual Meeting, University of Chicago Work-in-Progress Workshop, and the Notre Dame Law & Economics Workshop for helpful comments. Hubbard thanks the Jerome F. Kutak Faculty Fund for research support. The view expressed in this article is solely of the authors, not of Judicial Yuan of Taiwan, Academia Sinica, or any other government agencies.

Table of Contents

I. Introduction

How much can theories of suit and settlement developed in the context of one, specific legal system (the US) tell us about other legal systems, and how much can data from other legal systems tell us about theories of suit and settlement? In this paper, we explore these questions, with application to the use of unusually comprehensive data on civil court cases from Taiwan to explore empirical predictions of the canonical Landes-Posner-Gould and Priest and Klein (1984) models of litigation and settlement. In so doing, we generate new evidence of common patterns in litigation across widely diverse institutional settings, but also identify some of the challenges that arise in translating a model of litigation to a new context.

There are, of course, major differences in the design of court systems across jurisdictions, and a robust body of theoretical work examines some of the better known and perhaps most consequential differences (such as the American and English rules for allocation of attorney fees). Nonetheless, some theories of litigation and settlement are sufficiently general that they should hold across legal systems, even quite distinct ones. For example, the Landes-Posner-Gould model of litigation predicts that higher-stakes disputes will be less likely to settle than lower-stakes disputes, and the Priest and Klein model predicts that as the trial rate in any set of cases approaches zero, the plaintiff's rate of prevailing at trial will converge on 50 percent. These predictions rest on a model of human behavior in the context of a legal dispute, not on any distinctive features of US litigation. They should be true everywhere.

But are they? We don't know, largely because there has been little or no evidence collected on this question to date. What little we have in terms of empirical studies of these hypotheses almost always uses data from US courts. Yet, we might think that data from courts outside the US present a *better*—i.e., more rigorous—test of models like Priest and Klein (1984). These were developed with the US legal system in mind, and so by looking for the same phenomena in a different legal system, we can conduct an "out of sample" test; we can assess the "external validity" of a theory that, by its own internal logic, should apply equally in civil law or common law countries, and in unitary or federal court systems.

Nor is the empirical testing of economic models of litigation an exercise of purely academic interest. These models shape how we evaluate the legal system (both positively and normatively). The Priest and Klein model, for example, predicts that

changes in legal standards may favor plaintiffs, but the selection of cases into settlement and trial will mean that we cannot detect the change by looking for a rise in plaintiffs winning at trial. (Instead, more cases that plaintiffs would win may settle instead.) If this insight is correct, then we cannot—as lawyers and commentators often do—infer from high win rates for defendants that the law favors defendants.

Models also guide the formation of agendas for improving the legal system. These models imply prescriptions for changing the system, depending on our goals. For example, let's say that we want to encourage settlement, in order to reduce the expenses associated with litigation, and allow parties to obtain relief faster. If this is our goal, different models of litigation identify different culprits for why parties fail to settle:

The Priest and Klein model assumes that settlement failure is the product of parties who are equally (yet incompletely) informed about the merits of a case, but who are optimistically biased about their chances of winning. (For this reason, the Landes-Posner-Gould model and the Priest and Klein model are called "divergent expectations" models.) Because both parties overestimate their chances of winning, they cannot agree on a settlement amount. The difference in their perceived chances of winning outweighs the fact that they both save litigation costs by settling.

The policy prescriptions that might flow from this include identifying sources of optimism bias in clients and lawyers, and designing legal procedures to incorporate de-biasing techniques, informed by psychology and behavioral economics. To some extent, the US legal system may do this already—legal representation is commonplace in US courts, and a growing body of evidence shows that a person acting on behalf of another (i.e., an attorney) is less prone to cognitive bias than a person acting on their own behalf (i.e., the client). See, e.g., Arlen and Tontrup (2015). Alternatively, or in addition, reducing sources of mutual uncertainty would reduce the range in which biases could affect behavior. A tort system with highly predictable, rather than discretionary, judgments about liability and damages might accomplish this (see, e.g., Ramseyer 2015: 10–70).

Other canonical models, such as Bebchuk (1984) and Reinganum and Wilde (1986) assume that settlement failure is the product of asymmetric information between the parties—one party knows something that the other does not. (For this reason, these models are called "asymmetric information" models.) For example, the defendant knows the extent of its negligence, but the plaintiff does not. Because of information asymmetry, the less-informed party does not know how much to settle for, and thus

that party will sometimes risk trial rather than agree to a lowball settlement. One obvious policy prescription that might flow from this is to eliminate information asymmetry. Of course, this is exactly the policy that the US federal courts have followed since the implementation of the Federal Rules of Civil Procedure in 1938. Asymmetric information may exist when a complaint is filed, but broad discovery should eliminate information asymmetries before trial.

The US legal system is, in many ways, unique. In Taiwan, the country from which our data comes, attorney representation is much less common, and there is no such thing as broad discovery. Further, settlement rates in Taiwan are very low—below 25%—which is far below settlement rates in US federal courts—which exceed 60%. But why aren't cases settling? Is it the lack of lawyers? The lack of discovery? Or something else? Should courts in Taiwan become more like courts in the US? Or vice versa?

Thus, the answer to the question "What should we do?" depends on which model or models we think effectively describe the real world. In other words, while all of these models are theoretically sound, determining how well they predict real-world empirical outcomes is an important step toward transforming theoretical insights about law and economics into concrete proposals for legal reform.

This paper takes precisely this step, examining which theoretical model passes muster in the real world. We draw empirical predictions from the theoretical literature on litigation and settlement and test those predictions against our data. We seek to get a grip on which theoretical models of litigation and settlement seem to work tolerably well as an approximation of what is happening in the real world. To the extent that a model's predictions are supported by the data, we can be more confident that the model may be a useful guide in prescriptive analysis. And if its predictions are rejected, more caution is warranted.

Of course, we know *a priori* that none of the theories are perfect. Human behavior in general, and the legal system in particular, are too complex for any simple theory to work with high accuracy. But in our view, a rough guide is better than no guide, and empirical data on the success or failure of a model's predictions also provides fodder for newer and better models. And while no single empirical study can provide conclusive evidence for or against a model, this paper presents a large set of robust empirical results, based on unusually rich data.

In sum, this paper contributes to the literature testing the empirical predictions of theoretical models of suit and settlement. We take advantage of a dataset of civil cases from Taiwan that is more comprehensive in its coverage of cases than most US

datasets and that provides detailed and consistent measurement of stakes, a quantity that is particularly hard to find in extant US datasets. We test the two predictions of the Priest and Klein and Landes-Posner-Gould models noted above, as well as several corollary or competing predictions drawn from the empirical literature that has sought to test these models, including Kessler, Meites, and Miller (1996), Siegelman and Waldfogel (1999), Waldfogel (1995), and Waldfogel (1998). Our main result is that we find both qualitative and quantitative evidence supporting the Priest and Klein (1984) model, including its famous 50% prediction noted above. Our analysis finds mixed support for competing hypotheses, and rejects the hypothesis that asymmetric information will push plaintiff win rates away from 50% as trial rates approach zero.[1]

Taken together, these results tend to confirm the predictions of the Priest and Klein model, while providing less support for predictions consistent with existing models of asymmetric information. The findings are certainly not conclusive, but they are strongly suggestive, and we believe that further empirical testing of these same models using other hypotheses and data is merited. Based upon our results, we very tentatively conclude that models based on behavioral factors, such as optimism bias, may have more empirical relevance in explaining litigation and settlement than models based on one-sided asymmetric information.[2]

The rest of this article is as follows: Part II summarizes relevant portions of the prior literature on litigation and settlement, and concludes with a discussion of important caveats from the literature about the unique challenges faced by empiricists attempting to distinguish between theoretical models of litigation and settlement. Part III gives an overview of the Taiwanese legal system. Part IV describes the pertinent data. Part V lays out the hypotheses crystalized from the prior literature and reports our empirical results. Part VI concludes.

---

[1] In addition, in Appendix A, we utilize this data to test other predictions from the literature on litigation and settlement, which are distinct from predictions that would distinguish the Priest and Klein model from its alternatives. Contrary to US experience, the presence of legal representation (rather than pro se parties) is not associated with higher settlement rates, and we discuss possible reasons for this difference. Also, we find that corporations have higher win rates than individuals, consistent with the claim that repeat players will have higher win rates, and represented parties win more often than pro se parties.

[2] Note that the canonical asymmetric information models assume one-sided private information. Because the Priest and Klein model can be rationalized as a model of two-sided private information (Lee and Klerman 2015; Lee and Klerman 2016), our results should be understood as tending to reject models of one-sided asymmetric information, not models of asymmetric information generally.

II. Literature on Litigation and Settlement

A. The Canonical Models and the Empirical Literature

What drives parties to a dispute into litigation, and what leads them to settlement, have long been studied in the economic analysis of law. Canonical models were first developed in the 1970s. Landes (1971: 102) points out that the differences in parties' estimates of liability and damages, risk-aversion, and the relative costs of settlement and trials affect whether parties settle or litigate. Gould (1973: 296) emphasizes the import of parties' degree of agreement on the plaintiff's probability of winning at trial. Posner (1973) discusses the effect of judicial error costs, judicial delay, costs of litigation and settlement, and mutual optimism. The synthesis of these papers is often referred to as the Landes-Posner-Gould (or LPG) model. This model treats parties as wealth-maximizing actors whose decisions maximize their perceived net payoffs in litigation. A plaintiff will file suit if the probability of winning at trial, times the amount in dispute, is greater than the cost of litigation. Each party will be willing to settle if the settlement amount is higher than the expected verdict at trial.

In this framework, if the parties have the same expectation of the probability that the plaintiff wins at trial, settlement is inevitable. Because litigation is costly for both parties, there is a range of settlement amounts that both parties would prefer to trial. For example, a settlement for the plaintiff's probability of winning times the amount in dispute would replicate the expected outcome from trial, but both parties would save litigation costs. For this reason, the LPG framework rationalizes trial—i.e., failure to settle—as the product of "divergent expectations" about the plaintiff's likelihood of winning. If each party is optimistic about their chances, then the plaintiff will tend to demand more, and the defendant will tend to demand less, in settlement. If the stakes are high enough, or the cost of going to trial is low enough, divergent expectations about the parties' odds at trial may make settlement impossible. Thus, in this framework, settlement becomes more likely as litigation costs rise relative to stakes or as mutual optimism diminishes.

The LPG model have been developed, extended, and criticized over more than three decades. Its prediction that settlement rates fall as stakes rise relative to costs has been challenged in theoretical grounds. Cooter and Rubinfeld (1989: 1077) and Kessler, Meites, and Miller (1996: 246) argue that risk-averse parties will tend to

settle rather than litigate when stakes are high.[3]

Building on the LPG model, Priest and Klein (1984) describe a model that generates a number of dramatic predictions about litigation. They posit, if parties have expectations about their chances at trial that are unbiased and fairly accurate, then mutual optimism will lead to settlement failure only in relatively "close" cases—cases where the plaintiff's probability of winning is close to 50%. Importantly, this logic generates empirically testable hypotheses. Lee and Klerman (2016: 60) enumerate six concrete hypotheses originating from Priest and Klein (1984):

First, the "trial selection hypothesis" states that litigated cases are not a random or representative sample of all disputes. This is widely considered as the most important contribution of Priest and Klein (1984) (Lee and Klerman 2016). There has been little theoretical or empirical challenge to this hypothesis, although Helland, Klerman, and Lee (2017 forthcoming) find that, in a data set from New York state courts, the distributions of damage awards in judicial decisions and in settlements are very similar, suggesting that there is at least one occasion where there is no selection into settlement based on expected value of the claim.

Second, the strong selection during the litigation process produces the "no inferences hypothesis", suggesting that "no inferences can be made about the law or legal decision-makers from the plaintiff win rates" (Lee and Klerman 2016). This claim has been challenged by Klerman and Lee (2014).

Third, there is the conjecture that plaintiff trial win rates converge toward 50% as the trial rate falls. This is the most famous of the predictions and the one to which the empirical literature has devoted the most attention.

Fourth, regardless of the legal standard (favoring plaintiffs, defendants, or neither), the plaintiff trial win rate will demonstrate a strong tendency toward 50%, as compared to the percentage of cases plaintiff would have won if all cases went to trial.

Fifth, the plaintiff trial win rate will be unrelated to the shape of the distribution of disputes.

Sixth, Priest and Klein (1984: 24–26) contends that asymmetric stakes will push

---

[3] Landes (1971: 99), in the criminal law context, points out that the greater the defendant's aversion to risk, the more likely that a settlement is to take place. Huang (2016: 385) uses data from face-to-face surveys in Taiwan to study out-of-court settlement behaviors, and finds that the amount of claim (X-axis) and estimated settlement probability (Y-axis) have an inverted U-shape relationship. That is, settlement rates are highest when the stakes are medium, neither high nor low.

plaintiff trial win rates away from 50%; that is, parties that systematically have larger stakes will tend to win more often.

The most important challenge for any model of human behavior, of course, is to generate empirical predictions that survive the rigorous testing of such predictions. The Priest and Klein (or PK) model, and the LPG framework upon which it builds, have attracted its share of attention in this regard, as studies have both questioned and supported its empirical prediction (among others) that, as the share of cases that are litigated as opposed to settled goes to zero, the fraction of cases won by plaintiffs converges to 50%.

Most of the empirical articles to date use observational data to test the PK model. The literature has generated mixed results in terms of how closely litigation outcomes fit the PK hypotheses. Eisenberg (1990: 341) fits the distribution of plaintiff win rates over a binomial distribution (with p=0.5) to test the 50% limit hypothesis. Using administrative data on US federal courts, Eisenberg (1990: 348–51) finds that the distribution of plaintiff win rates across districts and case types does not support the Priest-Klein model. He concludes that the 50% limit hypothesis should be rejected (see also Eisenberg 1991).

A number of studies, however, find support for the PK model when testing other predictions of the model. Siegelman and Waldfogel (1999) use administrative data from the US federal courts to generate proxy measures for the degree of asymmetry of stakes, the amount of uncertainty in parties' estimates of the strength of the plaintiff's case, the settlement rate, and the plaintiff win rate for several categories of cases. They then use a structural model to predict settlement rates and plaintiff win rates based on the measures of asymmetry of stakes and uncertainty. By comparing the predictions of the structural model with the observed settlement rates and plaintiff wins, they attempt to test the predictive power of the PK model. They find mixed support for the model in the data.

Waldfogel (1998) uses the same administrative data from Siegelman and Waldfogel (1999) to compare predictions of the PK model with predictions of asymmetric information (or AI) models. Bebchuk (1984) presents the canonical screening model and Reinganum and Wilde (1986) present the canonical signaling model. In these canonical models of litigation and settlement in an environment of one-sided asymmetric information, better-informed parties with stronger cases are more willing to go to trial, and thus as trial rates appraoch zero, only the better-informed parties with the strongest cases litigate. Thus, as cases become more highly selected due to settlement, Priest and Klein predict that plaintiff win rates

will converge to 50%, while AI models predicts that the win rate of the better-informed party with converge to 100%. Comparing plaintiff win rates across case categories in the early stages of litigation with win rates in the late stages of litigation, Waldfogel (1998) finds a clear convergence toward 50% as litigation progresses, and concludes that this supports the Priest and Klein 50% prediction but rejects the AI models.

Waldfogel (1995; 1998) also tackles the fifty-percent limit hypothesis by examining the relationship between settlement rates and plaintiff win rates, using the same data. As the 50% win rate will be attained in the limit as the number of trials falls, higher settlement rates should be associated with plaintiff win rates closer to 50%. Waldfogel (1995: 244–45) finds that the prediction holds when settlement rates and win rates are computed by judge and case type. Case types that have low settlement rates produced either very high (close to 1) or very low (close to 0) plaintiff win rates, whereas those that have high settlement rates led to plaintiff win rates around 50%. He argues that the data is consistent with the PK model but not AI models.

Siegelman and Donohue (1995: 431) use administrative data from employment discrimination cases in US federal court and find support for two predictions of the PK model: first, the strongest and weakest cases are more likely to settle, and second, the party with higher stakes will tend to have higher win rates in litigated cases. (These are the first and sixth predictions listed above.) Siegelman and Donohue also note that the business cycle predicts patterns in the filing and settlement of employment discrimination cases. This result, while intuitive, does weaken one of the stronger (i.e., more ambitious) predictions of the PK hypothesis, which is that plaintiff win rates will bear no relation to changes in the underlying distribution of disputes. This finding lends support to Klerman and Lee's (2014) argument that, except in the limiting case of trial rates going to zero, selection effects attenuate but do not eliminate the effect of legal changes and changes in the underlying distribution of cases on plaintiff win rates at trial.

One recent paper attempts to test the 50% prediction with data from Chinese courts. Cui and Wang (2017) use administrative data on litigation against Chinese administrative agencies and, in the spirit of Waldfogel (1995), compare plaintiff win rates and adjudication rates across agency types and years. They find a strong negative relationship between adjudication rates and plaintiff win rates, a result that is consistent with Priest and Klein, given that plaintiff win rates tend to be below 50% in their data—thus, as the adjudication rate tends towards zero, win rates tend

toward 50%.

Ramseyer and Nakazato (1989) examine Japanese data on litigation over wrongful deaths in automobile crashes. (They use 26 years of data, comprising 766,948 Japanese court decisions, which makes their data set the largest in the literature.) They note an 80% plaintiff win rate but search for plausible theoretical reasons to explain why the finding could be consistent with the Priest-Klein model. In this way, Ramseyer and Nakazato do not so much test the PK model as incorporate factors such as asymmetric litigation costs and asymmetric risk aversion to fit the model to the data.

Some data sets have brought otherwise unobservable quantities into the analysis, something that we also do below with the Taiwanese administrative data. Most notable in this respect is Klerman (2012), who utilizes data on thirteenth-century English jury verdicts in criminal cases to test predictions derived from the PK model.[4] The most distinctive feature of Klerman's data set is that it includes jury verdicts from *settled* cases. (For the cases in Klerman's data, juries operated independently of the parties, akin to a modern grand jury, rather than petit jury.) Unlike any modern data set, this data set thus provides (what would otherwise be) the "counterfactual" outcome that would have obtained in the absence of settlement. Klerman (2012) focuses on the sixth of the PK hypotheses, i.e., that the party with the higher stakes will have the higher win rate in litigated (in Klerman's data, non-settled) case. In the criminal context, this is the defendant, and Klerman's results support this prediction of the PK model.

Finally, some experimental work has been done on the Landes-Posner-Gould framework. Stanley and Coursey (1990) conducted a laboratory experiment in which subjects engaged in simulated pretrial bargaining. The experiment varied litigation costs and the precision of the parties' estimates of the trial outcome. Consistent with the Landes-Posner-Gould framework, settlement rates rose as litigation costs rose and as parties' expectations about the strength of the plaintiff's case were based on more precise estimates of the trial outcome.[5]

---

[4] Because criminal cases were privately prosecuted at that time, Klerman argues that this data set is appropriate for testing theories developed to explain civil litigation.

[5] Other noteworthy extensions and criticisms include: Lee and Klerman (2015) propose modifications to the original Priest-Klein model, such as considering whether parties use Bayes' rule to refine their estimates. Loewenstein et al. (1993: 138–40) challenge the psychological foundation of the Priest-Klein model, as their experiments found that egocentric human nature induces parties to make biased estimate of case values and the magnitude of the bias is a strong predictor of nonsettlement. It is worth noting that Loewenstein et al. did

One difficulty faced by this literature is that many of the relevant variables in the models (such as party optimism, stakes, or costs) are not observable in the data, and empirical researchers must rely upon rough proxies. Consequently, what counts as a "rejection" of a model is open to debate. As Siegelman and Donohue (1995: 431) argue (and present evidence to support), in realistic settings the selection effects predicted by Priest and Klein will only be partial, so some variables may explain variations in plaintiff win rate (contrary to the second and fifth predictions listed above). If these variables are not observable to the researcher, however, it will be difficult to distinguish a failure of the model from lack of sufficiently granular data. Take asymmetric stakes as an example. No administrative data to date has included direct measures of parties' stakes. Asymmetry of stakes, however, could explain systematic deviations in plaintiff win rates away from 50%.

B. To What Extent Are the Canonical Models Empirically Falsifiable at All?

Magnifying this concern are theoretical critiques that ask whether models of suit and settlement, including the PK model, are falsifiable at all. These critiques are associated with Shavell (1996: 493), who noted that models of litigation and settlement can easily be tweaked so that they predict any plaintiff win rate at trial. Strictly speaking, though, this claim is not a rejection of the empirical content of Priest and Klein or other models, but rather a claim that different assumptions about the (usually unobservable) information structure of litigation will lead to different empirical predictions.

More recent work by Gelbach (2016), however, directly questions the empirical content of the PK model. Gelbach shows that the Priest and Klein (1984), Bebchuk (1984), Reinganum and Wilde (1986), and other important models of litigation and settlement can be represented by a single, common, "reduced form" theoretical framework that is consistent with the basic postulates of the PK model. If so, Gelbach argues, discriminating empirically between the PK model and other, competing models is virtually impossible, because any result, no matter how seemingly consistent with another model, can be rationalized with the PK model for some

---

not test predictions of the Priest and Klein model; rather, they tested an assumption of the model, which is unbiased prior estimates of the strength of the plaintiffs' case. Their laboratory experiment, which simulated pretrial settlement bargaining, found significant self-serving bias by subjects. Although couched as a rejection of Priest and Klein, this result seems to reinforce Priest and Klein's predictions, as well as the earlier divergent-expectations literature. Self-serving bias is a source of divergent expectations about trial outcome.

parameter values in the "reduced form" model.

In principle, this concern could be overcome with sufficiently rich data on the parameters of the model; for example, if we knew not only settlement rates and plaintiff win rates, but also had some information on the extent of asymmetric information, asymmetric stakes, and litigation costs, our ability to shoehorn data into the PK model would be meaningfully constrained. Such rich data, however, is unheard of currently.

We recognize the force of Gelbach's critique. We treat it not as a reason to despair in our empirical exercise, but rather as an all-important reminder of the need to discipline any attempt to test competing models of litigation and settlement. As Gelbach shows, it can be easy—not just rhetorically, but even as a matter of formal theory—to rationalize seeming rejections of the PK model to fit the model by adjusting our claims about unobserved quantities such as asymmetric stakes and asymmetric information. Thus, to say that we can cleanly reject or support a model, we must be sufficiently precise in defining ex ante the model we seek to reject or support.

In this paper, we impose this discipline by testing the hypotheses we identify under a strong *ceteris paribus* assumption: that unless we have a priori reasons for suspecting systematic differences in asymmetric stakes, asymmetric information, or the underlying distribution of case strength across categories of cases or parties, we assume that these factors cannot explain major discrepancies between the predictions of the original PK model and the data. We impose the same assumption to the extent that we test predictions based on other, competing theories, too.

In other words, we make explicit the assumptions implicit in the original versions of the canonical models: that unobservable parameter values are not so peculiar that that the models will generate empirical results *opposite* from the results implied by the original papers introducing the models. For example, we assume that case quality is distributed such that the distribution of plaintiff's true chance of prevailing has full support on the (0,1) interval.[6] (In English: we assume that there is a full spectrum of case strength from slam-dunks for the plaintiff to slam-dunks for the defendant.) We find this assumption to be uncontroversial, but it is just that: an

[6] This assumption is essential for the prediction, noted above in our discussion of Waldfogel (1995; 1998), that AI models predict that as settlement rates rise, plaintiff win rates will tend to move away from 50%. Obviously, if case strength is bounded above or below by 50%, then in the limit, plaintiff win rates can be 50% in AI models as well as in the PK model. But such a strict assumption seems implausible *a priori*.

assumption.

Given that reality is never as elegant as a simplified theoretical model, of course, there is nothing unreasonable about recognizing that discrepancies could be due to unobserved factors, but we take this approach to ensure that our empirical exercise involves a meaningful test of competing models. In doing so, we deliberately make it more difficult to find empirical support for *any* given model (including the PK model). First, to the extent that our results tend to support the predictions of the PK model (and we shall see that they tend to), stacking the deck against the PK model helps increase our confidence that our findings are due to genuine robustness of that model to variation in real-world settings, and not because we gave the model too much "wiggle room" in letting it fit the data. Second, and more fundamentally, we want to ensure that, if we do find that the data favors one model over another, that the choice of models matters. A model that predicts that any outcome is possible is useless for positive or prescriptive analysis; it is equivalent to having no model at all.

III. Background on Taiwanese Courts and Civil Litigation

Taiwan is a civil-law country. Most of its private laws resemble those in Germany with influences from Japanese laws (Wang 2002; Huang 2009: 251; Chang 2016: 227–228; Chang, Chen, and Wu 2017; Chang, Garoupa, and Wells 2017). American laws have only occasionally been referenced in legal reforms. Taiwan has a unitary legal system, unlike the United States' federal system. In addition to a constitutional court in charge of abstract, centralized, ex post constitutional review, there are administrative courts that deal with disputes between citizens and governments regarding public-law matters; ordinary courts that handle civil and criminal cases; and IP courts. Our data regard civil law disputes (excluding those related to family and inheritance issues, but including commercial issues) rendered by courts of first and second instances within the ordinary courts, where there are 22 district courts and 6 high courts. The Taiwan Supreme Court only reviews questions of law; our data do not include its decisions.[7]

As is typical for a civil-law system, there is no sharp distinction between pre-trial and trial phases litigation; the entire lawsuit could be considered the "trial." The public trial is not intensive. Rather, judges and attorneys meet once every few weeks

[7] For empirical studies of case selection by the Taiwan Supreme Court, see Eisenberg and Huang (2012).

for a few times.

There is no American-style discovery in Taiwan, although there are rules for disclosure of information early in the process. The Taiwan Code of Civil Procedure was reformed in 2000 to have two stages of hearings, and any evidence not presented at the first, preliminary stage hearing could be precluded in the second stage of hearings (Huang 2009: 251–54). This has the effect of making evidence disclosure in Taiwan something roughly equivalent to Rule 26(a) disclosures under the Federal Rules of Civil Procedure; i.e., the parties must disclose documents and materials that they intend to submit as evidence to the court. Anecdotally, however, Taiwanese judges tell us that they often allow evidence in the second stage of hearings when there is a good reason. No one has studied empirically how often judges have exerted their power to preclude new evidence and to what extent parties and their representatives have been incentivized to present evidence in the preliminary stage. Attorneys we interviewed corroborate the stories told by the judges, as they still tried to refrain from presenting evidence unless they have to. That is, the "not showing your hand until you have to" strategy described by Huang (2009: 252) appears to be still popular today.

Most attorneys charge flat fees, collected beforehand. Contingent fees are legal but are used in less than 5% of the cases (Chang and Tu 2018). Bigger law firms charge by the hour about half of the time (Chang and Tu 2018). Most attorneys charge 1,500 USD to 2,500 USD per case, per court instance (i.e., appeal is a separate fee).[8]

There are three types of court procedures: small-claim procedure, summary procedure, and ordinary procedure. Small-claim procedures apply when the amount at stake is below 100,000 NTD (approximately 3,333 USD). Small-claim cases are handled by a single judge in one of the 22 district courts. The median time between filing and termination for cases ended in adjudication is 50 days. This kind of case can be appealed to a three-judge district court panel on questions of law.

Summary procedures apply when the amount at stakes is below 500,000 NTD (approximately 16,666 USD) or in certain types of cases specified in Article 427 of Taiwan Code of Civil Procedure. This type of cases is also handled by a single judge. The median time between filing and termination for cases ended in adjudication is 58 days. Cases can be appealed to a three-judge district court panel for trial de novo, and

[8] Original research; data on file with authors.

the court may permit appeal to Supreme Court.

A case applying ordinary procedures is handled by a single judge in the district court. The median time between filing and termination for cases ended in adjudication is 166 days.[9] Appeals are handled by a three-judge panel in one of the six high courts. The review of facts and laws is de novo. Cases with more than 1,000,000 NTD (approximately 33,333 USD) at stake may be appealed to the Supreme Court for questions of laws. Appellants have to be represented by attorneys in the Supreme Court. The Taiwan Supreme Court has discretionary jurisdiction (Eisenberg and Huang 2012).

Settlement, at least its Chinese term in Taiwan law, can occur many ways. Parties can settle out of court after filing a lawsuit. If they prefer to keep the settlement terms strictly confidential, plaintiffs can withdraw their cases. Hence, "withdrawal" as a case outcome can represent settlement. Of course, plaintiffs may withdraw when finding out that they (at the time) do not have a winning case. Parties can also settle before the judge assigned to handle their case. Judges, in fact, are encouraged by the Taiwan Code of Civil Procedure to actively promote settlement. There is also mediation. During the mediation procedure, the dispute is first assigned to a mediator (or multiple ones) chosen by the judge (parties may object to the choice; mediators are chosen from a list of mediators prepared by each district court), and a judge in the summary court or the original judge who handled the dispute will step in at the final stage. Judges may also administer the mediation themselves throughout. Either party is free to reject any mediation proposal by mediators. We treat all of these forms of dispute resolution as "settlements" for our purposes. All of them contrast with a judicial judgment in which the court renders a verdict that declares a winner and awards (or declines to award) damages or issues injunctions.

The filing fee is pro rata (about 1% of the amount claimed in the first instance and 1.5% of the amount claimed in the second instance), meaning the higher the stakes, the more expensive (in absolute terms) it is to use litigation to resolve the dispute. Filing fees are two-thirds refunded if parties settle (including through mediation) in court before adjudication. Filing fees have to be paid in full by plaintiffs before any court procedure starts. If plaintiff wins partially, filing fees are allocated on pro rata basis between the parties. If plaintiffs win entirely, filing fees are the sole

[9] As a point of comparison, in the US federal courts, civil cases on average last approximately 326 days (median duration is 230). Authors' calculations for all U.S. district courts, civil cases filed 1999–2004 (data on file with authors).

responsibility of defendants. Nonetheless, filing fees are not refunded to plaintiffs by the court. Rather, a plaintiff enforces the award of filing fees against the defendant.

IV. Data and Summary Statistics

Our data are administrative data on the Taiwan courts collected by the Judicial Yuan of Taiwan and made available to us for research purposes. The Judicial Yuan of Taiwan is the administrative organ for judicial matters. It is independent of all courts, including the Supreme Court, and is responsible for making judicial policies and supervising the day-to-day operations of all courts. The Constitutional Court of Taiwan is part of the Judicial Yuan, and the Chief Justice is the President of the Judicial Yuan. All but one variable in the raw data were provided to one of us (Chang) by the Judicial Yuan based on a codebook designed by Chang.

One key variable, case types (nature of disputes such as contract, torts, property, etc.), was not contained in the raw data, but the administrative data included information that we used to construct case types. Each civil case filed in Taiwan is accompanied with a one-incomplete-sentence description of the nature of the plaintiff's claims, written by the clerks working in the filing office in each district court. Norms and conventions—and the similar nature of many disputes—result in frequent uses of several template descriptions. We used these descriptions as the major source of information for case types. The Judicial Yuan of Taiwan also assigns certain types of cases (such as maritime, international trade, medical malpractice, labor, consumer, state liability, etc.) to separate tracks. These tracks serve two functions: first, some of these cases will be handled solely by a specialized division within a court. Second, some cases, such as medical malpractice cases and financial cases, are expected to be more complicated than other ordinary cases. While cases have been randomly assigned to judges, judges are still concerned that misfortune will result in one judge being assigned all the knotty cases. Thus, random assignments are conducted within each track, so that judges' case burdens are more likely to be even. The track of a case is indicated in the case name. We used it to construct case types as well.

The original dataset includes 798,801 cases, each of which appears as a single observation in the data. For our analyses below, we excluded cases/observations in which the central or a local government is a party, remanded cases, cases coded as disputes arising after defective settlement or mediation agreements, and cases in which the number of plaintiffs or defendants was erroneously coded as 0. We also

exclude cases in which the parties did not litigate over a monetarily quantifiable claim. We then trimmed observations in the first and last percentiles of stakes, to minimize undue effects from extreme outliers, and we removed cases with more than one party in either side (for which patterns of settlements and verdicts may be complex and poorly documented in the data). And because the plaintiff win rate is one of the key outcome variables, we omitted cases in which defendants counter-claimed.[10] This left about 500,000 cases in our data that we used for the analysis below.

The data contain information on case outcomes, separating the outcomes of "plaintiff win" (plaintiff obtains all requested relief) and "partial win" (plaintiff obtains relief but less than the total amount claimed). In our regression analysis, we follow the prior literature and treat both of these categories as victories for the plaintiff. Thus, we recode "partial win" as "plaintiff win."[11] In Appendix C, we treat a partial win as "half a win," based on the notion that a partial win may reflect the type of close case that the PK hypothesis would predict goes to trial.[12] All results, except where specifically noted below, are the same in terms of sign and statistical significance.

Tables 1–3 provide summary statistics on our data. As Table 1 indicates, most cases are resolved quickly and involve modest stakes. (The mean stakes of about

---

[10] In Appendix B, we report regression models that include cases with counter claims and those with multiple parties.

[11] In the prior literature, an outcomes is usually coded as a plaintiff win if plaintiff received any benefit at trial or appeal (Kessler, Meites, and Miller 1996: 243; Eisenberg and Farber 1997: S100), and most prior articles, except Klerman (2012) and Huang (2009), did not distinguish between plaintiff partial wins and plaintiff full wins. The Japanese administrative data used in Ramseyer and Nakazato (1989: 283) and the American administrative data used in other articles do not distinguish between full and partial wins by the plaintiffs.

[12] Note that a partial win can arise in two ways. The first way is that the plaintiff could prevail on one claim and lose on a second claim. For purposes of testing the PK model, this could be coded as "0.5," because the plaintiff is winning 50% of the time. But the second, and we suspect more prevalent way that a partial win arises is when the plaintiff prevails but receives less than the full amount demanded. For example, if a plaintiff claims for, say, 100 dollars as damages, then the Judicial Yuan of Taiwan defines as a partial win any court award between 1 cent and 99.99 dollars. Only 100 dollars is a full win, and only a 0 dollar award is a full loss. (An entirely different approach would be to extend the PK model to allow for partial wins, which would occur for cases with case strength below the threshold for a win, but above a second, somewhat lower threshold. See Klerman (2012). Because this approach would be applicable for the former scenario above—winning on one claim and losing on another claim—but not the latter scenario above—a plaintiff winning on the claim but receiving less than the full amount demanded, we do not explore this option further.)

840,000 NTD is equivalent to approximately USD 28,000; the median of about 168,000 NTD is equivalent to about USD 5,600.) Stakes represent the amount claimed, which is usually the damages demanded by the plaintiff. Note that because filing fees are a positive function of the plaintiff's *ad damnum*, the plaintiff's claimed stakes do not represent "cheap talk" and thus constitute a more credible measure of stakes.[13]

Three facts jump out from Table 2. First, very few litigants are represented—about 26 percent of plaintiffs and 22 percent of defendants. Note however, that a business represented by in-house, rather than retained, counsel will be coded as unrepresented in this data, so this measure is somewhat noisy. (If corporate parties are excluded, 28 percent of plaintiffs and 21 percent of defendants are represented.) Second, most plaintiffs, but a minority of defendants, are corporations. This reflects the fact than many lawsuits involve debt collection efforts by corporations against individuals.[14] Third, most filed lawsuits are litigated to a court judgment rather than ending in settlement. In this paper, we treat several outcome categories in the data as "settlement": these include judge-approved settlement, judge-administered successful mediation, and plaintiff or appellant withdrawal. As noted in the previous part, we treat successful mediation and judge-approved settlement as settlement in the sense envisaged by literature on litigation versus settlement. Some unknown fraction of cases coded as ending in "withdrawals" involve private settlements. In this article, we follow Huang (2009: 256–257) and include all cases coded as ending in plaintiff or appellant withdrawal as "settled." In every withdrawal, the plaintiff voluntarily gives up the possibility of a court verdict, and, as soon as defendants plead their defense on the merits, plaintiffs can only withdraw if defendants agree. Thus, even if no money changes hands, this outcome is, to our eyes, fairly characterized as settlement.

---

[13] In contrast, US court data rarely has reliable information on stakes. Most states do not require the complaint to specify a damages amount (although plaintiffs often do), and the federal courts require only an allegation that stakes exceed USD 75,000 in cases based on diversity jurisdiction, which make up about one quarter of the federal civil docket. Thus, cases with damages specified in the complaint are a selected subsample of all cases. Further, because filing fees do not depend on the amount claimed in federal court (and in many state courts), incentives to accurately quantify stakes tend to be weaker for US cases.

[14] In 2015, for instance, official Judicial Statistics Yearbook indicates that 51,163 of the 116,601 non-commercial private-law filed cases (44%) are monetary loan cases. In 39,334 of the filed cases (77%), courts reached a substantive decision and ruled for plaintiffs in full in 35,507 of them (90%). The 90% plaintiff win rate is substantially higher than that in other types of cases.

Table 2 drives home the relative rarity of settlement in Taiwan. We see settlements in less than a quarter of all cases.[15] This is far less than in the U.S., where settlement rates in federal court tend to exceed 60% (Hadfield 2004; Eisenberg and Lanvers 2009: 143).

Judgment outcomes after trial in Taiwan also differ from the corresponding patterns in the US. We also see that plaintiffs in Taiwan (like those in Japan, see Ramseyer and Nakazato (1989)) win more than 50% of the time at trial. This is in contrast to the U.S., where various empirical articles have found that, at least for many categories of cases, plaintiffs win less than 50% at trial (Kessler, Meites, and Miller 1996: 236). Finally, Table 3 provides the distribution of the cases in our data across levels of the court system and across geographic regions.

One important terminology note: Our data include courts of first and second instance. To avoid cluttering in writing, "plaintiffs" refer to both "plaintiffs" in the court of first instance and "appellants" in the court of second instance.

V. Hypotheses and Results

In this section, we identify a set of hypotheses that we draw from the literature and then test them in our data. These hypotheses allow us to test predictions of the original Priest and Klein (1984) model, extensions to the model that introduce risk aversion, and competing, canonical models based on one-sided asymmetric information.

In this section, we initially organize our discussion around the two types of empirical tests implicated by the hypotheses: first, identifying factors that predict the settlement rate; second, identifying factors associated with plaintiff win rates closer to 50%. After conducting these tests, we then reorganize our results: we sort them based on which theories or models of litigation they shed light on. This allows us to make progress is assessing the empirical relevance of canonical models of suit and settlement. (To preview, these results are summarized in Table 8, which we'll discuss in greater detail at the end of this section.)

---

[15] The high trial rates and low settlement rates were about at the same levels in 1996–2006 (Huang 2009: 262–263).

A. Hypotheses

1. *Hypotheses related to rate of trial and settlement*

We have identified a pair of testable predictions about the determinants of the rate of trial and the rate of settlement. The canonical LPG divergent expectations model predicts that higher stakes make trial more likely. To be more precise, as stakes rise relative to the cost of litigating, small differences in the parties' beliefs about their likelihood of winning are more likely to render the defendant's settlement willingness to pay less than the plaintiff's willingness to accept. To the contrary is the view that risk-averse parties will tend to settle rather than litigating when stakes are high. This latter view does not reject the LPG wholesale, but rather hypothesizes that risk aversion will outweigh the effect of mutual optimism on the relationship between stakes and settlement. Thus, we state our first hypothesis in the alternative:

H1(a): Higher stakes will be associated with lower settlement rates.

H1(b): Higher stakes will be associated with higher settlement rates.

We note that it is possible that both effects will be empirically relevant, and for some levels of stakes, one effect may dominate the other.

2. *Hypotheses related to the 50% prediction from the PK model*

We have identified three testable predictions about the tendency of plaintiff win rates to approach or diverge from 50%.

First, in the spirit of Waldfogel's (1995; 1998) conceptual approach, we hypothesize that when comparing courts and case types, those with lower trial rates will have plaintiff win rates closer to 50%. We also note the contrary prediction from AI models, that lower trial rates will tend to be associated with plaintiff win rates closer to 0% or 100% (see supra Part II). This gives us our second hypothesis, stated in the alternative:

H2(a): When divided by court and nature of suit type, categories of cases with higher settlement rates will have plaintiff win rates closer to 50%.

H2(b): When divided by court and nature of suit type, categories of cases with higher settlement rates will have plaintiff win rates farther from 50%.

Second, Siegelman and Waldfogel (1999: 114) argue that information asymmetry is less likely to be a major factor in contract, trademark, and labor cases, since relevant actions by defendants are observable by both parties. By contrast, in the absence of discovery (which is not available in Taiwan), tort defendants know the amount of care they took, while plaintiffs do not. This suggests that the Priest-Klein prediction should be stronger in contract, trademark, and labor cases than in tort cases. As a consequence, plaintiff win rates in contract cases should be closer to 50% than those in tort cases. This argument gives us our third hypothesis:

H3: Plaintiff win rates will be closer to 50% in contract cases than in tort cases.

Finally, as noted above, Kessler, Meites, and Miller (1996: 246) argue that high stakes cases will involve more settlement, as risk aversion will push the parties towards settlement when stakes are higher. This implies that in high stakes cases, win rates of plaintiffs should be closer to 50%, as compared to low-stake cases. This gives us our fourth and last hypothesis:

H4: Plaintiff win rates will approach 50% as stakes rise.

B. Results

In this section, we present our regression results, accompanied by graphical illustrations, that we use to assess how these hypotheses fare in the data.

1. *Hypotheses related to rate of trial and settlement*

Table 4 presents a series of regressions that test our first pair of hypotheses. We use linear probability models (i.e., OLS) with robust standard errors clustered by court regions. The key variables of interest are stakes (log transformed) and party identity (corporation or individual). Also included in the regressions is a standard set of fixed effects that we use for all of the regressions in this paper: 3 types of proceedings (small-claim, summary, and ordinary); 2 court levels (first instance and second instance, i.e., appeal); 12 case types, 7 court regions, and 18 years of case filing (1998–2015).[16]

[16] In unreported results, we have re-run all regressions for first-instance courts only. Unsurprisingly, given that second-instance court represent less than 10% of the sample, all qualitative results are identical, and point estimates and p-values change very little. We have

Panel A of Table 4 includes cases with the full range of stakes, while Panel B incudes only high-stakes cases arguably comparable to US federal-court cases, which tend to involve far higher stakes than the typical case in Taiwan (or in US state court). Here, we limit our sample to cases with stakes in excess of USD 75,000, which corresponds to the amount-in-controversy requirement for diversity jurisdiction cases in US federal court.[17]

Specifications across columns vary the inclusion of covariates and the inclusion of loan contract and debt payment cases (hereinafter debt collection cases). Debt collection cases represent a disproportionate share of all cases, and we have *a priori* reasons to doubt that prevailing theories of suit and settlement apply to them: most of these actions involve efforts to collect on unpaid debts (most often, unpaid credit card debts) in which there is little or no uncertainty or asymmetric information about the existence of the debt or the amount due. Further, unlike other types of litigation, which tends to involve individual plaintiffs, debt collection cases overwhelmingly involve corporate plaintiffs suing individual defendants.[18] Finally, based on statistics from the Judicial Yuan on "money borrowing" cases, a category that corresponds to what we have labeled "debt collection" cases, roughly 70% of money borrowing cases end with default judgments, whereas only about 17% of other types of cases ended this way.[19] Of the money borrowing cases, fully 95% of them were decided 100% in favor of the plaintiff. For these reasons, we consider our results excluding debt collection cases to be our preferred specifications for regressions, and our figures also exclude these cases. Nonetheless, we note that the results are not particularly sensitive to this choice.[20]

Our results here provide some support for the basic LPG prediction, but not the

---

also tested the robustness of our results to our exclusion of cases with counterclaims or multiple parties. These exclusions reduced our sample size from about 620,000 to about 450,000 observations. Nonetheless, all results herein are robust to this change in specification, with virtually no changes in statistical significance levels or overall magnitudes (in fact, many results are stronger). See Appendix B.

[17] See 28 U.S.C. § 1332(a). We use an exchange rate of 30 NTD to 1 USD, which is (rounded to the nearest NTD), current as of January 3, 2018.

[18] In our data, 86% of the debt collection cases involve corporate plaintiffs suing natural-person defendants, while only 22% of other cases do.

[19] We acquired the default judgment statistics from the Statistics Department of the Judicial Yuan. Both we and the Department use the same raw data to combine cases into dispute types. We are confident that money borrowing cases as identified by the Department highly overlapped with the debt collection cases defined by us.

[20] Specifications including these cases are reported herein, and figures including these cases are available from authors upon request.

prediction that risk aversion will lead to higher settlement rates as stakes rise. The relationship between stakes and settlement appears to be negative. We illustrate in Figure 1, which groups cases by their stakes, and then shows, for each range of stakes, the average settlement rate among those cases. As Figure 1 makes remarkably clear, within each proceeding type (small, summary, and ordinary procedure), settlement rates fall as stakes rise.[21] Recall that the LPG model predicts that settlement rates fall as stakes rise as compared to costs. Stakes in Figure 1 rise steadily as one moves left-to-right, while, within each proceeding type, costs remain roughly constant (recall that most attorneys charge a flat, fixed fee, with rates keyed to the proceeding type and court instance), but costs rise discretely as one moves from small to summary, or summary to ordinary, proceedings.

More specifically, in the summary and small-claim procedures, plaintiffs are allowed to make oral claims. Essentially, plaintiffs can externalize the costs of formulating their claims formally to the court; therefore, they do not need to hire attorneys. By contrast, for plaintiffs in the ordinary procedure, written claims are required. In addition, in the summary and small-claim procedures, in principle, parties will meet only once for trials. If one party fails to show up, judges have discretion to issue default judgments. By contrast, in the ordinary procedure, there is no limit on trial days, and the use of default judgments is more constrained. These differences again ensure that the summary and small-claim procedures are less costly than the ordinary procedure for parties. Furthermore, the small-claim procedure is even less costly than the summary procedure, because in the former, plaintiffs may use a tabular form to make their claims; trials may be conducted in the evening or in weekends; and under certain circumstances, judges may skip evidence investigation and act more like arbitrators.

In the regression results, the overall relationship between stakes and settlement is fairly consistently negative across specifications. In most specifications, especially among higher-stakes cases, the negative relationship is statistically significant (at the 1% level) and consistent in magnitude.[22] Thus, we consider H1(a) supported by

---

[21] Unreported graphs show that the same pattern (settlement rates decreasing within proceedings) holds in all 7 court regions.

[22] Because risk aversion may affect individuals more so than corporations, we re-run these results including only cases between individuals. See Table B2. Our results hold even for this subset of the data. In our robustness check including counterclaim cases and multiple party cases, the results are even more consistently negative and statistically significant. See Table B1.

the regression results. Conversely, we do not find support for Hypothesis 1(b) in the data.

2. *Hypotheses related to the 50% prediction of the PK Model*

We now test our hypotheses related to the famous 50% prediction of the PK model and competing hypotheses generated by other theoretical approaches.

We set out to test H2 by dividing our data into groups. Many grouping are possible; for consistency, we create groups using the categories that served as controls throughout the paper.[23] We divide all our cases into bins based on court procedure (three procedures), appellate level (first and second instance), court (7 regions), and case category (12 subject-matter categories). This yields 504 bins (some of which may be empty in the data), each of which contains all cases with a unique combination of procedure, level, court, and case type. Within each bin, we computed the settlement rates and plaintiff win rates. Settlement rates are the number of settlements divided by the number of filed cases that were either settled or decided by judges. Plaintiff win rates are the number of plaintiff wins divided by the number of court decisions (i.e., non-settlements).

We hypothesized (H2(a)) that plaintiff win rates will approach 50% as settlement rates rise. To test this graphically, we plotted settlement rates against win rates. In Figure 2, each circle represents one bin, with the size of the circle proportional to the number of court decisions in the bin.

Given the low overall settlement rates in Taiwan compared to those in the US, it is not surprising that even bins with relatively high settlement rates have a wide range of plaintiff win rates. Nonetheless, Figure 2 reveals the predicted pattern: as settlement rates increase, plaintiff win rates do, in fact, converge towards 50%. We must remark that these qualitative results reveal a pattern quite similar to the graphical results that Waldfogel (1995) finds with US data.

Regression analysis supports these qualitative conclusions. Table 5 reports results of regressions where each observation corresponds to a bin defined by the same set of controls used for other regressions: procedure, level, court, case type, and

---

[23] In our figures, we group cases by court procedure, court level, court region, and case category (all of which correspond to the fixed effects we include in our regressions), but for graphical clarity, we do not also group by filing year (which we do include as fixed effects in our regressions). This would multiply the number of possible categories by a factor of 18, making most individual observations impossible to distinguish in a scatterplot.

filing year. The outcome variable for the regression is the absolute difference between the plaintiff win rate for that observation and 0.50. As the settlement rate rises, H2(a) predicts that plaintiff win rates will converge to 0.50, and thus we predict a negative relationship between settlement rates and the outcome variable. This is what we find, at least for regressions including all levels of stakes. For high stakes cases, we have a null result.[24] We interpret this as some support for H2(a): the famous Priest-Klein 50% hypothesis. Conversely, this evidence tends to reject H2(b), the prediction based on asymmetric-information, rather than divergent-expectation, models.

We apply a similar approach to test H3, that plaintiff win rates will be closer to 50% in contract, trademark, and labor cases (where one might conjecture that asymmetric information is less) than in tort cases (where one might conjecture that asymmetric information is greater). Notably, these assumptions seem *more* plausible in the Taiwanese context than the US context, because limited discovery in Taiwan means that asymmetries of information can persist throughout the litigation process, unlike in the US, where broad discovery makes claims about the strong impact of asymmetric information in the decision to go to trial much less plausible. Figure 3 employs the same strategy as Figure 2, except that it additionally distinguishes cases categorized as tort from those categorized as contract. Figure 3, to our eyes, is hard to interpret. Table 6 reports regression results, and here we see relationships consistent with the hypothesis that tort cases (our proxy for AI cases) have plaintiff win rates farther from 50%. Our inference from these results, however, is qualified by the fact that these results—unlike our other results—are not robust to changes in the definition of plaintiff win. In fact, if we treat partial wins as half-wins, the results flip (see Table C2, but compare Table B2, which is consistent with Table 6). Thus, we treat these results as inconclusive, or perhaps offering H3 some support.

Finally, we test H4, that the plaintiff win rate will approach 50% as stakes rise. This hypothesis is derivative of H1(b), which we have rejected. In Figure 4, as in Figure 1, we divide the data by stakes into groups. In each group, we separate out the first-instance and the second-instance (appellate) cases, and then compute the plaintiff win rate for each group. As Figure 4 makes clear, plaintiff win rates converge to 50% as stakes rise.

Figure 5 replicates Figure 2, except that the horizontal axis now measures the average (logged) stakes in each bin, rather than the settlement rate. Here, too, we see

[24] Note that the findings remain qualitatively similar in our robustness checks. See Tables B4 and C1.

clear convergence towards 50% win rates as stakes rise.

Table 7 provides regression results. Although not all specifications yield statistically significant results, all of them predict that plaintiff win rates will approach 50% as stakes rise.[25] This supports the conclusion that there is a negative relationship between stakes and deviations from a 50% win rate. Taken together, this evidence supports H4.

3. *Summary of Results*

Table 8 summarizes our findings in this paper. Rather than organizing the results around the outcomes studies (as we did in the previous sections), here we organize the hypotheses and results based on the models that were sources of the predictions. In this way, our results allow us to evaluate these models empirically.

Viewed in this light, the PK model gains support. We see support for the prediction of the LPG model (whose logic underlies the PK model) that higher stakes will be associated with lower settlement rates (H1(a)). Most notably, Priest and Klein's famous 50% prediction (H2(a)) holds up. We see a tendency toward 50% win rates as settlement rates rise, even in data from a court system where the trial rate never gets close to zero.

We see that the hypotheses based on canonical asymmetric information models garnered less support. The competing prediction that selection will tend to drive win rates away from 50% (H2(b)) is based on canonical models of litigation and settlement under one-sided asymmetric information. This prediction is rejected by the data. Hypothesis 3, which looked for evidence that cases with symmetric information (contract rather than tort) exhibit patterns closer to the PK model fared better, although the results were sensitive to how partial wins were coded. While not a prediction of asymmetric information model, Hypothesis 3 nonetheless implied that the PK model would have less empirical relevance in contexts with asymmetric information. Our results here, however, are tentative at best.

Perhaps the most intriguing result is the very strong pattern that plaintiff win rates approach 50% as stakes rise (H4). In the PK model, win rates approaching 50% are correlated with higher settlement rates, and this is what we find in the data. Further, in the PK model, higher stakes are correlated with lower settlement rates, and this, too, is what we find in the data. If both of these are true, why then do we see

---

[25] Our findings are qualitatively similar in our robustness checks. See Tables B6 and C3.

in our data that higher stakes are correlated with win rates approaching 50%?

We cannot definitively answer this question with the data. One conjecture is that the relationship between higher stakes and win rates closer to 50% is being driven primarily by the negative relationship between stakes and win rates in first-instance cases. (See the top series in Figure 4, and recall that over 90% of the sample is first-instance cases.) This negative relationship, rather than being a puzzle, may simply be yet more evidence in favor of the LPG model.[26] Here is why: in the LPG model, a plaintiff files suit only if the expected value of the suit (stakes times probability of winning) exceeds the cost of litigation. For low-stakes cases, this will likely be the case only when likelihood of winning is high. For high-stakes cases, expected value may exceed cost even if the plaintiff's chances of winning are low. This logic predicts exactly the pattern we see in the top series in Figure 4.

This reasoning rationalizes nearly all of our results, but not quite all. The same logic would predict that the lower series in Figure 4—win rates among cases brought in second-instance courts—should also have the same negative relationship. After all, the decision to appeal entails the same cost-benefit analysis as the decision to file in the first instance. (And keep in mind that lawyers in Taiwan general charge separately for appeals.) Thus, we are left with intriguing, but hardly conclusive, explanations for the overall pattern of results in our data. Nonetheless, we see the results as largely consistent with the predictions of the LPG and PK models.

VI. Conclusion

Using a new, comprehensive dataset of civil cases from Taiwan, we tested a series of hypotheses reflecting the predictions of the PK model and the related literature. Some hypotheses were strongly supported by the data, some were weakly supported, and some hypotheses were rejected. Overall, though, we interpret our results as largely finding support for the predictions of the PK model. We interpret our results as less consistent with predictions based on canonical models of asymmetric information.

These results augur favorably for the PK model and come at a time when the theoretical literature on litigation and settlement, after years of focusing on models of asymmetric information, is experiencing something of a "renaissance" for the PK model (see, e.g., Klerman and Lee 2014; Gelbach 2016) and for symmetric

[26] We must credit Dan Klerman for this insight.

information models more generally (see, e.g., Hubbard 2016). In contrast, predictions based on AI models (H2(b) and H3) generate more mixed results. This does not undermine the conceptual soundness of these models, but it does undercut their empirical relevance to explaining broad patterns of court practice.

To be sure, these results are hardly conclusive, and they represent only a test of a few implications in the context of one legal system. Before we can translate these results into normative or prescriptive conclusions, further empirical testing and model development is necessary.

Nonetheless, we emphasize two reasons to give weight to our results. *First*, they reinforce the generally favorable overall results of earlier empirical studies on the PK model, and do so using entirely new data in an entirely new context.[27] In this way, our data amplify earlier findings because they provide a stronger test of the external validity of the PK model than most prior studies, many of which have relied on subsets of the same US federal court administrative data. *Second*, one would expect the Taiwanese context to be especially favorable for AI models. In the US context, broad pre-trial discovery imposes an environment of nearly symmetric information upon the parties well in advance of trial. This gives one reason to doubt, a priori, the empirical relevance of models whose central feature is one-sided information asymmetry. In contrast, the much more limited requirements of disclosure in Taiwanese litigation suggest an environment in which asymmetric information could persist throughout most of the civil-law trial process. For this reason, too, our results present a tougher, more informative test of the PK model.

Importantly, though, the limited support of the AI models tested here does not reject asymmetric information models generally. Indeed, as noted earlier, the PK model itself can be understood as a model of two-sided private information. Thus, one implication of our results is that theoretical work on litigation may benefit from a focus on less elegant, but perhaps more descriptively powerful, models based on two-sided incomplete or private information.

---

[27] We note that unlike the PK model, which has been aggressively challenged through empirical tests, AI models of litigations have generated few testable empirical prediction and even fewer empirical tests. Indeed, one drawback of many existing models is that they were developed as tools for developing conceptual insights about the legal system, not as positive theories of how the legal system works. Thus, these models are well suited as pedagogical tools, but less suited as guides to practical policy. One of the challenges for theoretical work in this area is to develop models that generate explicit, crisp empirical predictions that can be tested. Models that succeed at better describing reality should have greater weight in policy-relevant analysis.

We believe these results contribute the growing literature that seeks to identify the empirical implications of influential theories and then test those predictions against data. And we believe that empirical tests of theories of litigation versus settlement with non-U.S. data may serve as especially rigorous tests of the predictions of theories developed in the context of the US legal system.

Tables

TABLE 1. SUMMARY STATISTICS, CONTINUOUS VARIABLES

Variable	Mean	Median	Max	Min	N
Stakes (000s NTD)	843.8	168.5	33,950	6.57	476,709
Year Filed	2012.3	2012	2015	1998	473,229
Year Terminated	2012.5	2013	2015	2010	476,709
Case Duration (Days)	97.04	58	5167	0	476,709

TABLE 2. SUMMARY STATISTICS, DUMMY VARIABLES

Variable	Mean	N=476,709
Any Party Appeals	0.10	
Corporate Plaintiff/Appellant	0.63	
Corporate Defendant/Appellee	0.13	
Represented Plaintiff/Appellant	0.26	
Represented Defendant/Appellee	0.22	
Outcomes		N=476,709
Settled	0.24	
Plaintiff/Appellant Win	0.53	
Plaintiff/Appellant Partial Win	0.12	
Plaintiff/Appellant Loss	0.11	

Note. Outcomes (including "procedural dismissal", "transfer jurisdiction", and some other procedural management) not reflecting settlement or a judgment on the merits are excluded from this table.

TABLE 3. SUMMARY STATISTICS, COURT VARIABLES

Variable	Mean	N
Court Procedure and Level		476,709
Small Claim, 1st Instance	0.34	
Small Claim, 2d Instance	0.004	
Summary Procedure, 1st Instance	0.39	
Summary Procedure, 2d Instance	0.03	
Ordinary Procedure, 1st Instance	0.19	
Ordinary Procedure, 2d Instance	0.05	
Court Region		476,709
Taipei	0.35	
North	0.26	
Central	0.16	
South: Tainan	0.09	
South: Kaohsiung	0.13	
East	0.02	
Island	0.002	

TABLE 4. PREDICTORS OF SETTLEMENT RATE

Panel A. All Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.011	–0.013	–0.015**	–0.015**
	(0.007)	(0.006)	(0.002)	(0.003)
Corporate Plaintiff		–0.047**		–0.013
		(0.010)		(0.010)
Corporate Defendant		0.023*		0.001
		(0.009)		(0.004)
Include "debt collection" cases	Y	Y		
N	453,045	453,045	214,973	214,973

Panel B. High Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.019**	–0.019**	–0.020**	–0.021**
	(0.002)	(0.002)	(0.002)	(0.002)
Corporate Plaintiff		–0.008		0.019
		(0.006)		(0.008)
Corporate Defendant		0.003		–0.007
		(0.007)		(0.008)
Include "debt collection" cases	Y	Y		
N	34,115	34,115	29,558	29,558

Note. High stakes cases are cases with stakes exceeding USD 75,000. All columns include fixed effects for procedure, level, case type, court, and filing year. Linear regressions including cases with one plaintiff against one defendant. Robust standard errors (in parentheses) clustered on court region. ** p<0.01, * p<0.05.

TABLE 5. RELATIONSHIP BETWEEN SETTLEMENT RATES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%

Variable	(1)	(2)	(3)	(4)
Settlement Rate	–0.220**	–0.115**	0.037	0.035
	(0.016)	(0.020)	(0.032)	(0.033)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	2,893	2,360	1,204	1,022

Note. High stakes cases are cases with stakes exceeding USD 75,000. Outcome variable is the absolute value of the difference between the plaintiff win rate and 0.5. Each observation includes the average settlement rate and the absolute deviation from 0.5 of the average plaintiff win rate for a category of cases. Observations (categories) weighted by number of court decisions within the category. Categories are defined by court, procedure, level, and case type. All columns include fixed effects for procedure, level, case type, court, and filing year. Robust standard errors in parenthesis. ** $p<0.01$, * $p<0.05$.

TABLE 6. RELATIONSHIP BETWEEN CASE CATEGORY (PROXY FOR AI) AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%

Variable	(1)	(2)	(3)	(4)
Tort (Asymmetric Info)	0.035**	0.036**	0.086**	0.090**
	(0.005)	(0.004)	(0.010)	(0.010)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	1,896	1,363	799	617

Note. High stakes cases are cases with stakes exceeding USD 75,000. Outcome variable is the absolute value of the difference between the plaintiff win rate and 0.5. Each observation includes an indicator for a proxy of asymmetric information (tort cases as opposed to contract, trademark, and labor cases), the average settlement rate and the absolute deviation from 0.5 of the average plaintiff win rate for a category of cases. Observations (categories) weighted by number of cases within the category. Categories are defined by court, procedure, level, and case type. Settled cases and cases other than tort, contract, trademark, and labor excluded. All columns include fixed effects for procedure, level, case type, court, and filing year. Robust standard errors in parenthesis. ** $p<0.01$, * $p<0.05$.

TABLE 7. RELATIONSHIP BETWEEN STAKES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%

Variable	(1)	(2)	(3)	(4)
Ln Stakes	–0.096**	–0.016**	–0.002	–0.008
	(0.005)	(0.006)	(0.017)	(0.018)
Include "debt collection" case type categories	Y		Y	
High Stakes Cases Only			Y	Y
N	2,893	2,360	1,204	1,022

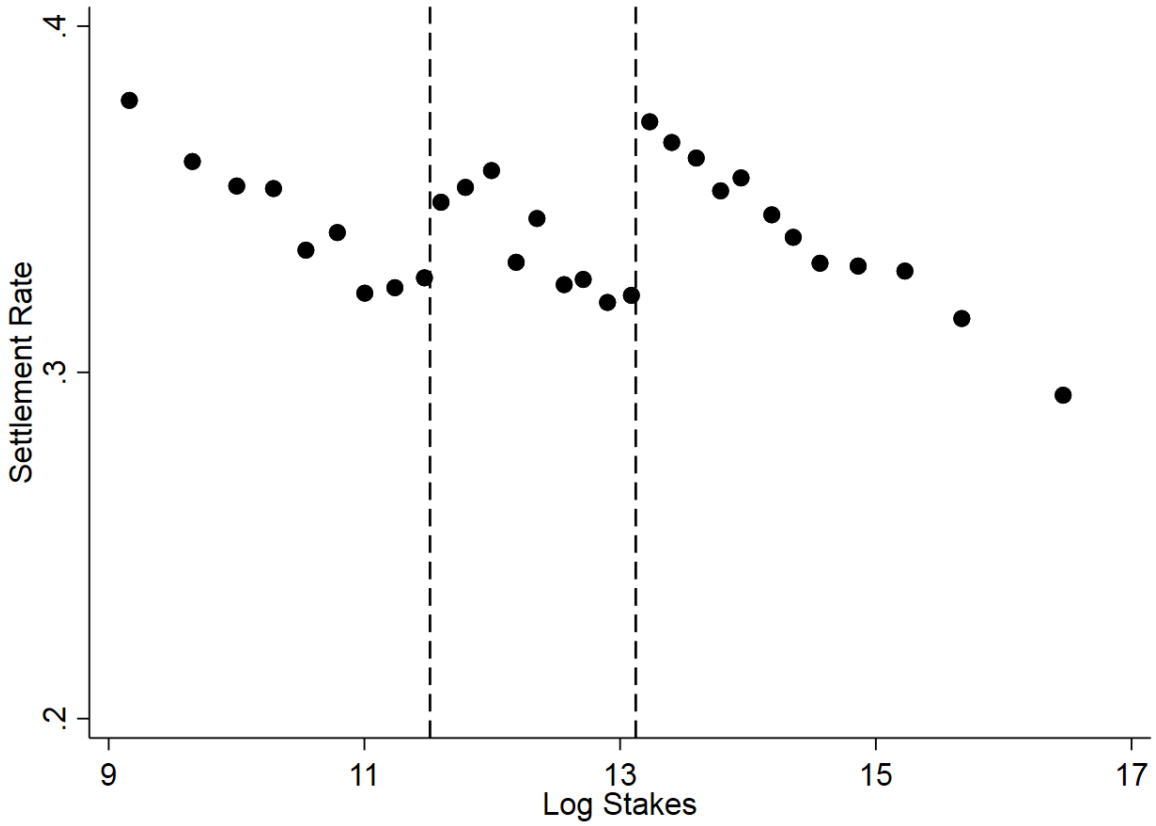
Note. High stakes cases are cases with stakes exceeding USD 75,000. Outcome variable is the absolute value of the difference between the plaintiff win rate and 0.5. Each observation includes average log stakes, the average settlement rate and the absolute deviation from 0.5 of the average plaintiff win rate for a category of cases. Observations (categories) weighted by number of cases within the category. Categories are defined by court, procedure, level, and case type. Settled cases excluded. All columns include fixed effects for procedure, level, case type, court, and filing year. Robust standard errors in parenthesis. ** $p<0.01$, * $p<0.05$.

TABLE 8. REVIEW OF RESULTS OF HYPOTHESIS TESTING

Model	Hypothesis	Result
PK Model	H1(a): Higher stakes ↔ lower settlement rates	Some Support
	H2(a): Higher settlement rates ↔ plaintiff win rates closer to 50%	Some Support
Competing Model: AI	H2(b): Higher settlement rates ↔ plaintiff win rates farther from 50%	Rejected
	H3: Contract (compared to tort) cases ↔ plaintiff win rates closer to 50%	Some Support
Competing Model: Risk Aversion	H1(b): Higher stakes ↔ higher settlement rates	Rejected
	H4: Higher stakes ↔ plaintiff win rates closer to 50%	Supported

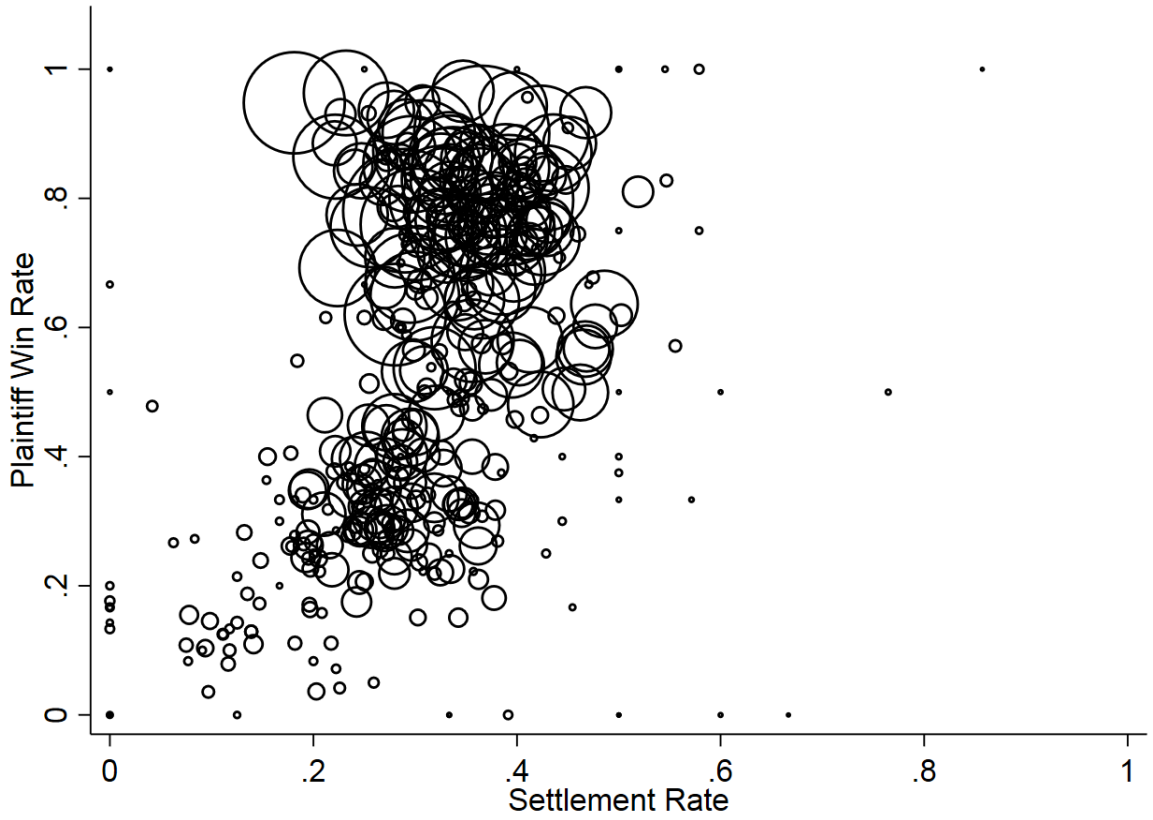
Figures

FIGURE 1. RELATIONSHIP BETWEEN SETTLEMENT AND STAKES



Note. Binned scatterplot representing the average log stakes and settlement for 30 quantiles of stakes. Vertical lines indicate the cutoffs for eligibility for summary procedure (NTD 100,000; logged value 11.5) and for ordinary procedure (NTD 500,000; logged value 13.2). All figures exclude loan contract and debt collection cases.

FIGURE 2. RELATIONSHIP BETWEEN SETTLEMENT RATE AND PLAINTIFF WIN RATE, BY PROCEDURE, LEVEL, COURT, AND CASE CATEGORIES

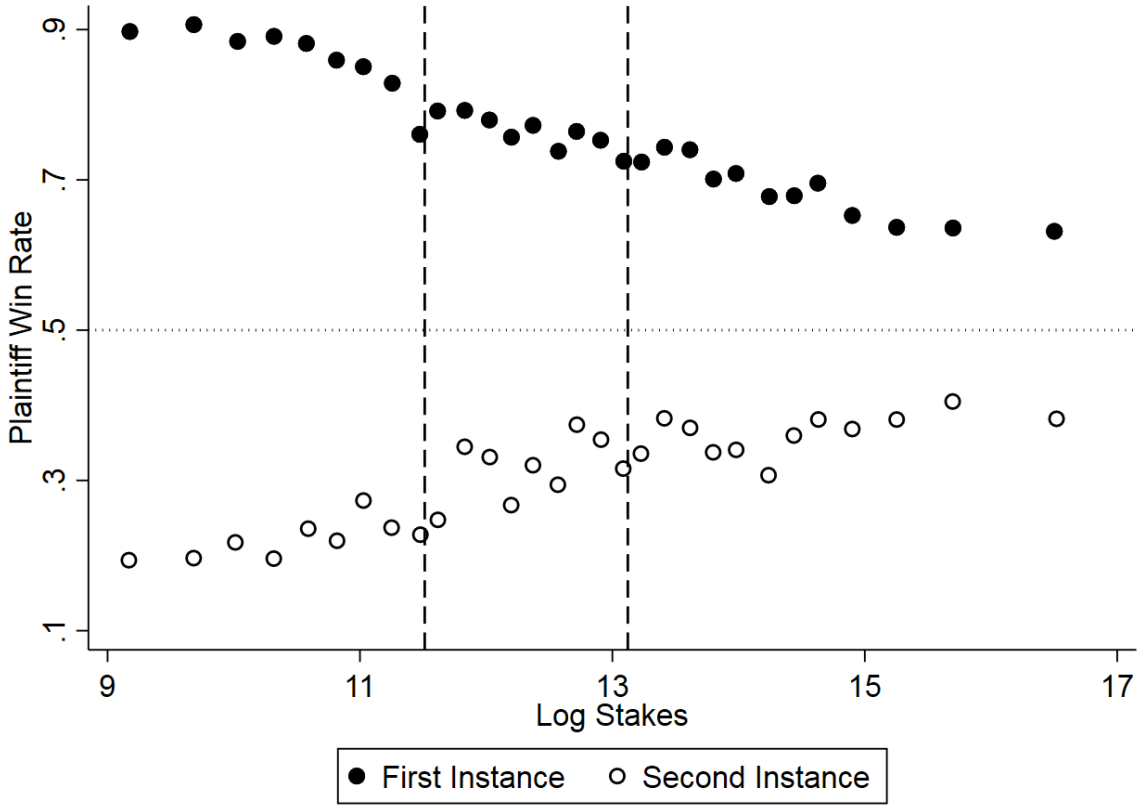


Note. Scatterplot of 422 observations of the settlement rate and plaintiff win rate among cases within the procedure/level/case-category. Size of circle is proportional to the number of court decisions represented by each observation.

FIGURE 3. RELATIONSHIP BETWEEN SETTLEMENT RATE AND PLAINTIFF WIN RATE, COMPARING CONTRACT (LOW AI) AND TORT (HIGH AI) ACTIONS BY REGION, PROCEDURE, LEVEL, AND CASE CATEGORIES

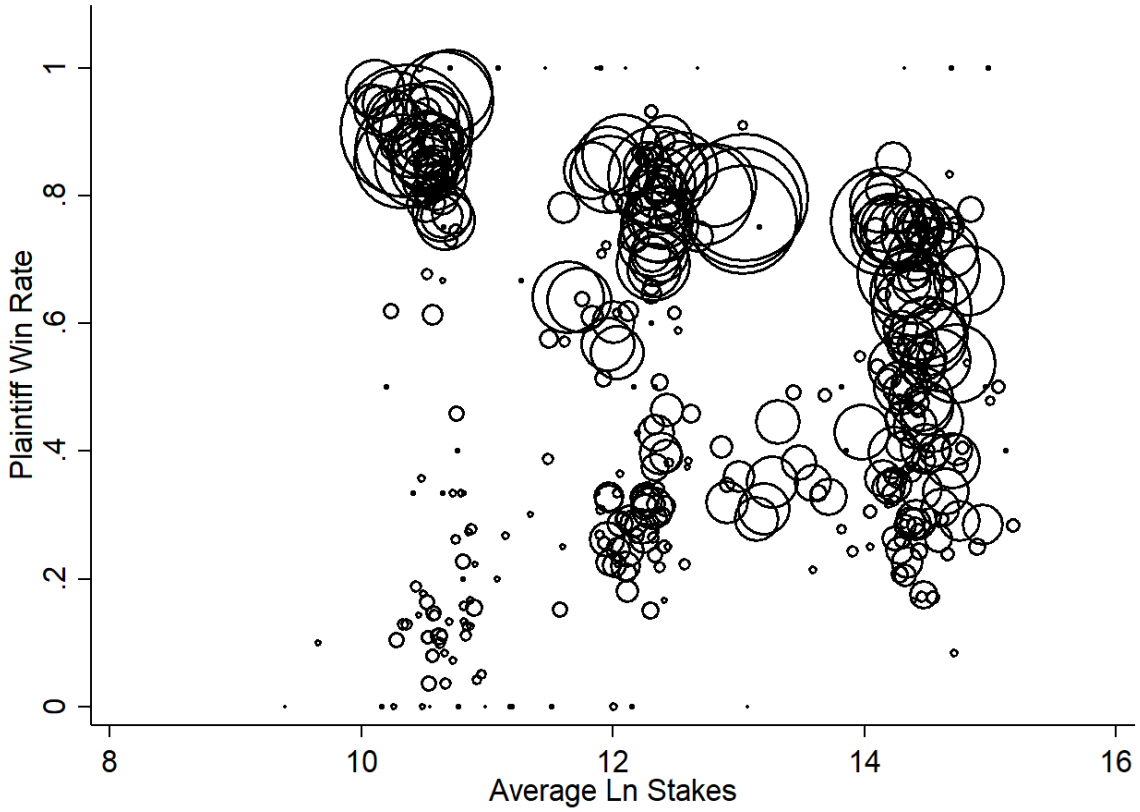
Note. Scatterplot of 204 observations of the settlement rate and plaintiff win rate among cases with and without both parties represented, by region/procedure/level/case-category. Size of circle is proportional to the number of court decisions represented by each observation.

FIGURE 4. RELATIONSHIP BETWEEN STAKES AND PLAINTIFF WIN RATE, BY STAKES AND COURT LEVEL



Note. Binned scatterplot representing the average log stakes and settlement for 30 quantiles of stakes. Within each quantile, averages for observations from first instance courts are represented separately from averages for observations from second instance courts.

FIGURE 5. RELATIONSHIP BETWEEN LOG STAKES AND PLAINTIFF WIN RATE, BY REGION, PROCEDURE, LEVEL, END YEAR, AND CASE CATEGORIES



Note. Scatterplot of 414 observations of the settlement rate and plaintiff win rate among cases within the procedure/level/case-category. Size of circle is proportional to the number of cases represented by each observation.

Appendix A: Additional Hypotheses Regarding Party Status and Attorney Representation

In this Appendix, we present additional hypotheses grounded in the empirical literature on litigation and settlement, but which do not bear directly on testing the PK model or the canonical AI models. These additional hypotheses explore the empirical relationships between party status (individual, corporation) or attorney representation, on the one hand, and settlement rates or win rates, on the other hand.

A. Hypotheses

1. *Hypotheses related to rate of trial and settlement*

First, Eisenberg and Farber (1997: S93) argue that because plaintiffs are drawn from individuals with the highest tastes for litigation (or, equivalently, the lowest psychological cost of litigation), conditional on the expected value of claim, individual plaintiffs, as compared to corporate plaintiffs, have higher trial rates and lower win rates. Like H2, this prediction reflects an extension of the PK model to incorporate an additional behavioral assumption—in this case, that individual plaintiffs have a higher degree of optimism about their chances than corporate plaintiffs. This gives us our hypothesis A1:

> **H(A1): Individual plaintiffs will have lower settlement rates than corporate plaintiffs.**

Second, Siegelman and Waldfogel (1999: 114–116) posit that when both parties to a suit are represented, they face lower legal uncertainty and thus are more likely to reach a settlement. This is because, in the PK framework, legal uncertainty increases the variance in the parties' estimates of the plaintiff's probability of winning at trial, which in turn makes it more likely that both parties will be optimistic about their chances at trial. This gives us our hypothesis A2:

> **H(A2): Settlement rates will be higher when both parties are represented than when at least one party is pro se.**

We note, however, that testing this hypothesis in our data faces potentially confounding factors. Unlike stakes and the individual or corporate status of the parties, the presence of legal representation in the case is not exogenous to the

dispute. Rather, the decision to hire attorneys is in some sense an *outcome* of a dispute, in the same way that an agreement to settle is an *outcome* of a dispute—it is something that is determined by the parties after the dispute arises. Thus, even if, in a world where parties are randomly assigned lawyers, those with lawyers would be more likely to settle, other factors may lead to an opposite correlation in our data. For example, parties may only hire lawyers in cases that they perceive as "close" cases, or only in cases where the parties see little hope for settlement. Further, the presence of lawyers in a case is likely a strong proxy for the parties' perceived stakes of the case. If so, the presence of lawyers may be associated with less settlement, rather than more.

2. *Hypotheses related to which party has higher win rates*

We have identified two testable predictions about the determinants of higher plaintiff win rates for different categories of parties or claims.

First, we hypothesize that party representation will be associated with higher win rates. The presence of lawyers could proxy for party stakes (parties with more to lose are more likely to retain lawyers) or could proxy for party sophistication (see Kessler, Meites, and Miller (1996: 242–43) and Hylton (1993: 205))—or both. As a consequence, this hypothesis is consistent with both the PK model and AI models with one-sided asymmetric information.

To the extent that lawyers are a proxy for stakes, parties with lawyers will tend to have higher win rate than unrepresented parties because those parties have higher stakes, consistent with an extension to the PK model that adds the assumption that lawyers proxy for stakes. As Siegelman and Waldfogel (1999: 109) and Eisenberg and Farber (1997: S98) argue, parties with higher stakes will invest more in winning, and thus will tend to have higher win rates. Hiring lawyers is an obvious investment in this regard. In the same vein, Kessler, Meites, and Miller (1996: 237–242) argue that the party with the greater stakes is likely to win more often, as it has more to lose from litigation and will offer enough to avoid a larger loss at trial. Thus, the litigated cases are likely to be ones in which the party with the greater stakes is more likely to win than would be the case when the stakes are equal.

To the extent that lawyers are a proxy for party sophistication, the presence of lawyers on one side of a case indicates information asymmetry, and AI models predicts that the party with the information advantage will win more than 50% of the

time.[28] This is our hypothesis A3:

> **H(A3): Represented parties will have higher win rates than unrepresented parties.**

Second, Siegelman and Waldfogel (1999: 109) and Eisenberg and Farber (1997: S98) add the assumption that, on average, corporations have higher stakes in litigation because, as repeat players, their stakes include not merely the damages claimed, but reputational concerns as well. In contrast, they argue that individual plaintiffs are more likely to be "one-shotters." Corporations, however, may also be more sophisticated parties, as noted above, and have the benefit of information asymmetries. Thus, as with HA3, our prediction that corporations will tend to have higher win rates in litigated cases is consistent with versions of both the PK and AI models. This is our hypothesis A4:

> **H(A4): Corporations will have higher win rates than individuals.**

3. *Hypothesis related to the 50% prediction from the PK model*

As noted earlier, Siegelman and Waldfogel (1999) argue that cases where both sides are represented, as opposed to pro se, are more likely to settle. This yields the corollary that plaintiff win rates should be closer to 50% when both sides are represented than when at least one party is pro se. When both parties have professionals on their side, they are more likely to bring only cases of high legal uncertainty to trials. Pro se parties, on the other hand, may litigate over easy cases with predictable outcomes (i.e., outcomes predictable only to lawyers) (Siegelman and Waldfogel 1999: 114). This yields our hypothesis A5:

> **H(A5): Plaintiff win rates will be closer to 50% in cases where both parties are represented than cases where at least one party is pro se.**

---

[28] We have set aside the most obvious potential reason for this claim, which is that lawyers help one win. The empirical implications of this assumption are ambiguous; parties may only seek a lawyer when they have a tough case, thereby attenuating the observed correlation between represented parties and win rates. Further, retaining a lawyer could also change the likelihood of settlement (as hypothesized above), such that the net effect of retaining a lawyer depends on whether lawyers are more likely to settle strong cases or weak cases. Thus, regardless of whether lawyers have a causal effect on the plaintiff win rate, the empirical prediction depends crucially on the assumptions discussed above about the extent to which the presence of lawyers proxies for factors relevant to models of the selection of disputes for litigation or settlement.

B. Results

In this section, we present our regression results, accompanied by graphical illustrations, that we use to assess how these hypotheses fare in the data.

1. *Hypotheses related to rate of trial and settlement*

First, Tables 4 and Table A1 provide little evidence for HA1. Coefficients for corporate plaintiff are mostly negative, implying that individual plaintiffs have higher settlement rates, contrary to HA1, although most specifications yield no statistically significant result for this hypothesis. Overall, we consider this to be a null result, and conclude that Hypothesis A1 is unsupported by the data.

Second, the presence of lawyers on both sides of a lawsuit is consistently associated with lower settlement rates. Across all specifications in Table A1, the coefficient for "both parties represented" is large, negative, and statistically significant. This counterintuitive result strongly rejects HA2. Importantly, this hypothesis fairs equally poorly with high-stakes cases, for which one might otherwise expect lawyers would be most necessary to bring parties to a meeting of the minds on settlement. (These are also the cases most comparable to US federal court cases, which is the context that gave rise to the hypothesis in the first place). Panel B of Table A1 makes clear that this is not so.

Instead, this result may be driven by selection effects, i.e., parties only hire lawyers when the cases is particularly close, high-stakes, or contentious—all factors that make the case less likely to settle. In support of this view, Huang (2008: 217) reviews Taiwanese court data from 2000–2006 and argues that "it is a party's initial decision to litigate the case that leads to him or her retaining counsel."

2. *Hypotheses related to which party has higher win rates*

Table A2 presents a series of regressions that test our second set of hypotheses. We again report linear probability model (i.e., OLS) regression results with fixed effects. The dependent variable for these regressions is "plaintiff win." As noted in the main text, for our regressions, we recode partial wins as wins for the plaintiffs, such that both "plaintiff win" (plaintiff obtains all requested relief) or "partial win" (plaintiff obtains relief but less than the total amount claimed) are coded 1, and "plaintiff loss" is coded as 0. In unreported results (available from authors), we treat a partial win as "half a win," based on the notion that a partial win may reflect the

type of close case that the PK hypothesis would predict goes to trial. All results, except where specifically noted below, are the same in terms of sign and statistical significance.

These hypotheses fare better. First, HA3 predicts that represented parties will have higher win rates than unrepresented parties. In Table A2, we see that plaintiffs facing represented defendants have much lower win rates than plaintiffs facing unrepresented defendants; the estimates for "represented defendant" are negative, large in magnitude, and statistically significant. For represented plaintiffs, the estimates for "represented plaintiff" are consistently positive, and sometimes statistically significant. These results support HA3.

Second, we also see that corporations have higher win rates than individuals across all specifications, regardless of whether the corporation is the plaintiff or the defendant. These estimates are statistically significant, and all have the predicted sign; corporate plaintiffs have higher win rates than individual plaintiffs, and plaintiffs facing corporate defendants have lower win rates than plaintiffs facing individual defendants. This pattern of results supports HA4, i.e., the prediction that corporations (presumably the parties with greater sophistication and higher stakes) will have higher win rates than individuals (presumably the parties with less sophistication and lower stakes).

3. *Hypothesis related to the 50% prediction of the PK Model*

Finally, we test Hypothesis A5, that plaintiff win rates will be closer to 50% in cases where both parties are represented, using a similar approach. We divide all cases in groups to visualize the relationship between the settlement rate and the plaintiff win rate, but this time distinguish between cases where both parties are represented and those where they are not both represented. Figure A1 illustrates this relationship when dividing cases into bins by procedure, level, court, and case type. Circles represent bins where both parties are represented, and triangles represent bins where one or both parties are pro se. (Thus, for each combination of court region, procedure, appellate level, and case type, there are two bins: one for cases where both parties are represented, and one for other cases.) Our hypothesis predicts that the circles will tend to have plaintiff win rates closer to 50%; Figure A1 suggests this may be true.

We then apply the same regression approach to HA5 that we applied to H2 in the main text. Table A3 reports our results, and we find a consistently negative

relationship between both parties being represented and the deviation of the plaintiff win rate from 50%. In other words, the presence of lawyers is associated with win rates closer to 50%. This is consistent with the picture drawn by Figure A1. The strength of our inference from these results is tempered, however, by the fact that these results are not robust to changes in the definition of plaintiff win; when partial wins are counted as half-wins, the results are largely null. Thus, we conclude that these results provide some, but not unqualified, support for HA5, which predicts that litigated cases with represented parties will be a more highly selected, and "closer" set of cases than litigated disputes involving unrepresented parties.

4. *Summary of Results*

The prediction that individual plaintiffs would have lower settlement rates (HA1) was not supported. In other words, this data shows no patterns consistent with the claim that individuals are systematically more litigious than corporations. Indeed, one could have formed the opposite hypothesis, on the assumption that individuals are more risk-averse than corporations, and thus more likely to settle—but there is not much evidence for this, either.

Likewise, results based on theories about the role lawyers play are potentially confounded by the fact that the presence of lawyers may proxy for the complexity, value, or contentiousness of the dispute. This is likely why the results on lawyers were mixed. The prediction that both parties being represented leads to higher settlement rates (HA2) was rejected by the data, but the prediction that both parties being represented leads to plaintiff win rates closer to 50% (HA5) had some support in the data.

The remaining hypotheses were supported. To the extent that retaining a lawyer (HA3) or having corporate status (HA4) proxies for a party having higher stakes, the prediction of the PK model that the party with higher stakes will have a higher win rate is supported. (Of course, it could simply be that retaining a lawyer is not merely a proxy for stakes, but lawyers have a direct, causal impact on win rates!) We do not count these results as weighing against AI models, however, as they generate identical predictions in this context.

Appendix A Tables and Figures

TABLE A1. PREDICTORS OF SETTLEMENT RATE, WITH ATTORNEY REPRESENTATION

Panel A. All Stakes	(1)	(2)
Ln Stakes	–0.004	–0.004
	(0.006)	(0.003)
Corporate Plaintiff	–0.037**	–0.009
	(0.009)	(0.010)
Corporate Defendant	0.023*	0.006
	(0.007)	(0.005)
Both Represented	–0.170**	–0.191**
	(0.008)	(0.003)
Include "debt collection" cases	Y	
N	453,045	214,973

Panel B. High Stakes	(1)	(2)
Ln Stakes	0.004	0.002
	(0.003)	(0.003)
Corporate Plaintiff	0.003	0.024*
	(0.005)	(0.008)
Corporate Defendant	0.013	–0.001
	(0.008)	(0.009)
Both Represented	–0.241**	–0.243**
	(0.007)	(0.007)
Include "debt collection" cases	Y	
N	34,115	29,558

Note. See Table 4.

TABLE A2. PREDICTORS OF PLAINTIFF WIN RATE

Panel A. All Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.013**	–0.010*	–0.008*	0.002
	(0.002)	(0.004)	(0.002)	(0.002)
Corporate Plaintiff	0.129**	0.134**	0.094**	0.097**
	(0.013)	(0.014)	(0.010)	(0.010)
Corporate Defendant	–0.076**	–0.075**	–0.055**	–0.049**
	(0.010)	(0.011)	(0.005)	(0.007)
Plaintiff Represented		0.071**		0.030**
		(0.012)		(0.006)
Defendant Represented		–0.127**		–0.180**
		(0.011)		(0.001)
Include "debt collection" cases	Y	Y		
N	348,390	348,390	142,991	142,991

Panel B. High Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.007	0.011	–0.007	0.008
	(0.008)	(0.007)	(0.008)	(0.008)
Corporate Plaintiff	0.096**	0.090**	0.074**	0.079**
	(0.015)	(0.015)	(0.014)	(0.015)
Corporate Defendant	–0.062**	–0.055**	–0.050**	–0.048*
	(0.013)	(0.014)	(0.013)	(0.014)
Plaintiff Represented		0.002		0.017
		(0.011)		(0.013)
Defendant Represented		–0.200**		–0.194**
		(0.004)		(0.004)
Include "debt collection" cases	Y	Y		
N	28,378	28,378	20,174	20,174

Note. Cases in which courts did not give substantive verdicts (including those ending in settlement) are omitted. All columns include fixed effects for procedure, level, case type, court, and filing year. Robust standard errors (in parenthesis) clustered on court region. ** p<0.01, * p<0.05.

TABLE A3. RELATIONSHIP BETWEEN BOTH PARTIES REPRESENTED AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%

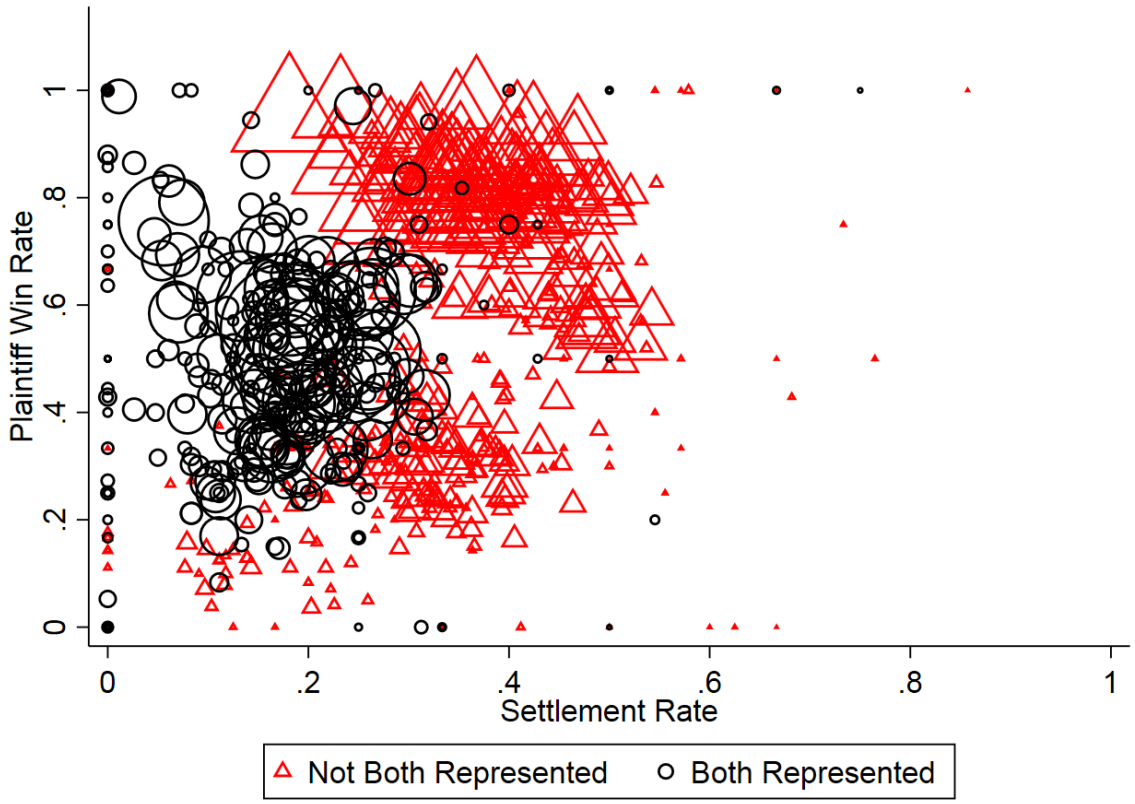
Variable	(1)	(2)	(3)	(4)
Ln Stakes	–0.106**	–0.012**	0.001	–0.002
	(0.004)	(0.005)	(0.012)	(0.013)
Both Parties Represented	–0.005*	–0.092**	–0.135**	–0.123**
	(0.003)	(0.005)	(0.006)	(0.006)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	4,656	3,830	2,012	1,708

Note. High stakes cases are cases with stakes exceeding USD 75,000. Outcome variable is the absolute value of the difference between the plaintiff win rate and 0.5. Each observation includes an indicator for whether in all cases within the observation/group both parties are represented by attorneys, the average settlement rate and the absolute deviation from 0.5 of the average plaintiff win rate for a category of cases. Observations (categories) weighted by number of court decisions within the category. Categories are defined equivalently to Figure 3. Settled cases excluded. All columns include fixed effects for procedure, level, case type, court, and filing year. Robust standard errors in parenthesis. ** $p<0.01$, * $p<0.05$.

TABLE A4. REVIEW OF RESULTS OF HYPOTHESIS TESTING

Model	Hypothesis	Result
Extension of PK: Degree of Optimism	H(A1): Individual (non-corporate) plaintiffs ↔ lower settlement rates	Unsupported
	H(A2): Both parties represented ↔ higher settlement rates	Rejected
	H(A5): Both parties represented ↔ plaintiff win rates closer to 50%	Some Support
Extension of PK: Asymmetric Stakes *and/or* Competing Model: AI	H(A3): Represented parties ↔ higher win rates	Supported
	H(A4): Corporations, not individuals ↔ higher win rates	Supported

FIGURE A1. RELATIONSHIP BETWEEN SETTLEMENT RATE AND PLAINTIFF WIN RATE, COMPARING CASES WITH AND WITHOUT BOTH PARTIES REPRESENTED, BY REGION, PROCEDURE, LEVEL, AND CASE CATEGORIES



Note. Scatterplot of 918 observations of the settlement rate and plaintiff win rate among cases with and without both parties represented, by region/procedure/level/case-category. Size of circle is proportional to the number of court decisions represented by each observation.

Appendix B: Robustness Checks: Regressions Including Cases with Counterclaims and Multiple Parties

TABLE B1. PREDICTORS OF SETTLEMENT RATE, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Panel A. All Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.015*	–0.016*	–0.018**	–0.018**
	(0.006)	(0.005)	(0.002)	(0.002)
Corporate Plaintiff		–0.010		0.027*
		(0.011)		(0.008)
Corporate Defendant		0.011		–0.006
		(0.005)		(0.004)
Include "debt collection" cases	Y	Y		
N	618,917	618,917	340,046	340,046

Panel B. High Stakes	(1)	(2)	(3)	(4)
Ln Stakes	–0.015**	–0.015**	–0.015**	–0.015**
	(0.001)	(0.001)	(0.003)	(0.002)
Corporate Plaintiff		–0.001		0.016*
		(0.004)		(0.006)
Corporate Defendant		–0.014		–0.001
		(0.009)		(0.009)
Include "debt collection" cases	Y	Y		
N	74,198	74,198	64,132	64,132

Note. See Table 4.

TABLE B2. PREDICTORS OF SETTLEMENT RATE, INCLUDING ONLY CASES BETWEEN INDIVIDUALS

Panel A. All Stakes	(1)	(2)
Ln Stakes	–0.013**	–0.013**
	(0.003)	(0.003)
Include "debt collection" cases	Y	
N	132,793	116,061

Panel B. High Stakes	(1)	(2)
Ln Stakes	–0.022**	–0.019*
	(0.006)	(0.006)
Include "debt collection" cases	Y	
N	18,994	16,646

Note. See Table 4.

TABLE B3. PREDICTORS OF PLAINTIFF WIN RATE, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Panel A. All Stakes	(1)	(2)	(4)	(5)
Ln Stakes	–0.006*	–0.003	–0.001	0.008**
	(0.002)	(0.004)	(0.002)	(0.002)
Corporate Plaintiff	0.093**	0.099**	0.046*	0.049*
	(0.016)	(0.017)	(0.015)	(0.015)
Corporate Defendant	–0.045**	–0.045**	–0.037**	–0.030**
	(0.007)	(0.008)	(0.005)	(0.006)
Plaintiff Represented		0.085**		0.054**
		(0.011)		(0.004)
Defendant Represented		–0.146**		–0.186**
		(0.011)		(0.002)
Include "debt collection" cases	Y	Y		
N	461,931	461,931	225,960	225,960

Panel B. High Stakes	(1)	(2)	(4)	(5)
Ln Stakes	–0.001	0.017**	–0.001	0.014*
	(0.003)	(0.005)	(0.004)	(0.006)
Corporate Plaintiff	0.075**	0.064**	0.036	0.040
	(0.016)	(0.015)	(0.017)	(0.017)
Corporate Defendant	–0.020	–0.016	–0.038**	–0.030*
	(0.009)	(0.011)	(0.010)	(0.011)
Plaintiff Represented		0.003		0.031**
		(0.006)		(0.007)
Defendant Represented		–0.203**		–0.195**
		(0.004)		(0.003)
Include "debt collection" cases	Y	Y		
N	52,121	52,121	44,494	44,494

Note. See Table A2.

TABLE B4. RELATIONSHIP BETWEEN SETTLEMENT RATES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Variable	(1)	(2)	(3)	(4)
Settlement Rate	–0.241**	–0.152**	0.007	0.068
	(0.016)	(0.021)	(0.034)	(0.035)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	3,088	2,540	1,409	1,203

Note. See Table 5.

TABLE B5. RELATIONSHIP BETWEEN BOTH PARTIES REPRESENTED AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Variable	(1)	(2)	(3)	(4)
Ln Stakes	–0.097**	0.018**	0.029*	0.014
	(0.004)	(0.005)	(0.012)	(0.013)
Both Parties Represented	–0.015**	–0.111**	–0.153**	–0.135**
	(0.003)	(0.004)	(0.005)	(0.005)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	5,078	4,208	2,402	2,056

Note. See Table A3.

TABLE B6. RELATIONSHIP BETWEEN CASE CATEGORY (PROXY FOR AI) AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Variable	(1)	(2)	(3)	(4)
Tort (Asymmetric Info)	0.028**	0.027**	0.078**	0.079**
	(0.005)	(0.004)	(0.008)	(0.008)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	1,983	1,435	905	699

Note. See Table 6.

TABLE B7. RELATIONSHIP BETWEEN STAKES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, INCLUDING CASES WITH COUNTERCLAIMS AND MULTIPLE PARTIES

Variable	(1)	(2)	(3)	(4)
Ln Stakes	–0.077**	0.026**	–0.001	0.003
	(0.005)	(0.006)	(0.018)	(0.019)
Include "debt collection" case type categories	Y		Y	
High Stakes Cases Only			Y	Y
N	3,088	2,540	1,409	1,203

Note. See Table 7.

TABLE B8. RELATIONSHIP BETWEEN STAKES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, INCLUDING ONLY CASES BETWEEN INDIVIDUALS

Variable	(1)	(2)	(3)	(4)
Ln Stakes	–0.013*	–0.013*	0.011	0.004
	(0.006)	(0.006)	(0.017)	(0.019)
Include "debt collection" case type categories	Y		Y	
High Stakes Cases Only			Y	Y
N	2,480	2,018	1,003	840

Note. See Table 7.

Appendix C: Robustness Checks: Regressions Treating Plaintiff Partial Win as Half a Win

TABLE C1. RELATIONSHIP BETWEEN SETTLEMENT RATES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN

Variable	(1)	(2)	(3)	(4)
Settlement Rate	-0.389**	-0.180**	-0.019	-0.025
	(0.018)	(0.023)	(0.028)	(0.029)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	2,893	2,360	1,204	1,022

Note. See Table 5.

TABLE C2. RELATIONSHIP BETWEEN CASE CATEGORY (PROXY FOR AI) AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN

Variable	(1)	(2)	(3)	(4)
Tort (Asymmetric Info)	-0.134**	-0.138**	-0.004	-0.002
	(0.006)	(0.005)	(0.009)	(0.009)
Include "debt collection" cases	Y		Y	
High Stakes Cases Only			Y	Y
N	1,896	1,363	799	617

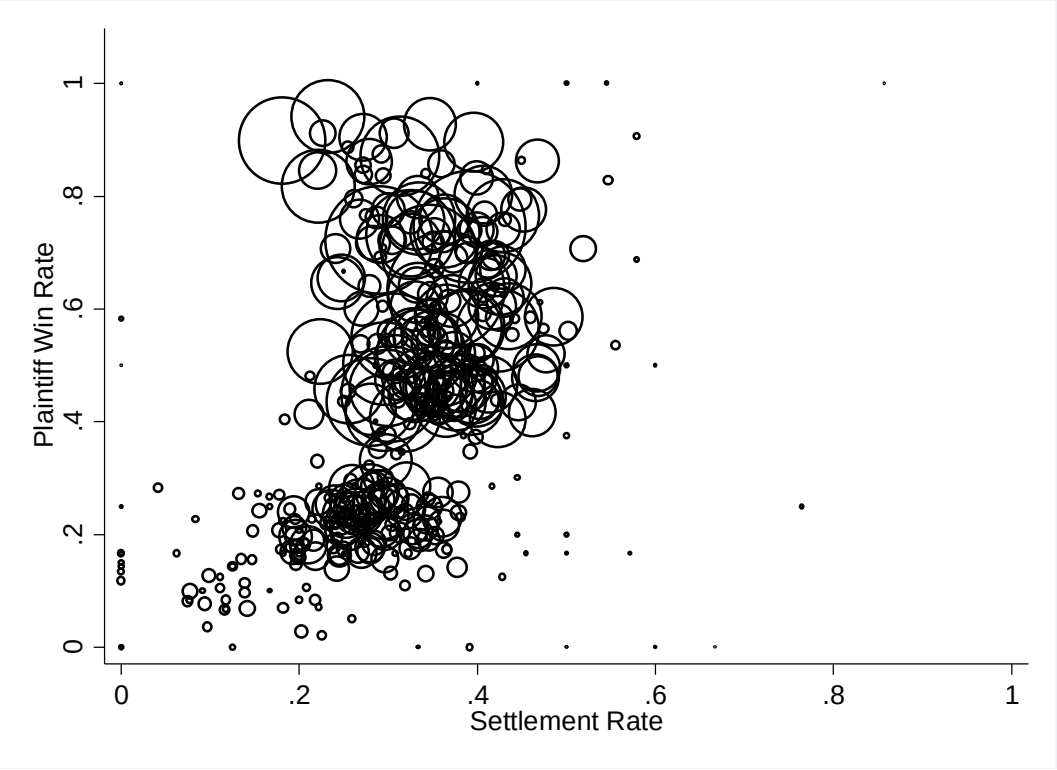
Note. See Table 6.

TABLE C3. RELATIONSHIP BETWEEN STAKES AND ABSOLUTE DEVIATION OF PLAINTIFF WIN RATES FROM 50%, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN

Variable	(1)	(2)	(3)	(4)
Ln Stakes	-0.103**	-0.074**	-0.014	-0.018
	(0.005)	(0.007)	(0.015)	(0.016)
Include "debt collection" case type categories	Y		Y	
High Stakes Cases Only			Y	Y
N	2,893	2,360	1,204	1,022

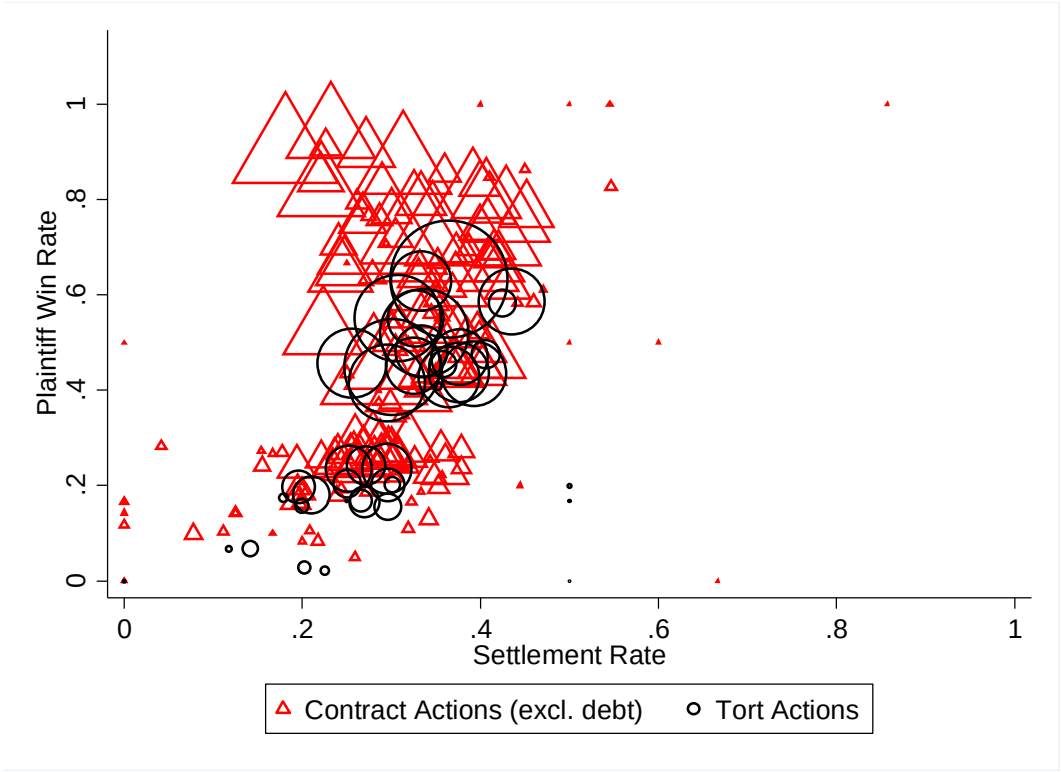
Note. See Table 7.

FIGURE C1. RELATIONSHIP BETWEEN SETTLEMENT RATE AND PLAINTIFF WIN RATE, BY PROCEDURE, LEVEL, COURT, AND CASE CATEGORIES, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN



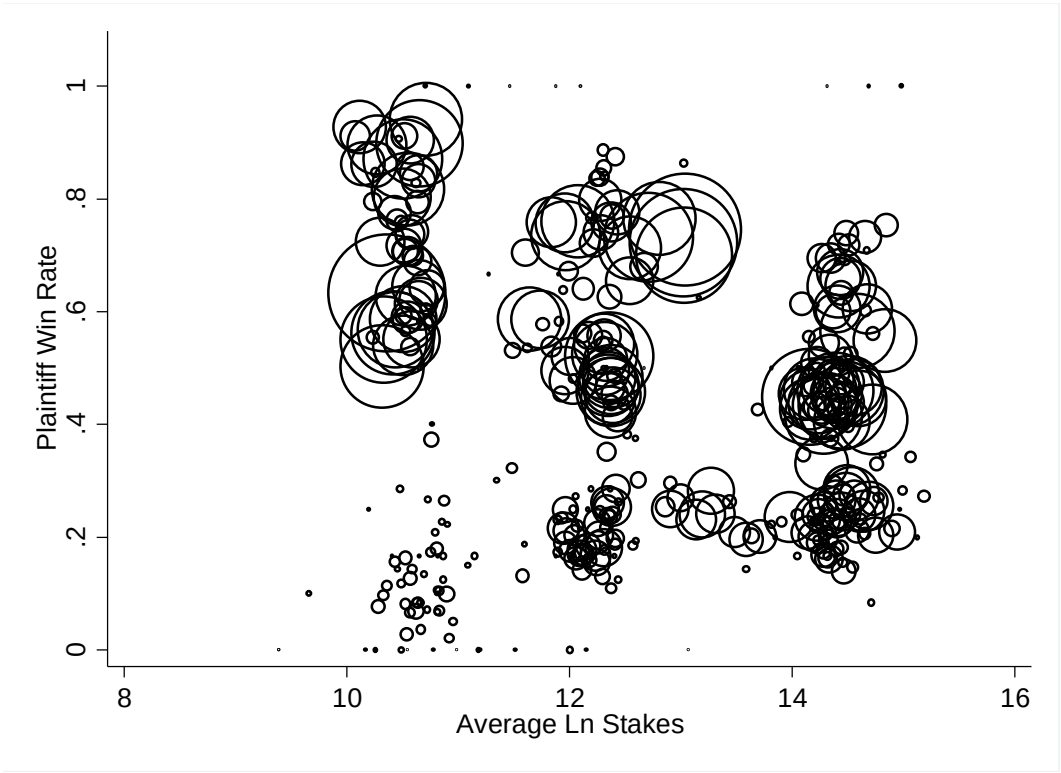
Note. See Figure 2.

FIGURE C2. RELATIONSHIP BETWEEN SETTLEMENT RATE AND PLAINTIFF WIN RATE, COMPARING CONTRACT (LOW AI) AND TORT (HIGH AI) ACTIONS BY REGION, PROCEDURE, LEVEL, AND CASE CATEGORIES, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN



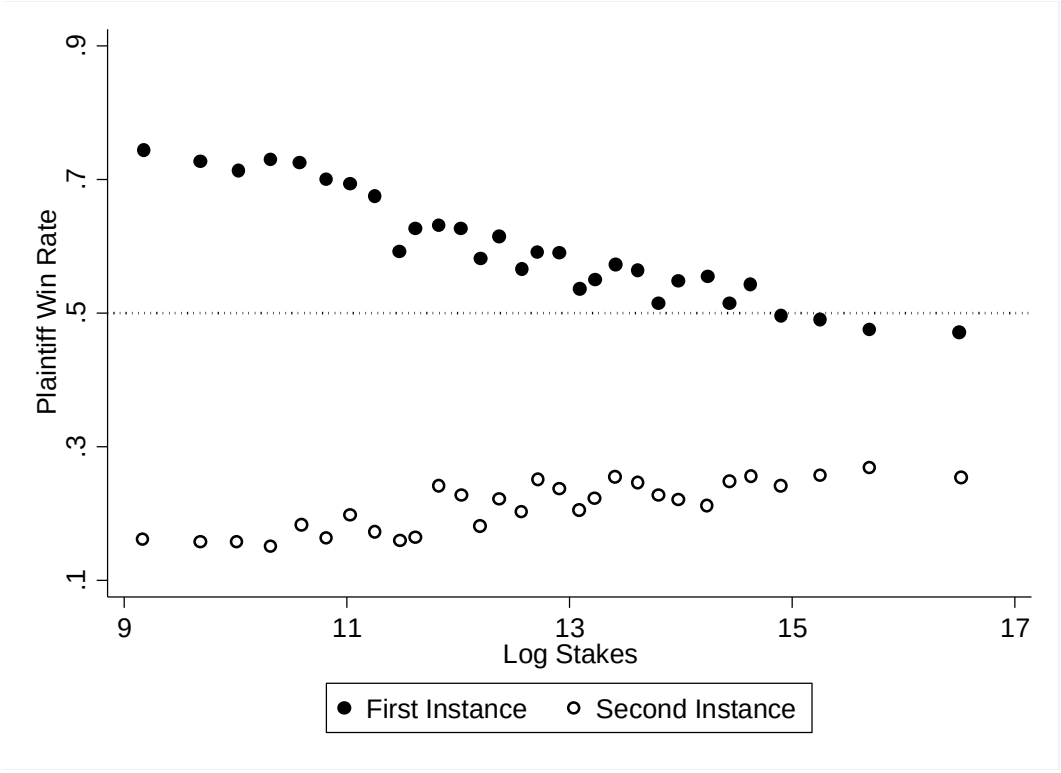
Note. See Figure 3.

FIGURE C3. RELATIONSHIP BETWEEN STAKES AND PLAINTIFF WIN RATE, BY STAKES AND COURT LEVEL, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN



Note. See Figure 4.

FIGURE C4. RELATIONSHIP BETWEEN LOG STAKES AND PLAINTIFF WIN RATE, BY REGION, PROCEDURE, LEVEL, END YEAR, AND CASE CATEGORIES, TREATING PLAINTIFF PARTIAL WIN AS HALF A WIN



Note. See Figure 5.

.

References

Arlen, Jennifer, and Stephan Tontrup. 2015. Does the Endowment Effect Justify Legal Intervention? The Debiasing Effect of Institutions. *The Journal of Legal Studies* 44 (1):143–182.

Bebchuk, Lucian Arye. 1984. Litigation and Settlement Under Imperfect Information. *The RAND Journal of Economics* 15 (3):404–415.

Chang, Yun-chien. 2016. The Evolution of Property Law in Taiwan: An Unconventional Interest Group Story. In *Private Law in China and Taiwan: Economic and Legal Anaylses* 212–244, edited by Yun-chien Chang, Wei Shen and Wen Yeu Wang. Cambridge: Cambridge University Press.

Chang, Yun-chien, Weitseng Chen, and Ying Chieh Wu. 2017. *Property and Trust Law in Taiwan*. Alphen aan den Rijn: Wolters Kluwer.

Chang, Yun-chien, Nuno Garoupa, and Martin Wells. 2017. Redrawing the Legal Family Trees: An Empirical Approach. *working paper*.

Chang, Yun-chien, and Su-hao Tu. 2018. Two-way Selections Between Flat-fee Attorneys and Litigants: Theoretical and Empirical Analyses. *European Journal of Law and Economics*.

Cooter, Robert D., and Daniel L. Rubinfeld. 1989. Economic Analysis of Legal Disputes and Their Resolution. *Journal of Economic Literature* 27 (3):1067–1097.

Cui, Wei, and Zhiyuan Wang. 2017. The Selection of Litigation against Government Agencies: Evidence from China. *Review of Law & Economics* 13 (3):1–41.

Eisenberg, Theodore. 1990. Testing the Selection Effect: A New Theoretical Framework with Empirical Tests. *The Journal of Legal Studies* 19 (2):337–358.

———. 1991. The Relationship Between Plaintiff Success Rates Before Trial and at Trial. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 154:111–116.

Eisenberg, Theodore, and Henry S. Farber. 1997. The Litigious Plaintiff Hypothesis: Case Selection and Resolution. *RAND Journal of Economics* 28:S92–S112.

Eisenberg, Theodore, and Kuo-Chang Huang. 2012. The Effect of Rules

Shifting Supreme Court Jurisdiction from Mandatory to Discretionary—An Empirical Lesson from Taiwan. *International Review of Law & Economics* 32 (1):3–18.

Eisenberg, Theodore, and Charlotte Lanvers. 2009. What Is the Settlement Rate and Why Should We Care? *Journal of Empirical Legal Studies* 6 (1):111–146.

Gelbach, Jonah B. 2016. The Reduced Form of Litigation Models and the Plaintiff's Win Rate. *Available at SSRN: https://ssrn.com/abstract=2837154*.

Gould, John P. 1973. The Economics of Legal Conflicts. *The Journal of Legal Studies* 2 (2):279–300.

Hadfield, Gillian K. 2004. Where Have All the Trials Gone? Settlements, Nontrial Adjudications, and Statistical Artifacts in the Changing Disposition of Federal Civil Cases. *Journal of Empirical Legal Studies* 1 (3):705–734.

Helland, Eric, Daniel Klerman, and Yoon-Ho Alex Lee. 2017 forthcoming. Maybe There's No Bias in the Selection of Disputes for Litigation. *Journal of Institutional and Theoretical Economics*.

Huang, Kuo-Chang. 2008. How Legal Representation Affects Case Outcome: An Empirical Perspective from Taiwan. *Journal of Empirical Legal Studies* 5 (2):197–238.

———. 2009. Does Discovery Promote Settlement? An Empirical Answer. *Journal of Empirical Legal Studies* 6 (2):241–278.

———. 2016. The Effect of Stakes on Settlement: An Empirical Lesson from Taiwan. In *Research Handbook on Comparative Law and Economics* 369–392, edited by Theodore Eisenberg and Giovanni Battista Ramello. Northampton, Mass.: Edward Elgar.

Hubbard, William H.J. 2016. Sinking Costs to Force or Deter Settlement. *The Journal of Law, Economics, and Organization* 32:545–585.

Hylton, Keith N. 1993. Asymmetric Information and the Selection of Disputes for Litigation. *The Journal of Legal Studies* 22 (1):187–210.

Kessler, Daniel, Thomas Meites, and Geoffrey Miller. 1996. Explaining Deviations from the Fifty-percent Rule: A Multimodal Approach to the Selection of Cases for Litigation. *The Journal of Legal Studies* 25 (1):233–259.

Klerman, Daniel. 2012. The Selection of Thirteen-Century Disputes for

Litigation. *Journal of Empirical Legal Studies* 9:320–346.

Klerman, Daniel, and Yoon-Ho Alex Lee. 2014. Inferences from Litigated Cases. *The Journal of Legal Studies* 43 (2):209–248.

Landes, William M. 1971. An Economic Analysis of the Courts. *The Journal of Law & Economics* 14 (1):61–107.

Lee, Yoon-Ho Alex, and Daniel Klerman. 2016. The Priest-Klein Hypotheses: Proofs and Generality. *International Review of Law and Economics* 48:59–76.

Lee, Yoon-Ho Alex, and Daniel M. Klerman. 2015. Updating Priest and Klein. *https://ssrn.com/abstract=2619856*.

Loewenstein, George, Samuel Issacharoff, Colin Camerer, and Linda Babcock. 1993. Self-serving Assessments of Fairness and Pretrial Bargaining. *The Journal of Legal Studies* 22 (1):135–159.

Posner, Richard A. 1973. An Economic Approach to Legal Procedure and Judicial Administration. *The Journal of Legal Studies* 2:399–458.

Priest, George L., and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *The Journal of Legal Studies* 13 (1):1–56.

Ramseyer, J. Mark. 2015. *Second-Best Justice: The Virtues of Japanese Private Law*. Chicago, IL: The University of Chicago Press.

Ramseyer, J. Mark, and Minoru Nakazato. 1989. The Rational Litigant: Settlement Amounts and Verdict Rates in Japan. *The Journal of Legal Studies* 18 (2):263–290.

Reinganum, Jennifer F., and Louis L. Wilde. 1986. Settlement, Litigation, and the Allocation of Litigation Costs. *The RAND Journal of Economics*:557–566.

Shavell, Steven. 1996. Any Frequency of Plaintiff Victory at Trial Is Possible. *The Journal of Legal Studies* 25 (2):493–501.

Siegelman, Peter, and John J. Donohue. 1995. The Selection of Employment Discrimination Disputes for Litigation: Using Business Cycle Effects to Test the Priest-Klein Hypothesis. *The Journal of Legal Studies* 24 (2):427–462.

Siegelman, Peter, and Joel Waldfogel. 1999. Toward a Taxonomy of Disputes: New Evidence through the Prism of the Priest/Klein Model. *The Journal of Legal Studies* 28 (1):101–130.

Stanley, Linda R., and Don L. Coursey. 1990. Empirical Evidence on the Selection Hypothesis and the Decision to Litigate or Settle *Journal of*

Legal Studies 19:145–172.

Waldfogel, Joel. 1995. The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory. *Journal of Political Economy* 103 (2):229–260.

———. 1998. Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation. *The Journal of Law and Economics* 41 (2):451–476.

Wang, Tay-sheng. 2002. The Legal Development of Taiwan in the 20th Century: Toward a Liberal and Democratic Country. *Pacific Rim Law & Policy Journal* 11:531.