

LAW AND ECONOMICS SEMINAR
Autumn Quarter 2017

Professor Polinsky

Thursday, October 5, 2017
4:15 - 5:45 p.m.
Stanford Law School
Room 320D

“Contingent Fee Litigation in New York City”

by

Daniel Klerman

(University of Southern California Gould School of Law)

Note: It is expected that you will have reviewed the speaker’s paper before the seminar. As you will see, there are actually two papers attached – one with the title announced above and a second one that builds on that paper. You are only asked to review the first paper, though the second, more technical, paper may of interest to some participants.

September 20, 2017

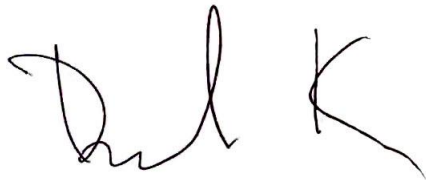
Dear Stanford Law & Econ Workshop participants,

Attached please find two papers. The first, "Contingent Fee Litigation in New York City," describes a data set that Eric Helland and I are working with, including some interesting preliminary results. Section III contains a short discussion of the implications of the data for the selection of disputes for litigation. The second paper, "Maybe There's No Bias in the Selection of Disputes for Litigation," expands on that discussion.

In addition to discussing the results in these papers, Eric Helland and I would welcome your ideas for other papers to write with these data.

I look forward to the workshop and our discussion.

Sincerely,

Handwritten signature of Daniel M. Klerman, consisting of the letters 'DK' in a cursive, flowing script.

Contingent Fee Litigation in New York City

Eric Helland*, Daniel Klerman**, Brendan Dowling***, and Alexander Kappner****

Draft 9/20/17

Abstract

Since 1957, New York courts have required contingent fee lawyers to file “closing statements” that disclose settlement amounts, lawyers’ fees, an accounting of expenses, and other information. This article provides preliminary analysis of these data for the period 2004-2013. Among this article’s findings are that settlement rates in New York state courts are very high (84%) relative to previous studies, that very few cases are resolved by dispositive motions, that litigated cases and settled cases have almost exactly the same average recovery, that median litigation expenses, other than attorney’s fees, are 3% of gross recovery, that claims are disproportionately from poor neighborhoods, and that attorneys’ fees are almost always one third of net recovery, which is the maximum allowed by law.

Since 1957, most contingent fee lawyers in New York City have been required to file a “closing statement” with the clerk of the Appellate Division when a case is resolved, whether the case is resolved by settlement, judgment, or abandonment by the client. The closing statement includes the amount of any settlement or judgment, the amount paid to the lawyer, and an itemization of the lawyer’s expenses. Because they provide information on issues not generally available elsewhere, closing statements provide a unique window into contingent fee litigation. This article aims to provide a preliminary analysis of the data in the closing statements

In the next section we provide information on the origins and contents of the data. In section II we provide evidence on rates at which claims are abandoned, settled or adjudicated. We find that the settlement rate in the New York data is significantly higher than previous

* William F. Podlich Professor of Economics, Claremont McKenna College and Senior Economist, RAND. The authors thank Scott Altman, Andrew Daughety, Jennifer Reinganum, Chris Robertson, and participants in the Vanderbilt State Courts Symposium and University of Arizona QuantLaw conference for helpful comments, suggestions, and criticisms. The authors also thank Paul Moorman, USC Law Librarian, for excellent assistance.

** Charles L. and Ramona I. Hilliard Professor of Law and History.

*** J.D. student, USC Law School. Brendan Dowling refined and implemented the OCR (optical character recognition) strategy to extract information from the closing statement PDFs.

**** Associate, Knobbe Martens. Alexander Kappner developed the strategy and began implementing the OCR (optical character recognition) strategy to extract information from the closing statement PDFs.

estimates. In section III we provide information plaintiff recovery rates. We find that once we account for the riskiness of going to trial there is relatively little difference between settlement amounts and amounts received in adjudication. In section IV we provide evidence on the size and composition of expenses and fees. In section V we over some evidence on the demographic distribution of tort claims. Specifically we find that litigation rates are negatively correlated with income. Finally, section VI concludes.

I. Retainer and Closing Statements

In the 1920s, the bar and bench in New York City became increasingly concerned about the conduct of contingent fee lawyers. In 1928, the bar associations for New York City, Manhattan, and the Bronx petitioned the Appellate Division of the First Judicial Department of the New York Supreme Court, which had supervisory powers over state courts in Manhattan and the Bronx, to conduct an investigation. The Appellate Division ordered Justice Wasservogel to produce a report.

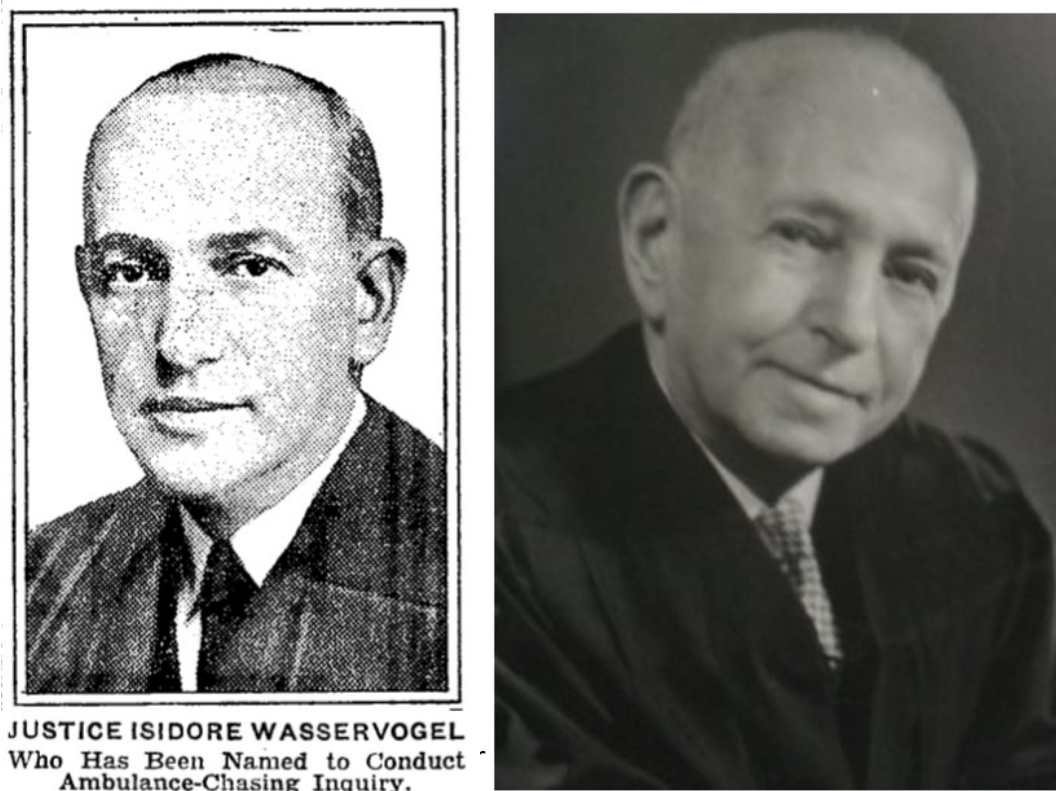


Figure 1. Pictures of Isidore Wasservogel. Left: New York Times (February 10, 1928). Right: undated from nycourts.gov.

Judge Wasservogel held hearings, and, in 1929, issued a report that concluded:

The evidence adduced before me bears out the truth of the allegations contained in the petition of the three bar associations, to the effect that there exists in the Judicial Department a practice commonly known as “ambulance chasing.”

Personal injury cases have, in the main, come into the hands of relatively few lawyers, some of whom have conducted their practice purely as a business, to the detriment of the public and the profession.¹

To “prevent a recurrence of the improper practices by which attorneys secure retainers from injured persons,” the report recommended that attorneys be required to file “a copy of the retainer by which the attorney for the plaintiff was engaged, and also an affidavit by such attorney stating that the case was not solicited directly or indirectly, and setting forth how the retainer was obtained.”² The report also recommended that all settlements be approved by the court.³

In 1929, the First Department implemented some of the recommendations of the report. In particular, they required plaintiffs’ lawyers to file “retainer statements” with the court within 10 days of signing of contingent fee agreements. The retainer statement sets out “the terms of compensation.”⁴ Regulations also required lawyers to mail their clients a written statement accounting for any judgment or settlement within 10 days of receiving the money.⁵ If the lawyer was unable to find the client, the lawyer was required to send the accounting to the court, but otherwise the accounting was sent only to the client.⁶

In 1955, Judge Wasservogel, then retired from his judgeship and back in private practice, was again asked to produce a report on contingent fee lawyers, this time with a charge to consider capping contingent fees. This report was commissioned not only by the First Department (Manhattan and the Bronx), but also by the Second Department, which covers the rest of New York City as well as Long Island and five counties immediately north of New York City (Dutchess, Orange, Putnam, Rockland, and Westchester). This report used the retainer statement required by the 1929 regulations to ascertain that 60% of retainers filed with the First and Second Judicial Departments were “50% Retainers.” That is, over half of retainer contracts specified that 50% of any judgment or settlement went to the lawyer.⁷ The report considered, but

¹ ISIDOR WASSERVOGEL, REPORT TO APPELLATE DIVISION, FIRST JUDICIAL DEPARTMENT 1, 4 (1928).

² *Id.* at 8–9.

³ *Id.* at 12.

⁴ Special Rules Regulating the Conduct of Attorneys and Counselors-at-Law in the First Judicial Department, Rule 4-A, *reprinted in* PARSONS’ PRACTICE MANUAL OF THE STATE OF NEW YORK (Austin B. Griffin & John T. Fitzpatrick eds., 7th ed. 1930).

⁵ *Id.* 4-B.

⁶ *Id.*

⁷ ISIDOR WASSERVOGEL, REPORT, FINDINGS AND RECOMMENDATIONS IN THE MATTER OF THE HEARINGS ORDERED BY THE APPELLATE DIVISIONS OF THE SUPREME COURT IN AND FOR THE FIRST AND

did not recommend, caps on fees, although it did mention the possibility of “a sliding scale of fees” whereby lawyers would get 50% of damages up to \$2000, 40% of damages between \$2000 and \$10,000, and one third of amounts over \$10,000.⁸ So that the court could ensure that clients received their proper share, the report recommended that “closing statements” be filed with the court specifying how a case was terminated, the gross amount of recovery or settlement, the amounts paid to the lawyer and client, and an itemization of costs.⁹

The First Department responded with new regulations in 1957. These regulations capped the contingent fee in “actions for personal injury or wrongful death” at one-third or a sliding scale similar to the one set out in Judge Wasservogel’s report, although the scale was a little less generous to the lawyer.¹⁰ The regulations also required that the percentage that goes to the lawyer, whether a simple flat percentage or a percentage based on the sliding scale, must be computed “on the net sum recovered after deducting taxable costs and disbursements and expenses for legal, medical, investigative or other services properly chargeable to the claim.”¹¹ Most importantly for the purposes of this article, the regulations required that lawyers file with the court a “closing statement” within fifteen days of receiving any money on behalf of a client, whether in judgment or settlement. The closing statement records “[t]he gross amount of the recovery, ... [t]he taxable costs and disbursements, ... [t]he net amount of the recovery actually received by the client, ... [t]he amount of the compensation actually received or retained by the attorney,” and some additional information not important to this article.¹² The Second and Fourth Departments issued similar rules.¹³ The First and Second Departments continue to require closing statements in much the same form, although the Fourth Department repealed its requirement in 2003.

The First Department regulations were modified periodically between 1957 and 2002. By 2002, the requirement to file retainer and closing statement had expanded to include tort cases involving personal injury, property damage, or wrongful death, as well as condemnation and change of grade proceedings.¹⁴ The 2002 regulations also make clear that a closing statement

SECOND JUDICIAL DEPARTMENTS REGARDING A PROPOSED RULE TO LIMIT COMPENSATION OF PLAINTIFFS’ ATTORNEYS IN PERSONAL INJURY AND WRONGFUL DEATH ACTIONS 1, 38 (1955).

⁸ *Id.* at 43–44.

⁹ *Id.* at 39–40.

¹⁰ Special Rules Regulating the Conduct of Attorneys and Counselor-at-Law in the First Judicial Department, Rule 4(b), *reprinted in* CLEVINGER’S ANNUAL PRACTICE OF NEW YORK 21-10 (Jos. R. Clevenger ed., 1957).

¹¹ *Id.* Rule 4(c). A rule change in 2014 allowed the lawyer’s percentage to be based on gross recovery. *See infra* note [16+15](#).

¹² *Id.* Rule 4(e).

¹³ Because this article uses data only from the First Department, it does not analyze rules from other departments in any detail.

¹⁴ *See* N.Y. COMP. CODES R. & REGS. tit. 22, § 603.7(a)(1) (2002) (“Every attorney who, in connection with any action or claim for damages for personal injuries or for property damages or for death or loss of services resulting from personal injuries, or in connection with any claim in condemnation or change of grade proceedings, accepts a retainer or enters into an agreement, express or implied, for compensation for services rendered . . . whereby his compensation is to be dependent or contingent in whole or in part upon the successful prosecution or settlement thereof . . .”). Section (a)

must be filed even if the plaintiff recovers nothing, whether due to an adverse judgment or because the plaintiff abandoned the claim.¹⁵ There have been no significant changes to the regulations since 2002.¹⁶

The First Department's retainer and closing statement requirements apply to all attorneys who practice in Manhattan or the Bronx, regardless of the court they file a case in and regardless of whether they practice regularly in New York City.¹⁷ This means that filing statements should be filed for cases brought in federal court. In fact, there are over sixteen thousand closing statement relating to federal cases. Nevertheless, the overwhelming majority of closing statements involve cases filed in state court, and by far the largest number (129,269) involve cases filed in the New York Supreme Court, which is the trial court of general jurisdiction. The Supreme Court is the state trial court one would expect would hear tort cases of any significance. A few thousand closing statement relate to county court and surrogates court (probate) proceedings, and a handful of cases were filed in small claims court, justice or police court, or out of state.

Retainer and closing statements are considered "confidential" and information may not be "divulged or made available.... except upon written order of the presiding justice of the Appellate Division."¹⁸ The authors of this article obtained an order from the Presiding Justice in 2014 allowing us access to retainer and closing statements from the First Department, albeit with severe restrictions to protect confidentiality. The court administrator then provided us with PDFs of retainer and closing statements filed in 2004-2013, as well as a "flat file" compiled by court personnel in which some (but not all) of the relevant information had been typed into a database similar to an Excel spreadsheet. Unfortunately, the flat file does not contain key information, such as expenses or amounts paid to lawyer and client. The authors of this article have hand coded a random sample of the PDFs and use that hand-coding to generate the results presented in Section IV. For future projects, the authors are also utilizing optical character recognition (OCR) software to extract this information from the PDFs. This software converts information that was originally handwritten or typed into a form that statistical software can analyze.

technically only applies to retainer statements, but section (b)(1) states that "[a] closing statement shall be filed in connection with every claim . . . in which a retainer statement is required." There are relatively few cases involving condemnation or change of grade. We are looking for ways to analyze such cases separately or exclude them.

¹⁵ *See id.* § 603.7(b)(1), (b)(2) Item 6 (stating that a "closing statement shall be filed in connection with every claim, action or proceeding in which a retainer statement is required").

¹⁶ On February 19, 2014, the rules were changed to allow the lawyer's percentage to be computed "on the gross some recovered before deduct[ing] expenses and disbursements," if the client agreed to that method in the retainer agreement. N.Y. SUP. CT. R. § 691.20(e)(3) (LexisNexis 2017). Because the data analyzed here comes from 2004-13, that change is not relevant to this article.

¹⁷ *See* N.Y. COMP. CODES R. & REGS. tit. 22, § 603.1(a) ("This Part shall apply to all attorneys who are admitted to practice, reside in, commit acts in or have offices in this judicial department, or who are admitted to practice by a court of another jurisdiction and who practice within this department.").

¹⁸ *See id.* § 603.7(c)(1).

A typical retainer statement is displayed below, with identifying information removed, as required by the court order.

RETAINER STATEMENT

TO: OFFICE OF COURT ADMINISTRATION
Retainer and Closing Statements
Post Office Box 2016
Church Street Station
New York, NY 10008

For office use:
NYC OCA [REDACTED] 13

1. **Date of agreement as to retainer.**
[REDACTED] 2013

2. **Terms of compensation.**
33 1/3% OF ANY SUMS RECOVERED TO [REDACTED] & [REDACTED]

3. [REDACTED]

4. ***If engaged by an attorney, name and office address of retaining attorney**

5. **If claim for personal injuries, wrongful death or property damage, date and place of**
[REDACTED] 2010 at [REDACTED] New York, NY 10021

6. **If condemnation or change of grade proceeding:**
(a) **Title and description:** _____

7. **Name, address, occupation and relationship of person referring client:**
[REDACTED]

Dated: NEW YORK _____ NY, [REDACTED]

Figure 2. Sample Retainer Statement

Retainer statements contain relatively little information—the date of agreement, a summary of the terms of compensation (nearly always one third of recovery to the lawyer), the date and address of the accident (or address of property condemned or subject to a grade proceeding), and information about referrals.

The next two pages display a typical closing statement in a relatively large case, with identifying information removed, as required by the court order.

CLOSING STATEMENT

FOR OFFICE USE:

TO: OFFICE OF COURT ADMINISTRATION
Retainer and Closing Statements
25 Beaver Street - 8th Floor
New York, New York 10004

NYC OCA

1. [REDACTED]
2. [REDACTED]
3. [REDACTED]
4. [REDACTED]
5. (a) If action commenced, state date: [REDACTED], Supreme Court, Queens County; and
(b) Was the action disposed of in open Court - No. If not and a Request for Judicial
Intervention was filed, state the date the Stipulation or Statement of Discontinuance was
filed with the Clerk of the Part to which the action was assigned - [REDACTED] If not and
an Index number was assigned, but no Request for Judicial Intervention was filed, state
the date the Stipulation or Statement of Discontinuance was filed with the Clerk - N/A.
6. Check items applicable: Settled [] or claim abandoned by client [] or Judgment
[]. Date of payment by carrier or defendant - [REDACTED] Date of payment to client -
[REDACTED]
7. Gross amount of recovery (if judgment entered, including any interest, costs or
disbursements allowed): \$975,000.00.
8. Name and address of insurance carrier or person payment judgment or claim and carrier's
file number, if any: [REDACTED]
9. Net amounts: TO CLIENT - \$451,043.09; COMPENSATION TO UNDERSIGNED - \$143,229.18. Also,
include the name(s), address(es) and amount(s) paid to attorneys participating in the
contingent compensation: [REDACTED] in
the amount of \$95,486.12. [REDACTED] substituted for [REDACTED]
[REDACTED] which received \$79,229.18 for its attorney's fees. Their retainer
number is unknown to the undersigned.
10. Compensation fixed by: Retainer Agreement []; under scheduled []; or by Court
[].
11. If compensation fixed by Court: Name of Judge: N/A; Court: N/A; County: N/A; Index
No.: N/A; and date of Order: N/A.
12. Itemized statement of payments made for hospital, medical care or treatment; liens'
assignments; claims; and expenses on behalf of client(s), which would have been charged
against the client's share of the recovery, together with the name, address, amount and
reason for each payment: [REDACTED] in
the amount of \$97,554.85 for worker's compensation and [REDACTED],
[REDACTED] in the amount of \$87,976.19 for reimbursement of plaintiff's loan.
13. Itemized statement of the amounts of expenses and disbursements paid or agreed to be paid
by others for expert testimony; investigative; or other services properly chargeable to
the recovery of damages, together with the name, address and reason for each payment:
SEE ATTACHED RIDER.
14. Date on which a copy of this Closing Statement has been forwarded to the client:
[REDACTED]

DATED: New York, New York
[REDACTED]

Figure 3. Page 1 of Sample Closing Statement

[REDACTED] DISBURSEMENTS

Explanation	Class	Disbs
Court Calendar Watching Services	Calendar Watching	\$ 20.23
Supreme Queens, Summons and Complaint, RJJ, Note of Issue and Motion	Court Filing Fees	\$ 445.70
[REDACTED]	Court Reporters	\$ 373.75
[REDACTED] M.D., [REDACTED] D.O., [REDACTED] Ph.D., [REDACTED]	Expert Witnesses	\$ 11,500.00
[REDACTED] M.D., [REDACTED] M.D.		
[REDACTED]	Investigation	\$ 898.40
Jams	Mediation Fee	\$ 875.00
[REDACTED] Hospital, [REDACTED]	Medical Records	\$ 1,479.13
[REDACTED] M.D., [REDACTED] Physical Therapy, [REDACTED] Surgical Center, [REDACTED] M.D.		
Misc Expenses: Xerox Copies, Express and Certified Mail, Scanning, Faxes, travel expenses	Misc	\$ 3,228.73
[REDACTED]	Process Service	\$ 763.60
[REDACTED] M.D. and [REDACTED]	subpoena	\$ 30.00
IRS	Tax returns	\$ 228.00
[REDACTED]	Comp Records	\$ 296.26
		\$ 20,138.80

Figure 4. Page 2 of Sample Closing Statement

Item 1 on the first page (blanked out) is a code that identifies the corresponding retainer statement

Item 2 (blanked out) is the name and address of the client.

Item 3 (blanked out) is the name of the plaintiffs.

Item 4 (blanked out) is the name of the defendants

Item 5 on the first page indicates what court the case was filed in. In some statements, the case was resolved before any court filing, in which case this line is blank. Item 5 also provides other information about court proceedings.

Item 6 specifies whether the case was settled, abandoned, or resolved by judgment.

Item 7 indicates the gross recovery. Note that an amount is specified, \$975,000, even though Item 6 indicates that the case was settled.

Item 9 indicates the amount paid to the client and to the various lawyers in the case. In the case described in the closing statement, the lawyers' fee was split among three law firms.

Item 12 indicates that the plaintiff's recovery was partially paid to the workers compensation fund, which had subrogation rights for amounts it had previously paid the plaintiff. In addition, Item 12 shows that the plaintiff had received a loan from a litigation financing company, which was entitled to repayment out of the plaintiff's recovery.

Expenses were too large to fit on Item 13 of the form, so the lawyers attached a separate page or rider. This page indicates that expenses totaled \$20,138.80, of which over half was paid to doctors who served as expert consultants or witnesses.. \$1,479.13 was paid for medical records, \$3,228.73 was paid for mailing, photocopying, travel and other miscellaneous expenses, and smaller amounts were paid for service of process, court fees, mediation, and other expenses. Together, the information allows calculation of how the settlement was divided:

975,000.00	Gross Recovery
20,138.80	Expenses (2% of gross recovery)
954,861.20	Net Recovery (gross recovery minus expenses)
636,574.13	Total amount paid to the client or to reimburse entities that previously paid the client (66.67% of net recovery) This is the sum of:
	451,043.09 paid directly to the client
	97,554.85 paid to workers compensation fund
	79,229.18 paid to litigation financing company
\$317,944.48	Total amount paid to lawyers (33.30% of net recovery) This is the sum of:
	143,229.18 paid to 1 st law firm
	95,486.12 paid to 2 nd law firm
	79,229.18 paid to 3 rd law firm
342.59	Unaccounted for (0.03% of net recovery)

The division of the recovery is relatively typical. Expenses are a low in relationship to gross recovery, reducing net recovery by only 2%. The client received exactly two-thirds of net recovery, which is consistent with a one-third contingent fee. Lawyer fees were shared among three firms and totaled 33.30%, slightly less than the one-third standard contingent fee. The numbers on the form do not add up properly, as there is \$342.59 unaccounted for. This amount probably represents amounts that were paid to lawyers that were somehow omitted from the form. If this amount were added to the amounts paid to lawyers specified on the form, the total amount paid to lawyers would be exactly one third, as would accord with the standard contingency fee.

While the closing statement discussed above is typical of a case in which the plaintiff received a substantial settlement, in many cases the plaintiff recovered nothing, either because the plaintiff abandoned the case or because the plaintiff pursued the case to trial and lost. Closing statements are filed even when cases were abandoned or result in defense victories. In such cases, plaintiff's lawyer receive no fee and no reimbursement of expenses.

II. Settlement Rates

Eisenberg and Lanvers have produced the most careful study of settlement rates. Using a sample of filed cases they conclude that, in the two federal district courts they analyzed, settlement rates were 57.7% and 71.6%. They note, however, that settlement rates in tort cases were 63.8% and 87.2%.¹⁹ Nevertheless, it should be noted that Eisenberg and Lanvers probably overstate the settlement rate for the districts they study, because they cannot accurately distinguish between settled and abandoned cases. They rely on "case disposition codes" in docket sheets. Unlike the closing statements studied in this article, docket sheets do not cleanly divide cases into abandoned, settled, or adjudicated categories. While about 22% of cases in docket sheets are explicitly marked as "settled," settlement must be inferred for other cases from "voluntary dismissals." This is problematic, because voluntary dismissals could also be cases abandoned by the plaintiff without settlement. This is especially likely, because the number of cases explicitly categorized by Eisenberg and Lanvers as "other withdrawn, no evidence of settlement" or "dismissal/withdrawal to allow refileing" is less than 1%. That is, Eisenberg and Lanvers categorize less than 1% of cases as abandoned, which is unrealistically low. Some of the cases they classify as settled were almost certainly abandoned.²⁰ Eisenberg and Lanvers also find that a substantial fraction (about 13%) of federal cases are resolved by dispositive motions, such as summary judgment or motions to dismiss for failure to state a claim.

There is also a conventional wisdom that 95% of cases settle. Although there is remarkably little evidence for this settlement rate, it has been well known among scholars that the settlement rate is well below 95%, the conventional wisdom seems to be remarkably persistent. One possible source of the 95% rate is Ross's classic study of closed auto insurance claims from 1962.²¹ Although that study is sometimes cited for the idea that 95% or more of cases settle, that is a misreading of Ross's findings. It is true that only 4% of claims resulted in trial, but that does not mean that 96% settled, because 32% were resolved with no recovery and no trial. Using the case disposition categories employed here for the New York data, the 32% resolved without recovery and without trial would be classified as abandoned. That would mean

¹⁹ Theodore Eisenberg & Charlotte Lanvers, *What is the Settlement Rate and Why Should We Care?*, 6 J. EMPIRICAL LEGAL STUD. 111, 130 (2009).

²⁰ *Id.* at 127–28.

²¹ See H. LAURENCE ROSS, SETTLED OUT OF COURT: THE SOCIAL PROCESS OF INSURANCE CLAIMS ADJUSTMENT (Rev. 2d ed. 1980) The figures in this paragraph are calculated by the authors of this article using data from p. 217 of Ross's book. Of course, Ross's data is different from that analyzed here, because it involves only auto accidents, whereas the cases in the New York data cover a wide range of torts, including medical malpractice and premises liability.

that the settlement rate was only 64%. Unlike the New York data studied here, Ross’s data included cases resolved without a lawyer. If only cases involving a lawyer are analyzed, the settlement rate in Ross’s data would rise to 73%, and the fraction of represented cases in Ross’s data that went to trial is 15%.

Because the New York closing statements clearly distinguish between abandoned and settled cases, and because these data are much more recent than Ross’s, they provide new insights into modern settlement rates in an important state court. Item 6 of the closing statement, with a few adjustments,²² allows categorization into three rough “case dispositions” – abandoned, settled, or adjudicated.

Table 1. Case Dispositions

	Unknown	Abandoned	Settled	Adjudicated
Number	3766	38,532	246,542	3849
% of all cases	1.3%	13.2%	84.2%	1.3%
% of cases for which disposition known		13.3%	85.3%	1.3%
% of all cases not abandoned			98.5%	1.5%
Total Cases: 292,689				

As shown in Table 1 for the New York data about 84% of all cases settled, 13% were abandoned, and less than 2% went to trial or were otherwise adjudicated. The 84% settlement rate is clearly lower than the 95% or 98% settlement rate that is sometimes mentioned but within the range of settlement rates found by Eisenberg and Lanvers for tort cases. As noted above however Eisenberg and Lanvers’s estimate counts abandonments as settlements. If abandoned New York cases were also counted as settled, or if abandoned cases are excluded, the settlement rate would be over 98%, which is extremely high.²³

New York contingent fee litigation also differs from the federal litigation studied by Eisenberg and Lanvers in that dispositive motions must be far less frequent in New York than the substantial fraction (about 13%) of federal cases Eisenberg and Lanvers find were resolved by dispositive motions. In the New York data, these would be categorized as “adjudicated.” Since

²² Cases that the closing statement says were “abandoned by client,” but in which there was a positive recovery (Item 7) were classified as of “unknown” disposition. Similarly, cases that were “settled,” but in which the recovery (Item 7) was zero were also classified as “unknown.” Fortunately, there were less than a thousand of such cases. Most cases listed as “unknown” in Table 1 were cases where Item 6 was not recorded in the court’s flat file, presumably because it was not marked on the closing statement.

²³ These calculations exclude unknown cases. If unknown cases were included the settlement rate would still be above 97%. If one analyzed the closing statements of cases that were filed in federal court, the settlement rate is 99.4%. The extremely small number of cases abandoned (47) or adjudicated (44) suggests that lawyers who filed cases in federal court did not consistently file closing statements.

less than 1.5% of New York cases are adjudicated, it is clear that many fewer were resolved by dispositive motions, such as summary judgment or motions to dismiss. This accords with the general view that federal courts are more likely to grant summary judgment than state courts, although the difference is more stark than one might have expected.

It is also helpful to distinguish between cases that were resolved before a lawsuit was filed and those that were resolved after. Obviously, adjudicated cases were resolved after a lawsuit was filed, but cases can be settled or abandoned before a lawsuit is filed. Fortunately, lawyers are required to file closing statements even in such cases. Tables 2 and 3 show case dispositions depending on whether a suit was filed or not.

Table 2. Case Dispositions, No lawsuit Filed

	Unknown	Abandoned	Settled	Adjudicated
Number	2981	28,689	100,761	0
% of all cases	2.3%	21.7%	76.1%	0%
% of cases for which disposition known		22.2%	77.8%	0%
% of all cases not abandoned			100%	0%
Total Cases: 132,431				

Table 3. Case Dispositions, Lawsuit Filed

	Unknown	Abandoned	Settled	Adjudicated
Number	785	9,843	145,781	3849
% of all cases	0.5%	6.1%	91.0%	2.4%
% of cases for which disposition known		6.2%	91.4%	2.4%
% of all cases not abandoned			97.4%	2.6%
Total Cases: 160,258				

Interestingly, almost half of cases were resolved before a lawsuit was filed. Note that these are cases in which the plaintiff had hired a contingent fee lawyer (thus triggering the obligation to filing a closing statement), so these resolutions do not include cases resolved by insurance companies or defendants without the involvement of a plaintiff's lawyer. Cases resolved without filing a lawsuit are obviously very different. Many more of them are abandoned (22% versus 6%), fewer are settled (76% versus 91%), and, of course, none are adjudicated. Interestingly, the settlement rate among filed cases is over 90%, which is higher

than in the federal data studied by Eisenberg and Lanvers, which was also restricted to cases in which lawsuits were filed.²⁴

Settlement rates in the New York data are also much higher than that reported in Ross’s study of closed auto insurance claims mentioned above. As noted above if only cases involving a lawyer are analyzed, the settlement rate in Ross’s data is 73%, which would still be much lower than in the New York data. Also, the fraction of represented cases in Ross’s data that went to trial is 15%, which is much higher than in the New York data. If one looked at cases which resulted in a lawsuit in Ross’s data, the contrast to the New York data would be even more extreme. In such cases, Ross’s data show that only 51% settled, 25% were tried, and 24% were abandoned.²⁵

III. Amounts Recovered

Table 4 summarizes the amounts recovered.

Table 4. Recovery Amounts

	Settled Cases	Adjudicated Cases	Adjudicated Cases with Non-Zero Recoveries
Mean (average)	\$89,168	\$92,534	\$315,187
10th percentile	1,534	0	3,500
25th percentile	4,796	0	10,000
50th percentile (median)	12,000	0	30,000
75th percentile	37,500	6,000	148,153
90 th percentile	125,000	78,845	659,772
95 th percentile	300,000	300,000	1,385,527
99 th percentile	1,305,681	2,010,000	5,000,000
N	246,540	3,849	1,130

²⁴ Figures in Table 3 may slightly overstate the settlement rate because some cases may have been misclassified as not filed when they were actually filed. Cases were classified as “not filed” if no court is listed in Item 5 of the closing form. *Id.* at 126. It is a reasonable assumption that such cases were not filed in court, because, if they were, one would expect the court to be mentioned. Nevertheless, about 20% of adjudicated cases also do not list a court in Item 5, suggesting that some cases which do not list a court in Item 5 were cases in which a lawsuit was filed.

²⁵ See H. LAURENCE ROSS, SETTLED OUT OF COURT: THE SOCIAL PROCESS OF INSURANCE CLAIMS ADJUSTMENT (Rev. 2d ed. 1980) The figures in this paragraph are calculated by the authors of this article using data from p. 217 of Ross’s book. Of course, Ross’s data is different from that analyzed here, because it involves only auto accidents, whereas the cases in the New York data cover a wide range of torts, including medical malpractice and premises liability.

The average recovery among settled and adjudicated cases is remarkably similar, both around \$90,000. The similar averages, however, mask large differences in distributions. Plaintiffs win only 29% of adjudicated cases, so most adjudicated cases result in no recovery for the plaintiff. In contrast, a settlement, by definition, involves some payment to the plaintiff. Most recoveries in the dataset are relatively small. 75% of settlements are \$37,500 or less, and 75% of adjudicated cases result in judgments of \$6000 or less. On the other hand, the distribution of adjudicated cases is dominated by the fact that over 70% of cases result in no recovery at all. If attention is restricted to adjudicated cases with non-zero recoveries, the seventy-fifth percentile jumps to \$148,153.

It is also helpful to look at the distributions graphically.²⁶

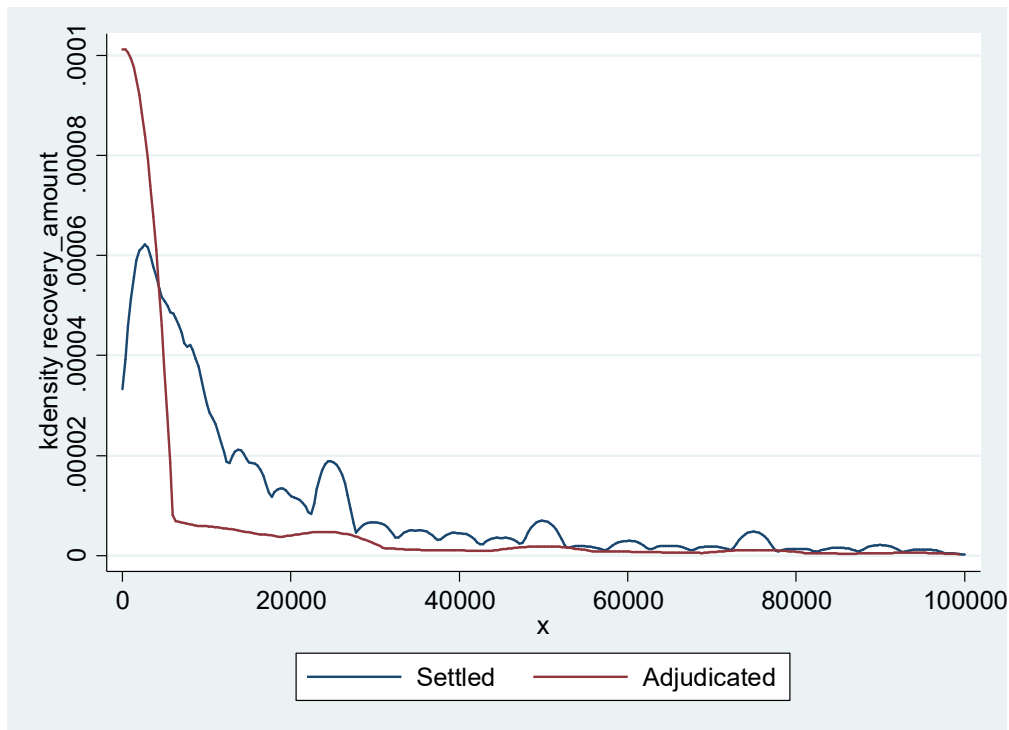


Figure 5. Distribution of Recoveries in Settled and Adjudicated Cases

²⁶ These graphs use kernel density estimation to create curves from data that is inherently discrete and lumpy.

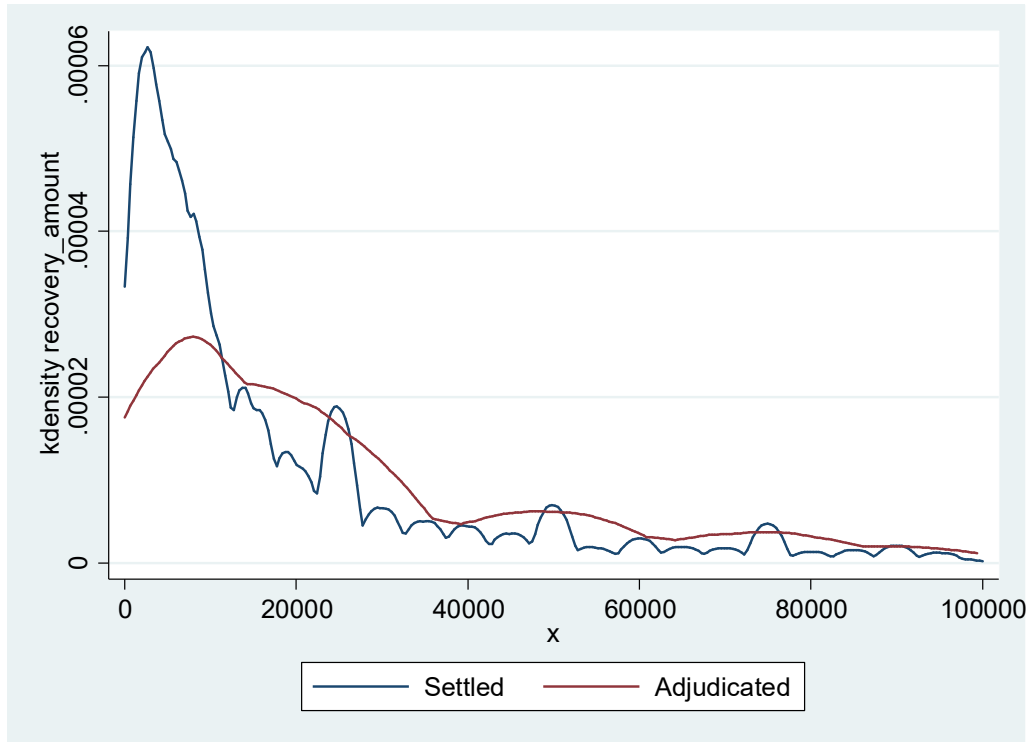


Figure 6. Distribution of Recoveries in Settled and Adjudicated Cases with Non-Zero Recoveries

In both graphs, the horizontal axis has been truncated at \$100,000. If the graphs included higher amounts, they would be nearly unintelligible because of the long “tail” to the right caused by a few very high settlements and judgments. If the horizontal axis included all recoveries, there would be a spike at the left next to the vertical axis, and then a nearly flat line just above the horizontal axis.

Figure 5 shows the large fraction of adjudicated cases that result in no plaintiff recovery, but doesn’t reveal much about the rest. Settled cases mostly result in relatively low recoveries, and, in general, the higher the recovery the less likely it is. Figure 6 shows that, if one restricts attention to adjudicated cases with positive recoveries, adjudicated recoveries tend to be higher than settled recoveries. There are fewer very low recoveries and more recoveries larger than about \$20,000.

The distribution of recoveries sheds some light on the selection of disputes for litigation.²⁷ Priest and Klein’s famous 1984 article asserted that litigated cases are a non-random

²⁷ See Eric Helland, Daniel Klerman & Yoon-Ho Alex Lee, *Maybe There is No Bias in the Selection of Disputes for Litigation*, J. INSTITUTIONAL & THEORETICAL ECON. (forthcoming) (discussing in detail the implications of the New York data for the selection of disputes for litigation).

sample of all cases.²⁸ This article has spawned a large theoretical and empirical literature.²⁹ Most of the empirical literature focuses on Priest and Klein's prediction that plaintiffs will generally win 50% of tried cases. The New York data analyzed here allows additional insight, because it reveals settlement amounts. The fact that mean recoveries for settled and adjudicated cases are nearly identical might suggest that there is relatively little selection. That is, settled and adjudicated cases are similar, at least in their average value. This is inconsistent, for example, with simple asymmetric information models of litigation and settlement, which would suggest that litigated cases would be skewed towards those that favor the more informed party.³⁰ So, if, as is usually assumed in tort cases, the defendant is better informed, one would expect defendants to settle when the plaintiff has a strong case and adjudicate when the plaintiff's case is weak. This would imply that settlements should, on average, be higher than adjudicated recoveries.

Nevertheless, the fact that average recoveries are similar in litigated and settled cases could be made consistent with selection theory, if additional assumptions are made. For example, if plaintiffs were usually better informed, that would lead to adjudicated cases with higher average recoveries. So, if plaintiffs were sometimes better informed, and, other times defendants were better informed, the net result might be that average settlements and adjudicated recoveries were the same, as observed.

The fact that plaintiffs win only 29% of litigated cases is inconsistent with Priest and Klein's famous 50% prediction, but is consistent with their broader theory, which provides many explanations for deviations from 50%, including asymmetric stakes.³¹ If, for example, defendants suffer reputational harm from an adverse judgment, they have more at stake and one would expect defendants to settle all but the cases they are most likely to win, leading to plaintiff trial win rates less than 50%.

IV. Fees and Expenses

The analysis of attorneys' fees and expenses is, at the moment, tentative, because New York court administrators did not include this information in their flat file, so we must extract it

²⁸ George L. Priest & Benjamin Klein, *The Selection of Disputes for Litigation*, 13 J. LEGAL STUD. 1 (1984).

²⁹ See Daniel Klerman & Yoon-Ho Alex Lee, *Inferences from Litigated Cases*, 43 J. LEGAL STUD. 209, 210 n.1 (2014) (noting that the 1984 Priest and Klein article is "one of the most influential law articles of all time" and citing some of the relevant literature).

³⁰ See Lucian Arye Bebchuk, *Litigation and Settlement under Imperfect Information*, 15 RAND J. ECON. 404 (1984); Jennifer F. Reinganum & Louis L. Wilde, *Settlement, Litigation, and the Allocation of Litigation Costs*, 17 RAND J. ECON. 557 (1986); Abraham L. Wickelgren, *Law and Economics of Settlement*, in RESEARCH HANDBOOK ON THE ECONOMICS OF TORTS 330, 336 (Jennifer Arlen ed., Edward Elgar Publishers 2013).

³¹ See Daniel Klerman & Yoon-Ho Alex Lee, *The Priest-Klein Hypotheses: Proofs and Generality*, 48 INT'L REV. L. & ECON. 59 (2016).

from the PDFs. The information in this section is based on hand coding a random sample of one hundred cases.

Expenses are a heterogeneous category that includes court filing fees, expert fees, record fees, postage, transportation, service of process, deposition transcription, and other services. Lawyer fees are, of course, distinct from expenses. Median expenses were 3% of recoveries. Average expenses were somewhat higher (5%), because of a small number of cases with very high expenses. In general, while expenses went up as recoveries went up, the percentage of money paid in expenses went down. As would be expected, the expense percentage also went up when cases were filed (as opposed to resolved before suit) and as the case went through more procedural stages (such as motions), although even if a suit was filed and a motion or other procedure requiring a “request for judicial intervention” was initiated, the expense percentage only rose to 6%. Experts were hired in 21% of cases, and, if an expert was hired, expert fees averaged \$3,654 per case. Court fees averaged \$321 in the cases where suit was filed.

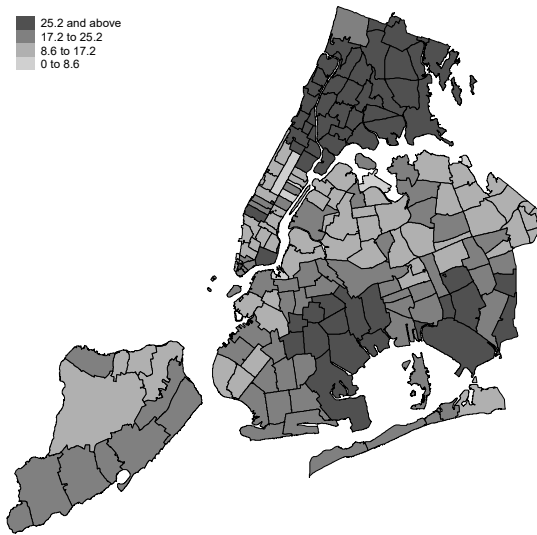
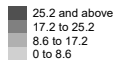
Attorneys’ fees were exactly one third of net recovery in most cases. Net recovery is total recovery minus expenses. One third is the maximum allowed by the New York courts, except when a sliding scale fee is used, which is rare. Even when there were deviations from one-third, the deviations were usually small (e.g. lawyers’ fees between 30% and 40%). In only three cases were fees lower than 30%, and in only two cases were fees higher than 40%. One possible explanation for fees lower than a third is that the retainer contract specified a lower percentage. Unfortunately, we were not able to confirm that by looking at the retainer statements, because retainer statements were missing for those cases, probably because they were signed before 2004 and thus not turned over to us. Another reason attorneys’ fees are sometimes lower than one third is that, in medical malpractice cases, a lower percentage is required by New York Law.³² Attorneys’ fees higher than one third are more puzzling, because they are forbidden by law. In one case, it appears that the lawyers may have received judicial permission to go over the cap.

V. The Demographics of Tort Claims

Because the data set includes the plaintiff’s address, it is possible to map claims by zip code:

³² N.Y. JUD. § 474-a (LexisNexis 2017).

Claims Per 1000 Population



Median Household Income

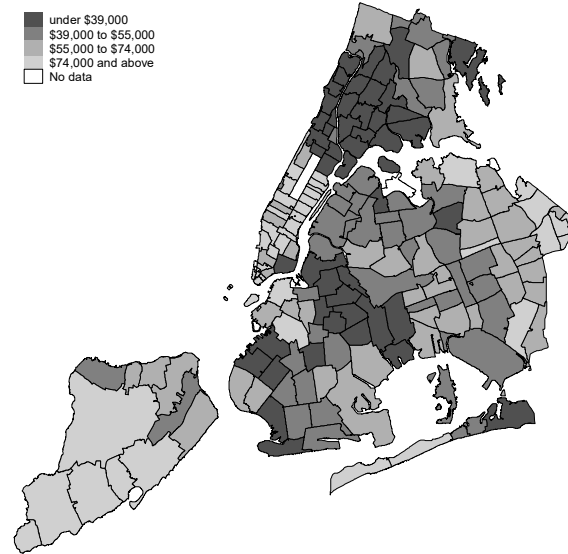
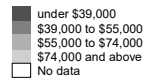


Figure 7. Suits per capita (by zip code)

Figure 8. Income per Capita (by zip code)

The figure on the left shows claims per capita. Zip codes with more suits are darker. The map on the right shows income per capita.³³ Zip codes with higher median household income are darker. The two maps are almost identical. Places with many claims per capita (dark areas on the left map), such as Harlem and the Bronx, tend to be places with low per capita income (dark areas on the right map). Conversely, places with relatively few suits (light areas on the left graph), such as the Upper East and Upper West Sides of Manhattan, tend to be places with high incomes (light areas on the right map).

There are a number of possible explanations for these patterns that we intend to explore in more depth in a future article. One possibility is that poor people are more likely to be injured, perhaps because they are more likely to live or work in less safe conditions, more likely to use unsafe products, and/or more likely to receive inferior medical care. Another possibility is that poor people are more likely to hire a lawyer and make a claim if injured, perhaps because they are less likely to have insurance or because recoveries of a few thousand dollars, which make up a substantial fraction of all recoveries, would mean more to a poor person than to a rich person.

VI. Conclusion

The New York courts have collected a wealth of data on contingent fee litigation. This article barely scratches the surface of what the data contain. Among this article’s findings are that settlement rates are higher than in federal court, that very few cases are resolved by

³³AMERICAN FACTFINDER, <https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml> (Figures created using Data from the 2010 US Census) [<https://perma.cc/J994-DYV5>].

dispositive motions, that litigated cases and settled cases have almost exactly the same average recovery, that poor people are more likely to make a tort claim, and that expenses average 3% of gross recovery, while attorneys' fees are almost always one third of net recovery, the maximum allowed by law.

Maybe There's No Bias in the Selection of Disputes for Litigation

Eric Helland, Daniel Klerman, and Yoon-Ho Alex Lee¹

Forthcoming in *Journal of Institutional and Theoretical Economics*

Draft 9/19/2017

ABSTRACT

New York “closing statement” data provide unique insight into settlement and selection. The distributions of settlements and adjudicated damages are remarkably similar, and the average settlement is very close to the average judgment. One interpretation is that selection effects may be small or non-existent. Because existing litigation models all predict selection bias, we develop a simple, no-selection-bias model that is consistent with the data. Nevertheless, we show that the data can also be explained by generalized versions of screening, signaling, and Priest-Klein models.

1. Introduction

For more than three decades, legal scholars have been fascinated and frustrated by the selection of disputes for litigation. The theory tantalizingly suggests that reality may be very different from what we observe. Nevertheless, because this conclusion relies on the contrast between observed judgments and usually unobservable settlements, it is vexingly difficult to test. In addition, because selection theory piggybacks on models of litigation and settlement, it reflects both the power of those models and their limitations, including unobservable parameters and unrealistic simplifying assumptions.

This article attempts to contribute to the empirical and theoretical literature on litigation, settlement, and selection by analyzing a unique dataset. For the last few decades, New York courts have required contingent fee lawyers to file “closing statements” that include a wealth of information, including, most importantly, settlement amounts. Because settlement amounts are usually confidential, these data provide an unusual and potentially powerful opportunity to test theories of litigation and selection.

Most “tests” of selection theory have focused on the plaintiff trial win rate. Priest and Klein (1984) famously hypothesized that, among the disputes that go to trial, plaintiffs would generally win fifty-percent of the time. No serious scholar, however, tests just for conformity with the fifty-percent prediction, because even Priest and Klein pointed out that the plaintiff trial win rate would often deviate from fifty percent for a variety of reasons, most importantly

¹ The authors are grateful for comments and suggestions to Jonah Gelbach, Urs Schweizer, Martin Wells, and participants in Seminar on Empirical Methods for Law sponsored by the Max Planck Institute for Research on Collective Goods and University of Bonn.

asymmetric stakes. In addition, Shavell (1996 p. 493) pointed out that, under Bebchuk's (1984) screening model, "any frequency of plaintiff victory at trial is possible." Nevertheless, by measuring the effect of factors that theory predicts affect the plaintiff trial win rate, one can test at least some versions of selection theories. See, e.g., Kessler, Meites & Miller (1996).

More sophisticated tests of selection theories, such as Waldfogel (1995, 1998), examine both the plaintiff trial win rate and the settlement rate. Nevertheless, as Gelbach (2016) has pointed out, nearly any combination of win rates and settlement rates can be accommodated under a "generalized Priest-Klein model." This article similarly shows that nearly any combination of win rates and settlement rates can be consistent with generalized versions of the standard screening and signaling models. So even empirical work using both win rates and settlement rates is unlikely to "prove" which settlement selection model (if any) is correct.

Other tests of selection theories have used unusual datasets that include information on settled cases. Klerman (2012) examines thirteenth-century privately-prosecuted criminal cases, where, for a short period of time judges took jury verdicts even in settled cases, so one can compare the prosecutor's win rate in settled and non-settled cases. Similarly, Studdert and Mello (2007) had experts score insurance files in medical malpractice cases, so they can compare the strength of litigated and settled cases. Both Klerman (2012) and Studdert and Mello (2007) confirm the existence of selection effects with respect to liability, the probability that the plaintiff will prevail. Both show that strong cases (where the plaintiff has a high probability of prevailing) are more likely to settle, whereas weak cases (where the plaintiff has a low probability of prevailing) are more likely to result in litigation.

The analysis in this article is in the spirit of Klerman (2012) and Studdert and Mello (2007) in that it takes advantage of a dataset with information on settled cases. Unlike Klerman (2012) and Studdert and Mello (2007), the data examined here are about settlement amounts, rather than the strength of settled cases.

The data on damages suggests, surprisingly, that there may be little or no selection effects. The average settlement is almost identical to the average judgment in litigated cases. The distributions of settlements and judgments are also similar. The simplest interpretation of the data is that there is no discernible selection bias. Litigated cases look very similar to settled cases. This is surprising, because all existing models of litigation and settlement suggest that there should be selection effects. This article, therefore, sets out a simple model that predicts no selection bias and is consistent with the New York data. Klerman and Lee (2014) had argued that selection effects are partial. This article goes further and presents data and theory that suggest that, at least sometimes, there may be no selection effects at all.

Moreover, the fact that average damages equal average judgments is inconsistent with simple versions the three canonical litigation models: Bebchuk's (1984) screening model, Reinganum and Wilde's (1986) signaling model, and Priest and Klein's (1984) model. Nevertheless, we also show that all three of these models can be generalized to be consistent with a wide range of observed data, including the New York data analyzed in this article.

Thus, even with the unusually rich data analyzed here, it is not possible to reject any of the standard models. The best one can say is that some models require more restrictive assumptions than others. The "no selection bias" model requires the fewest parameter assumptions, and, unlike the screening and signaling models, does not require litigation costs to

be low. Occam's razor might, therefore, suggest that the no-selection bias model is the most plausible explanation of the New York data.

The theoretical results in this article—that standard models can be generalized to explain a wide range of data—are in line with other papers that show non-falsifiability. For example, Shavell (1996) shows that the Bebchuk screening model can generate any plaintiff trial win rate, and thus that that model cannot be falsified using win-rate data. Similarly, as discussed further in Section 4.C.ii, Gelbach (2016) shows that a generalized version of the Priest-Klein model can generate nearly any combination of plaintiff trial win rates and settlement rates, so such data cannot be used to show the superiority of asymmetric information models. Relatedly, Gelbach (2012, 2012a) shows that changes in pleading standards are consistent with both increases and decreases in motions-to-dismiss filing and grant rates.

Section 2 describes the data and five facts about settled and litigated disputes in New York: (1) the settlement rate is very high (98.5%), (2) plaintiffs win 29% of litigated cases, (3) the average settlement amount equals the average judgment amount, (4) judgment amounts range widely, and lower judgments are more common than higher ones, and (5) settlement amounts also range widely, lower settlements are more common than higher ones, and the distribution of settlements roughly matches the distribution of expected judgments. Section 3 sets out a simple model with no selection bias that is consistent with the data. Section 4 shows that the data are inconsistent with simple versions of all three canonical settlement models. It also shows, however, that generalizations of these models allow all of them to generate predictions consistent with the data. Section 5 concludes.

2. Data and Five Facts to be Explained

The data used in this article are described in some detail in Helland, Klerman Dowling & Kappner (2017). As such, this section will highlight only a few especially relevant aspects. Since 1957, courts in New York City have required contingent fee lawyers to file “closing statements” in cases involving personal injury, property damage, or wrongful death. The main purpose of these closing statements is to ensure that the lawyer does not overcharge the client, so these statements include an accounting of payments received from the defendant or defendant's insurer (whether by judgment or settlement), expenses paid by the lawyer (such as filing fees or payments to experts), the lawyer's fee, liens on the judgment (such as amounts paid by workers' compensation or litigation financing), and the amount actually paid to the client. This article focuses on amounts paid by the defendant, whether as the result of settlement or court judgment.

Closing statements must be filed whenever a lawyer retained on contingent fee resolves a case, whether the case settled or went to judgment, and even if the case settled without ever having been filed in court. That is, even settlements that were reached wholly without judicial intervention, or even without contact with the courts, still trigger the obligation to file a closing statement. In addition, unlike most insurance closed claim data, these statements are filed even if the plaintiff recovered nothing, so they also provide information on the plaintiff trial win rate.

Table 1 and Figure 1 describe the distribution of settlement and judgment amounts:²

	Settled Cases	Adjudicated Cases	Adjudicated Cases with Non-Zero Recoveries
Mean (average)	\$89,168	\$92,534	\$315,187
Standard deviation	1,723,505	633,879	1,139,840
10th percentile	1,534	0	3,500
25th percentile	4,796	0	10,000
50th percentile (median)	12,000	0	30,000
75th percentile	37,500	6,000	148,153
90 th percentile	125,000	78,845	659,772
95 th percentile	300,000	300,000	1,385,527
99 th percentile	1,305,681	2,010,000	5,000,000
N	246,540	3,849	1,130

Table 1. Recovery Amounts

² Thirteen percent of cases in the New York data set were abandoned. Those cases are excluded from the analysis in this article. While such cases could be classified as settlements for \$0, that is not a plausible way of analyzing abandoned cases, because none of the standard models discussed in this article predict a large fraction of settlements for \$0. Rather, abandoned cases probably reflect nuisance suits the defendant refused to settle or cases where investigation or discovery revealed that the plaintiff's case was very weak. Analysis of abandoned cases would require more complicated models that take into account nuisance suits and/or the revelation of information over time.

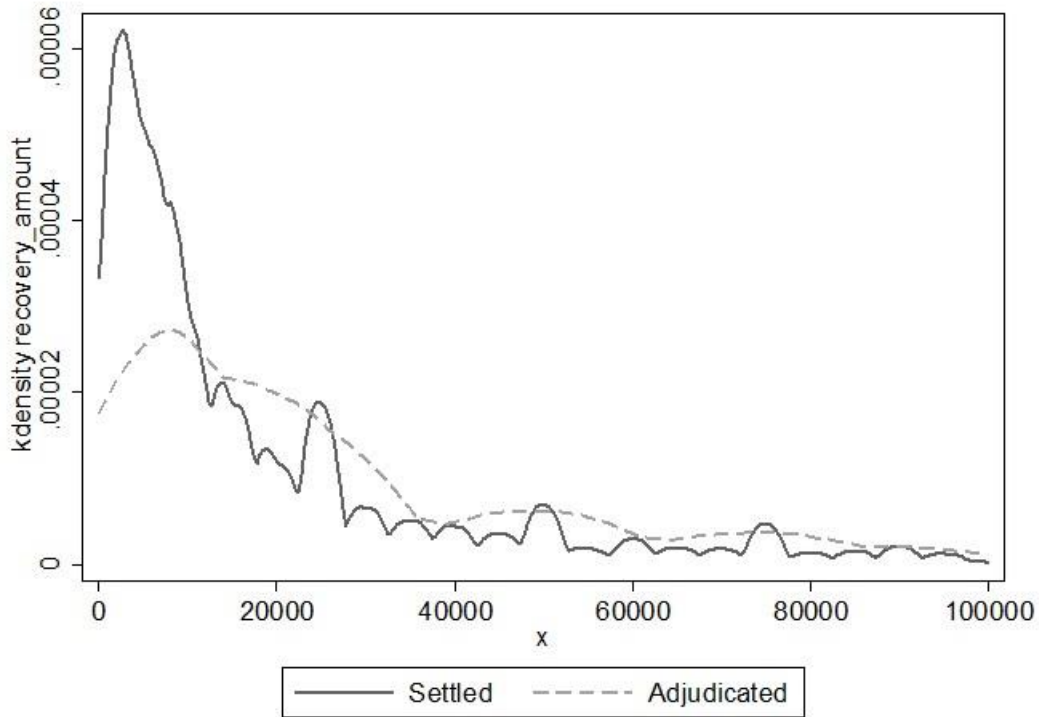


Figure 1. Distribution of Recoveries in Settled and Adjudicated Cases with Non-Zero Recoveries

It is especially illuminating to graph settlements and *expected* judgments. Graphing expected judgments requires taking into account judgments for the defendant, where the plaintiff recovered nothing. Of course, we do not know what the judgment would have been in those cases, so it is impossible to know whether defendant judgments are more common among low-damage cases, high-damage cases, or equally common among all levels of damages. Figure 2 shows the distributions of settlements and expected judgments, if judgments for the defendant (zero judgments) were equally common among all levels of damages.

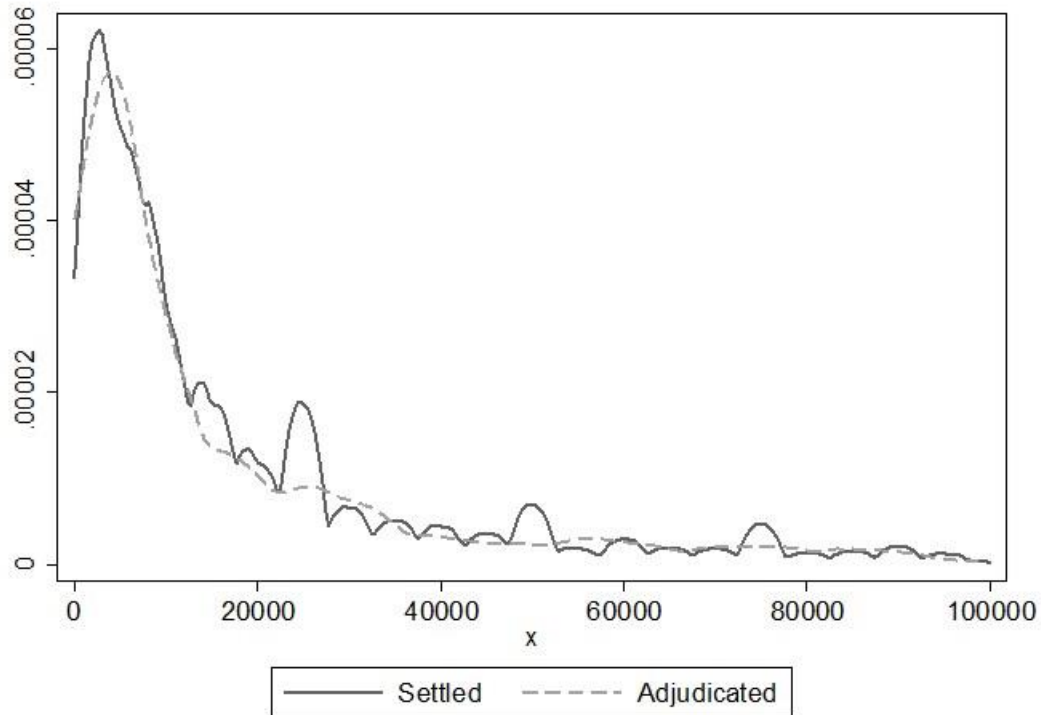


Figure 2. Distribution of Settlement Amounts and Expected Trial Judgments
(Assuming Plaintiff Trial Win Rate is 29% for All Damage Levels)

The two distributions are remarkably similar. Not only are average settlements and judgments nearly the same, but the two may share nearly identical distributions. Of course, there are some differences. For example, the standard deviation of settled cases is very high. This reflects the disproportionate influence of a few outliers.³ In addition, settlements seem to clump together at round numbers, such as \$25,000, \$50,000, and \$75,000, causing humps in the distribution of settled cases.⁴ Although the graphs above are truncated at \$100,000, the similarity continues for higher recoveries. The similarity for the entire distribution is best seen in a graph where both settlement and expected judgment are expressed in logarithms.

³ So, for example, if the top 0.1% of settled and adjudicated cases are dropped, the standard deviation of recoveries in adjudicated cases actually drops below that of adjudicated cases (266,369 versus 445,473 for all adjudicated cases and 960,486 for adjudicated cases with non-zero recoveries). Similar results hold for trims between 0.1% and 1.0%. Of course, trimming the outliers also affects the mean. The effect on the mean varies. When 0.3% are trimmed, the means of settled and adjudicated cases are closer together. At other trim levels, the means are farther apart.

⁴ Primarily because of this clumping, statistical tests fail to reject the hypothesis that the distribution of settled and adjudicated case is different. A Kolmogorov-Smirnov equality-of-distributions test rejects the equality of the two distributions at a high level of significance (p -value = .000). A Wilcoxon rank-sum test also rejects that the two distributions are the same with a z statistics of 77.008 (p -value=.0000).

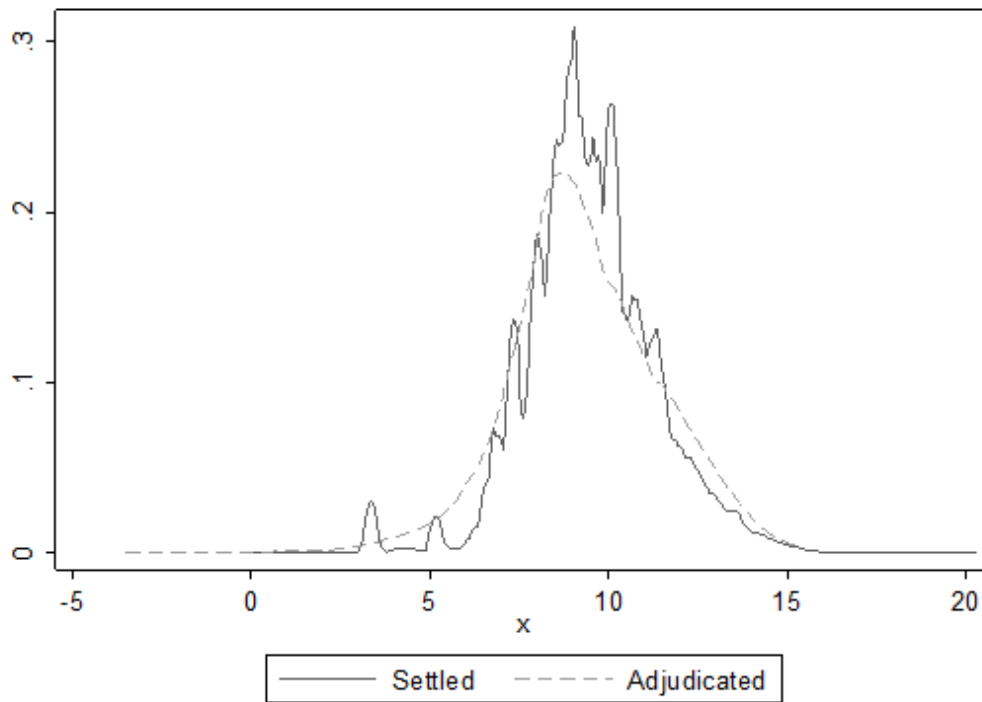


Figure 3. Distribution of Logged Settlement Amounts and Logged Expected Trial Judgments (Assuming Plaintiff Trial Win Rate is 29% for All Damage Levels)

As can be seen, the overall distributions remain similar when recoveries are logged, although the distribution of settled cases is less smooth and somewhat more concentrated near the center of the logged distribution.⁵

The similarity in the means and distributions suggests that there may be very little selection bias. The sections that follow show how that could be consistent with theory – both with a simple (and new) model with no selection effects, and with generalized versions of the three most influential existing models. Of course, other explanations are possible. For example, it might be that high damages cases are more likely to settle (which would make the average settlement higher than the average judgment, if bargaining power were equal), but defendants have greater bargaining power (bringing the average settlement down to the average judgment).⁶

For the purposes of this article, the data reveal five salient characteristics of contingency-fee litigation in New York City:

- 1) The settlement rate is 98.5%, which is very high.

⁵ The authors thank Martin Wells for suggesting graphing the log of settlements and expected recoveries.

⁶ Standard models predict that settlement rates go down with damages, because litigation costs go up less than proportionally. On the other hand, standard models assume risk neutrality, and, if parties are risk averse, settlement rates might go up with damages.

- 2) The plaintiff trial win rate is 29%. That is, of cases that were resolved by judgment, the defendant won 71% of the time, and the plaintiff received a non-zero judgment 29% of the time.⁷
- 3) The average settlement is almost identical to the average judgment. The average payout in settled cases is \$89,168, and the average judgment (including zero judgments for the defendant) is \$92,534.
- 4) Judgment amounts follow the distributions illustrated in Table 1 and Figure 1. In particular, non-zero judgments range widely, from less than a dollar to well over a million, and the density of judgments falls as judgment amounts increase, except for very low judgments.
- 5) Like judgments, settlements range from one dollar to well over a million, and the density of settlement amounts falls as settlements amounts increase, except for very low settlements. More particularly, the distribution of settlements is very similar to the distribution of the expected value of judgments, i.e., to the product of the plaintiff trial win rate times actual judgments, as depicted in Figure 2.

The rest of this article shows how these five facts are or are not consistent with theoretical models of settlement, including the most common models, generalizations of them, and a new no-selection-bias model.

3. *A Simple Model with No Selection Bias*

Consider a simple inconsistent-priors model where the parties' estimates of damages are, on average, unbiased and equally accurate.⁸ Let $J > 0$ be the true damages if the case were to go to trial and plaintiff were to win. The parties, who are risk-neutral, do not know J and have inconsistent priors about the true value of J . The plaintiff estimates the damages to be $J_p = J(1 + \varepsilon_p)$, and the defendant as $J_d = J(1 + \varepsilon_d)$, where ε_p and ε_d are independent errors, uncorrelated with J , have mean zero, and are uniformly distributed⁹ over the interval $[-k, k]$, where $0 < k < 1$.¹⁰ For example, if the true damages are 100 and $k = 0.8$, then the plaintiff's or defendant's estimate damages can be as low as 20 or as high as 180. The fact that ε_p and ε_d are always greater

⁷ It should be noted that the plaintiff trial win rate is very low and would make the plaintiff's threat to go to trial non-credible if one made the standard assumption that litigation costs were one-third of damages. The rest of this article assumes litigation costs are low enough to solve the plaintiff's threat-credibility problem. See also Bebchuk (1996) on why plaintiff credibility might not be a problem even if litigation costs were higher.

⁸ Urs Schweizer's comments shows that similar results can be achieved by assuming that the parties receive private signals. See ___.

⁹ Any mean zero, symmetric distribution, such as truncated normal, would work equally well. Uniform distribution makes the math and proofs easier.

¹⁰ Urs Schweizer's comment shows that the model can be further simplified by assuming errors take three values rather than being uniformly distributed. See ___. This assumption would make the proof of Proposition 1 much simpler and more elegant. In his model, average judgments would converge to average settlements as parties became more accurate (e.g. as λ_p and λ_d go to zero). Similarly, the distribution of settlements would converge to the distribution of expected judgments.

than negative one assures that the parties' estimates of J are always positive. For another model featuring party estimates with bounded errors, see Schweizer (2016). Note that as k approaches zero, the parties' estimates of J will become more and more accurate, and when k is small enough, all cases will settle. Importantly, each party knows the other's prior, and yet they do not learn from the other's estimates and cannot come to an agreement on the value of J . For a model of this type, see Spier & Prescott (2016).

The probability that the plaintiff will prevail is p , which is fixed, invariant to J , and common knowledge. The parties estimate that their litigation costs will be proportional to their estimates. Thus, the plaintiff expects her litigation cost to be $C_p = \gamma J_p$, and the defendant estimates his to be $C_d = \gamma J_d$. We assume $0 < \gamma < p$ to ensure that litigation costs are positive and that the plaintiff's threat to go to trial is credible. As is conventional, settlement is assumed to be costless and to require no litigation expenditures.¹¹ The parties settle if and only if there is a settlement amount, S , that both plaintiff and defendant would perceive as making them better off than if the case went to trial: $pJ_p - C_p \leq S \leq pJ_d + C_d$.¹² Parties have equal bargaining power, so, if they settle, they settle at the mid-point of the bargaining range, $S = [(pJ_p - C_p) + (pJ_d + C_d)]/2$. Under this model, we show that there is no selection bias.

PROPOSITION 1. (THE NO-SELECTION-BIAS MODEL). Under the model set out in the prior paragraph, all case types (levels of damages) are equally likely to be litigated, so there is no selection bias in disputes that go to trial. In addition, as the settlement rate approaches one (e.g., as k gets sufficiently small or as γ approaches p), the average settlement amount approaches the average judgment. In particular, $E[pJ] < E[S] < E[pJ]/t$, where t is the settlement rate.

The proof is included in the Appendix. The main no-selection-bias result, however, is easy to see. With a little rearranging, the necessary and sufficient condition for settlement can be rewritten as $\varepsilon_p \leq \varepsilon_d(p + \gamma)/(p - \gamma) + 2\gamma/(p - \gamma)$. Since ε_p , ε_d , γ , and p are uncorrelated with J , this means that each case type (damage level) is equally likely to go to trial, so there are

¹¹ These assumptions about litigation costs may seem inconsistent with the fact that the plaintiff in the New York data had hired a contingent fee lawyer who would take his/her percentage even if the case settled. Nevertheless, as discussed in Klerman and Lee (2014) pp. 233-34 n. 21, it is not appropriate to model litigation costs as one-third the settlement or judgment. This would yield the untenable implications that (a) trial is no more costly than settlement and (b) that there are zero litigation costs when the defendant wins at trial. Instead, it is best to assume that the plaintiff's lawyer decides (or has great influence on) whether a settlement offer is made or accepted and that, for the lawyer, litigation costs are proportional to time spent on the case, which may have very little to do with the amount actually received in judgment or settlement in a particular case. For simplicity, we assume that lawyer time to litigate a case to judgment is proportional to the amount at stake, J . It is also important to note that, because settlement costs are normalized to zero, litigation costs, C_p and C_d , represent the incremental cost of trial over settlement, rather than total litigation costs.

¹² Because inconsistent-priors models, like this one, assume that the priors are common knowledge, settlement occurs whenever the plaintiff's expected gain is smaller than the defendant's expected loss, net of litigation costs. There is no need to worry about the inefficiency of bargaining, Myerson & Satterthwaite (1983), or incentive-compatible bargaining mechanisms (such as take-it-or-leave-it offers). See Spier & Prescott (2016).

no selection effects. Litigated cases will look like a random sample of all cases. The proof in the Appendix establishes the relationship between average settlement and average judgment.

With this model, it would be relatively simple to replicate the five facts about the New York data from the prior section. The settlement rate can be 98.5% simply by making the parameter related to the size of party errors, k , only slightly larger than γ/p . (If $k = \gamma/p$, then all cases would settle). The plaintiff trial win rate can be 29% (or any other value) simply by setting p equal to the plaintiff trial win rate. With a settlement rate of 98.5%, the model predicts that the average settlement will be within 1.5% of the average judgment, which means that average settlement would be very close to average judgments, as observed.¹³ By setting the distribution of true damages, J , to match the observed distribution of judgments, the distribution of judgments will match the observed distribution, and settlements will follow a similar distribution.

4. Consistency with Canonical Litigation Models

A. Screening Model

i. Simple version

Under Bebchuck's (1984) original screening model, a risk-neutral plaintiff sues a risk-neutral defendant. The damages, J , that the plaintiff would win if the case went to trial is assumed to be known by both parties, but the defendant knows the probability that the plaintiff will prevail, p , whereas the plaintiff knows only the probability density function of p , $f(\cdot)$ defined over $[p, 1]$. This model will not generate any dispersion in damage amounts, so it is not appropriate for the New York data.

Nevertheless, the model can be easily modified to produce judgments with a distribution of damage amounts. Instead of assuming that the damage amount is fixed and common knowledge, one could assume that damages follow a probability distribution, $g(\cdot)$, defined over $[J, \infty]$, that the plaintiff knows the true damages, but that the defendant knows only the distribution. Similarly, whereas in Bebchuk's original model, the amount of damages was assumed to be fixed and common knowledge, in this modified model, the plaintiff's probability of prevailing, p , is assumed to be fixed and common knowledge.¹⁴ The defendant makes a take-it-or-leave-it settlement offer, S , to the plaintiff, which the plaintiff accepts or rejects depending on litigation costs and its private information about damages. The plaintiff and defendant are assumed to have litigation costs C_p and C_d , respectively, where $C_p < pJ$, so plaintiff always has

¹³ The average settlement would be within 1.5% of the average judgment according to Proposition 1, because $1/t = 1/0.985 \approx 1.015$. Of course, the model predicts that average settlements will be slightly *higher* than average judgments, where in the data they are slightly *lower*. This may reflect the fact that the model assumed equal bargaining power, whereas, in reality, defendants may have more bargaining power. Note also, that, depending on what percentage of cases are trimmed, the average settlement may be higher than the average judgment. See note _.

¹⁴ The probability that the plaintiff prevails could also take on a distribution of values, as long as the mean was known to both parties and did not vary with damages, J .

a credible threat to go to trial. Plaintiff accepts the settlement offer if $S \geq pJ - C_p$. That is, plaintiff accepts if its type is below a threshold value, $J^* = (S + C_p)/p$, or, equivalently, if the expected value of the case (disregarding litigation costs) is less than pJ^* .

The defendant's optimal settlement offer, S , depends on the distribution of disputes, $g(\cdot)$, which is the distribution of damages if all cases were litigated. The distribution of disputes will also determine the fraction of cases that settle or litigate. If most cases have damages very close to \underline{J} , but there are some cases with very high damages, then the settlement rate will be very high because the optimal settlement offer, S , will be set to extract low settlements from the large number of low damage cases, resulting in litigation with the small number of plaintiffs who know that the damages are high. Conversely, if a fair number of cases have low damages, but a significant number have much higher damages, the defendant's settlement offer will be set to settle with the low cases but not the larger number high damage cases, and the settlement rate will be lower. The selection of disputes under the screening model for a distribution of damages similar to that in the New York data is depicted graphically in Figure 4.

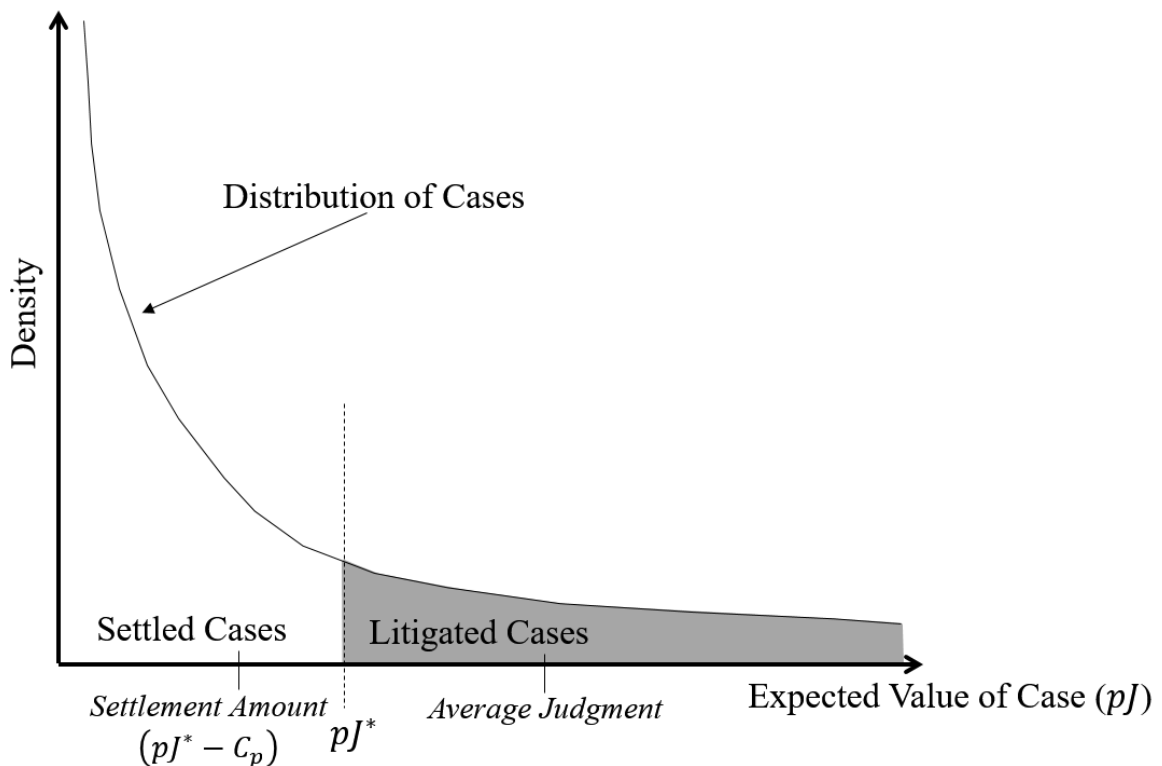


Figure 4. Selection under the Simple Screening Model
(Assuming Plaintiff Has the Informational Advantage)

All cases with expected value above the threshold, pJ^* , go to trial. This means that the average judgment is larger than the average settlement. This suggests that the simple screening model

could yield results that predict only two of the observed facts about the New York cases in the dataset.

- 1) The settlement rate could be very high (98.5%) by assuming a distribution of disputes with many, clustered low value suits and a small number of high value disputes.
- 2) The plaintiff trial win rate could be 29% by simply assuming that p , the unvarying probability that plaintiff prevails, is 0.29.
- 3) Average judgments will *not* equal average settlements. As noted above, the plaintiff only rejects the offer, S , if $S < pJ - C_p$. Since pJ is the expected value of litigated disputes, and S is the settlement value, the (average) settlement amount will be significantly lower than the expected value of litigated disputes.
- 4) It will also be impossible to get the model to match the distribution of judgments, because litigated judgments should never go below the threshold established by the settlement offer. It is clear from the data that there are many judgments lower than settlement offers. In fact, the lowest judgment in the data is 10 cents.
- 5) Settlement values will not range widely (or at all) or be similar to the distribution of judgments, because, under the screening model, there is only a single settlement amount which is offered to all defendants.

Reversing the information structure, and assuming that defendant knows the damages and plaintiff knows only the distribution, would not help. The results would essentially just flip the shaded and unshaded areas in Figure 4, and it would still be impossible for the model to generate a range of settlement amounts or for the average settlement to equal the average judgment.

ii. Generalized Screening Model

Although the screening model in its original version is not consistent with the five observed facts about New York litigation, the model can be generalized to be consistent with them (and, in fact, with a wide array of data on settlement rates, plaintiff trial win rates, and the distribution of settlement and judgment amounts). To generalize the model, we start with Bebchuk's original version of the screening model, where damages are fixed and there is asymmetric information about the plaintiff's probability of prevailing. Then we set the distribution of disputes, the fraction of cases in which plaintiff (or defendant) has superior information, and other parameters so that the model generates the desired settlement rate and plaintiff trial win rate. Then we assume that potential damages (damages in all cases, including damages in cases that settle or result in judgments for the defendant), in fact, vary in ways consistent with the observed judgments, that damages are common knowledge to the parties, and that the parties play the Bebchuk screening game with the parameters just mentioned for each level of damages.

More formally, consider the following definitions, proposition, and proof.

DEFINITION 1. A data set is described by *aggregate data profile* $(t, p_T, \beta, g(x), s(x))$ if:

- 1) The settlement rate is $t \in (0,1)$.

- 2) The plaintiff trial win rate is $p_T \in (0,1)$.
- 3) $\beta = [\text{the average settlement}]/[\text{the average judgment}]$, where $\beta \in \mathbf{R}^+$
- 4) Positive judgment amounts vary according to the distribution, $g(x)$.
- 5) Settlement amounts vary according to the distribution, $s(x)$.

DEFINITION 2. Given an aggregate data profile $(t, p_T, \beta, g(x), s(x))$, a model is *non-falsifiable* if there exists a set of parameters such that the model predicts the aggregate data profile $(t, p_T, \beta, g(x), s(x))$ or can approximate $(t, p_T, \beta, g(x), s(x))$ arbitrarily closely. In the latter case, we require only that there exists a sequence of parameters such that the model generates a sequence of aggregate data profiles that converges to $(t, p_T, \beta, g(x), s(x))$ in a pointwise manner.

The idea behind non-falsifiability is that a given litigation dataset does not invalidate a particular model unless there is no set of model parameters that predict the aggregate data profile. Of course, an aggregate data profile may be generated by more than one set of parameters. Thus, we only consider whether an aggregate data profile can be generated by at least one set of parameters.

In our dataset, we have $(t, p_T, \beta, g(x), s(x)) = (0.985, 0.29, 1, g(x), s(x))$, where $g(x)$ and $s(x)$ are the observed distributions of judgments and settlements. Furthermore, we noted that the distribution of settlements appears to be in a particular relationship with $g(x)$. Namely, $s(x) \approx g(x/\beta p_T)/\beta p_T$. For this reason, we will focus our results to the non-falsifiability of generalized screening (and signaling) models with respect to $(t, p_T, \beta, g(x), g(x/\beta p_T)/\beta p_T)$.¹⁵ Note that β can be greater than, equal to, or less than 1. For ease of exposition, we will assume $\beta p_T < 1$ and $p_T > 0$. In other words, it cannot be too much greater than 1.

In the case of generalized screening models, we find that the aggregate data profile can be generated exactly. In the case of generalized signaling models, we show only that the aggregate data profile can be generated arbitrarily closely. It should be noted that our results are sufficient but not necessary conditions for establishing non-falsifiability. The fact that some of the required conditions do not hold does not mean that either of these models has been falsified.

Under this set-up, we have the following result.

PROPOSITION 2. (THE NON-FALSIFIABILITY OF THE GENERALIZED SCREENING MODEL). Suppose $\beta p_T < 1$ and $p_T > 0$. Let the cost of going to trial be γJ for both the plaintiff and the defendant, and let J be the damage amount if the plaintiff prevails. Then the generalized screening model is non-falsifiable for aggregate data profile $(t, p_T, \beta, g(x), g(x/\beta p_T)/\beta p_T)$ as long as $\gamma \in [tp_T|1 - \beta|/(2 - t), \min\{p_T, \beta tp_T/2, t(1 - \beta p_T)/(2 - t)\}]$. In particular, $(t, p_T, \beta, g(x)) = (0.985, 0.29, 1, g(x), g(x/p_T)/p_T)$ is non-falsifiable as long as $\gamma \leq \beta tp_T/2 \sim 0.143$.

The proof is included in the Appendix. The main idea is the following. Given a particular damage amount, J , we construct a defendant-offer screening model in which weak plaintiffs all settle for a single settlement amount, and strong plaintiffs all go to trial with the same probability of trial victory. Likewise, we construct a plaintiff-offer screening model that forms a mirror image of the defendant-offer screening model. If the defendant-offer models and the plaintiff-offer models arise in the proper ratio, the average plaintiff trial win rate of the

¹⁵ We have not shown the non-falsifiability results for general $s(x)$.

combined models is p_T and the settlement amount is always $\beta p_T J$. By allowing J to be distributed according to $g(x)$, we can also ensure that the settlement values are distributed according to $g(x/\beta p_T)/\beta p_T$.

Remark 2.1. The proof requires either γ to be small (smaller than p_T , $\beta t p_T/2$, and $t(1 - \beta p_T)/(2 - t)$) or that plaintiffs are willing to go to trial without credible threats. Although the assumption of small γ may seem implausible, it should be remembered that γ represents the *incremental* cost of litigation over settlement. While it is true that total litigation costs are usually more than 14% of judgments, much of those costs—pre-filing investigation, pleading, and, in many cases, much discovery—would be usually be incurred even if the case settled. So it is not implausible to think that the incremental cost of litigation (e.g. trial) is less than or equal to 14% of the judgment. In addition, as mentioned above, these are merely sufficient conditions for establishing non-falsifiability. The fact that we observe higher γ in practice does not mean the model has been falsified.

Remark 2.2. The set of feasible γ values will be non-empty for a wide set of parameters. For example, it is clear that when $\beta = 1$, the set is necessarily non-empty. When $\beta < 1$, we can also show that the set is non-empty as long as $\beta > 2/(4 - t)$, regardless of p_T .

B. Signaling Model

i. Simple Signaling Model

The signaling model is different from the screening model in that the party with private information makes the offer and thereby signals her information to the receiver. Under Reinganum & Wilde (1986)'s original signaling model, a risk-neutral plaintiff sues a risk-neutral defendant. The probability that the plaintiff will prevail at trial, p , is known by both parties, but only the plaintiff knows the extent of true damages, J . The defendant only knows the distribution of damages (plaintiff types). J is assumed to be distributed according to $f(\cdot)$ defined over $[\underline{J}, \bar{J}]$. The plaintiff makes a take-it-or-leave-it settlement demand, and the defendant probabilistically accepts the settlement demand based on the amount. In equilibrium, the plaintiff's settlement demand fully and accurately reveals its type, which is true damages. The settlement amount increases monotonically with true damages, and the defendant's probability of rejection also increases monotonically with the settlement amount so that weak plaintiffs do not bluff and demand higher amounts than the equilibrium allocates to them. More specifically, in equilibrium, each plaintiff demands $pJ + C_d$, the expected value of trial plus the defendant's cost of going to trial.

Under this model, the observed plaintiff trial win rate will be p and the observed distribution of judgments will range from \underline{J} from \bar{J} , but will be skewed toward higher amounts as compared to the underlying distribution of damages. The observed settlement value will range from $p\underline{J} + C_d$ to $p\bar{J} + C_d$, but the distribution will be skewed toward lower amounts as compared to the shape of the underlying distribution of damages. Although settlement values are distributed with a skew toward left, because the plaintiff can extract C_d from the defendant, it is

possible that the average settlement may match or even exceed the expected value of trial for the plaintiff. That is, although the average pJ for settled cases will be lower than the average pJ for litigated cases, since settlements are $pJ + C_d$, the average settlement might be the same or higher than the average judgment. Nevertheless, this will occur only when the dispersion of damages is low compared to litigation costs. The selection of disputes under the signaling model for a distribution similar to that in the New York data is depicted graphically in Figure 5.

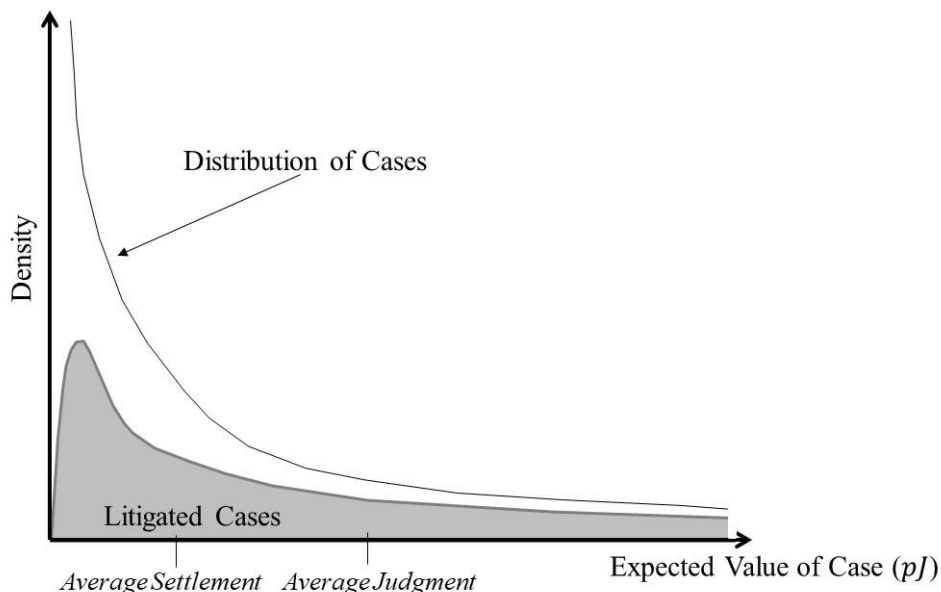


Figure 5. Selection under the Simple Signaling Model
(Assuming Plaintiff Has the Informational Advantage)

When, as in the New York data, judgments can range from a few hundred dollars to several million, the average settlement will be much lower than the average judgment. The simple signaling model is, therefore, inconsistent with the third characteristic of the data set out in Section 2. In addition, settlement rates will be lower than 98.5%, because, in order for the plaintiff to have an incentive to make low settlement offers when it knows damages are low, the defendant must reject with positive probability all settlement offers above the most minimal amount, and the probability of rejection (litigation) must increase with the amount offered in settlement. Calculations with the known distribution of judgment damages suggest that, if the New York judgments were generated by the simple signaling model, the average settlement would be \$2 (compared to an average judgment of \$92,534) and the settlement rate would be less than 1% (as opposed to the 98.5% actually observed).¹⁶

¹⁶ This calculation assumed that litigation costs were proportional to true damages, as in the no-selection-bias and generalized screening and signaling models, that each party's litigation costs were one quarter true damages, and that the lowest potential damage was 9 cents. In the signaling model, cases with the lowest possible damages always settle, so it is necessary to make some assumption about what they would have been. The lowest judgment in the data set is 10 cents, so we assumed the lowest settlement is a little lower. We could not derive the lowest potential judgment from the observed settlements, because the lowest observed settlement is \$1, which would suggest that the lowest observed judgment should be greater than \$3.45 ($1/0.29$), which would also be inconsistent with the model. If the

Switching the information structure and assuming that the defendant has superior information would not help. As with the screening model, switching the information structure would essentially flip the shaded and unshaded areas in Figure 5. The settlement rate would be even lower, and average settlements would far exceed average judgments.¹⁷

ii. Generalized Signaling Model

Like the simple screening model, the simple signaling model can be generalized to make predictions consistent with the New York data. As with the screening model, the key is to assume that the parties know the damages¹⁸ and that the informational advantage and probability of prevailing are distributed in a way that generates the desired settlement rate and plaintiff trial with rate.¹⁹ Then, one can easily construct a distribution of damage amounts to match the observed distribution of settlements and judgments. More formally, we show the following.

PROPOSITION 3. (THE NON-FALSIFIABILITY OF GENERALIZED SIGNALING MODEL). Suppose $\beta p_T < 1$ and $p_T > 0$. Let the cost of going to trial be γJ for both the plaintiff and the defendant, and let J be the damages if plaintiff prevails. Then the generalized signaling model is non-falsifiable for aggregate data profile $(t, p_T, \beta, g(x), g(x/\beta p_T)/\beta p_T)$ as long as $\gamma \in (|1 - \beta| p_T, \min\{\beta p_T/(4 - 2t), p_T, (1 - \beta p_T)/(3 - 2t)\})$. In particular, aggregate data profile $(0.985, 0.29, 1, g(x), g(x/p_T)/p_T)$ is non-falsifiable if $\gamma \leq \beta p_T/(4 - 2t) \sim 0.143$.

The proof can be found in the appendix. The main idea is similar to that of Proposition 2. Given a particular damage amount, J , we construct a plaintiff-offer signaling model in which a

lowest potential damage amount is set differently, results change somewhat, but are qualitatively similar. We also made calculations excluding low judgments from the data set (e.g. judgments for less than \$1000 or \$10,000) and varying the assumption about how much lower the lowest potential judgment was than the lowest actual non-excluded judgment (10%, 1% or 0.1% lower). The average settlements increase to between \$688 and \$9912, and the settlement rate increases to between 3% and 93%, depending on which judgments are excluded and on the assumption about the lowest potential damage amount. Nevertheless, the wide divergence between average settlements and average judgments remains, as does the fact that settlement rates are below the observed 98.5%. If one followed Reinganum & Wilde's (1986) original assumption that litigation costs are fixed, the model would still produce results that are far from the observed data. One would also have to assume absurdly low litigation costs (e.g. 2 cents) in order to preserve the credibility of the plaintiff's threat to take each case to trial, or one would need to delete from the data set all cases with low damage amounts.

¹⁷ When the defendant has superior information, the probability of settlement equals one for the highest potential damage case, and decreases for lower damage cases. So settlement rates would be high for the high damage cases, but low for low damage cases. Because there are a small number of very high damage cases and a large number of low damage cases in the New York data set, the Reinganum and Wilde model would predict an extremely low overall settlement rate.

¹⁸ Although Reinganum and Wilde (1986) assume that the parties have asymmetric information about damages and that the probability that the plaintiff prevails is common knowledge, Klerman and Lee (2014) show that the model can be relatively easily modified so that asymmetric information is about the probability the plaintiff prevails.

¹⁹ Gelbach (2016, p.18) discusses briefly how the Klerman-Lee version of the signaling model could be generalized by randomizing between the plaintiff-offer and defendant-offer versions so as to generate any plaintiff trial win rate.

weak group of plaintiffs settle at a constant settlement value, strong plaintiffs sometimes settle and sometimes litigate with a particular probability of prevailing. Likewise, we construct a defendant-offer signaling model that forms a mirror image of the plaintiff-offer signaling model. If the defendant- and plaintiff-offer models arise in the appropriate ratio, the average plaintiff trial win rate of the combined models is p_T . We then again allow J to be distributed according to $g(x)$. However, there is one importance difference as compared to the screening model: there will always be at least two different settlement values for each value of J for both the plaintiff- and defendant-offer models, which means that the distribution of settlements and expected judgments will not be identical. Nevertheless, as the population of weak plaintiffs (defendants) in the plaintiff-offer (defendant-offer) model goes to 0, the resulting settlement values will be distributed nearly according to $g(x/\beta p_T)/\beta p_T$, and will approach $g(x/\beta p_T)/\beta p_T$ in a pointwise manner.

Remark 3.1. As with the generalized screening model results, the proof for the signaling model requires either γ to be small or otherwise assumes that plaintiffs are willing to go to trial without credible threats. Here, too, when $\beta = 1$, we need p_T to be at least twice γ .

Remark 3.2. The set of feasible γ values will be non-empty for a wide set of parameters. If $\beta = 1$, for example, the set will be trivially non-empty. For $\beta < 1$ and t close to 1, the set will be non-empty as long as β is greater than $2/3$.

C. The Priest and Klein Model

A number of scholars have formulated models with two-sided incomplete information. Priest and Klein's (1984) model is usually classified as a "divergent expectations" model, but Lee and Klerman (2016 & 2017) show that the Priest-Klein model can also be interpreted as a model with two-sided incomplete information, and that most of Priest and Klein's original results can be proven rigorously under such a model. Schweizer (1989) set out the first game-theoretically rigorous model of two-sided incomplete information. The model in that article, however, assumes that information and thus outcomes are discrete, and it would be difficult to adapt it to a situation, like that observed in the New York data, where outcomes (damages) are essentially continuous over a wide interval. Daughety and Reinganum (1994) develop an ingenious model where the plaintiff has superior information about damages, but the defendant has superior information about liability. Nevertheless, their model assumes that damages and the plaintiff's probability of prevailing are both distributed uniformly over some interval, and we have not figured out how to adapt their model to a situation, like that observed in the New York data, where lower damages seem to be more common than higher damages. Gelbach (2016) generalizes the Priest and Klein model, and his generalization is particularly helpful here.

i. Simple Priest-Klein Model

The simple version of the Priest-Klein model assumes that damages do not vary, and that incomplete information is about case merit, which is closely related to the probability that the plaintiff prevails. Unlike the screening and signaling models, the Priest-Klein model changes

dramatically when incomplete information is about damages rather than the probability the plaintiff will prevail. Whereas the screening and signaling models work almost identically when probability of prevailing is fixed and damages vary and when damages as fixed and probability of prevailing varies, that is not true of the Priest-Klein model.

A key to the Priest-Klein model is that the parties get signals about case merit, Y , and that there is a threshold value, y^* , the legal standard, that divides cases into those that the plaintiff wins, $Y > y^*$, and those that the plaintiff loses, $Y \leq y^*$. The parties must estimate the probability that the plaintiff will prevail based on how close their signals are to y^* and based on the accuracy of the signals. If, instead, the parties know case merit, Y , or if both receive the same signal about case merit, and if the parties receive noisy signals about damages, J , the model will, essentially, be the no-selection-bias model discussed in Section 3. In that model, the probability of prevailing is common knowledge, and the parties receive unbiased information about damages.²⁰

ii. Generalized Priest-Klein Model

Gelbach (2016) generalizes the Priest-Klein model in a way that is particular helpful and allows one to replicate almost any combination of plaintiff trial win rate, settlement rate, and distribution of damages and settlements. Gelbach shows that all of the standard litigation models, including Priest and Klein (1984), can be represented in “reduced form.” In reduced form, each case is represented by the potential damages, J , and a pair of beliefs, the plaintiff’s belief about its probability of prevailing, Q_p , and the defendant’s belief about the plaintiff’s probability of prevailing, Q_d . A reduced form consists a distribution of party beliefs about the plaintiff’s probability of prevailing, a litigation rule stating which cases will result in litigation, and a conditional win rate function stating the probability the plaintiff would win each case, if it were litigated. The first two aspects of the reduced-form approach are best illustrated in a graph, such as the one below, which is a possible representation of the New York data:

²⁰ If the parties in the Priest-Klein model, as formalized in Lee and Klerman (2016), had incomplete information about damages, there would be some differences with the model in Section 3, because the model in Section 3 is a model with inconsistent priors, and the model assumed in most of Lee and Klerman (2016) is a model with consistent priors and two-sided, correlated incomplete information. Nevertheless, as discussed in Lee and Klerman (2016), the difference between inconsistent priors and consistent priors in a model like Priest-Klein is very small. See Lee and Klerman (2016) pp. 61, 62 n. 9, 68.

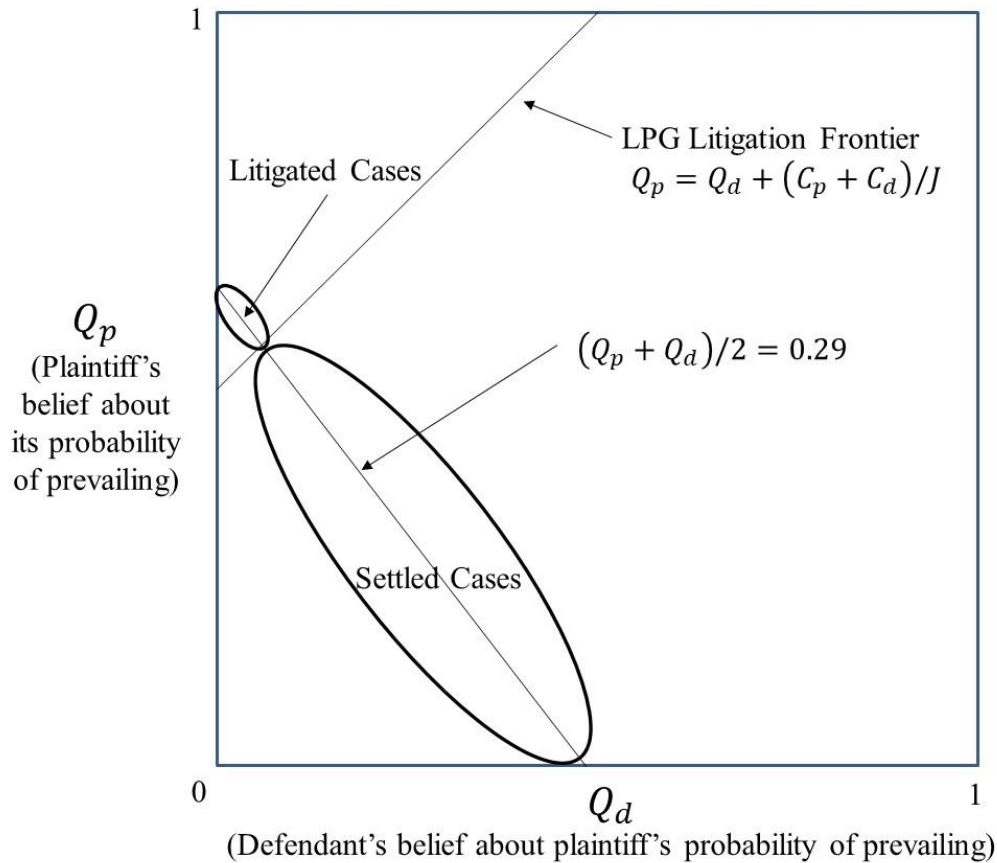


Figure 6. Reduced Form Representation of the New York Data

Damages, J , are assumed to be fixed. Defendants' beliefs about plaintiffs' probability of prevailing are along the horizontal axis, and range from zero to 1. Plaintiffs' beliefs are along the vertical axis. The Priest-Klein model follows the LPG (Landis-Posner-Gould) litigation rule, under which cases litigate if and only if plaintiff are sufficiently more optimistic to offset the litigation savings from settlement, $Q_p > Q_d + (C_p + C_d)/J$.²¹ That means that the plaintiff litigates only if cases fall in the upper left triangle of the above graph. Although it is not represented in the graph, it is easy to concoct a plausible conditional win rate—the average of the two parties' beliefs, $(Q_p + Q_d)/2$. The two ellipses in the above graph represent the distribution of cases. Cases within the ellipses have density one. Cases outside the ellipses have density zero.²² The fact that the upper ellipse is symmetric around the line $(Q_p + Q_d)/2 = 0.29$ means that the plaintiff trial win rate is 29%, although any other plaintiff trial win rate between $(1 - (C_p + C_d)/J)$ and $(C_p + C_d)/J$ could be represented by moving that line up or down. If

²¹ Like Priest and Klein's original model, Gelbach also considers settlement costs, S_p and S_d . That is a complication not necessary here, so we have omitted it.

²² The elliptical shapes and binary densities are for convenience only. Any densities symmetric across the line $(Q_p + Q_d)/2 = 0.29$ would work equally well.

litigation costs are equal, $C_p = C_d$, and parties have equal bargaining power, the fact that the lower ellipse is also symmetric around the line $(Q_p + Q_d)/2 = 0.29$ means that the average judgment will equal the average settlement. Any settlement rate can be achieved merely by changing the relative size of the two ellipses. In Figure 6, the ellipse representing settled cases is much larger, because in the New York cases nearly all cases (98.5%) settle. Nevertheless, by making the ellipsis representing settled cases larger or smaller (or by changing the size of the ellipsis representing litigated cases), any plaintiff trial win rate can be represented.

Gelbach (2016) shows that any reduced form can be represented as a generalized Priest-Klein model.²³ A generalized Priest-Klein model is the original Priest-Klein model in which various parameters take on a distribution of values. So, some cases might have asymmetric stakes while others might have equal stakes. Litigation costs might also vary. That means that the reduced form illustrated above, where the plaintiff trial win rate is 29%, the settlement rate is 98.5%, and average settlements equal average judgments, can be achieved by manipulating the distribution of parameters in the generalized Priest-Klein model.

All that remains, as in the other generalized litigation models, is to replicate the generalized Priest-Klein model (with the distribution of parameters assumed in the prior paragraph) for multiple damages levels, with a distribution of damages matching the distribution of non-zero judgments observed in the New York data. Then all five litigation facts set out in Section 2 will be predicted by the model.

5. Conclusion

The New York data provide unique insight into the distribution of damages and settlements amounts in ordinary tort litigation. The distribution of damages and settlements are remarkably similar, and the average settlement is very close to the average judgment. This suggests that selection effects may be small or non-existent. Although all existing litigation models predict selection bias, it is relatively easy to create an inconsistent-priors model that results in no selection effects. The New York data are inconsistent with simple versions of the standard screening and signaling models, and with the Priest-Klein model. Nevertheless, those three models can be generalized so as to be consistent with the New York data. Thus, the New York data do not allow us to choose among the various models. One might, however, prefer the no-selection-bias model as the simplest model consistent with the data.

²³ Gelbach's proof of the representation theorem assumes a distribution of a parameter which Gelbach calls K , where $K = (C_p + C_d - S_p - S_d)/J$. Nevertheless, it is not necessary that damages J , vary, as long as the other components of K , litigation and settlement costs, can vary. So it is appropriate to construct a reduced form for each damage level, J , and then to sum over possible values of J . Similarly, Gelbach's proof assumes a distribution of J_d/J_p , but that is just the asymmetry of the stakes, which can vary while holding J constant.

6. References

- Bebchuck, Lucian Arye. 1984. Litigation and Settlement under Imperfect Information. *RAND Journal of Economics* 15:404-15.
- Bebchuk, Lucian Arye. 1996. A New Theory Concerning the Credibility and Success of Threats to Sue. *Journal of Legal Studies* 25:1-25.
- Daughety, Andrew and Jennifer Reinganum. 1994. Settlement Negotiations with Two-Sided Asymmetric Information: Model Duality, Information Distribution, and Efficiency. *International Review of Law and Economics* 14:283-298.
- Gelbach, Jonah B. 2012. Locking the Doors to Discovery? Assessing the Effects of *Twombly* and *Iqbal* on Access to Discovery. *Yale Law Journal* 121: 2270-2345.
- Gelbach, Jonah B. 2012a. Selection in Motion: A Formal Model of Rule 12(b)(6) and the *Twombly-Iqbal* Shift in Pleading Policy,” Available at SSRN: <http://ssrn.com/abstract=2138428>.
- Gelbach, Jonah B. 2016. The Reduced Form of Litigation Selection Models and the Plaintiff’s Win Rate. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2837154
- Helland, Eric, Daniel Klerman, Alexander Kappner and Brendan Dowling. Contingent Fee Litigation in New York City. Forthcoming in *Vanderbilt Law Review*
- 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. *Journal of Legal Studies* 25:233-259.
- Klerman, Daniel. 2012. The Selection of 13th-Century Disputes for Litigation. *Journal of Empirical Legal Studies* 9:320-346.
- Klerman, Daniel and Yoon-Ho Alex Lee. 2014. Inferences from Litigated Cases. *Journal of Legal Studies* 43:209-248.
- 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 Klerman. 2017. Litigation and Selection with Correlated Two-Sided Incomplete Information. Unpublished. Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2927248
- Myerson, Roger B. and Mark A. Satterthwaite. 1983. Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory* 29:265-81.
- Priest, George L. and Benjamin Klein. 1984. The Selection of Disputes for Litigation. *Journal of Legal Studies* 13:1-55.
- Reinganum, Jennifer F. and Louis L. Wilde. 1986. Settlement, Litigation, and the Allocation of Litigation Costs. *RAND Journal of Economics* 17:557-566.
- Schweizer, Urs. 1989. Litigation and Settlement under Two-Sided Incomplete Information. *Review of Economic Studies* 56:163-177.
- Schweizer, Urs. 2016. Litigated Cases: The Selection Effect Revisited. *Journal of Institutional and Theoretical Economics* 172:409-416.
- Shavell, Steven. 1996. Any Frequency of Plaintiff Victory at Trial is Possible. *Journal of Legal Studies* 25:493-501.
- Spier, Kathryn E. and J. J. Prescott. 2016. Contracting on Litigation. Unpublished. Available at SSRN: <http://ssrn.com/abstract=2765033> or <http://dx.doi.org/10.2139/ssrn.2765033>

Studdert, David M. and Michelle M. Mello. 2007. When Tort Resolutions are “Wrong”: Predictors of Discordant Outcomes in Medical Malpractice Litigation. *Journal of Legal Studies* 36:547-578.

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

Waldfogel, Joel. 1998. Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation. *Journal of Law and Economics* 41:451-476.

Appendix

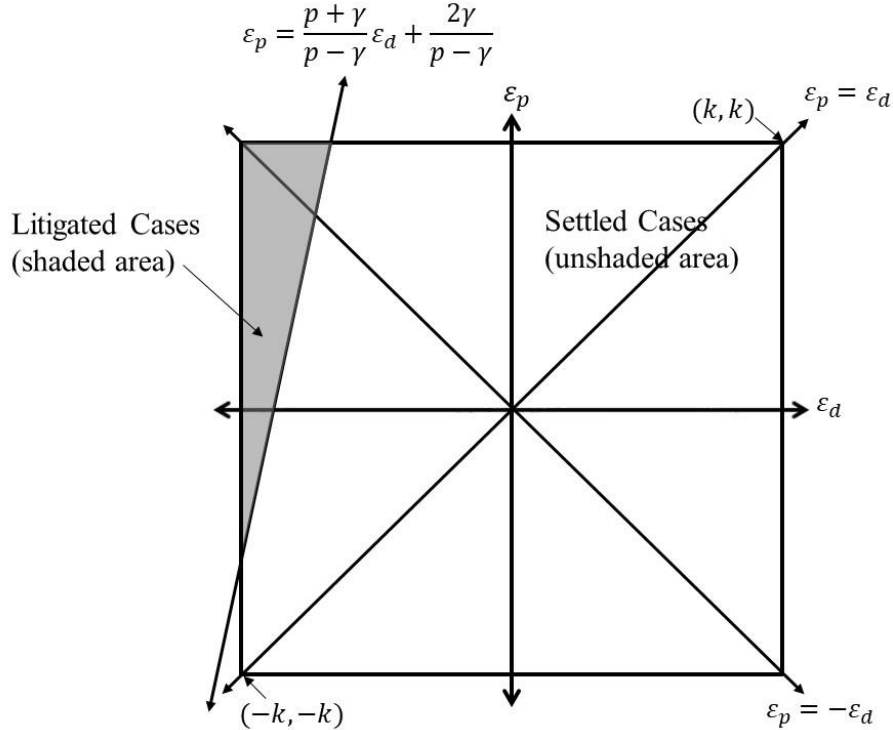
PROOF OF PROPOSITION 1. We already established the no-selection-bias result in the main text. It remains to show that, as the settlement rate approaches one (e.g. as k gets sufficiently small or as γ approaches p), the average settlement amount approaches the average judgment. In particular, $E[pJ] < E[S] < E[pJ]/t$, where t is the settlement rate. We have

$$S = \frac{(pJ_p - C_p) + (pJ_d + C_d)}{2} = \left(\frac{(p - \gamma)(1 + \varepsilon_p) + (p + \gamma)(1 + \varepsilon_d)}{2} \right) J$$

Thus, the average settlement is $E[S] = E \left[\left(\frac{(p - \gamma)(1 + \varepsilon_p) + (p + \gamma)(1 + \varepsilon_d)}{2} \right) J \right] / 2[\text{settlement}] = E[J] \left[p + \{ (p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d \} / 2 \mid \varepsilon_p \leq (p + \gamma)\varepsilon_d / (p - \gamma) + 2\gamma / (p - \gamma) \right]$. The key is to focus on the average value of the weighted average of the error terms given settlement:

$$E \left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma} \varepsilon_d + \frac{2\gamma}{p - \gamma} \right] \quad (1)$$

If the weighted average value of the error terms is zero, then the average settlement will equal the average judgment. If the weighted average is positive, the average settlement will be larger. Consider the graph below.



Figures A1. The Selection of Disputes under the No-Selection Bias Model, $\gamma/p < k < 1$

The square bounded by $(-k, -k)$ and (k, k) represents all cases. Because ε_p and ε_d are independent and uniformly distributed, all points in the square have equal density. Litigated

cases are those in the shaded triangle on the left side. This makes sense, because cases are more likely to litigate when the parties are very optimistic (ε_p is very positive and/or ε_d is very negative). It is easy to prove that the line separating litigated and settled cases, $\varepsilon_p = (p + \gamma)\varepsilon_d/(p - \gamma) + 2\gamma/(p - \gamma)$, will always intersect the left and top sides of the square, as long as $\gamma/p < k < 1$. We can ignore $k \geq 1$, because we assumed above that $k < 1$ (otherwise parties' estimates of J might be negative). If $k \leq \gamma/p$, then all cases settle, which is uninteresting empirically and irrelevant to an analysis of selection bias.

The fact that the line separating litigated and settled cases always intersects the left and top sides of the square means that the shaded area (litigated cases) is always above the line $\varepsilon_p = \varepsilon_d$. This allows us to prove that $E[S] > E[pJ]$. First, rewrite (1), the weighted average error terms given settlement, as follows:

$$E \left[\frac{p(\varepsilon_p + \varepsilon_d)}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma} \varepsilon_d + \frac{2\gamma}{p - \gamma} \right] + E \left[\frac{\gamma(\varepsilon_d - \varepsilon_p)}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma} \varepsilon_d + \frac{2\gamma}{p - \gamma} \right] \quad (2)$$

Consider the first term of (2). The line $\varepsilon_p = -\varepsilon_d$ bisects the square. If the unshaded area (settled cases) were symmetric across that line, then the average value of $\varepsilon_p + \varepsilon_d$ would be zero, and the average settlement would equal the average judgment. Nevertheless, it is apparent that the areas are not symmetric, because much more of the shaded area (litigated cases) is below the line $\varepsilon_p = -\varepsilon_d$. This is true because the slope, $(p + \gamma)/(p - \gamma)$, is always greater than one, because litigation costs are positive ($\gamma > 0$). This means that, in the unshaded area (settled cases), the average value of $\varepsilon_p + \varepsilon_d$ is greater than zero, so $E[(\varepsilon_p + \varepsilon_d)/2] > 0$, and the first term of (2) is greater than 0. Now consider the second term of (2). Because, as mentioned above, the shaded area (litigated cases) lies entirely above the line $\varepsilon_p = \varepsilon_d$, the average value of $\varepsilon_d - \varepsilon_p$ among settled cases is positive, so the second term of (2) is also positive. Since both terms of (2) are positive, the entirety of (2) is positive, which means $E[S] > E[pJ]$.

Furthermore, an upper bound can be placed on the average settlement as follows. Because ε_p and ε_d are independent and uniformly distributed, the fraction of the area of the square bounded by $(-k, -k)$ and (k, k) that is unshaded is the settlement rate, t . Because the square bounded by $(-k, -k)$ and (k, k) is symmetric across both the horizontal and vertical axes:

$$\begin{aligned} 0 &= E \left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \right] \\ &= tE \left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma} \varepsilon_d + \frac{2\gamma}{p - \gamma} \right] \\ &\quad + (1 - t)E \left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p > \frac{p + \gamma}{p - \gamma} \varepsilon_d + \frac{2\gamma}{p - \gamma} \right] \end{aligned} \quad (3)$$

$E[(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d/2 \mid \varepsilon_p > (p + \gamma)\varepsilon_d/(p - \gamma) + 2\gamma/(p - \gamma)]$ is the average value of $((p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d)/2$ among cases that are litigated. If the shaded area (litigated cases),

were an infinitesimally narrow vertical sliver at the left side of the square bounded by $(-k, -k)$ and (k, k) , then, in that sliver, $E\left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2}\right] > -(p + \gamma)/2$, because the average value of ε_p in that sliver would be greater than zero, and ε_d would be greater than -1 , because $k < 1$. It follows that in the entire shaded area, $E\left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2}\right] > -(p + \gamma)/2$, because the average value of both ε_p and ε_d in the shaded area is greater than in the sliver. Since to assure plaintiff threat credibility, $0 < \gamma < p$, it must be that $-(p + \gamma)/2 > -p$. So:

$$E\left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p > \frac{p + \gamma}{p - \gamma}\varepsilon_d + \frac{2\gamma}{p - \gamma}\right] > -p \quad (4)$$

Combining (3) and (4), we know that $E\left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma}\varepsilon_d + \frac{2\gamma}{p - \gamma}\right] < p(1 - t)/t$. Plugging that into the equation for average settlement at the beginning of the proof, $E[J] \left[p + E\left[\frac{(p - \gamma)\varepsilon_p + (p + \gamma)\varepsilon_d}{2} \mid \varepsilon_p \leq \frac{p + \gamma}{p - \gamma}\varepsilon_d + \frac{2\gamma}{p - \gamma}\right] \right]$, we get $E[S] < E[pJ]/t$. Combining that with $E[S] > E[pJ]$ as shown above, we get $E[pJ] < E[S] < E[pJ]/t$. As the settlement rate approaches one—either because parties become increasingly accurate in estimating J (i.e., k approaches γ/p) or litigation costs become very high (i.e., γ approaches p)—the average settlement, $E[S]$, approaches the average judgment, $E[pJ]$. \square

PROOF OF PROPOSITION 2. Consider first a defendant-offer model with a fixed damage level J , but variable p , the plaintiff type (i.e., the probability that the plaintiff will prevail at trial). Let $p_1 = \beta p_T + \gamma$ and $p_2 = \beta p_T + \gamma(2/t - 1)$. Given $\gamma \leq t(1 - \beta p_T)/(2 - t)$, it is easy to check that $\gamma \leq p_1 < p_2 \leq 1$. The first inequality ensures that the plaintiff has a credible threat to go to trial. Note also that $p_2 - p_1 = 2\gamma(1 - t)/t$. Consider now a discrete distribution where t plaintiffs are type p_1 (weak plaintiffs) and $1 - t$ plaintiffs are type p_2 (strong plaintiffs).

The defendant has only two plausible offers: $(p_1 - \gamma)J$ or $(p_2 - \gamma)J$. No other offer will strictly dominate these two. With an offer of $(p_1 - \gamma)J$, all weak plaintiffs will settle, all strong plaintiffs will go to trial, and the defendant's expected payout is $[(p_1 - \gamma)t + (p_2 + \gamma)(1 - t)]J$. With an offer of $(p_2 - \gamma)J$, all plaintiff types will settle, and the defendant's expected payout is simply $(p_2 - \gamma)J$. Therefore, there will be a separating equilibrium as long as $[(p_1 - \gamma)t + (p_2 + \gamma)(1 - t)]J \leq (p_2 - \gamma)J$,²⁴ which comes out to $p_2 - p_1 \geq 2\gamma(1 - t)/t$, which is satisfied. Within this separating equilibrium, the settlement rate is t , the plaintiff trial win rate is p_2 , the average amount of judgment awarded at trial is p_2J (gross of attorney fees) and the observed settlement amount is $(p_1 - \gamma)J = \beta p_T J$. The only observed positive judgment award is J . Call this Game Δ_J .

Consider now Game π_J , the plaintiff-offer model. Defendants come in two types, p_3 and p_4 , where $p_3 = \beta p_T - \gamma(2/t - 1)$ and $p_4 = \beta p_T - \gamma$. These are the probabilities that the plaintiff will prevail at trial. So p_3 represents a strong defendant and p_4 represents a weak

²⁴ We assume when the defendant is indifferent between offering low or high, he will offer the lower amount. Similarly, when the plaintiff is indifferent between demanding low or high, he will demand the higher amount.

defendant. We then have $\gamma \leq p_3 < p_4 \leq 1$ as long $\gamma \leq \beta t p_T / 2$, which holds by assumption. In addition, $p_4 - p_3 = 2\gamma(1 - t)/t$. Here, suppose now that t defendants come with type p_4 (weak defendants) and $1 - t$ defendants come with type p_3 (strong defendants).

The plaintiff has only two rational demands: $(p_3 + \gamma)J$ or $(p_4 + \gamma)J$. With a high demand of $(p_4 + \gamma)J$, all weak defendants settle, all strong defendants go to trial, and the plaintiff's expected payoff is $[(p_3 - \gamma)(1 - t) + (p_4 + \gamma)t]J$. With a demand of $(p_3 + \gamma)J$, all defendants settle, and the plaintiff's expected payoff is $(p_3 + \gamma)J$. Therefore, there will be a separating equilibrium as long as $[(p_3 - \gamma)(1 - t) + (p_4 + \gamma)t]J \geq (p_3 + \gamma)J$, which comes out to $p_4 - p_3 \geq 2\gamma(1 - t)/t$, which is satisfied. Within this separating equilibrium, the settlement rate is t , the plaintiff trial win rate is p_3 , the average judgment awarded at trial (including a judgment of 0) is p_3J (gross of attorney fees) and the observed settlement amount is $(p_4 + \gamma)J = \beta p_T J$. The only positive judgment awarded is J .

Now suppose the informational advantage is distributed such that Game π_J and Game Δ_J appear in an α to $1 - \alpha$ ratio, where $\alpha = \left((\beta - 1)p_T / (\gamma(2/t - 1)) + 1 \right) / 2$. Thus, the defendant has the informational advantage with probability α . The fact that $\frac{t p_T |1 - \beta|}{2 - t} \leq \gamma$ ensures us that $\alpha \in [0, 1]$. Then in the combination of the two game types, the settlement rate is t , the plaintiff trial win rate is $\alpha p_3 + (1 - \alpha)p_2 = \beta p_T + (1 - 2\alpha)\gamma(2/t - 1) = p_T$, the average judgment awarded at trial is $p_T J$, the average value (or the *only* value) of settlement is $\beta p_T J$, and the only observed positive judgment is J .

Finally, let J vary according to the observed distribution of positive judgments, $g(x)$, and let there be Game π_J and Game Δ_J for each J , with the parameters set out in the previous paragraph. Then the distribution of judgments will be $g(x)$, the distribution of settlements will be $g(x/\beta p_T)/\beta p_T$, and we can replicate the five facts we observe.

For $(t, p_T, \beta, g(x), g(x/\beta p_T)/\beta p_T) = (0.985, 0.29, 1, g(x), g(x/p_T)/p_T)$, consider $\gamma = 0.14$, $\alpha = 0.5$. Let $p_1 = p_T + \gamma = 0.43$, $p_2 = p_T + \gamma(2/t - 1) = 0.434$, $p_3 = p_T - \gamma(2/t - 1) = 0.146$, and $p_4 = p_T - \gamma = 0.15$. The model then generates $(0.985, 0.29, 1, g(x), g(x/p_T)/p_T)$ exactly. \square

PROOF OF PROPOSITION 3. As with the generalized screening model, consider first a plaintiff-offer signaling model with a fixed damage level J . Suppose there are only two plaintiff types, p_1, p_2 where $\gamma < p_1 = p_2 - \epsilon < p_2 = \beta p_T - \gamma$ for some $\epsilon > 0$. We will return to check the inequality conditions at the end.

Suppose τ plaintiffs are type p_1 (weak plaintiffs) and $1 - \tau$ plaintiffs are type p_2 (strong plaintiffs). Consider the separating equilibrium in which the plaintiff of type p_1 demands $(p_1 + \gamma)J$ and the plaintiff of type p_2 demands $(p_2 + \gamma)J$. The defendant accepts all settlement demands up to $(p_1 + \gamma)J$, accepts settlement demand of $(p_2 + \gamma)J$ with probability ρ , and rejects all other settlement demands.

In order to make sure the plaintiff of type p_1 will not be incentivized to mimic type p_2 , we need $(p_1 + \gamma)J \geq (p_2 + \gamma)J\rho + (p_1 - \gamma)J(1 - \rho)$ or $\rho \leq 2\gamma/(2\gamma + \epsilon)$.²⁵ Likewise, the plaintiff of type p_2 will not have any incentive to deviate as long as $(p_1 + \gamma)J \leq (p_2 + \gamma)J\rho + (p_2 - \gamma)J(1 - \rho)$ or $\rho \geq 1 - \epsilon/2\gamma$. Since $1 - \epsilon/2\gamma < 2\gamma/(2\gamma + \epsilon)$, any $\rho \in [1 -$

²⁵ As is conventional in models of this kind, we assume that if the plaintiff is indifferent, he will make a demand based on his type.

$\epsilon/2\gamma, 2\gamma/(2\gamma + \epsilon)]$ will work. For simplicity, we let $\rho = 1 - \epsilon/2\gamma$. The defendant's belief function is given by the following: $b(S) = p_1$ for $S < (p_2 + \gamma)J$ and p_2 for $S \geq (p_2 + \gamma)J$.

Under this equilibrium, the settlement rate is $t = \tau + \rho(1 - \tau)$. Note that as τ approaches 1, t will approach 1, and as τ approaches 0, t will approach ρ . The plaintiff trial win rate is p_2 , and thus the expected value of trial is p_2J . The average settlement amount is $((p_1 + \gamma)\tau + (p_2 + \gamma)\rho(1 - \tau))/(\tau + \rho(1 - \tau))J = (\beta p_T - \epsilon\tau/(\tau + \rho(1 - \tau)))J$. Note that as τ approaches 0, *nearly* all observed settlement values are $(p_2 + \gamma)J = \beta p_T J$.

Consider now a defendant-offer signaling model with a fixed damage level J , and suppose there are only two defendant types, p_3, p_4 , where $p_3 = \beta p_T + \gamma < p_4 = p_3 + \epsilon$. p_3 and p_4 are the probabilities that the plaintiff will prevail at trial. Suppose $1 - \tau$ defendants are type p_3 (strong defendants) and τ defendants are type p_4 (weak defendants). Take the separating equilibrium in which the defendant of type p_3 demands $(p_3 - \gamma)J$ and the defendant of type p_4 demands $(p_4 - \gamma)J$. Proceeding as above, we obtain an equilibrium in which the settlement rate is $t = \tau + \rho(1 - \tau)$. The plaintiff trial win rate is p_3 , and thus, the expected value of trial is p_3J . The average settlement amount is $((p_4 - \gamma)\tau + (p_3 - \gamma)\rho(1 - \tau))/(\tau + \rho(1 - \tau))J = (\beta p_T + \epsilon\tau/(\tau + \rho(1 - \tau)))J$. Note that as τ approaches 0, *nearly* all observed settlement values will simply be $(p_3 - \gamma)J = \beta p_T J$.

Now suppose the informational advantage is distributed such that the plaintiff-offer and defendant-offer signaling models appear in an $\alpha: 1 - \alpha$ ratio, where $\alpha \in [0,1]$. That is, plaintiff has the informational advantage with probability α . Then in the combination of the two game types, the settlement rate is $\tau + \rho(1 - \tau)$, the plaintiff trial win rate is $\alpha p_2 + (1 - \alpha)p_3 = \alpha(\beta p_T - \gamma) + (1 - \alpha)(\beta p_T + \gamma) = \beta p_T + (1 - 2\alpha)\gamma$, the average judgment awarded at trial is therefore $(\beta p_T + (1 - 2\alpha)\gamma)J$. If we set $\alpha = (1 - (1 - \beta)p_T/\gamma)/2$, then the plaintiff win rate becomes p_T . Note that $\alpha \in (0,1)$ as long as $|1 - \beta|p_T < \gamma$, which holds by assumption. Therefore, the average judgment awarded at trial is $p_T J$. The average value of settlements is $\left(\beta - \epsilon\tau(1 - \beta)/((\tau + \rho(1 - \tau))\gamma)\right)p_T J = \beta p_T J - \tau\left(\epsilon(1 - \beta)p_T J/((\tau + \rho(1 - \tau))\gamma)\right)$. As τ approaches zero, *nearly* all settlement values observed are in fact $\beta p_T J$ and the second term will vanish.

Finally, let J vary according to the observed distribution, $g(x)$, and let the plaintiff-offer and defendant-offer signaling models appear in α to $1 - \alpha$ ratio for each J . Let $m(x)$ be the distribution of settlement values. In this case, for each J , we will observe 3 settlement values, $\beta p_T J$, $(\beta p_T - \epsilon)J$, and $(\beta p_T + \epsilon)J$, which appear in a $(1 - \tau)\rho: \alpha\tau: (1 - \alpha)\tau$ ratio, where τ is very small and $(1 - \tau)\rho + \alpha\tau + (1 - \alpha)\tau = t$. The probability density function of the resulting set of settlement values when $J \sim g(x)$ is given by $m(x) = ((1 - \tau)\rho g(x/\beta p_T))/\beta p_T + \alpha\tau g(x/(\beta p_T - \epsilon))/(\beta p_T - \epsilon) + (1 - \alpha)\tau g(x/(\beta p_T + \epsilon))/(\beta p_T + \epsilon)/t$. It is clear that as τ approaches zero, $m(x)$ will approach $g(x/\beta p_T)/\beta p_T$ in a pointwise manner.

Therefore, given an observed settlement rate t , choose ρ to be arbitrarily close to (but smaller than) t . In other words, let $\rho = t - \delta$, for a small $\delta > 0$. The corresponding τ value is $\delta/(1 - t + \delta)$, which will be arbitrarily small as δ becomes arbitrarily small. Let $\epsilon = 2\gamma(1 - \rho) = 2\gamma(1 - t + \delta)$. Then $p_1 = \beta p_T - \gamma - 2\gamma(1 - t + \delta) < \beta p_T - \gamma = p_2$. Note that since $\gamma < \beta p_T/(4 - 2t)$, we can also find a suitable $\delta > 0$ such that $\gamma < \beta p_T/(4 - 2t + 2\delta)$, which will ensure that $\gamma < p_1$. In addition, $\beta p_T - \gamma = p_2 < 1$ is guaranteed since $\beta p_T < 1$. Furthermore, $p_3 = \beta p_T + \gamma > \gamma$ holds trivially and $p_4 = \beta p_T + \gamma + 2\gamma(1 - t + \delta) = \beta p_T + \gamma(3 - 2t + 2\delta) < 1$ for a sufficiently small δ since $\gamma < (1 - \beta p_T)/(3 - 2t)$. In particular, for $(t, p_T, \beta, g(x), g(x/\beta p_T)/\beta p_T) = (0.985, 0.29, 1, g(x), g(x/\beta p_T)/\beta p_T)$, consider $\gamma = 0.14$,

$\alpha = 0.5$, and $\epsilon = 2\gamma(1 - t + \delta) \sim 0.0042$. Then we can let $\rho \sim t = 0.985$, $\tau \sim 0$, and $p_1 = p_T - \gamma - \epsilon = 0.146$, $p_2 = p_T - \gamma = 0.15$, $p_3 = p_T + \gamma = 0.43$, and $p_4 = p_T + \gamma + \epsilon = 0.434$. And we can approximate $(0.985, 0.29, 1, g(x), g(x/p_T)/p_T)$ arbitrarily closely pointwise. \square

Eric Helland
Claremont McKenna College
Robert Day School of Economics and Finance
888 Columbia Avenue
Claremont, CA 91711
U.S.A.
eric.helland@claremontmckenna.edu

Daniel Klerman
University of Southern California
Gould School of Law
University Park, MC-0071
Los Angeles, CA 90089-0071
U.S.A.
dklerman@law.usc.edu

Yoon-Ho Alex Lee
Northwestern Pritzker School of Law
375 E Chicago Ave
Chicago, IL 60611
alex.lee@law.northwestern.edu