**The Selection of Thirteenth-Century
Disputes for Litigation**

Daniel Klerman

USC Legal Studies Research Paper No. 00-10

USC
GOULD
SCHOOL OF
LAW

LEGAL STUDIES
RESEARCH PAPER SERIES

University of Southern California Law School
Los Angeles, CA 90089-0071

The Selection of Thirteenth-Century
Disputes for Litigation

Daniel Klerman

USC Law School

Daniel Klerman, Charles L. and Ramona I. Hilliard Professor of Law and History, USC Law School, University Park MC-0071, 699 Exposition Blvd, Los Angeles, CA 90089-0071. dklerman@law.usc.edu. Phone: 213 740-7973. Fax: 213 740-5502.

Abstract

Priest and Klein's seminal 1984 article argued that litigated cases differ systematically and predictably from settled cases. This article tests the Priest-Klein selection model using a data set of thirteenth-century English cases. These cases are especially informative because juries rendered verdicts even in settled cases, so one can directly compare verdicts in settled and litigated cases. The results are consistent with the predictions of the Priest-Klein article, as well as with the asymmetric-information selection models developed by Hylton and Shavell.

1. Introduction

George Priest and Benjamin Klein's 1984 article, "The Selection of Disputes for Litigation,"[1] has proven to be one of the most influential articles in legal scholarship generally, and it has been especially influential in empirical work.[2] In the nearly three decades since its publication, hundreds of researchers have relied upon and augmented its principal hypothesis, that litigated cases differ systematically and predictably from settled cases.[3] Priest and Klein's analysis has been especially important to empirical legal studies, because it implies that quantitative analysis is incomplete and probably misleading if it does not take into account disputes that never reached final judgment.

The intellectual power of the Priest-Klein hypothesis has, in some respects, outpaced its empirical verification. While legal scholars have mounted a sustained campaign to test the hypothesis in a number of legal settings, ranging from torts to tax,

---

[1] George Priest and Benjamin Klein, "The Selection of Disputes for Litigation," 13 J. Legal Stud. 1 (1984).

[2] Fred R. Shapiro, "The Most-Cited Law Review Articles Revisited," 71 Chicago-Kent L. Rev. 751, 771 (1996) (ranking it 99th of all law articles); James E. Krier & Stewart J. Schwab, "*The Cathedral* at Twenty-Five: Citations and Impressions," 106 Yale L. J. 2121, 2145 (1997) (ranking it 81st of all law articles); William M. Landes & Richard A. Posner, "Heavily Cited Articles in Law," 71 Chi-Kent L. Rev 825, 838 (1996) (ranking it 28th of all articles in "predicted 'lifetime' citations").

[3] See, e.g., Keith Hylton, "Asymmetric Information and the Selection of Disputes for Litigation," 22 J. Legal Stud. 187 (1993); Steven Shavell, "Any Frequency of Plaintiff Victory at Trial is Possible," 25 J. Legal Stud. 493 (1996); Joel Waldfogel, "The Selection Hypothesis and the Relationship between Trial and Plaintiff Victory, 103 J. Pol. Econ. 229 (1995); Theodore Eisenberg, "Testing the Selection Effect: A New Theoretical Framework with Empirical Tests," 19 J. Legal Stud. 337 (1990); Theodore Eisenberg & Henry S. Farber, "The Litigious Plaintiff Hypothesis: Case Selection and Resolution," 28 RAND J. Econ. S92 (1997); Samuel R. Gross & Kent Syverud, "Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial, 90 Mich. L. Rev 319 (1991); J. Mark Ramseyer and Minoru Nakazato, "The Rational Litigant: Settlement Amounts and Verdict Rates and Japan," 15 J. Legal Stud. 263 (1989); Joel Waldfogel, "Reconciling Asymmetric Information and Divergent Expectations Theories of Litigation," 41 J. Law & Econ. 451 (1998); Kuo-Chang Huang, "How Legal Representation Affects Case Outcomes: An Empirical Perspective from Taiwan," 5 J. Empirical Legal Stud. 197 (2008).

the results have been mixed.[4] This article endeavors to test the Priest-Klein hypothesis

using a data set of private prosecutions of crime from thirteenth-century England. While

the Priest-Klein analysis is usually applied only to civil cases, it is appropriate to test it on

these criminal cases, because, unlike modern criminal cases, they were privately

prosecuted by the victim or a relative, who could settle the case and retain the entirety of

the proceeds. I find that the data are largely consistent with the theoretical predictions.

My results are particularly notable for at least two reasons. First, this is the only

existing study to evaluate the Priest-Klein selection theory using data from before 1870.[5]

Perhaps more importantly, the data set of medieval private prosecutions facilitates a

uniquely powerful test of selection theories, because courts at the time elicited jury

verdicts *even after settlement*, thereby allowing for direct comparison of litigated and

settled cases. As discussed more fully in section 2, judges in mid-thirteenth-century

England decided to disregard settlements in criminal cases and to send settling defendants

to trial. This practice was made possible by the fact that juries at this time were "self-

informing" and so did not need the prosecutor's cooperation at trial to learn about the

case.[6] Because modern data never include post-settlement verdicts, previous empirical

---

[4] For a review of the literature through 1995, see Daniel Kessler, Thomas Meites, and Geoffrey
Miller, "Explaining Deviations from the Fifty-Percent Rule: A Multimodal Approach to the Selection of
Cases for Litigation," 25 J. Legal Stud. 233 (1996), 236-341. See also the articles by Waldfogel,
Eisenberg, Gross & Syverud, and Ramseyer & Nakazato cited in the previous footnote, the two articles
cited in the next footnote, and Peter Siegelman & John J. Donohue III, "The Selection of Employment
Discrimination Disputes for Litigation: Using Business Cycle Effects To Test The Priest-Klein
Hypothesis," 24 J. Legal Stud. 427, 429 (1995); Peter Siegelman & Joel Waldfogel, "Toward a Taxonomy
of Disputes: New Evidence through the Prism of the Priest/Klein Model," 28 J. Legal Stud. 101 (1999).

[5] In fact, with one exceptions, all empirical studies that I am aware of use data from after 1950.
The exception is Stanton Wheeler et al., "Do the 'Haves" Come Out Ahead? Winning and Losing in State
Supreme Courts, 1870-1970," 21 Law & Soc. Rev. 403 (1987).

[6] Daniel Klerman, "Was the Jury Ever Self-Informing? 77 S. Cal. L. Rev. 123 (2003).

tests of selection theories could test only predictions regarding litigated cases.[7] Such tests

have been inconclusive, in part, because the predictions about litigated cases necessitate

assumptions about the underlying distribution of all cases (i.e. both settled and litigated

cases). As such, a critic of selection theories could always argue that prior empirical tests

were flawed, because the settled cases, if they could be observed, might have

characteristics identical to the litigated cases. Because the medieval data studied here

include jury verdicts in both settled and litigated cases, they can provide a more

conclusive test of the Priest-Klein hypothesis.

Priest and Klein argued that litigated cases are a biased sample of all legal

disputes, because most cases settle and because settled cases are likely to differ from

litigated cases. In general, they argued that litigation is more likely when the case is

close and both sides have a roughly equal chance of prevailing at trial. Thus, their most

heralded prediction is that litigated cases will tend, on average, to result half in verdicts

for the plaintiff and half in verdicts for the defendant. They are careful to acknowledge,

however, that a number of factors will cause deviations from this prediction, including

the underlying distribution of all cases (settled and litigated), the proportion of cases

which settle, and most importantly for this article, asymmetric stakes. When defendants

have more at stake in the litigation (perhaps because they fear bad publicity or an adverse

precedent), they will be more cautious, settling even close cases and litigating only when

---

[7] There are, however, two notable exceptions. In their article, Linda Stanley and Don Coursey
used an experimental design, which gave them information on the settled cases. Linda R. Stanley & Don
L. Coursey, "Empirical Evidence on the Selection Hypothesis and the Decision to Litigate or Settle," 19 J.
Legal Stud. 145 (1990). In her research, Leandra Lederman compared characteristics of settled and
litigated tax cases, including the size of the stakes and the decade the judge was appointed to the bench, but
not (for the obvious reason) the verdict or judicial decision. Leandra Lederman, "Which Cases Go to Trial:
An Empirical Study of Predictions of Failure to Settle," 49 Case W. Res. L. Rev. 315 (1999). It should also
be noted that Waldfogel's articles, cited in the previous two footnotes, examine not only the percentage of
plaintiff victories in litigated cases, but also the percentage of cases that were litigated.

they are confident they will prevail. As a result, when defendants have more at stake, litigated cases will contain more pro-defendant verdicts. Conversely, if plaintiffs have more at stake, litigated cases will contain more pro-plaintiff verdicts.

The implications of asymmetric stakes are important for the analysis of my data because it is doubtful that plaintiffs (i.e. prosecutors) and defendants valued conviction equally. Although society may have had a large stake in conviction, the prosecutors' and defendants' valuations are what matter here.   Because prosecution was private, only the prosecutor's and defendant's personal stakes affected the decision to settle. I argue below in Section 2 that defendants in thirteenth-century private prosecutions had more at stake than prosecutors, because guilty verdicts usually resulted in a substantial fine levied against the defendant, which was paid exclusively to the king. In contrast, the prosecutor obtained more attenuated benefits from the conviction, consisting of the avoidance of a small fine for false prosecution and, perhaps, reputational benefits and retributive satisfaction. Under such circumstances, the Priest-Klein model predicts that litigated will contain a higher proportion of victories for the defendant.

The data largely conform to the theoretical predictions. Defendants were found guilty as charged more than twice as often in settled cases (82%) than in litigated cases (37%). While this difference would seem remarkably consistent with the Priest-Klein hypothesis, the analysis is complicated by partial guilt verdicts and missing data. Partial guilt verdicts are verdicts in which the defendant was acquitted of the accusation made by the prosecutor, but still found guilty of a less severe, related offense. For example, defendants accused of mayhem (causing serious bodily injury) were often found guilty of simple battery, and defendants accused of homicide were sometimes found guilty only of

being an accessory. I also classify as partial guilt verdicts situations where the prosecutor accused the defendant of multiple crimes (e.g. beating and theft), but the jury found the defendant guilty of only one (e.g. just the theft). Existing theoretical work on the selection of disputes for litigation assumes only two possible verdicts (liable or not liable) and does not address situations where three verdicts (innocent, partially guilty, guilty-as-charged) are possible. Thus, in order to conduct a genuine test of the Priest-Klein hypothesis using the medieval data, it will be necessary to generalize their theory to incorporate the possibility of partial verdicts. After making such modifications, I still find that the Priest-Klein framework generates predictions that are consistent with the medieval data. This modification might also have some relevance to modern civil cases involving multi-count complaints. For example, if a complaint alleged both fraud and negligence, a jury verdict finding liability for negligence but not fraud could be analyzed as a partial verdict within the framework developed here.

While the practice of rendering verdicts even in settled cases makes the thirteenth-century data especially valuable for testing selection theories, such verdicts were not rendered and recorded in every case. Moreover, for an overlapping set of cases, it is unclear whether a case actually settled or was merely dropped by the prosecutor for other reasons. The number of cases lacking information on either or both of these dimensions (guilt and/or settlement) is large enough that it could, under some assumptions, erase the differences between settled and litigated cases. For this to occur, it would be necessary for the missing data to be drawn from a population of disputes that diverged sharply and implausibly from the recorded data. Such a divergence is possible. Indeed, a key insight of the Priest-Klein analysis is the necessity of exploring the plausibility of such

divergences between observed and unobserved cases. Nevertheless, as will be discussed in section 5.A, under the most plausible assumptions about the characteristics of the missing data, the difference between settled and litigated cases remains large and statistically significant.

The data set of thirteenth-century private prosecutions is also unique in containing jury verdicts in cases which the prosecutor initiated but then dropped before trial without settlement. As relatively simple economic models of litigation would predict, juries reported that defendants were innocent in the overwhelming majority (74%) of such cases. Prosecutors rationally decided that it was not worth their time to litigate these cases, which were likely to lead neither to a negotiated settlement nor to the retributive satisfaction of a trial conviction.

Section 2 of this article provides general background on thirteenth-century private criminal prosecutions and argues that the stakes in these cases were probably larger, on average, for the defendant than for the prosecutor. Section 3 reviews the Priest-Klein theory of the selection of disputes for litigation, modifies it to take into account the possibility of partial-guilt verdicts and of trials after settlement, and generates empirical predictions. Section 4 describes the data, tests the empirical predictions, and discusses the results. Section 5 addresses some limitations of the data, including the implications of missing data and the possibility that jurors inferred guilt from settlement. Section 5 also investigates asymmetric information, differences between crimes, and alternative explanations for differences between settled and litigated cases.

2. Thirteenth-Century Private Prosecutions of Crime[8]

During the middle ages, and indeed until the twentieth century, much if not most crime in England was prosecuted privately by victims or their relatives.[9] In the thirteenth century, such prosecutions were usually called "appeals." Such appeals should not be confused with modern appeals, because they had nothing to do with the correction of errors by a superior court. To "appeal" someone meant simply to prosecute him for crime. Appeals could be brought for a wide array of offenses, most typically for assault (including beating, wounding and mayhem), homicide, theft (including simple larceny, robbery and burglary), and rape. Lawyers seldom represented either party.

The best way to understand a thirteenth-century private prosecution is to read a case. Some of the more important parts are underlined.

> John son of Benedict appealed Ivo Quarel, Osbert Cokel and Henry Wyncard in county court of [breach of the] king's peace, wounds and imprisonment etc. And he [John] now comes and does not want to prosecute them. Therefore let him be committed to jail and his sureties, Ayltrop Balliol and Walter son of Odo, are in mercy [fined]. And Ivo and the others come [to court]. And the jurors testify that they [John, Ivo, Osbert and Henry] have settled and they [the jurors] say that, in truth, the aforesaid Ivo and the others came to the property of Matthew of Leyham in Barford and fished there without Matthew's permission and contrary to his wishes. The aforesaid John came along and asked them for a pledge,[10] and the aforesaid Ivo would not give him one, but instead [Ivo] struck the aforesaid John in the head with a hatchet and made two wounds each three

---

[9] Private prosecutions could take two forms: prosecution by indictment and prosecution by appeal. In the thirteenth century, prosecution by indictment was not yet a significant source of private prosecution, so prosecution by appeal was the dominant form of private prosecution. On the transition from appeal to indictment, see Daniel Klerman, "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1, 5-8 (2001).

[10] John probably asked Ivo and the others for a pledge that they would show up at court if they were sued for fishing without permission.

inches long down to the crest of the head. And they [Ivo and the others] beat him badly. And afterwards they took him and bound him and put him in a boat and took him from this county [Bedfordshire] to the county of Huntingdonshire to Ivo's house at Buckden. There they dragged him with a rope to a window of Ivo's solarium and forced him to break the window with an ax. And they painted the wall near the window with the blood flowing from the wounds the aforesaid Ivo had given the aforesaid John, and they dragged him through the window and set upon him a blanket and some linen saying that he had stolen them.[11] And they raised the hue [and cry] and caused the men who responded to the hue [and cry] to understand that eighteen thieves had come to his house, and that all except the aforesaid John had gotten away. So they put the blanket and the linen on him and took him to Huntingdon and [they] gave him to the sheriff to be incarcerated. And he remained in prison until his tithing delivered him.[12] Therefore let the aforesaid Ivo and the others be taken into custody.[13] Later Ivo Quarel came and made fine for forty marks[14] [i.e. promised to pay the king forty marks to be released from custody] by sureties Ralf Ridel [and eleven others].[15]

In this case, John accused Ivo and others of wounding and false imprisonment.

Like nearly all prosecutors, John initiated the case in county court. In fact, potential

prosecutors were required to sue at the first county court meeting after the commission of

the alleged offense. County court met every four weeks, so prosecutors had a month or

less to initiate their cases. Trial was postponed until royal justices arrived in the county,

which usually meant a delay of several years. During that interval, John, Ivo and the

---

[11] When a thief was caught in the act, the apprehenders often tied the stolen goods to the thief before bringing him to local authorities. See R.C. van Caenegem, *English Lawsuits from William I to Richard I*, Selden Society vol. 107, 2.508 (1990).

[12] The tithing probably secured his release, pending trial, upon a promise that they would ensure his presence at trial. Every adult male was required to be in a tithing, a group whose most important function was producing its members' attendance in court when necessary.

[13] The justices had no intention of keeping Ivo in prison. Imprisonment, or the threat of it, was used not as punishment, but to induce convicts to pay fines.

[14] A mark was a unit of currency, equivalent to two-thirds of a pound.

[15] Public Record Office, London, JUST 1/4, m. 30 (Bedfordshire 1247). This case exists only in Latin in manuscript form. Transcription and translation by the author. Dr. Paul Brand helped with two difficult-to-read words.

other defendants settled the case for undisclosed consideration. As a result, when the case came to trial, John told the judges he did not want to prosecute.

Although it was not officially condoned, settlement was common. Judicial treatment of settled cases varied over the thirteenth century. During most of the early part of the century, settlements were tolerated. When judges were aware of settlement, they fined both parties. Nevertheless, such fines were uncommon, because judges seldom knew when a case a settled. They knew when a case was not prosecuted, but they did not know whether non-prosecution reflected settlement, prosecutorial realization that the case was weak, or some other factor. As a result, defendants who settled were, *de facto*, acquitted. Starting in 1234, however, judges began to ask jurors routinely whether the parties had settled, and thus began more often to fine those who had settled. The case quoted above reflects this new policy towards settlement. Although the judge's inquiry regarding settlement is not recorded, the jurors' response – "They have settled" – is. Because, as discussed below, the jury was local and "self-informing," it would have known whether the case had settled.

In addition to inquiring about settlement, during the period 1218-1222 and after 1239, judges took an interest in the non-prosecuted cases and usually sent the accused to jury trial in spite of the fact that these cases were often settled. As explained in the next paragraph, these cases were tried without a prosecutor. Or, as one thirteenth-century commentator put it metaphorically: in such cases, the king acted as prosecutor.[16] If convicted, the defendant would be fined or hanged for the offense, just as he would have

---

[16] *Bracton on the Laws and Customs of England*, ed. George E. Woodbine, trans. Samuel. E. Thorne (1968-77), 2:402, f. 142b (When an appellor defaults, the "the king may proceed *ex officio….*"). Of course, the king would not himself actually prosecute, and there were no professional or official prosecutors to prosecute on his behalf in ordinary cases.

been if he had not settled. This, indeed, is what happened in the case quoted above. In spite of the settlement, Ivo and the others were tried. The jurors reported that Ivo both wounded John and caused him to be falsely imprisoned. The jury thus rendered a guilty-as-charged verdict, which made Ivo liable for a forty mark fine. Forty marks was a very large amount, about $40,000 in 2011 dollars.

Because this was the period of the "self-informing jury," the prosecutor's participation at trial was not required for conviction. Little or no evidence was presented at trial. Instead, the jury, which was composed exclusively of local men, including men from the defendant's village, was expected either to have investigated the dispute before trial or to decide the case based on reputation and rumor. Trial was an opportunity for the jury to report what it knew to the judges, not for jurors and judges to hear evidence from witnesses and the prosecutor.[17] The record of the case quoted above reflects this practice. There is no indication that John testified. Indeed, no evidence at all seems to have been presented at trial. Instead, immediately after statement of the accusation and ascertainment of the presence or absence of the parties, the jurors "testify" to what happened, thus rendering a guilty verdict. Guilty verdicts were common in settled cases. As discussed in section 4, defendants were convicted of all charges in eighty-two percent of settled cases, even though the prosecutor usually contributed nothing at trial to these cases, and may not even have shown up. The possibility that juries inferred guilt from settlement is discussed in section 5.2.

As a result of the two inquiries described above (inquiry into whether non-prosecuted defendants had settled and inquiry into whether they were guilty), after 1239, it is possible to ascertain both how far the prosecutor went with the case and the guilt of

the defendant. For example, in the case above we know that the prosecutor settled and that the defendants were guilty. In other cases, prosecutors proceeded to trial or dropped their cases. "Dropped" means initiated prosecution in county court, but neither settled nor prosecuted to trial.

Figure 1 illustrates the key decisions made by each actor—prosecutor, defendant, and judge— and the sequence of their actions[18]:
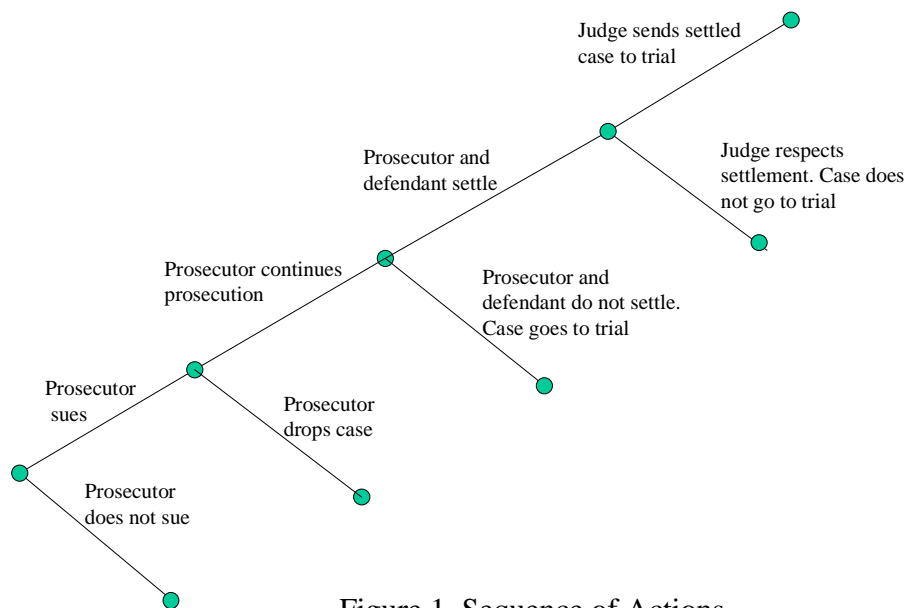
Figure 1. Sequence of Actions

The fact that defendants would be sent to jury trial even if they settled, of course, severely undercut their incentive to settle.[19] Why settle if one is to be tried and punished anyway? The fact that any defendants settled during periods when settling defendants

---

[17] Daniel Klerman, "Was the Jury Ever Self-Informing? 77 S. Cal. L. Rev. 123 (2003).

[18] Note, however, that although the prosecutor's decision to drop or continue the case is depicted as taking place prior to the decision to settle, the prosecutor might drop the case after the parties failed to settle.

[19] In addition, as fully explored in Daniel Klerman, "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1 (2001), the dilution of the defendants'

were tried thus requires some explanation. There are three principal reasons defendants

might still settle. First, sending settled defendants to trial was a new and unannounced

policy, so some defendants settled with the expectation that their settlements would be

respected and then were surprised to be sent to trial. This is an especially plausible

explanation for the period 1239-1249, when the policy was relatively new. Because royal

judges visited the countryside to hear criminal cases only every few years, it would have

taken about a decade for the inhabitants of every county to become aware of the new

policy. Second, until the 1260's, judges did not apply the new policy with absolute rigor.

They obtained jury verdicts in only about sixty percent of non-prosecuted cases. As a

result, settlement could still be rational even for well-informed defendants, because there

was a significant probability it would protect them from trial and punishment. The data

analyzed in this article are restricted to the period for which these two explanations are

most plausible, that is from 1239 to the early 1260's. Third, the fact that a defendant

might be sent to trial even if he settled *increases* the prosecutor's incentive settle.

Because the prosecutor could obtain the benefits of both settlement and a conviction, the

prospect of trial after settlement lowered the prosecutor's minimum settlement demand.

This, in turn, made it possible for the defendant to settle even if he made a low offer.

This phenomenon is discussed in greater depth and is modeled at the end of section 3.

Two other reasons may also help explain why defendants continued to settle:

punishments imposed on settling defendants may have been less severe than on those

who contested their cases, and settlement may have provided non-legal benefits to the

defendant, such as partial repair of the reputational damage conviction might impose.

---

incentive to settle also reduced victims' incentive to prosecute, and thus significantly reduced the number
of private prosecutions brought.

While most verdicts were simply "guilty" or "not guilty," jurors could also give more nuanced verdicts indicating that the defendant was partially guilty. For example, a jury could report that a defendant accused of beating and theft was guilty of the beating but not the theft, or a jury could report that a defendant accused of mayhem (maiming bodily injury) was guilty only of battery (simple, non-maiming beating). The sanctions imposed when juries returned these partial verdicts tended to be much smaller than the sanctions when the defendant was found guilty as charged. The defendant was never hanged, and the fines were, on average, only a third as large as the fines imposed when the defendant was found guilty as charged.

As in the suit between Ivo and John, guilty defendants were usually fined, with the fines paid to the royal treasury. Those convicted of homicide and theft were sometimes hanged, and, when hanged, forfeited their lands and chattels. Nevertheless, a prosecutor who secured a conviction received neither damages from the defendant nor a bounty from the state. An unsuccessful prosecution, however, would result in imposition of a small fine for having made a false accusation. Dropped prosecutions also resulted in a small fine, and settled prosecutions might result in a somewhat larger fine. These fines were usually much smaller than the fines paid by guilty defendants. For example, although the fine is not recorded in the case quoted above, John probably paid a fine of one mark or less. In contrast, Ivo paid forty marks.

As the previous paragraph suggests, the stakes in thirteenth-century private prosecutions were probably asymmetric and larger for defendants, because convicted defendants usually paid a fine to the royal treasury, while prevailing prosecutors received no monetary benefit from conviction, except avoidance of the small fine which would

otherwise be imposed for making a false accusation. Because prosecution was private, only the prosecutor's and defendant's stakes mattered for settlement. Although society's stake in a conviction might have been large, there was no institutional mechanism which forced the parties to take it into account in their settlement negotiations.[20] That defendants' stakes were larger, at least on average, is crucial to testing the Priest-Klein theory, because, as discussed in the next section, the Priest-Klein theory yields radically different predictions depending on whether the stakes are symmetric or if prosecutors have larger stakes.

If one considers only prosecutors' monetary stakes, the conclusion that defendants' stakes were larger would be nearly unassailable. Nevertheless, the analysis is complicated by the fact that prosecutors probably received nonmonetary benefits from conviction. A conviction might, for example, deter others from committing offenses against the prosecutor in the future.[21] A rape victim might also vindicate her reputation through a successful prosecution. In addition, prosecutors might derive satisfaction from retribution, that is, from the punishment of the defendant. The magnitude of these nonmonetary benefits is difficult to measure. It is possible that they were so large as to make prosecutors' stakes equal to or larger than defendants'. Nevertheless, as explained in the next paragraph, it seems most plausible to infer that the nonmonetary benefits were significant but small compared to the sanctions a convicted defendant would bear.

---

[20] Of course, the fact that judges began to put defendants to trial in spite of settlement reflected society's stake in the case.

[21] See David Friedman, "Making Sense of English Law Enforcement in the Eighteenth Century," 2 U. Chi. Roundtable 475 (1995).

A rough sense of the importance of the nonmonetary benefits can be garnered by examining how potential prosecutors changed their behavior when settlement policy changed. As noted above, sometimes settlements were respected, while at other times they were not. When settlement was respected, victims could be motivated to prosecute by the prospect of a monetary (or in-kind) settlement, as well as the nonmonetary benefits mentioned above (deterrence, reputation, and retribution). When judges took jury verdicts in settled cases, however, defendants became reluctant to settle, because settlement did not provide protection against conviction and punishment.[22] Thus, when judges took jury verdicts in settled cases, prosecutors who continued to prosecute were motivated almost exclusively by the nonmonetary benefits of prosecution. When judges took jury verdicts in settled cases, potential prosecutors brought only one-third as many prosecutions.[23] This suggests that the nonmonetary benefits of prosecution were relatively small compared to the benefits of settlement, but still large enough to motivate some prosecutions. The fact that defendants entered into settlements voluntarily suggests that the benefits of settlement were smaller than the sanctions upon conviction. Thus, the nonmonetary benefits of successful prosecution to the prosecutor were probably, on average, smaller than the sanctions imposed on the defendant upon conviction. Since the prosecutor's stakes were principally the nonmonetary benefits of prosecution and the defendant's stakes were principally the sanctions, it is reasonable to conclude that the stakes were, on average, asymmetric and larger for the defendant.

---

[22] Daniel Klerman, "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1 (2001).

[23] Id.

It should also be noted that, although the discussion so far has focused on the nonmonetary aspect of the prosecutor's stakes, the defendant's stakes also had a nonmonetary component, even when the formal sanction was a fine. Conviction almost certainly injured the defendant's reputation. This reputational injury would increase the defendant's stakes at trial and thus lend further plausibility to the idea that the defendant's stakes were larger than the plaintiff's.

3. Theory and Predictions

In their seminal 1984 article, George Priest and Benjamin Klein analyzed the decision to settle or litigate and showed that litigated cases are not a random sample of all disputes. They argued that the cases which are most likely to be litigated are those where liability is highly uncertain. When liability is certain, both plaintiff and defendant will find it advantageous to settle in order to avoid litigation costs. Similarly, when the defendant is clearly not liable, the plaintiff will not even bother to sue, or, if she does sue, she will be willing to settle for a small amount. Thus, only the "hard" cases, in which plaintiff and defendant can come to very different assessments of the outcome, are likely to be litigated. Because only the close cases will be litigated, Priest and Klein argue that, under certain, special circumstances, plaintiffs and defendants should each prevail in about fifty percent of the litigated cases.

Priest and Klein's 1984 article also analyzed several factors which might cause deviations from the fifty-percent tendency, most importantly asymmetric stakes. When the stakes for the defendant are larger, Priest and Klein showed that defendants should prevail more often in litigated cases. In such situations, litigation is still more likely in

close cases, but defendants have strong incentives to settle cases in which they are more likely to be found liable and to litigate only those in which they are more likely to be found not liable. Because their losses from an adverse verdict are larger than plaintiffs' gains from a favorable verdict, defendants are willing and able to make settlement offers that plaintiffs will find attractive, even when defendants think a plaintiff victory highly uncertain. Conversely, only when the facts are clearly in the defendant's favor, will a defendant find the risk of an adverse judgment more attractive than settlement. As a result, Priest and Klein predict that, when the stakes are larger for the defendant, defendants should prevail more often than plaintiffs in litigated cases. Conversely, if plaintiffs have more at stake, litigated cases will contain more pro-plaintiff verdicts.

Figure 2 illustrates the selection effect when the defendant has larger stakes.
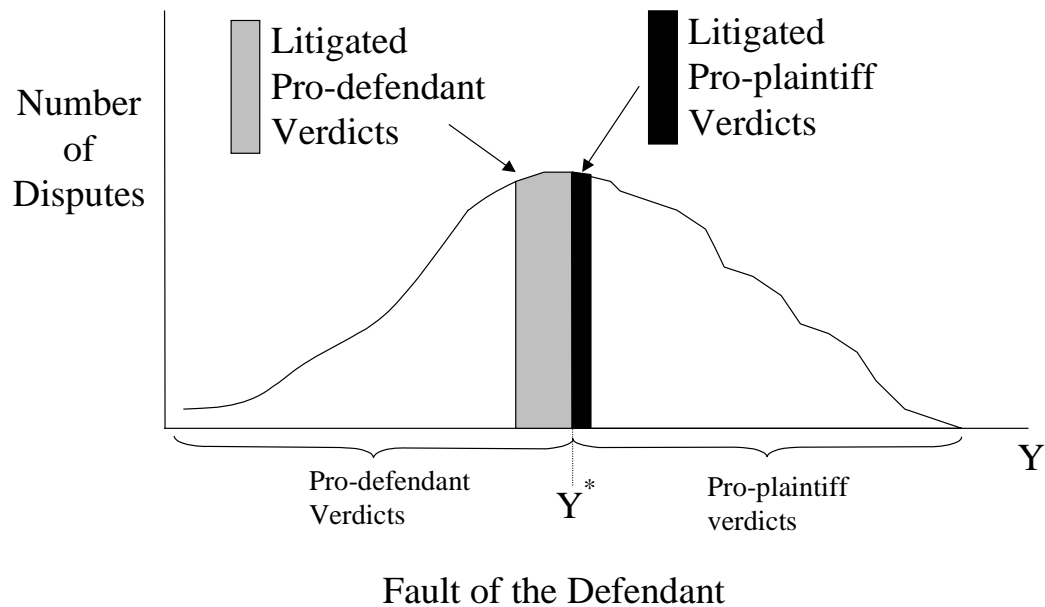
Figure 2. Priest & Klein Model

Y measures the degree of the defendant's fault, and Y^* is the decision standard. If the defendant's fault exceeds Y^*, the defendant is found liable. If the defendant's fault is less than Y^*, the defendant is found not liable. Cases resulting in litigation are shaded. Litigated cases resulting in verdicts of liability are shaded heavily, while those resulting in no liability are shaded lightly. The unshaded area under the curve represents settled cases. The figure illustrates two important aspects of the Priest-Klein model. First, litigation is most likely in close, hard cases. Thus, the shaded areas flank the decision standard, Y^*. Second, because the stakes are asymmetric and larger for the defendant, defendants are more likely to settle cases which they expect to lose, and conversely more likely to litigate only those cases that they are likely to win. Thus, the shaded area to the left of the decision standard is larger than the shaded area to the right. Of course, this graph is a simplification. Sometimes there will be litigation even when the defendant's fault is far from Y^*, and sometimes the parties will settle even when the defendant's fault is close to Y^*. Nevertheless, the figure helps to illustrate the general tendency in the selection of disputes for litigation.

The analysis in the preceding paragraph is complicated by the possibility that prosecutor and defendant could disagree not only about whether the defendant was guilty, but also about the degree of the defendant's guilt. As noted above in section 2, juries did not always simply report that the defendant was guilty as charged or completely innocent. Rather, jurors sometimes returned verdicts of partial guilt, declaring the defendant guilty of a less severe offense than charged or of some but not all charged offenses. Priest and Klein do not discuss the possibility of such partial verdicts, but their model can be

adapted to take them into account. This modification might also have some relevance to

modern civil cases involving multi-count complaints. For example, if the complaint

alleged both fraud and negligence, a jury verdict finding liability for negligence but not

fraud could be analyzed as a partial verdict within the framework developed here. The

biggest change is that one needs two decision standards, Y^p and Y^g, rather than just one,

Y^*. Y^p represents the degree of culpability just sufficient to lead a jury to render a

partial guilty verdict, while Y^g represents the degree of culpability just sufficient to lead a

jury to render a guilty-as-charged verdict. Figure 3 illustrates the Priest-Klein model

modified to take into account three possible verdicts.
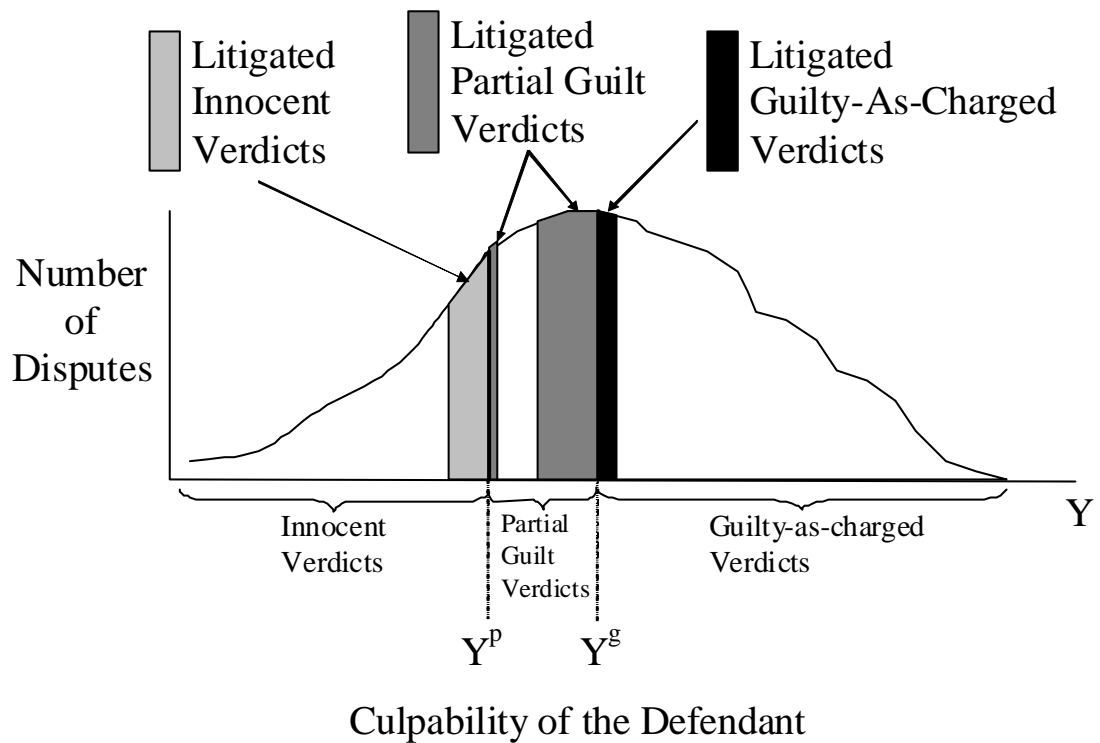

Figure 3. Modified Priest & Klein Model

As in Figure 2, litigated cases are shaded, while unshaded areas under the curve represent

settled cases. As in the original Priest-Klein model, disputes are more likely near the

decision standards, that is, near Y^p and Y^g. Since the stakes are asymmetric, one expects

more partial guilt verdicts in disputes near Y^g and more innocent verdicts in disputes near

Y^p.

In Figure 3, Y^p and Y^g have been placed so that the proportion of innocent,

partially guilty, and guilty-as-charged verdicts approximates the proportion in the dataset,

as described in the next section.[24] That is, Y^p and Y^g are close together, because there

are relatively few partial-guilt verdicts in the dataset, and Y^p is rather far to the left,

because there are fewer innocent than guilty-as-charged verdicts in the dataset. It is thus

possible to make some rough predictions from the figure about the settled and litigated

cases. Guilty-as-charged verdicts will be relatively uncommon among the litigated cases,

because there are many guilty-as-charged cases in the full data set, but disputes near

Y^g will result in relatively few such verdicts. Conversely, partial guilt verdicts will be

disproprortionately common among the litigated cases, because such verdicts are

uncommon in the full data set, but disputes near Y^g will result in many such verdicts,

while disputes near Y^p will contribute a few more. Finally, there will be more not guilty

verdicts among the litigated cases, because disputes near Y^p will result in large numbers

of innocent verdicts, while such verdicts are somewhat under-represented in the full data

set.

---

[24] For this paragraph, the dropped cases have been omitted from the analysis. As discussed below,
they require an additional modification of the Priest-Klein framework.

While the discussion above helps to develop intuitions, it is also useful to analyze the problem more formally. The formal analysis tracks Priest and Klein's article, modifying the mathematics to take into account partial verdicts.[25] In each dispute, the defendant's true culpability is Y'. Prior to trial, each potential litigant forms an estimate of Y'. Let $\hat{Y}'_\pi$ be the prosecutor's estimate of the defendant's culpability, and let $\hat{Y}'_\Delta$ be the defendant's estimate of the defendant's culpability.[26] Thus:

$$\hat{Y}'_\pi = Y' + \varepsilon_\pi$$

and

$$\hat{Y}'_\Delta = Y' + \varepsilon_\Delta,$$

where ε_π and ε_Δ are independent random variables with zero expectation and identical standard errors, σ_ε, where σ_ε is know to the parties. Given σ_ε and their estimates of Y', each party can estimate the probability of a partial guilt verdict and a guilty-as-charged verdict. The prosecutor's and defendant's estimates of the probability of a partial verdict are:

$$\hat{P}^p_\pi = P(Y^p \leq Y' < Y^g \mid \hat{Y}'_\pi)$$

and

$$\hat{P}^p_\Delta = P(Y^p \leq Y' < Y^g \mid \hat{Y}'_\Delta).$$

Similarly, the parties' estimates of the probability of guilty-as-charged verdicts are:

$$\hat{P}^g_\pi = P(Y' \geq Y^g \mid \hat{Y}'_\pi)$$

and

$$\hat{P}^g_\Delta = P(Y' \geq Y^g \mid \hat{Y}'_\Delta).$$

---

[25] Those unfamiliar with the Priest-Klein model may find it helpful to review pages 6-12 and 24-28 of "The Selection of Disputes for Litigation," 13 J. Legal Stud. 1 (1985).

[26] I use the Greek letters π and Δ, rather than p and d, to denote plaintiff and defendant to avoid confusion regarding p, which already denotes a partial guilt verdict.

Let J_π^p and J_Δ^p represent the value to prosecutor and defendant of a partial guilty verdict, and let J_π^g and J_Δ^g represent the value of a guilty-as-charged verdict. If, as I maintain, the stakes were asymmetric and greater for the defendant, then $J_\pi^p < J_\Delta^p$ and $J_\pi^g < J_\Delta^g$. For simplicity, I will express the prosecutor's valuation of partial guilty and guilty-as-charged verdicts as proportional to the defendant's valuations, so $J_\pi^p = \beta J_\Delta^p$ and $J_\pi^g = \beta J_\Delta^g$, where β represents the asymmetry of the stakes, $0 < \beta < 1$. In addition, since the sanction for a partial verdict was less than the sanction for a guilty-as-charged verdict, J_Δ^p and J_π^p and can be rewritten as αJ_Δ^g and $\alpha\beta J_\Delta^g$, respectively, where α represents the degree to which partial verdicts resulted in lesser sanctions than guilty-as charged verdicts, $0 < \alpha < 1$.

The parties will settle when the prosecutor's minimum settlement demand (A) is less than or equal to the defendant's maximum settlement offer (B).[27] A and B can be represented as follows:

$$A = \hat{P}_\pi^p J_\pi^p + \hat{P}_\pi^g J_\pi^g - C_p + S_p = \alpha\beta\hat{P}_\pi^p J_\Delta^g + \beta\hat{P}_\pi^g J_\Delta^g - C_p + S_p$$

and

$$B = \hat{P}_\Delta^p J_\Delta^p + \hat{P}_\Delta^g J_\Delta^g + C_\Delta - S_\Delta = \alpha\hat{P}_\Delta^p J_\Delta^g + \hat{P}_\Delta^g J_\Delta^g + C_\Delta - S_\Delta,$$

where C_π and C_Δ are the prosecutor's and defendant's litigation costs, and S_π and S_Δ are their respective settlement costs. Since the parties settle when $A \leq B$, the conditions for litigation can be simplified to:

$$\beta\hat{P}_\pi^g - \hat{P}_\Delta^g + \alpha(\beta\hat{P}_\pi^p - \hat{P}_\Delta^p) > (C - S)/J \tag{1}$$

---

[27] This is a strong and somewhat unrealistic assumption, as it assumes that bargaining never breaks down. Nevertheless, I have made the assumption, because Priest and Klein make it. I briefly discuss the implications of relaxing this assumption in the last paragraph of section 4.

where $C = C_\pi + C_\Delta$, $S = S_\pi + S_\Delta$, and $J = J_\Delta^g$.

It is difficult to make analytic predictions about what percentage of litigated cases will result in innocent verdicts, partial verdicts, or guilty-as-charged verdicts, because these percentages will vary with the underlying distribution of all cases (both settled and litigated) and with the values of α, β, σ_ε, and $(C\text{-}S)/J$. Some predictions can be made about the limit case, as σ_ε, the standard deviation of the parties' errors, goes to zero.[28] Nevertheless, since parties were probably unable to predict outcomes with precision, it is most helpful to consider a range of values of σ_ε. Simulations are most appropriate for this task. Table 1 shows some simulation results under various assumptions.

---

[28] See the next three footnotes.

Table 1. Simulation Results

Assumptions			Results						
			Litigated cases				Settled cases		
(C-S)/J	β	σ_ε	% litigated	% innocent	% partially guilty	% guilty as charged	% innocent	% partially guilty	% guilty as charged
0.5	0.75	1.0	2	23	35	42	25	23	52
		0.5	1	17	50	33	25	23	52
		0.1	0	*	*	*	25	23	52
	0.50 & 0.25	1.0	0	*	*	*	25	23	52
		0.5	0	*	*	*	25	23	52
		0.1	0	*	*	*	25	23	52
0.2	0.75	1.0	14	27	31	42	24	22	54
		0.5	9	21	47	31	25	20	54
		0.1	1	6	59	36	25	23	52
	0.50	1.0	6	32	34	33	24	22	53
		0.5	3	25	52	23	25	22	53
		0.1	0	*	*	*	25	23	52
	0.25	1.0	0.2	34	32	33	25	23	52
		0.5	0.1	11	80	9	25	23	52
		0.1	0	*	*	*	25	23	52
-0.2	0.75	1.0	55	33	25	42	15	21	64
		0.5	52	41	27	31	7	19	74
		0.1	48	50	42	8	1	6	93
	0.50	1.0	40	42	28	31	13	20	66
		0.5	39	54	29	17	6	19	74
		0.1	43	57	40	3	1	11	89
	0.25	1.0	30	51	27	22	13	22	65
		0.5	31	65	27	8	7	21	72
		0.1	28	85	14	1	1	27	72
-0.5	0.75	1.0	88	26	22	52	18	30	53
		0.5	93	25	22	52	10	43	47
		0.1	99	25	23	52	0	43	57
	0.50	1.0	74	30	24	46	10	21	69
		0.5	74	32	25	43	3	18	79
		0.1	60	41	37	22	0	2	98
	0.25	1.0	54	40	28	32	7	17	76
		0.5	49	47	33	19	2	14	84
		0.1	48	51	44	4	0	4	96

All results were rounded to the nearest integer and are based on simulations using 50,000 observations (disputes). For all results, $\alpha = 0.33$, as discussed in Section 2. In addition, it was assumed that culpability was distributed normally among the population of litigated and settled (i.e. non-dropped) cases, and that the population of litigated and settled cases contained 25% innocent defendants, 23% partially guilty defendants, 52% guilty-as-charged defendants. These percentages were derived from the actual population of litigated and settled cases, as described in Table 2 below.

Although the table shows that many results change dramatically as the parameters are changed, a few results are robust across most specifications: 1) The percent guilty-as-charged is almost always higher among settled cases than among litigated cases.[29] 2) The percent innocent is usually higher among litigated cases than among settled cases.[30] 3) The percent partially guilty is almost always higher among litigated cases than among settled cases.[31]

The above three results give rise to three predictions about the data:

Prediction 1. There should be a greater proportion of innocent verdicts among the litigated cases than among the settled cases.

Prediction 2. There should be a greater proportion of partial guilt verdicts among the litigated cases than among the settled cases.

Prediction 3. There should be a greater proportion of guilty-as-charged verdicts among the settled cases than among the litigated cases.

These predictions are consistent with those derived informally above from analysis of Figure 3. The paragraphs before and after Figure 3 suggest intuitions supporting the predictions. Of course, while these predictions hold true almost regardless of the value of the various parameters, many results change dramatically with the parameters. For example, the percent litigated varies dramatically with the value of *(C-S)/J*. When *(C-*

---

[29] It can be proven that, in the limit, as σ_ε goes to zero, this result will hold as long as the proportion of guilty-as-charged verdicts in the entire population of disputes is greater than *(1-α)/2*. Since the data suggest that $\alpha = 0.33$ and that the proportion of guilty-as-charged verdicts was 52%, this prediction holds in the limit.

[30] This result can be proven, in the limit, as long as β is not very close to 1. Note that four of the five exceptions in the simulations occur when β is high *($\beta = 0.75$)*. Note that results for the parameters (C-S)/J=0.2, β=0.5, σ_ε=0.5, appear to be an exception only because of rounding. Without rounding, the percent innocent would be higher among the litigated than among the settled cases (25.09 versus 24.74), although the difference is not statistically significant.

[31] This result holds in the limit, as long as the proportion of partial guilt verdicts in the entire population is less than half.

S)/J is positive, litigation is very unlikely. This is similar to the simple Priest-Klein result with asymmetric stakes and reflects the fact that when litigation is more expensive than settlement (as nearly all modern scholarship assumes), settlement is mutually advantageous unless the parties' estimates of the probability of prevailing are extremely far apart. On the other hand, when *(C-S)/J* is negative, litigation is very common. A negative value of *(C-S)/J* means that settlement is more costly than litigation. While this is implausible in modern litigation, it may have been true for medieval private criminal prosecutions, because litigation involved few monetary costs (because there were no lawyers) but settlement could result in the imposition of significant fines. Of course, litigation had some costs, such as the cost of travel and the cost of the litigants' time, but it is at least possible that the settlement fine was larger than these costs.

Priest and Klein make no predictions about dropped cases, that is, about cases which the prosecutor initiated in county court but then neither settled nor pursued to trial. Priest and Klein seem to assume that all cases will either be litigated or settled, and thus leave no room for such cases. This oversight is partially explained by the fact that the theoretical part of their article was written with empirical testing in mind, and few modern data sets distinguish between settled and dropped cases. It also reflects Priest's view that, as a matter of theory, plaintiffs will be able to extract some settlement (albeit a small one) even from weak cases, except when litigating against "defendants who have

adopted a strategy of contesting every filed action to a verdict."[32] This view is

idiosyncratic and in no way essential to the broader theory.[33]

A more plausible approach to dropped cases builds on the models of litigation

developed by Gould, Landes, Posner, and Shavell, upon which Priest and Klein also

relied for most of their analysis. Under these models, as long as litigation is costly, if the

prosecutor estimates that the expected benefit of litigation is low, then the prosecutor will

not even initiate prosecution. If however, the prosecutor must initiate the prosecution

before he has sufficient information, and if he later learns that the expected benefit is low

(most plausibly, because he learns that the probability of conviction is low), then there

will be cases which the prosecutor initiates but then drops. Such cases will consist

predominately of weak cases. The fact that potential medieval prosecutors had to initiate

their cases in county court within four weeks suggests that they would often have had to

initiate prosecution before they had sufficient information. Litigation was costly, not so

much in monetary terms, but in time. Prosecutors were seldom represented by counsel,

so they did not have to worry about lawyer's fees, but litigation did require travel, time in

court, and perhaps time lobbying jurors. As a result, one would predict that prosecutors

would drop significant numbers of cases. This analysis leads to a fourth prediction:

Prediction 4. Innocent verdicts should predominate among the dropped cases.

So far, the analysis has assumed that settlement protected the defendant from trial

and punishment. Nevertheless, as discussed in Section 2, this assumption became

---

[32] Priest, "Reexamining the Selection Hypothesis: Learning from Wittman's Mistakes," 14 J. Legal Stud. 215, 225 (1985).

[33] Priest concedes that this assumption contradicts the model of litigation developed by Shavell and others, which was standard at the time. Id at 222, 224. Priest's position, however, has gained some

increasingly false. Judges began to send defendants to trial in spite of settlement. By the 1250's, they did so in sixty percent of the cases.[34] While at first this new policy probably took litigants by surprise, toward the end of the period studied here it is likely that the parties began to anticipate that there was a significant chance that defendants would be tried and punished in spite of settlement. It is thus important to adjust the model to take this into account. Let λ be the probability that the defendant will be tried in spite of settlement. The fact that the defendant may be tried in spite of settlement makes settlement less attractive to the defendant and lowers his maximum settlement offer (B) by the expected disutility of trial, $\lambda[\hat{P}_\Delta^p J_\Delta^p + \hat{P}_\Delta^g J_\Delta^g]$.[35] On the other hand, the fact that the defendant may be tried (and punished) in spite of settlement makes settlement more attractive to the prosecutor and lowers his minimum settlement demand (*A)* by $\lambda[\hat{P}_\pi^p J_\pi^p + \hat{P}_\pi^g J_\pi^g]$. As a result, A and B can be rewritten as follows:

$$A = (1-\lambda)[\hat{P}_\pi^p J_\pi^p + \hat{P}_\pi^g J_\pi^g] - C_p + S_p = (1-\lambda)[\alpha\beta\hat{P}_\pi^p J_\Delta^g + \beta\hat{P}_\pi^g J_\Delta^g] - C_p + S_p$$

and

$$B = (1-\lambda)[\hat{P}_\Delta^p J_\Delta^p + \hat{P}_\Delta^g J_\Delta^g] + C_\Delta - S_\Delta = (1-\lambda)[\alpha\hat{P}_\Delta^p J_\Delta^g + \hat{P}_\Delta^g J_\Delta^g] + C_\Delta - S_\Delta,$$

Since the parties settle when $A \leq B$, the conditions for litigation can be simplified to:

$$\beta\hat{P}_\pi^g - \hat{P}_\Delta^g + \alpha(\beta\hat{P}_\pi^p - \hat{P}_\Delta^p) > \frac{C - S}{(1-\lambda)J} \qquad (2)$$

---

support from Lucian Arye Bebchuk, "A New Theory Concerning the Credibility and Success of Threats to Sue, " 25 J. Legal Stud. 1 (1996).

[34] See "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1, 39 (2001).

[35] Note that I am assuming that the defendant incurred no litigation costs from trial after settlement. That is, the defendant's maximum settlement offer is not further decreased by λC_Δ. This is historically accurate, because trials after settlement often occurred without the presence of either party, a fact made possible by the self-informing jury. For similar reasons, I am also assuming the prosecutor incurred no litigation costs from trial after settlement. If one assumed that the parties incurred litigation costs from trial after settlement, the results would be only slightly different. See next footnote.

where $C = C_\pi + C_\Delta$, $S = S_\pi + S_\Delta$, and $J = J_\Delta^g$. The only difference between this

condition—which takes into account the possibility of trial after settlement—and

condition (1) above—which does not—is that the right hand side is divided by $(1 - \lambda)$.

Since the effect of asymmetric stakes comes through the left hand side, Predictions 1

through 3 remain valid. The principle effect of dividing the right hand side by the

$(1 - \lambda)$ is to make settlement more likely when $C - S$ is positive (i.e. when litigation is

more expensive than settlement), and to make litigation more likely when $C - S$ is

negative (i.e. when settlement is more expensive than litigation). The intuition behind

these effects is easiest to grasp in the extreme case, where trial after settlement is a

certainty (i.e. $\lambda = 1$). If litigation costs are higher than settlement costs, then the parties

should always settle, because settlement avoids litigation costs[36] but otherwise leads to

the same result (i.e. trial) as settling. For the same reason, when settlement costs are

higher than litigation costs, the parties should always litigate. For moderate probabilities

of trial after settlement, the effects are correspondingly moderate: an increased tendency

to settle when litigation costs are higher than settlement costs and an increased tendency

to litigate when settlement costs are higher than litigation costs. These effects are

confirmed by the simulations. During the 1250's, the probability (λ) that the defendant

would be tried in spite of settlement was sixty percent. Thus, taking into account the

possibility of trial after settlement means considering values of *(C-S)/J* that are two and a

---

[36] Note here that I am assuming that the parties incurred no litigation costs from trial after settlement. See previous footnote. If one assumed that the parties incurred litigation costs from trial after settlement, the right hand side of condition (2) would be $\dfrac{C - [S/(1-\lambda)]}{J}$. This would have the effect of making settlement always less attractive, irrespective of whether settlement costs were larger or smaller than settlement. Since, as discussed in section 4, it is most plausible to assume that settlement costs were greater than litigation costs, the predictions for the data are the same under either assumption.

half $(1/(1-0.6))$ times higher. By displaying results for *(C-S)/J* equal to ± 0.5 as well as

± 0.2, Table 1 already illustrates the effect of multiplying *(C-S)/J* by 2.5. Predictions 1

through 3 hold equally well for ± 0.5 as for ± 0.2.[37] Nevertheless, litigation is less likely

for 0.5 than for 0.2, and more likely for –0.5 than for -0.2. This analysis leads to a fifth

prediction:

> *Prediction 5. If litigation costs are higher than settlement costs (i.e. if (C-S)/J>0), then the proportion of litigated cases should go down as parties learn that settlement may be disrespected. Conversely, if litigation costs are lower than settlement costs (i.e. if (C-S)/J<0), then the proportion of litigated cases should go up as parties learn that settlement may be disrespected.*

The fact that judges might send a defendant to trial in spite of settlement should

have little direct effect on a prosecutor's decision to drop a case. Perhaps the increased

chance of settlement (when litigation is more expensive than settlement) might persuade

some prosecutors to continue prosecution of cases they might otherwise have dropped.

Conversely, the increased chance of litigation (when settlement is more expensive than

litigation), might induce some prosecutors to drop some cases they might otherwise have

continued. More importantly, however, at the same time that judges began sending

settled cases to trial, they also began sending dropped cases to trial. This could have

made dropping a case more attractive, as a prosecutor could still get the satisfaction of

conviction without the trouble of continued prosecution. On the other hand, dropping a

case involved forgoing the possibility of settlement and thus would have been an

unattractive option for all but the weakest cases. As a result, Prediction 4 – that innocent

---

[37] When *(C-S)/J* gets closer to one or negative one, the predictions begin to change, although I do not yet fully understand why.   Nevertheless, because there were few or no lawyers in these cases, both litigation costs and settlement costs were relatively low, so it is not necessary to consider such values of *(C-S)/J.*

verdicts should predominate among the dropped cases – should still hold even when prosecutors anticipate that judges will send settling defendants to trial.

4. Empirical Results

The data used in this article were originally gathered in order to calculate and explain the changing frequency with which appeals were brought in the period 1194-1294.[38] The data set contains over a thousand appeals from select districts in fourteen English counties, ranging from Kent in the south to Yorkshire in the north. The districts were chosen because a larger percentage of their records survive. Some of the cases have been printed and translated, while others exist only in Latin, parchment manuscripts stored in the Public Record Office in London. Assault (including beating, wounding and mayhem) was the most common crime, representing about thirty-nine percent of all appeals, but homicide (27%), theft (12%, including simple larceny, burglary and robbery), and rape (10 %) were also common. In twelve percent, the crime was rare (such as arson) or unrecorded. Most prosecutors were male, but a significant minority (36%) were female.[39]

The cases from the period 1239-1263 are the most useful to test selection theories, because, as mentioned above, they involve defendants caught by surprise by the new treatment of settled cases and defendants who settled before the new treatment of settled cases became routine. The precise ending point of 1263 was chosen, because that is the

---

[38] Daniel Klerman, "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1 (2001). This article contains a detailed discussion of the data and summary statistics.

[39] The large number of female prosecutors is discussed in Daniel Klerman, "Women Prosecutors in Thirteenth-Century England," 14 Yale J. L. & Hum. 271 (2002).

year an "eyre" (judicial circuit of the countryside) ended. The previous one had ended in

1258 and the next one did not end until 1277. Only cases from after 1239 are analyzed,

because only such cases usually contain jury verdicts even in non-litigated cases.

Nevertheless, even after 1239, there are many settled cases without jury verdicts. In

addition, there are a number of non-prosecuted cases for which jurors did not report

whether the case had been settled or dropped. The implications of the incomplete data

are discussed in section 5.A.

The table below summarizes the jury verdicts in those cases for which there are

data both on the defendant's guilt and on whether non-prosecuted cases were dropped or

settled:

Table 2. Jury Verdicts in Private Criminal Prosecutions, 1239-1263

	Number	% Innocent	% Partially guilty	% Guilty as charged
			Prediction 4	
Dropped (i.e. neither prosecuted nor settled)	57	74%	5%	21%
Settled	60	Prediction 1 { 10%	Prediction 2 { 8%	Prediction 3 { 82%
Litigated	123	33%	30%	37%
Total	240	37%	19%	45%

The table is consistent with the predictions of the previous section. On the most

basic level, the data strongly suggest that there is a selection process at work. The

litigated cases are very different from both the settled and the dropped cases in percent

guilty as charged, percent partially guilty, and percent innocent. The table is also

remarkably consistent with the first four predictions laid out in section 3. (1) Innocent verdicts are more common among the litigated cases than among settled cases. This difference is statistically significant with a p-value of 0.001. (2) Partial guilt verdicts are much more common among litigated than settled cases, and the difference is statistically significant with a p-value of 0.001. (3) Guilty-as-charged verdicts are much more common in settled cases than in litigated cases, and the difference is statistically significant with a p-value of less than 0.001. (4) Innocent verdicts constitute well over fifty-percent of the dropped cases, and the difference between the actual percentage (seventy-four percent) and fifty-percent is statistically significant with a p-value of less than 0.001.

It should be noted, however, that the percentage of litigated cases is very high. Of the cases in the table, fifty-one percent (123/240) were litigated. Arguably, the dropped cases should be excluded from the denominator, in which case, sixty-seven percent (123/183) were litigated. As will be discussed further below, when one takes into account the incomplete data, the true percent litigated is probably forty-seven percent. Nevertheless, no matter how the proportion is calculated, it is very high. These high litigation rates are not consistent with the predictions based on the modified Priest-Klein model, if one makes the normal assumption that litigation is more expensive than settlement (i.e. if *(C-S)/J* is positive). On the other hand, if settlement were more expensive than litigation (i.e. if *(C-S)/J* were negative), then litigation rates are in the range of values predicted by the simulations in Table 1. As discussed above in section 3, settlement was plausibly more expensive than litigation, because of the low pecuniary costs of litigation and the possibility that settlement would result in fines.

While the possibility that settlement was more expensive than litigation can account for the high litigation rate, it is worth considering some reasons that litigation rates might have been high, even if litigation were more expensive. One reason might be that Priest and Klein assume that parties will settle whenever it is mutually advantageous for them to do so. To the extent that negotiations may break down because of strategic behavior by either party or because of other factors, one would expect more litigated cases. In addition, to the extent that some defendants were poor and lacked the ability to pay substantial settlements or fines, even a moderate prosecutorial desire for retribution might cause prosecutors to litigate rather than settle, thus increasing the percentage of litigated cases. Both of these explanations, however, would also affect the composition of the litigated cases, causing them to look more like a random sample of all disputed cases.

So far this section has analyzed cases from the entire period 1239-1263. As discussed in sections 2 and 3, however, it is also useful to break down the data into two subperiods: (A) 1239-1249, when parties were probably surprised by judicial disrespect for settlements, and (b) 1250-1263, when parties could have anticipated that there was a good chance that judges would send defendants to trial even if they had settled. Tables 3 and 4 replicate Table 2 for these two subperiods.

Table 3. Jury Verdicts in Private Criminal Prosecutions, 1239-1249

	Number	% Innocent	% Partially guilty	% Guilty as charged
Dropped (i.e. neither prosecuted nor settled)	31	65%	6%	29%
Settled	38	11%	3%	87%
Litigated	61	26%	30%	44%
Total	130	31%	16%	53%

Table 4. Jury Verdicts in Private Criminal Prosecutions, 1250-1263

	Number	% Innocent	% Partially guilty	% Guilty as charged
Dropped (i.e. neither prosecuted nor settled)	26	85%	4%	12%
Settled	22	9%	18%	73%
Litigated	62	39%	31%	31%
Total	110	44%	22%	35%

The discussion at the end of section 3 suggested that predictions 1 though 4 should be equally valid for both subperiods. Indeed, the data for both subperiods is consistent with the four predictions: (1) innocent verdicts and (2) partial guilt verdicts are more common in litigated than in settled cases; (3) guilty-as-charged verdicts are more common in settled than in litigated cases; and (4) innocent verdicts predominate in dropped cases. Five of the eight results are statistically significant at the 0.05 level. Results for prediction 1 for 1239-1249 and prediction 4 for 1250-1263 are marginally

statistically significant (p-values of 0.07 and 0.10), and the result for prediction 2 for 1250-1263 is not statistically significant (p-value of 0.24). This reduction in statistical significance is primarily due to the fact that dividing the data set into two subperiods results in smaller samples.

Prediction 5 suggests that the percent litigated should change when parties can anticipate judicial disrespect for settlement. The sign of that change, however, depends on whether settlement is more expensive than litigation, or vice versa. As discussed above, the generally high levels of litigation suggest that it is more plausible to assume that settlement was more expensive than litigation. As a result, Prediction 5 suggests that litigation rates should be higher in the period 1250-1263 than in 1239-1249. The data are consistent with this prediction. Of the cases in the Table 3 and 4, forty-seven percent were litigated in the period 1239-1249, while fifty-six percent were litigated in the period 1250-1263. If, as seems most sensible, the dropped cases are excluded from the denominator, the increase is even larger, from sixty-two percent to seventy-four percent. As will be discussed further below, when one takes into account the incomplete data, the true percent litigated is somewhat lower for both subperiods. Nevertheless, the revised figures show a similar increase from forty-two to fifty-two percent litigated. None of these three increases, however, is statistically significant at the 0.05 level, although the latter two are marginally significant (p-values of 0.08 and 0.06).

5. Limitations and Extensions

A. Effect of cases lacking information on guilt and/or settlement

As noted in section 2, although juries after 1239 usually reported both the defendant's guilt and whether non-prosecuted cases had been settled or simply dropped, they did not always do so. Table 2 above analyzed only the cases in which jurors did report this information and thus excluded all cases without a verdict and all non-prosecuted cases without a jury report on settlement. Unfortunately, the number of excluded cases is large and could affect the conclusions derived from Table 2. Table 5 helps to evaluate the effect of the excluded data by describing the larger data set of 1239-1263 cases, of which the data in Table 2 is just a subset:

Table 5. Analysis of Cases Included and Excluded from Table 2, 1239-1263

	Number in whole data set	Number excluded from Table 2	% excluded from Table 2	Reason for exclusion from Table 2
Dropped	58	1	2%	Lack of jury verdict
Settled	128	68	53%	Lack of jury verdict
Litigated	128	5	4%	Lack of jury verdict
Uncertain (i.e. either dropped or settled)	100	100	100%	Lack of jury report on settlement. Also, lack of verdict for 43 (43%)
Total	414	174	42%	

The table suggests that the figures in Table 2 are affected by two kinds of problems: 1) lack of jury verdicts in cases that can be definitively categorized as dropped, settled or litigated, and 2) a large number of uncertain cases which might be either

dropped or settled. Neither of these problems significantly affects the litigated cases, so

the data in Table 2 regarding them are pretty solid. The settled cases, however, are

affected by both problems: there are no jury verdicts for over fifty percent of them, and

there are almost as many uncertain cases (which might be settled) as there are cases

which can be confidently categorized as settled.   The first problem (lack of jury verdicts)

is the more difficult problem, because the reason for the missing verdicts is unknown. It

is possible that verdicts of "innocent" were under-reported, because they generally

produced no fines. Documenting fines to-be-collected was an important reason for

keeping written records, so cases with no revenue implications might have been under-

reported. If all defendants in settled cases for which there is no recorded verdict were

innocent, settled cases would not be very different from litigated ones.[40] While this

under-reporting of not-guilty verdicts in settled cases is possible, it seems implausible,

because such reasoning is inconsistent with the rigorous reporting of innocent verdicts in

the dropped cases. More plausibly, jury verdicts were sometimes not recorded, because

judges did not ask for them, perhaps because the defendant was already to be fined for

having settled. If so, then the cases which lack verdicts may be approximately a random

sample of all settled cases, and the percentages in Table 2 would be largely accurate.

Because judges sat only in counties where they did not live, they would have had little

knowledge of the cases and thus could not have selectively sent defendants likely to be

guilty to trial.

The problem presented by the uncertain cases, which might be either dropped or

settled, is made somewhat more tractable by the fact that jury verdicts are available for

---

[40] Under these assumptions, defendants would have been in innocent in 58% percent of settled

most (57%) of them. Table 6 presents those verdicts and compares them to verdicts in

the cases known to have been dropped or settled.

Table 6. Jury Verdicts in Dropped, Settled, and Uncertain Cases, 1239-1263

	Number	% Innocent	% Partially guilty	% Guilty as charged
Dropped	57	74%	5%	21%
Settled	64	10%	8%	82%
Uncertain: either dropped or settled	57	61%	12%	26%

By comparing the uncertain cases to the dropped and settled cases, it is apparent

that the uncertain cases cannot simply be a random sample of all settled or dropped cases,

because, if they were, the percent guilty-as-charged would be much higher, and the

percent innocent much lower. For example, if half of the uncertain cases were dropped

and half were settled, and if these settled and dropped cases had the same percent guilty-

as-charged as in Table 2, then the overall percentage guilty-as-charged among the

uncertain cases would be fifty-two percent rather than the twenty-six percent actually

observed. More plausibly, if one assumed that the proportion of settled cases among the

uncertain cases matched the percentage of settled cases among those known to have been

settled or dropped (69%=128/(58+128)), then the percent guilty-as-charged among the

uncertain cases would be even higher. Thus, it must be the case either that dropped cases

are under-represented among the cases known to have settled or dropped, or that guilty-

cases, partially guilty in 4%, and guilty-as-charged in 38%.

as-charged verdicts are over-represented among the cases known to have settled, or both.[41]

It is difficult to imagine any reason guilty-as-charged verdicts would be over-represented among the cases known to have settled. Any plausible reason for over-reporting guilty verdicts among settled cases would also apply to dropped cases, and, from Table 2, it is apparent that guilty verdicts were not substantially over-reported among the dropped cases, because the proportion is already so low (21%). Thus, the most plausible assumption is that dropped cases are under-represented among the cases known to have settled or dropped. Perhaps clerks tended not to record the absence of settlement, because it did not result in revenue. If dropped cases are under-represented among cases known to have settled or dropped, dropped cases should form a disproportionate number of the uncertain cases. This would largely explain the verdicts among the uncertain cases. For example, if dropped cases were eighty-five percent of the uncertain cases, then the verdicts among the uncertain cases would be very close to what they are in Table 6.[42] Of course, if that were true, the percentages in Table 2 would remain valid with only minor modifications.

_____

[41] Over-representation of guilty-as-charged verdicts among the cases known to have dropped could not account for the overall low percentage of guilty-as-charged verdicts among the uncertain cases, because the percent guilty-as-charged among the known dropped cases is already so low (21%). Under- and over-representation, are, of course, relative. The implicit comparison here is between cases whose settlement status is know and those whose settlement status is uncertain. The last sentence in the text could, with equal validity, be rewritten to say: "Thus, it must be the case either that dropped cases are over-represented among the uncertain cases or that guilty-as-charged verdicts are under-represented among the uncertain cases which settled."

[42] The percent guilty-as-charged would be 30% = (85% x 21%) + (15% x 82%), the percent partially guilty would be 5%, and the percent innocent would be 64%. These percentages are off by 4%, 7%, and 3% respectively, and none of the discrepancies are statistically significant at the 0.05 level. The difference in the proportion partially guilty, however, has a p-value of 0.09.

In addition, if one assumed (as in the previous paragraph) that eighty-five percent of the uncertain cases were dropped, then the proportion of non-dropped cases (i.e. litigated or settled cases) which were litigated would be forty-seven percent.[43] Even so, the percent litigated is higher than predicted by Priest and Klein, if litigation is more expensive than settlement. Of course, if one did not assume that dropped cases were so over-represented among the uncertain cases, the percent litigated would be lower. Nevertheless, even if one assumed that all the uncertain cases were settled, the percent litigated would be thirty-six percent, which is still relatively high.

So far this analysis has grouped together cases from the entire period 1239-1263. Tables 5A, 5B, 6A, and 6B in the Appendix replicate Tables 5 and 6 for the subperiods 1239-1249 and 1250-1263. As is readily apparent, the data in the two periods are sufficiently similar that the analysis for the entire period applies equally well to the subperiods. The only relevant difference is that, when the percent settled is adjusted to take into account the uncertain cases, I estimate that forty-two percent of cases were litigated in the period 1239-1249, while fifty-two percent were litigated in the period with 1250-1263.[44] As noted at the end of section 4, this is consistent with prediction 5.

B. Jury inferences of guilty from settlement

An alternative way of explaining the fact that juries found settling defendants to be disproportionately guilty is that jurors may have inferred guilt from settlement. While

---

[43] The total number of settled cases would be $143 = 128 + (15\% \times 100)$. As a result, the proportion of litigated cases would be $47\% = 128/(143+128)$.

[44] To calculate these percentages, I first estimated the proportion of uncertain cases which settled. Following the method used for the entire period 1239-1263, I estimated that eighty-five percent of the uncertain cases from 1239-1249 and seventy-one percent of the uncertain cases from 1250-1263 were settled. These percentages were then combined with the numbers in Tables 5A and 5B in the Appendix to

this is a plausible argument, there are three reasons to believe that it does not undermine the selection-theory explanation.

First, it would be rational for jurors to infer guilt from settlement only if settling defendants were more likely to be guilty. But it would be rational to believe that settling defendants were more likely to be guilty, only if guilty defendants were more likely to settle, which itself is evidence of a selection effect. Thus, while jury inferences of guilt from settlement might reinforce selection effects, it is difficult to see why jurors would make such inferences if the selection effect were not, in fact, operating.

Second, jury inferences of guilt from settlement cannot explain why partial guilt verdicts are over-represented among the litigated cases.

Third, most historians would have predicted that jurors would have struggled (or perjured themselves) to find settling defendants innocent. The most influential recent historian of the jury, Thomas Green, argues that royal justice sought to impose strict ideas of criminal liability on the population and that jurors responded by nullification.[45] That is, when faced with law that would require them to convict someone they thought not deserving of criminal sanctions, jurors rendered "not guilty" or "partially guilty" verdicts. Royal justice was clearly stricter with regard to settling defendants than the population as a whole, so it would be expected that jurors would have used their nullification power to acquit defendants who had settled, even if such defendants were, in fact, guilty.[46] Thus,

---

calculate the percentage litigated. As with the corresponding figure for the entire period 1239-1263, dropped cases were excluded from the denominator.

[45] Thomas Andrew Green, *Verdict According to Conscience: Perspectives on the English Trial Jury, 1200-1800.*

[46] See, e.g., *id* at 9-10, 58; Roger D. Groot, "the Jury in Private Criminal Prosecutions Before 1215," 27 Am. J. Legal Hist. 113, 133 (1983).

to the extent that jurors slanted their verdicts in settled cases, it seems much more likely that they biased their verdicts towards more acquittals than that they increased the number of convictions by inferring guilt from settlement.

C. Asymmetric information

As one of law and economics' most influential theories, the Priest-Klein analysis has been the subject of intensive scrutiny. The strongest critique has come from Hylton and Shavell, who criticize Priest and Klein's reliance on litigation models which assume that parties do not systematically differ in their knowledge of relevant facts.[47] Instead, Hylton and Shavell employ litigation models based on asymmetric information to show that "any frequency of plaintiff victory at trial is possible" even when the stakes are symmetric and the parties agree on damages. Nevertheless, even these models generate some testable predictions. For example, the probability of victory at trial for the party with the informational advantage should be higher than the probability of victory for that party among settled cases.[48]

It is not clear how asymmetric information models would apply to thirteenth-century criminal disputes. It would be difficult to generalize as to which party had better information. Nevertheless, if one were to attempt to apply these models, there are two reasons that defendants in thirteenth-century private prosecutions might have had an informational advantage. First, defendants in private prosecutions tended to be of higher status than prosecutors. To the degree that many prosecutors were motivated by

---

[47] Keith Hylton, "Asymmetric Information and the Selection of Disputes for Litigation," 22 J. Legal Stud. 187 (1993); Steven Shavell, "Any Frequency of Plaintiff Victory at Trial is Possible," 25 J. Legal Stud. 493 (1996). For another kind of critique, see Donald Wittman, "Is the Selection of Cases for Trial Biased?" 14 J. Legal Stud. 185 (1985).

[48] See Shavell, 25 J. Legal Stud. 497, 498.

settlement, it didn't make sense to prosecute poor defendants. In addition, private prosecutions were often used by tenants and small landowners in their disputes against their feudal lords and more powerful neighbors. The more powerful were less likely to use the criminal process against their social inferiors, because they had other means of social control (physical violence, social sanctions, etc.), whereas law was often the only way those farther down on the social scale could control their more powerful neighbors.[49] Since jurors were generally chosen from among those of high social standing, it is likely that defendants, who were also of high social status, would have had a better understanding of jurors' values and ways of evaluating evidence. As a result, defendants were probably better able to predict likely jury verdicts than prosecutors. While legal advice might have helped prosecutors remedy their informational disadvantage, lawyers were seldom employed by either side.

In addition, the fact that defendants were usually eyewitnesses, whereas prosecutors often were not, meant that defendants, on average, had better information. If a defendant were guilty, he would obviously have had eyewitness information about the crime. If he were not guilty, he would at least have had first-hand information about whatever he was doing at the time. Assault and rape victims would have possessed comparable information, but prosecutors in homicide and many theft cases would usually have lacked any first-hand information. Of course, better information on the crime itself would not always translate into better prediction of likely jury verdicts. Nevertheless, this was the period of the "self-informing" jury, when parties and witnesses could approach and inform jurors of the facts privately and before trial. As a result, since

---

[49] This is discussed more fully in Daniel Klerman, "Settlement and the Decline of Private

defendants might have been able to provide jurors with more convincing information than prosecutors, and since only defendants would have known what they had told the jury, defendants might have been better able to predict jury verdicts. Since, for these two reasons (social status and eyewitness information), defendants might have had an informational advantage, asymmetric information models predict that defendants should have been more likely to prevail in litigated than in settled cases. This prediction is similar to the prediction of the Priest-Klein model under the assumption that the defendant had more at stake.

Of course, like the Priest-Klein model, the Hylton and Shavell models were developed for cases with two verdicts (liable or not liable) rather than three (innocent, partially guilty, guilty-as-charged). Adapting the Shavell model to take into account partial verdicts, however, is relatively easy.[50] Suppose there are two kinds of defendants, Low (L) and High (H).[51] If they litigate, Low defendants are found guilty as charged with probability p_L^g and are found partially guilty with probability p_L^p. Similarly, High defendants are found guilty as charged with probability p_H^g and are found partially guilty with probability p_H^p. J is the sanction imposed on guilty-as-charged defendants, αJ is the sanction imposed on partially guilty defendants, $\alpha < 1$, and C_Δ and S_Δ are the defendant's litigation and settlement costs respectively, as in section 3. High defendants are categorized as high, because the expected value of the sanction imposed on them is higher. Thus, $(p_H^g + \alpha p_H^p)J > (p_L^g + \alpha p_L^p)J$. The prosecutor will demand one of two

Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1 (2001).

[50] I have not tried adapting the Hylton model to take into account partial verdicts, because the Shavell model is basically an improved version of the Hylton model.

amounts: either $(p_L^g + \alpha p_L^p)J + C_\Delta - S_\Delta$, which will result in settlements with both kinds

of defendants, or $(p_H^g + \alpha p_H^p)J + C_\Delta - S_\Delta$, which will result in settlement with the Highs

and litigation with the Lows. Only the latter amount is interesting for the analysis of

selection, so we can assume that Highs will settle and the Lows will litigate. The

definition of Highs and Lows -- $(p_H^g + \alpha p_H^p)J > (p_L^g + \alpha p_L^p)J$ -- implies that

$p_H^g + \alpha p_H^p > p_L^g + \alpha p_L^p$. The asymmetric information model thus leads to the following

prediction:

> *Asymmetric Information Prediction. The percentage of guilty-as-charged*
> *verdicts plus α times the percentage of partial guilty verdicts will be larger*
> *among settled cases than among litigated cases.*

This prediction is easily satisfied by the data in Table 2. The percentage of guilty-as-

charged verdicts plus α times the percentage of partial guilty verdicts is larger among

settled cases (85% = 82% + [0.33 x 8%]) than among the litigated cases (47% = 37% +

[0.33 x 30%]). This difference is statistically significant with a p-value of less than

0.001.

The asymmetric information model does not make any predictions about the

percentage of cases that will be litigated. Since any proportion of the defendants can be

Highs, any proportion of the cases can result in litigation. The high litigation rates

observed in the data are thus less puzzling under the asymmetric information model.

The asymmetric information model does not address the issue of dropped cases,

but it can easily be adjusted to take them into account. Suppose prosecutors initiate

prosecution based on limited information and then learn some information about the case,

---

[51] As in Shavell's article, it is relatively easy to generalize from two types to many types.

but not enough to make them as informed as the defendants. One would then predict that, after acquiring their limited information, prosecutors would drop the least promising cases and make the settlement offers described above to the remaining defendants. As with the Priest/Klein model, the modified asymmetric information model, thus predicts that the dropped cases should consist primarily of innocent defendants.

It is very easy to modify the asymmetric information model to take into account the possibility that judges will send defendants to trial even if they have settled. The prosecutor's demands are lower— $(1-\lambda)(p_L^g + \alpha p_L^p)J + C_\Delta - S_\Delta$ to settle with both Highs and Lows, or $(1-\lambda)(p_H^g + \alpha p_H^p)J + C_\Delta - S_\Delta$ to settle only with the Highs—but the Asymmetric Information Prediction remains the same. The data for the subperiods in tables 3 and 4 are consistent with the prediction, and both differences are statistically significant at the 0.01 level.

D. Asymmetric stakes revisited

As discussed in sections 3 and 4, the data are consistent with the Priest-Klein selection model only if the defendants' stakes were larger than the prosecutors'. While this assumption of asymmetric stakes has historical support, it is interesting to note that even if it were false, Priest and Klein's most fundamental insight – that litigated cases differ systematically from settled cases – would still be vindicated by the data, because, as pointed out in the beginning of section 4, the litigated cases are remarkably different from the settled cases. In addition, while, the data would no longer be consistent with Priest and Klein's explanation for the difference between litigated and settled cases, the data could be explained by the asymmetric information model discussed in the prior

subsection, because that model does not assume anything about the symmetry or

asymmetry of the stakes.

E. Differences between crimes

Because aggregating over different offenses may produced spurious results, it is

important to break down the results by crime. Table 7 summarizes the analysis by

offense:

Table 7. Analysis by Offense, 1239-1263

	Prediction 1: More innocent in litigated cases	Prediction 2: More partially guilty in litigated cases	Prediction 3: More guilty-as-charged in settled cases	Prediction 4: Mostly innocent in dropped cases	Asymmetric Information Model Prediction
Assault	Consistent*	Consistent*	Consistent*	Consistent*	Consistent*
Homicide	Cannot test, because no settled homicide cases			Consistent	Cannot test[†]
Rape	Consistent	Consistent	Consistent	Consistent	Consistent
Theft	Consistent	Not consistent	Consistent	Consistent	Consistent
Other	Consistent*	Not consistent	Not consistent	Consistent	Consistent
All	Consistent*	Consistent*	Consistent*	Consistent*	Consistent*

"Consistent" means that the data for the offense conform to the prediction, although the result
   may not be statistically significant.
"*" means statistically significant at the 0.05 level.
"Assault" includes beating, wounding and mayhem.
"Theft" includes simple larceny, robbery, and burglary.
"Other" includes uncommon offenses, such as arson, and cases for which the offense is not
   ascertainable.
[†] Like Predictions 1, 2 and 3, the asymmetric information model prediction cannot be tested,
   because there are no settled homicide cases in the data set.
I did not try to test any of the predictions by both offense and subperiod, because the sample sizes
   would have been too small. For the same reason, I did not try to test Prediction 5 by offense.

Perhaps the most surprising result is that it is impossible to test most of the

predictions on homicide data by comparing settled and litigated cases, because there are

no settled cases in the data set. This anomaly may have been caused by the fact that

judges started routinely taking verdicts in settled homicide cases in 1218. As a result, by

1239, the starting date for the data set used in this article, defendants may already have

been unwilling to settle. Extending the analysis back before 1239 is not possible, because there are too few pre-1239 homicide cases with reports on whether the parties had settled for meaningful analysis. The inability to analyze homicide cases is especially unfortunate, because capital punishment was most likely in homicide cases, so the stakes would have been especially asymmetric. Thus, one would have predicted an especially stark difference between the settled and litigated cases.

Fortunately, analysis of the of the remaining four offense categories (assault, rape, theft, and other) is possible. For these offenses, there are twenty predictions—five per offense. The data are consistent with seventeen of the twenty predictions, although in only six are the results statistically significant. The lack of statistical significance is largely explained by the relatively small sample size once the data is broken down by crime. For example, the rape data would seem remarkable consistent with Prediction 3: two-thirds of the settled cases resulted in guilty verdicts, while only one-third of the litigated cases resulted in guilty verdicts. Nevertheless, since there are only six settled rape cases and six litigated rape cases in the data set, this difference is not statistically significant. I can think of no reason why the data were not consistent with three predictions, except for random variation given the small number of cases involved. For example, the "other" cases, which provide two of the three "not consistent" cells, are represented by only six settled cases and eleven litigated cases.

The fact that the data are almost entirely consistent with the predictions is, paradoxically, somewhat disappointing. If the data had been consistent with the Priest-Klein predictions, but not the asymmetric information prediction, or vice versa, the data might have been able to shed some light on the relative merits of the two theories.

Similarly, if the asymmetric information predictions had failed in assault and rape cases, this might have lent support to the asymmetric information model, because victim/prosecutor and defendant might have had equally good first-hand information about these crimes.[52]

One aspect of the theft data is particularly interesting. Fifty-six percent of litigated theft cases resulted in guilty verdicts. If the Priest-Klein theory were tested in the usual way, by looking only at the litigated cases, the theft data would be considered inconsistent with the theory. Under the usual interpretation, when the stakes are larger for the defendant, there should be less than fifty percent pro-plaintiff verdicts. Nevertheless, because of the ability to compare litigated to settled cases in the data set, it can be seen that the data really are consistent with the theory, because the percentage guilty among the settled cases (seventy-five percent) was even higher. Thus, there clearly was selection in exactly the way Priest and Klein predicted, but this effect is only visible because one can observe the settled cases.

F. Pre-appeal settlement

Although the data set used in the article is unique in containing jury verdicts even in settled and dropped cases, it is possible that one category of settled cases was not recorded. If the parties settled before the prosecutor brought the appeal in county court, the case would not appear in the surviving records. Fortunately, the number of such settlements was probably low for several reasons.[53] First, as noted in section 2, potential prosecutors had to bring their cases at the next county court. Because the county court

---

[52] See *supra*, p. __.

[53] Daniel Klerman, "Settlement and the Decline of Private Prosecution in Thirteenth-Century England," 19 Law & Hist. Rev. 1 (2001).

met every four weeks, potential prosecutors had at most a month to settle and potentially only a few days. Such quick settlements might have been difficult, especially in emotionally charged cases, such as homicide, rape and assault. In addition, during the period 1239-1249, the parties had only scant reason to think that settling defendants would be sent to trial. As a result, they had little incentive to settle so early.

Even if there were many potential appeals which settled before initiation in county court, there is no reason to think that such cases would be significantly different from those which settled after appeal in county court. In order to disturb the empirical findings of this article, the cases settled before appeal would have had to have consisted overwhelmingly of innocent defendants. There is no reason to think this to have been likely. In fact, once parties began to anticipate that settling defendants might be sent to trial, one would have expected that guilty defendants would have disproportionately settled before initiation of the appeal in county court, because doing so was the only certain way of avoiding trial on an appeal.

G. Alternative explanations

Although the data suggest that there was a strong selection effect and that the selection effect operated in a manner consistent with the Priest-Klein and asymmetric information theories, statistical analysis alone cannot show that these theories actually explain the data. Other explanations are certainly possible, although the author has not seen any in the literature. Nevertheless, conversations with historians and non-economists have suggested at least two other plausible explanations.[54] It is possible that guilty defendants felt guilty and settled in order to assuage their guilt, rather than because

they rationally calculated that they could settle for less than the expected value of the court-imposed sanction. In a similar vein, social pressure might also have persuaded guilty defendants disproportionately to settle.

Guilt and social pressure, however, are less clear in their predictions about those guilty of less than they were charged, i.e. those for whom juries rendered partial guilt verdicts. One might have thought that even these persons would feel guilty and thus that they would also disproportionately settle. Similarly, one might have thought that there would be social pressure on them to settle. The data, which show that partial verdicts are disproportionately represented among litigated cases, are not consistent with these predictions. On the other hand, one might predict that those unjustly accused of more serious crimes than they actually committed would feel more outrage than guilt, and thus not be willing to settle. Or, perhaps social pressure in such cases might be low, because the community would disapprove of the prosecutor's overcharging. If so, the data is consistent with these theories based on guilt and social pressure.

Although these theories would seem rather different from the Priest-Klein and asymmetric information models, they are closer than they might appear. At the most basic level, even the guilt and social pressure theories are selection theories. They too predict that litigated cases will differ systematically from settled cases. In addition, they could be modeled in almost exactly the same way. One would merely need to reinterpret J_Δ to be the defendant's valuation of guilt or social disapproval, rather than his valuation of the court imposed sanction.

---

6. Conclusion

Thirteenth-century private criminal prosecutions present a data set uniquely suitable for testing theories about the selection of disputes for litigation. Because thirteenth-century private criminal prosecutions often resulted in jury verdicts even in settled cases, this data set allows direct comparison of settled and litigated cases. Other tests of selection theories have had to rely on data about litigated cases exclusively. The analysis of this thirteenth-century data is consistent with the theoretical predictions both of Priest and Klein's original model and of more recent asymmetric information models.

---

alternative explanations.

Appendix

This Appendix provides support for the assertions in last paragraph section 5.A about the subperiods 1239-1249 and 1250-1263. Tables 5A and 5B were calculated in the same was as Table 5, and Tables 6A and 6B were calculated in the same way as Table 6.

Table 5A. Analysis of Cases Included and Excluded from Table 3, 1239-1249

	Number in whole data set	Number excluded from Table 2	% excluded from Table 2	Reason for exclusion from Table 2
Dropped	31	0	0%	Lack of jury verdict
Settled	80	42	53%	Lack of jury verdict
Litigated	64	3	5%	Lack of jury verdict
Uncertain (i.e. either dropped or settled)	64	64	100%	Lack of jury report on settlement. Also, lack of verdict for 29 (45%)
Total	239	109	46%	

Table 5B. Analysis of Cases Included and Excluded from Table 4, 1250-1263

	Number in whole data set	Number excluded from Table 2	% excluded from Table 2	Reason for exclusion from Table 2
Dropped	27	1	4%	Lack of jury verdict
Settled	48	26	54%	Lack of jury verdict
Litigated	64	2	3%	Lack of jury verdict
Uncertain (i.e. either dropped or settled)	36	36	100%	Lack of jury report on settlement. Also, lack of verdict for 14 (39%)
Total	175	65	37%	

Table 6A. Jury Verdicts in Dropped, Settled, and Uncertain Cases, 1239-1249

	Number	% Innocent	% Partially guilty	% Guilty as charged
Dropped	31	65%	6%	29%
Settled	38	11%	3%	87%
Uncertain: either dropped or settled	35	60%	17%	23%

Table 6B. Jury Verdicts in Dropped, Settled, and Uncertain Cases, 1250-1263

	Number	% Innocent	% Partially guilty	% Guilty as charged
Dropped	26	85%	4%	12%
Settled	22	9%	18%	73%
Uncertain: either dropped or settled	22	64%	5%	32%