10 Fee shifting

Avery Wiener Katz and Chris William Sanchirico

1. Introduction

In most Western legal systems, a party who prevails in litigation is generally entitled to indemnification from the losing party for at least part of his or her economic costs of prosecuting the lawsuit. The amount of litigation expenditures that can be recovered, however, varies substantially both among and within individual regimes. In the United States, the predominant rule awards a prevailing litigant what are officially termed ‘costs’ – typically defined by statute to include filing fees, court reporter charges, printing, copying, and witness fees, and the like – but does not entitle him or her to recover expenditures on attorneys’ fees, which are usually of far greater magnitude in the case. Consequently, US litigants can bear significant expense even when they are ultimately vindicated on the merits. In the other common-law countries, in contrast, and indeed in most of the rest of the Western world, winning litigants are entitled to recover attorneys’ fees as well as other out-of-pocket costs of litigation.

The substantial increase in expenditures on litigation and dispute resolution in the United States over the last two decades has led both policymakers and scholars to advocate a variety of substantive and procedural reforms in the legal system. The rules for allocating attorneys’ fees in civil litigation have drawn particular attention in this regard, with a number of influential commentators recommending a move in the direction of fuller indemnification – or what in the US is usually called, for historical reasons, the ‘English’ or ‘British’ rule. Such recommendations have had influence on both public and private lawmakers. In the mid 1990s one of the more prominent and widely supported provisions in the Republican Party’s ‘Contract with America’ platform would have adopted a modified form of the English rule for federal cases brought under diversity jurisdiction. In the late 2000s, as part of the ongoing debate over health care reform, several commentators and policymakers have proposed adopting a ‘loser pays’ rule as a way to curb medical malpractice costs.

But the ongoing political debate over litigation costs in the US does not seem to have assimilated the main lesson of the economic literature on the topic – that the effects of cost shifting on the amount and intensity of
litigation are substantially more complicated than a superficial consider-
ation of the matter might suggest. Indeed, the current state of economic
knowledge does not enable us reliably to predict whether a move to fuller
indemnification would raise or lower the total costs of litigation, let alone
whether it would better align those costs with any social benefits they
might generate.

The reason for this agnosticism is straightforward. In short, fee shifting
is too coarse a tool for the multifaceted problem that it is meant to solve.

The problem that fee shifting is meant to solve is, in essence, the exist-
ence of externalities in the litigation decision making. Most supporters of
fee shifting focus on a single negative externality: the costs that plaintiffs
impose on both defendants and the public at large in deciding to bring
suit. But the allocation of legal costs influences a long list of decision
points along the litigation flow, including the decision of how much effort
to expend in preparing for and participating in trial, the choice between
settlement and trial, the decision to initiate a lawsuit, and the choice to
modify one’s out-of-court behavior so as to avoid suit in the first place.
At each of these decision points, an individual litigant’s incentives do not
generally conform to the interests of society as a whole. And such diver-
gences of private and social interest vary both in magnitude and direction.

For example, the plaintiff’s private decision to bring suit is itself the
locus of several opposing externalities. A plaintiff’s decision to sue
imposes an obvious cost on the defendant and on taxpayers, who foot the
bill for public legal institutions. Less obviously, the suit affects litigants
in other filed cases by crowding the courts and delaying the resolution of
other disputes in the system. The suit also affects these and future litigants
by altering the state of legal precedent. Lastly, the suit affects the public at
large – both potential violators and their potential victims – by influencing
the perceived likelihood of sanctions for violating substantive legal duties.

The decision to pursue a lawsuit to trial rather than settling, as well as
the decision to litigate more rather than less intensively, generate analo-
gous external costs and benefits.

Litigation does not, therefore, present a single externality that might be
corrected with a single policy tool. It presents a complex bundle of posi-
tive and negative externalities whose correction requires an array of policy
instruments. Shifting legal fees may indeed repair some externalities but it
will fail to address others, and some it may well exacerbate.

This chapter surveys the effects of legal fee shifting on a variety of deci-
sions arising before and during the litigation process. Section 2 provides
a brief survey of the practical situations in which legal fee shifting does
and does not arise. Section 3 analyzes the effects of indemnification on the
incentives to expend resources in litigated cases. Section 4 examines how
indemnification influences the decisions to bring and to defend against suit, and Section 5 assesses its effects on the choice between settlement and trial. Section 6 addresses the interaction between the allocation of legal fees and the parties’ incentives for efficient primary behavior. Section 7 considers two important variants on simple indemnification: rules that shift costs based on the parties’ settlement negotiations (such as US Federal Rule 68 and the English practice of payment into court), and rules that shift costs based on the margin of victory (such as US Federal Rule 11 and the common-law tort of malicious prosecution). Section 8 reviews the empirical literature on legal cost shifting. Section 9 summarizes the discussion and offers conclusions.1

2. The Practical Extent of Legal Fee Shifting
This chapter does not attempt to survey the law governing fee shifting, either in the US or elsewhere. For a continually updated account of US state and federal law, see Martin (2005). It should be recognized, however, that there are significant areas of US legal practice that do not follow the traditional American rule. Most important among these are the various federal and state statutes that entitle a successful plaintiff, though not a successful defendant, to court-awarded attorneys’ fees as part of a recovery. Similar ‘one-way’ fee-shifting policies have also been established in both federal and state courts through a combination of statutory interpretation and common-law development, though the scope for such interpretations at the federal level was substantially limited by the Supreme Court in Alyeska Pipeline Service Co. v. Wilderness Society. Such provisions and policies, which make up a central part of litigation practice in such fields as civil rights, consumer, and antitrust law, have only some of the effects of the traditional two-way English rule. Second, both federal and state courts have authority to award indemnification to parties who are victimized by abuse of process, though such authority is typically exercised only in response to egregious behavior. Examples include the provisions in Federal Rule of Civil Procedure 11 dealing with frivolous or improper pleadings, and those in Federal Rule of Civil Procedure 37 relating to discovery abuse. US practice also provides litigants with an ‘offer-of-judgment’ procedure under which a defendant can make a settlement offer to the plaintiff which, if rejected and filed with the court, creates a trigger for partial indemnification. Both of these specialized types of provisions

---

1 The scholarly literature on fee shifting is vast and it is not possible to discuss every pertinent contribution. This entry describes the seminal papers on the topic as well as a sizable and representative sample of additional contributions.
Procedural law and economics

– sanctions for abuse and offers of judgment – are discussed separately in
Section 7 below.

Conversely, even in jurisdictions following the majority or ‘English’ rule,
indemnification for legal costs is substantially less than complete. Court-
awarded attorneys’ fees obviously do not compensate for the nonmon-
etary and psychic costs of litigation. Even the monetary amounts awarded,
furthermore, are limited by the judge’s view of what expenditures are
reasonable and, in some jurisdictions (for example, British Columbia), by
statutory schedule. Such judicial and statutory caps can and often do hold
fee awards below the going market rate for legal representation, forcing
winning litigants to pay the difference out of anticipated recoveries or their
own pockets (indeed, Leubsdorf 1984, presents evidence that such court-
imposed price ceilings were responsible for the historical development of
the American rule in the first place). Accordingly, the pure English and
American rules discussed below should be understood as ideal polar cases,
and the differences among actual jurisdictional practices as ones of degree
along a spectrum ranging from lesser to greater indemnification.

Additionally, Donohue (1991b) points out that the American rule is
a default rule rather than a mandatory one, in that parties are generally
free to provide for indemnification through private contract – either at the
time they begin their litigation or, for those disputes arising out of a con-
sensual relationship, in their original agreement. He presents anecdotal
evidence that such ex ante indemnification terms are widespread, though
the provisions he cites seem primarily to be drawn from standardized
form contracts and tend to operate asymmetrically in favor of the draft-
ing party: for instance, apartment leases that indemnify landlords but not
tenants for attorneys’ fees in the event of a dispute over unpaid rent. The
scope for fee shifting in the US, therefore, may be significantly greater than
is ordinarily supposed. Conversely, there is no apparent bar in England or
in the other jurisdictions following the English rule to a partial settlement
or stipulation in which the litigants agree in advance to give up their rights
to indemnification ex post.

Somewhat more problematic, however, are Donohue’s further con-
jectures that such contractual terms are likely to be efficiency-enhancing
and that the pattern of such terms will help reveal whether the English
or American rule is more efficient. To the extent that indemnification is
provided by a one-sided standardized term, there is no guarantee that it
promotes the joint interests of the parties. The nondrafting party may fail
to notice the indemnification provision at all; and if he does notice it, he
may avoid raising it as an issue for fear of revealing himself as someone
who anticipates a dispute. Even when such agreements arise out of arm’s-
length bargaining, furthermore, this does not imply that they are efficient.
As Bernstein (1993) and Shavell (1995) have observed in their respective analyses of alternative dispute resolution, because of the divergence of private and social incentives in litigation, the fact that a particular agreement is in the litigants’ ex post interest does not necessarily mean that it is socially efficient. The fact that the parties have come to litigation in the first place, moreover, casts doubt on the presumption that they are bargaining in a Coasian fashion.

3. The Effect of Fee Shifting on Trial and Pretrial Expenditures

The standard economic theory of litigation, as developed by Landes (1971), Posner (1973), and Gould (1973), models litigating parties as rational actors who seek to maximize their returns from the litigation process. From this perspective, amounts spent on trial preparation can be seen as a type of private investment. An additional hour of legal research or argumentation is profitable, in this view, only if the marginal return, measured by the change in the expected outcome of trial or settlement, outweighs the cost of the attorney’s time. Plaintiffs, accordingly, will choose to spend legal resources up to the point where their expected recoveries, net of expenses, are maximized; defendants will act so as to minimize total payouts. The precise outcome of this contest depends on how the parties react and adjust to each other’s decisions. One simple and natural assumption is that the litigants reach a Nash equilibrium in expenditure; that is, that each takes the other’s expenditure as given when choosing his own. Whatever the nature of the parties’ strategic interaction, however, the parties’ expenditures are determined in equilibrium by a host of economic and technological factors, including the stakes of the case, the marginal cost of legal resources, and the sensitivity of trial outcomes to the parties’ individual efforts. In high-stakes cases in which the outcome is heavily dependent on the parties’ work product, expenditures will be high; in petty cases where the outcome is largely predetermined by legal precedent, expenditures will be low.

As Braeutigam, Owen, and Panzar (1984) first proved, and Katz (1987) subsequently explained, it follows from the standard model that fee shifting encourages greater expenditure in litigated cases. The reasons are twofold.

First, fee shifting increases the stakes of the case by making legal expenditures part of the potential damages. Second, fee shifting lowers the expected marginal cost of legal expenditure: each party, when deciding whether to purchase an additional unit of legal services, will discount its cost by the probability with which she expects to win and to be reimbursed by her opponent.

More formally: if we let $p$ denote the probability of liability, $A$ the
amount awarded if the plaintiff wins, and \( x \) and \( y \) the amounts spent by the plaintiff and defendant respectively, then under the American rule a plaintiff will expect to recover \( p(x, y) A(x, y) - x \). Assuming risk-neutrality for the sake of simplicity, it follows she will choose \( x \) to satisfy the first-order condition, \( p_x A + pA_x = 1 \). The defendant, conversely, expects to pay out \( pA + y \), and will select \( y \) to satisfy his first-order condition, \( p_y A + pA_y = -1 \).

Under the English rule, in contrast, the plaintiff’s expected recovery is \( pA - (1-p)(x+y) \); so her first-order condition is \( p_x (A + x + y) + pA_x = 1 - p \). Similarly, the defendant’s expected payout is \( p(A + x + y) \), and his first-order condition is \( p_y (A + x + y) + pA_y = -p \).

In all of these equations, the left-hand side represents the marginal private benefit of expenditure, and the right-hand side its marginal cost. Inspection of the equations reveals that the marginal private cost of legal expenditure is lower for both parties under the English rule. If the parties’ expenditure affects the probability of liability (that is, if \( p_x \) and \( p_y \) are positive), the marginal private benefit is also higher; if expenditure affects only the amount awarded, and not also the probability of liability, marginal benefit is unchanged. Other things being equal, therefore, the English rule makes expenditure more attractive.

It should be noted that the marginal-cost effect depends not on the actual probability of liability, but on its perceived probability. It follows that the increase in expenditure under the English rule will be greater the more optimistic are the litigants. In the extreme, parties who regard themselves as very likely to win will perceive litigation as virtually costless and will increase their expenditures accordingly. To the extent that such efforts increase the probability of prevailing, therefore, such optimism will be partially self-fulfilling.

Similarly, in other than even cases, the marginal-cost effect will be stronger for the party with the stronger probability of prevailing \textit{ex ante}. For instance, if both parties regard the initial probability of liability as 90 percent, the plaintiff will discount the expected marginal cost of legal services to 10¢ on the dollar, while the defendant discounts it only to 90¢. Accordingly, if the English rule is adopted, the stimulus to the plaintiff’s expenditure will be ten times greater than the stimulus to the defendant’s. Thus, fee shifting reinforces the advantages of the party who is initially favored in litigation.

Because of the interaction between the parties’ expenditure decisions, it is not possible to prove that both sides will increase their expenditures under the English rule. The reason for this ambiguity is that a marginal increase in one side’s expenditure has an ambiguous effect on the other side’s expenditures. A marginal increase in the spending by one side could
either provoke the opponent to respond in kind, or intimidate him into reducing his own efforts.

Braeutigam, Owen, and Panzar, however, showed that in Nash equilibrium the sum of the parties’ expenditures must increase. The extent of the increase depends on how sensitive \( p \) and \( A \) are to litigation expenditure, as Plott (1987) has demonstrated. Using a Nash equilibrium model and making some simplifying technical assumptions regarding functional form, Plott found that if the case outcome depends entirely on factors out of the litigants’ control, the English rule has no effect on expenditure. If case outcome is determined solely by litigants’ efforts, conversely, the English rule will cause expenditure to increase without limit.

Such effects are mitigated in regimes that limit the amount of fees that can be shifted. For example, under both English and US practice, indemnification is limited to reasonable expenditures. Hyde and Williams (2002) study partial fee shifting, allowing for uncertainty regarding the quantity of costs that will be deemed shiftable by the court.

Similarly, some recent US proposals provide that a losing party need not pay any indemnification in excess of his or her own litigation costs. Hughes and Woglom (1996) show that the English rule operates as a tax on the weaker party’s expenditure, since increases in the weaker party’s spending raise the cap on the indemnification potentially payable to the stronger party.

Furthermore, while most of the economic literature on litigation expenditure has assumed a Nash equilibrium, a few authors (for example, Hersch 1990) have argued that it is not reasonable to expect litigants to ignore the effect on the other side’s expenditure when choosing their own. The Nash specification is most appealing when expenditure is simultaneous, when each side must choose how much to spend before learning the opponent’s decision, or when the expenditure decision is largely determined by one’s initial choice of an attorney; it is least appealing when one side can commit to a given level of expenditure and communicate that commitment to the opponent in advance. One can analyze the latter situation using the model of conjectural variations – so-called because it allows a party’s decision to depend upon his conjectures regarding how the opponent’s decision varies with his own. Formally, let \( v_x \) denote the rate at which the plaintiff expects the defendant to respond to her expenditures. This rate could be positive (in which case expenditure would be provocative), negative (in which case expenditure would be intimidating), or zero (as in the Nash model). The plaintiff’s first-order condition then becomes

\[
(px + vx py)A + p (Ax + vx Ay) = 1 \quad \text{under the American rule, and} \quad (px + vx py)(A + x + y) + p(Ax + vx Ay) = (1 - p)(1 + vx) \quad \text{under the English rule.}
\]

(The analysis for the defendant is symmetric and is omitted for the sake
of brevity.) Comparing the first term of each of the equations, one can see that the stakes effect is still present. The direction of the marginal-cost effect, however, is now ambiguous. Under the English rule, if the plaintiff spends an additional dollar on legal services it will cost her only \( 1 - p \), after she discounts for the probability of prevailing. But if the plaintiff loses, she will also have to pay the defendant’s costs, and the additional dollar of plaintiff spending induces the defendant to change his expenditures by \( v_x \). If the plaintiff’s expenditure is intimidating, this will lower her marginal cost even further. If her expenditure is sufficiently provocative, however, her marginal cost of legal resources will rise; if it is provocative enough to outweigh the stakes effect, her equilibrium expenditure will fall.

Lastly, most models of fee shifting and litigation expenditure assume that lawyer and client act as one. When this is not the case – when there are agency problems in the lawyer-client relationship – the effects of fee shifting on litigation spending become potentially subject to the nature of the contract for attorney services. Hyde (2006) studies the interaction between the fee allocation rule and two such contracts: contingent fee arrangements – wherein the victorious lawyer receives a percentage of the litigation award – and conditional fee arrangements – wherein a victorious lawyer is paid a multiple of the market value of her services.

The analysis in this section has focused on the amount of resources expended in litigated cases. Total expenditures on litigation, however, are the product of two factors: expenditures per litigated case, and the number of cases that are actually litigated. Fee shifting can influence the number of litigated cases in two ways: by influencing the decision to bring the dispute to court in the first place, and by influencing the parties’ incentive to settle cases before trial. The next section of this chapter discusses the former effect, and Section 5 discusses the latter.

4. Effects of Fee Shifting on the Decisions to File and Contest Lawsuits
Consider the case of a consumer who has purchased a defective ballpoint pen and who is in theory entitled to a refund. Because the value of the pen is exceeded by even the most streamlined judicial proceeding, the consumer’s threat to litigate is not credible; and absent procedural devices such as a class action that can allow aggregation of her claim with others, she will be forced to rely on nonlegal incentives such as the seller’s interest in its reputation. If the consumer can recover legal fees along with the value of her refund, however, her threat to sue becomes credible.

4.1 Basic Model
Shavell (1982a), extending the work of Landes (1971) and Gould (1973) on the incentives to sue, generalized this argument to show that the English
Fee shifting

rule, and indemnification in general, works to encourage lawsuits by plaintiffs with relatively small claims but relatively high *ex ante* probabilities of victory. The American rule, conversely, encourages plaintiffs with relatively large claims but lower probabilities of victory.

The formal logic of the argument is as follows: let $p$ represent the probability of a plaintiff victory, $A$ the expected award if the plaintiff wins, and $c$ the cost of litigation for each litigant. (To simplify the argument, suppose that this cost is the same for both sides; this will affect the specific point at which the incentives switch, but not the basic intuition of the argument.) Under the American rule, litigation is profitable if (and only if) $pA > c$; thus, a plaintiff will bring suit if she views her chances as better than the threshold probability $p_{US} c / A$. Under the English rule, however, the plaintiff’s expected litigation cost is not $c$ but $(1 - p)2c$, since she pays no costs if she wins, but $2c$ if she loses. She will accordingly wish to litigate if $pA > (1 - p)2c$, or equivalently, if she views her chances at better than $p_{ENG} 2c / (A + 2c)$. Algebraic manipulation reveals that $c > A/2$ implies $p_{US} > p_{ENG} > 1/2$, and $c < A/2$ implies $p_{US} < p_{ENG} < 1/2$. Thus, when costs are high or stakes low, the English rule encourages some better-than-average suits that would be deterred under the American rule; when costs are low or stakes high, the English rule discourages some worse-than-average suits that would be brought under the American rule.

An identical line of argument shows the effect of indemnification on the incentives to defend against a lawsuit once it has been brought. If it costs the defendant $c$ to put up a defense that will succeed with probability $p$, it is worthwhile to defend (rather than suffer a default) only if the expected savings $pA$ exceed the expected costs of litigation. Under the American rule these expected costs are $c$, and under the English rule they are $(1 - p)2c$. The logic is as before; the American rule encourages long-shot defenses in high-stakes and low-cost cases, while the English rule encourages high-probability defenses in high-cost and low-stakes cases.

Such arguments lend support to the frequently expressed view that the English rule is superior on grounds of corrective justice, since the claims and defenses that it promotes are relatively meritorious ones – at least when viewed from an *ex ante* perspective. Similarly, as Rosenberg and Shavell (1985) and Farmer and Pecorino (1998) argue, indemnification can help discourage certain frivolous or ‘strike’ suits. Rosenberg and Shavell argue that indemnification emboldens defendants to put forward costly defenses against strike suits (assuming that the frivolous nature of the suit is common knowledge; as Katz (1990) argues, the English rule may do little to discourage strike suits that cannot be identified as such without a trial). Farmer and Pecorino argue that indemnification reduces the frequency of strike suits by dampening the incentive of plaintiffs’ attorneys
to develop a reputation for carrying out the threat of pursuing frivolous litigation.

But there is a cost to this ostensible increase in justice. The claim and defenses encouraged by the English rule are low stakes and high cost – that is, expensive to try relative to their importance. The claims and defenses encouraged by the American rule may be relative longshots on the merits, but they are relatively cheap to resolve. Moreover, some suits, including those brought to test or clarify the law or to settle matters of principle, may be socially desirable notwithstanding a low \textit{ex ante} probability of success. Accordingly, legal policy in this area may present a tradeoff between justice and more narrow conceptions of efficiency.

Any conclusions regarding the effect of litigation fee shifting on incentives to sue must also take account of the litigants’ expected response to risk. As has been widely recognized, the English rule magnifies the private risk arising from litigation by increasing both the returns from success and the losses from defeat. Thus, it tends to discourage risk-averse parties from bringing or defending lawsuits, regardless of the merits of their positions – a factor that has been stressed by partisans of the American rule. What has been less well recognized, however, is that this same increase in variance can encourage more litigation by the risk-neutral. The reason is that most lawsuits are divided into a series of procedural stages, at each of which it is possible to decide whether to continue depending on how the case is going. Because of this flexibility, as Cornell (1990) has shown, the decision to litigate can be interpreted as the purchase of an option. Just as financial options can sell for a positive price even if the probability of exercising them is low, the option value of litigation can make it profitable to put forward claims with negative expected value. Because the value of an option increases with its variance, the English rule, by increasing both the upside and the downside of litigation, intensifies this incentive. Indeed, if parties can drop arguments before trial without penalty, such enhanced option value could increase litigation even by the risk-averse.

4.2 \textit{Multi-phase Models}

Additionally, as the previous section indicated, the English rule indirectly alters incentives to sue through its effects on the expected cost of the individual case. Because indemnification encourages parties to litigate their disputes more intensively, it increases the expected cost of bringing and defending suits \textit{ex ante}. This increase in the \textit{ex ante} expected cost of bringing suit will, in turn, affect parties’ decisions regarding whether to pursue their cases – that is, whether to file suit, in the case of a plaintiff, or refrain from defaulting in the case of a defendant. Adding litigation intensity to the model may thereby affect conventional conclusions regarding the
A number of recent papers on fee shifting have combined litigation intensity, filing decisions, and default decisions in a single model. In Farmer and Pecorino (1999), the plaintiff first decides whether to file suit. If the plaintiff does file suit, the defendant next decides whether to simply pay the default judgment. If the defendant decides not to default, the two parties engage in the kind of litigation expenditure game described in Section 3. Farmer and Pecorino ask which rule, English or American, imposes lower litigation costs, and which results in more accurate dispositions. They emphasize that the answer depends on the shape of the exogenous litigation expenditure technology that is assumed to be in place— that is, on the shape of \( p(x, y) \) and \( A(x, y) \), as defined above in Section 3. They show that for each fee allocation rule, there exist technology parameter values under which plaintiffs file only meritorious claims, defendants always pay the default judgment, and litigation costs are zero.

Baye, Kovenock, and de Vries (2005) also combine filing, default, and litigation expenditure decisions, though in a different manner. In their two-phase model, the parties first simultaneously decide whether to concede. If neither party concedes, the parties next simultaneously decide how much to spend on litigation. Baye, Kovenock, and de Vries assume that the party with the highest quality case wins the case with probability one. Furthermore, the quality of each party’s case is, in turn, a strictly increasing function of her litigation expenditure. Thus, the party spending the most wins the case with probability one. In this context—which resembles a first price auction as much as it does litigation—Baye, Kovenock, and de Vries study a wide variety of fee-shifting rules. With regard to the comparison of American and English rules, they obtain more definitive results than do Farmer and Pecorino. This is not, however, inconsistent with Farmer and Pecorino’s general point that the effect of fee shifting is sensitive to assumptions made about the exogenous litigation technology. The definitiveness of Baye, Kovenock, and de Vries’s results appears to be due in large measure to their having restricted attention to a specific litigation technology.

5. Effects of Fee Shifting on the Settlement of Litigation
Because the great majority of civil cases are settled rather than tried, and because trial substantially increases the cost of disputes, effects on settlement are a critical factor in any comparison of the English and American rules. As a result, the economic literature on fee shifting has focused on this issue more than any other. The conventional wisdom among practicing attorneys appears to be that a shift toward fuller indemnification
would encourage settlement. The conclusions of the scholarly literature, however, cannot be said to offer strong support for this proposition; at best the effects are ambiguous.

5.1 Relative Optimism Models
As a first approximation, fee shifting magnifies the effect of litigants’ optimism, making them less likely to settle. As Landes (1971) and Gould (1973) observed, since litigation is a negative-sum game ex post, parties who accurately assess their chances of victory have a strong collective incentive to avoid the costs of trial. Indeed, in a world of purely Coasian bargaining, there would be no trials at all, since full sharing of information would eliminate any differences of opinion. Because of random variations in information, judgment, and temperament, however, some fraction of litigants will inevitably overassess their chances, and it is these optimistic litigants who have an incentive to go to trial. Pessimistic or unbiased parties, in contrast, would prefer to settle. But the degree of optimism necessary for a trial to result depends on how litigation costs are allocated, as the following argument (suggested by Mause 1969, and formally demonstrated by Shavell 1982a) shows.

Under the American rule, a plaintiff who perceives the probability of liability as \( p_P \), her stakes as \( A_P \), and her costs as \( c_P \) will insist on receiving a settlement of no less than \( S_P = p_P A_P - c_P \). Similarly, a defendant who perceives the probability of liability as \( p_D \), his stakes as \( A_D \), and his costs as \( c_D \) will be willing to pay no more than \( S_D = p_D A_D + c_D \). (Notice for future reference that the parties’ stakes, and not just the probability assessments, may differ in this formulation.) Settlement is thus possible if (and only if) \( S_P \leq S_D \), or equivalently, if the total litigation costs, \( c_P + c_D \), exceed the difference between the parties’ reservation settlement values, \( p_P A_P - p_D A_D \). Parties whose litigation costs are below this cutoff level, conversely, will prefer to go to trial.

Under the English rule, however, the plaintiff’s reservation settlement value becomes \( S_P = p_P A_P - (1 - p_P)(c_P + c_D) \), and the defendant’s becomes \( S_D = p_D A_D + c_P + c_D \). Now settlement is possible only if \( c_P + c_D > (p_P A_P - p_D A_D)/(1 - p_P + p_D) \). If the plaintiff’s probability estimate that she will win, \( p_P \), exceeds the defendant’s probability estimate that the plaintiff will win, \( p_D \) — that is, if the parties are optimistic relative to each other — the aggregate litigation cost minimum for settlement under the English rule is equal to that under the American rule divided by a number that is less than one. The aggregate litigation cost minimum under the English rule is thus greater than the aggregate litigation cost minimum under the American rule. Settlements are thus less likely under the English rule.

Another way to see this is as follows: From the two expressions for
the plaintiff’s minimal acceptable settlement amount $S_p$, it is clear that changing to the English rule from the American rule causes $S_p$ to increase by 

$$-(1 - p_p)(c_p + c_d) - (c_p + c_d) = p_p(c_p + c_d) - c_d.$$ 

This change in $S_p$ is easily interpreted by triangulating with a hypothetical baseline in which the plaintiff pays both parties’ costs. Relative to this plaintiff-pays-all rule, the English rule pays the plaintiff back aggregate costs in the event that the plaintiff wins, a payment with expected value $p_p(c_p + c_d)$. Relative to the plaintiff-pays-all rule, the American rule always pays her back the defendant’s costs, $c_d$. Thus, switching from the American rule to the English rule increases the plaintiff’s expected costs by $p_p(c_p + c_d) - c_d$, as claimed.

We can conduct the same analysis for the defendant. From the above equations, we see that changing from the American rule to the English rule causes the defendant’s maximum acceptable settlement amount $S_d$ to increase by $p_d(c_p + c_d) - c_d$. Thus, triangulating with respect to a hypothetical baseline in which the defendant pays neither party’s costs, the English rule makes the defendant pay both parties’ costs in the event that the plaintiff wins, while the American rule always makes him pay his own costs.

Now, if $p_p = p_d$, fee shifting causes the plaintiff’s minimal acceptable settlement amount and the defendant’s maximum acceptable settlement amount to increase by the same amount, and there is no change in the minimum aggregate litigation cost $c_p + c_d$ for settlement. But if $p_p > p_d$, then fee shifting causes the plaintiff’s minimal acceptable settlement amount to increase by more than the defendant’s maximum acceptable settlement amount, and the minimum aggregate litigation cost for settlement increases.

The economic intuition underlying this result is that indemnification internalizes one externality while creating another. Under the English rule, a litigant is forced to take into account the other side’s litigation costs to the extent that she risks losing the case, making her more willing to settle. But conversely, she is freed of her own litigation costs to the extent that she hopes to win, making her less likely to settle. Since litigants are disproportionately drawn from the population of optimists (else they would settle however costs are allocated), the latter effect tends to outweigh the former. Indeed, in the limiting case when both parties are fully confident of winning, neither expects to pay any costs at all and settlement is impossible.

This line of argument, however, suggests an important exception to the basic result: in some cases, parties might choose to litigate due to a difference of opinion not over liability but over stakes. A plaintiff who regarded the stakes as sufficiently higher than did the defendant – for example, because she hoped to establish a favorable precedent that could be drawn on in later cases – might refuse all settlements even if the parties agreed on
the probability of liability or were both relatively pessimistic. If the parties’ relative optimism about the stakes were enough to outweigh their relative pessimism about probability, fee shifting would encourage settlement and discourage trial.

5.2 Structured Bargaining with Asymmetric Information

The Landes-Gould model of settlement bargaining (often called the ‘optimism model’ in subsequent literature) is open to the criticism that it assumes that each party knows the other’s reservation value in settlement negotiations, whereas such values are more naturally assumed to be private information. Accordingly, subsequent writers have often preferred to base their analyses on models of asymmetric information bargaining. In this formulation, as first proposed by Cooter, Marks and Mnookin (1982), trials are caused not by optimism but by uncertainty over the opponent’s reservation settlement value.

In most formal models of asymmetric information settlement bargaining (see, for example, Bebchuk 1984), one party has private information regarding the outcome of the case, and the other makes a single take-it-or-leave-it settlement offer. (In Reinganum and Wilde 1986, however, the offeror also has private information.) Thus, the offeror does not know whether her single offer will be accepted, and if it is not, she must pay the cost of trial. This generates a familiar tradeoff: a more generous settlement offer is more likely to be accepted, which is good, but, on the other hand, the acceptance itself is not as favorable to the offeror. The tradeoff is similar to that faced by the classic monopolist: lowering price increases purchases, but every purchase generates less revenue.

Whether cases settle in such a model depends on a number of factors, including the stakes in the case, the cost of litigation, and, most importantly, the extent of the offeror’s uncertainty regarding the reservation value of her opponent. Roughly speaking, greater dispersion in the probability distribution of the offeree’s reservation settlement values means a lower density at any given value, so that a decision by the offeror to take a marginally tougher position sacrifices fewer bargains. This increases the net marginal benefit of making a less generous settlement offer. If the effect is sufficiently uniform across the continuum of possible settlement offers, more uncertainty will mean less settlement.

Asymmetric information models of settlement bargaining tend to confirm the optimism model’s conclusion that the English rule generally discourages settlement in disputes revolving around liability, but not in disputes revolving around stakes. The reason is that indemnification magnifies uncertainty in the former set of cases but not in the latter. More precisely, uncertainty about opponents’ reservation values can stem from numerous
Fee shifting 285

sources: variations in the private cost and stakes of litigation, variation in attitudes toward risk and delay, and variation in private information relevant to the trial outcome. Differences in risk aversion, time preference and litigation stakes are not affected by fee shifting, but differences in private cost and in information relevant to liability are. Fee shifting thus increases the difference between the reservation values of parties with favorable private information and high litigation costs on the one hand, and parties with unfavorable information and low litigation costs on the other. This increase in uncertainty leads all types of parties to toughen their overall bargaining stance, thus lowering the probability of settlement. Ironically, as Polinsky and Rubinfeld (1997) point out, this implies that the English rule actually lowers the average quality of tried cases, since the marginal parties it sends to trial have relatively less favorable private information than those who would litigate absent the prospect of indemnification.

Even if fee shifting does not alter the probability of settlement, however, it can still influence its amount. As Cooter and Rubinfeld (1994) have argued in the context of legal discovery and Bebchuk and Chang (1996) have argued in the context of offers of judgment, fee shifting can, by equalizing the bargaining power of parties with asymmetric litigation costs, help to move the settlement amount closer to the expected trial outcome. To the extent that trial outcomes are deemed to be just, fee shifting thus may help promote equity; to the extent that trial outcomes reflect substantive legal norms, fee shifting helps promote incentives for proper primary behavior – a subject more fully explored in the following section.

In most asymmetric information models the subject of private information is the outcome of the case. But parties may also have private information regarding other variables, such as their own trial costs. Chopard, Cortade, and Langlais (2008) study this possibility (for the case of both one- and two-sided asymmetric information). They emphasize that changing the subject of private information markedly changes the predictions of the model.

5.3 Multi-phase Models

The foregoing discussion of both optimism and asymmetric information models of settlement bargaining has taken the cost of litigation as given. As previous sections have observed, however, indemnification generally raises litigation costs – through its effect on litigation intensity.

Recent research imbeds a model of settlement negotiation within a large multi-phase model of litigation that includes litigation intensity decisions. Doing so greatly complicates the analysis. Tractability is typically restored by adopting helpful function form assumptions.

Gong and McAfee (2000) combine a model of two-sided
asymmetric-information settlement negotiation with a model of litigation expenditure that is similar to that described in Section 3. In particular, for their litigation expenditure phase, Gong and McAfee adopt a simplified litigation technology in which the chance of plaintiff victory equals the plaintiff’s proportional contribution to total evidence. Gong and McAfee’s chief concern is the accuracy of the settlement amount. They find that the English rule shifts the settlement amount away from the best estimate of the proper settlement amount – where such estimate is based on aggregating the information that is privately received by each of the parties.

Chen and Wang (2007) also analyze a multi-stage model that includes both settlement negotiation and litigation spending. Chen and Wang’s model differs from Gong and McAfee’s in several important respects. First, Chen and Wang include the plaintiff’s filing decision. Second, the settlement negotiation phase in Chen and Wang’s model is the usual one-sided asymmetric model in which the uninformed party makes a take-it-or-leave-it offer. Lastly, the plaintiff’s lawyer, who is paid on a contingency fee basis, makes all litigation decisions for the plaintiff’s side. With the addition of several functional form assumptions, Chen and Wang find that the English rule results in fewer filed suits, higher settlement rates under certain distributional assumptions, greater trial costs, and a greater chance of defendant victory.

6. Effects of Fee Shifting on Substantive Behavior

The discussion thus far is in a fundamental sense incomplete, since it has focused largely on the procedural costs of litigation. If such costs were one’s only concern, of course, they could be eliminated entirely by abolishing the legal system and all publicly enforceable rights to relief. A central purpose of having a public system of courts, however, is to redress wrongs and to encourage compliance with primary substantive norms such as taking precautions against accidents and keeping one’s promises. Indemnity of legal fees, accordingly, must ultimately be judged on these latter criteria – or more accurately, on whether it increases the social value of substantive enforcement net of process costs.

Viewed from this perspective, the English rule initially appears attractive, since it tends to encourage high probability suits and discourage low probability ones. Assuming that the probability of liability is correlated with the actual violation of substantive norms, therefore, indemnification increases the net expected punishment for such violations and thus helps promote substantive compliance. This is easily seen in the case where courts’ liability determinations are error-free, as Rose-Ackerman and Geistfeld (1987) and Polinsky and Rubinfeld (1988) have shown. Consider a potential tortfeasor who can take precautions against an accident that
will cause an uncertain amount of damage. Suppose that the possible
damage ranges from zero to $A$, and that the cost of establishing liability
following an accident is $c$. Under the American rule, it follows that the
tortfeasor will have inadequate incentives for precaution. In the event that
damages turn out to be less than $c$, the victim will not sue, so the tortfeasor
will escape responsibility for a portion of the damages caused. Under the
English rule, however, the victim will always have the incentive to sue, so
that all accident costs will be fully internalized. Under a rule of negligence
as opposed to strict liability, indeed, complete cost internalization can be
achieved without incurring any litigation costs at all: defendants will be
induced to take optimal care by the threat of litigation, so plaintiffs will
never actually have to sue. Conversely, under the American rule, defend-
ants may rationally decide to take excess care – or to abstain from risky
though optimal activities – in order to avoid the greater expense of having
to defend their behavior in court. Indemnification protects them from such
expenses, thus preventing overdeterrence.

This happy outcome, however, depends on the assumption that deserv-
ing plaintiffs and defendants always win their cases. In the presence of legal
error, as P'ng (1987) and Polinsky and Shavell (1989) have shown, neither
the American nor the English rule provides incentives that are first-best
optimal. Optimal incentives, rather, require at least two separate policy
instruments – one to motivate efficient substantive behavior, and another
to promote an efficient amount of litigation. Polinsky and Che (1991)
demonstrate that, in general, this means decoupling the amounts paid by
losing defendants from those received by victorious plaintiffs. (Indeed, as
Polinsky and Rubinfeld (1996) show, decoupling is generally necessary
even to achieve the lesser goal of minimizing the litigation costs associ-
ated with achieving a given level of deterrence.) Devices combining fines,
punitive damages, and taxes (positive or negative) on litigation accomplish
such decoupling, but the English rule, which merely re-allocates costs
between the parties in zero-sum fashion, does not. Furthermore, it is not
even the case that the English rule is second-best efficient within the cat-
egory of zero-sum policy instruments. Kaplow (1993) shows that damage
multipliers, such as the treble damage provisions of US antitrust law,
provide a cheaper method of achieving any given amount of deterrence.
The reason is that damage multipliers provide incentives for private law
enforcement to be undertaken by those plaintiffs whose litigation costs
are lowest; fee shifting, in contrast, encourages plaintiffs to bring lawsuits
without regard to their costs of litigation.

One might still ask whether the English rule does better than the
American in promoting efficient substantive behavior, notwithstanding
the potential availability of alternatives that are superior to both. The
answer to this question, however, is ambiguous, as Gravelle (1993), Hylton (1993a, 1993b), and Beckner and Katz (1995) demonstrate in independent formal models. It is possible to draw generalizations regarding when the English rule improves matters, but they depend on the subtle interaction of a number of factors, including whether substantive precaution affects the magnitude of injury or just its probability, the extent to which precaution affects the probability of liability, whether damage awards are sufficient to compensate plaintiffs for their losses, and whether defendants have the opportunity to act strategically by taking just enough care to foreclose litigation. Hylton, for instance, concludes that a one-way fee-shifting rule operating in favor of plaintiffs would be best, but this conclusion depends upon several features of his model (including, perhaps most importantly, the assumption that plaintiff’s care does not affect the expected cost of accidents). Applying such generalizations to individual cases or categories of cases is probably beyond the capacity of either courts or legislatures. As Gravelle concludes, ‘[i]t seems more promising to pursue other, more direct means of correcting the inefficient incentives for care provided by a costly and imperfect legal system’.

More recently, Choi and Sanchirico (2004) investigate the effects of decoupling the amounts paid by losing defendants from the amounts received by victorious plaintiffs when substantive behavior, filing incentives, and the intensity of spending are all at issue. In their model, the defendant first chooses a substantive action. The plaintiff next decides whether to file suit. Finally, the parties choose their level of effort at litigation. Choi and Sanchirico investigate the problem of minimizing the social cost of a given level of deterrence by adjusting three policy instruments: an upfront filing fee, the amount that victorious plaintiffs recover, and the amount that losing defendants pay. Contrary to earlier work by Polinsky and Che (1991), whose model does not include litigation spending decisions, Choi and Sanchirico find that what plaintiffs recover should be no lower than what defendants pay when the conjunction of two conditions obtains: defendants are not wealth-constrained and the potential harm from defendants’ violations is sufficiently great. Choi and Sanchirico show that their analysis is invariant to whether the fee allocation rule is English or American.

Klement and Neeman (2005) investigate the effect of fee-shifting rules when the objective is to maximize the probability of settlement subject to the constraint that a given amount of deterrence is generated by the threat of litigation. Klement and Neeman model settlement negotiations, but not litigation spending decisions. They find that maximizing the probability of settlement requires adopting the English rule. The English rule creates the largest divergence between the trial payoffs of liable and non liable...
defendants. It, therefore, allows the greatest amount of deterrence per trial and so the fewest number of trials for a given amount of deterrence.

7. Variations on Simple Fee Shifting
The foregoing discussion has been premised on the assumption that ‘costs follow the event’ – that is, that any fee shifting that takes place is based solely on who wins the case. Much recent discussion in policy and scholarly circles, however, has focused on two more complicated forms of indemnification.

7.1 Fee Shifting Conditioned on Offers Made in Settlement
Both England and a number of American jurisdictions provide a mechanism through which a defendant who would otherwise be obliged to pay for legal expenses can partially avoid the obligation by making a suitable offer of settlement. In England this procedure is called ‘payment into court’ and requires the defendant to actually deposit funds with a court officer; while in the United States, Federal Rule of Civil Procedure 68 and similar rules² merely require the formal filing of what is labeled an ‘offer of judgment’. Under either provision, a defendant who makes such a formal offer is considered the prevailing party for purposes of cost allocation if the plaintiff rejects the offer and then is subsequently awarded a lesser amount at trial. In such an event, the defendant avoids having to pay any costs incurred by the plaintiff subsequent to the offer, and is entitled to indemnification for his own subsequent costs as well. By all accounts, defendants avail themselves of this procedure much more frequently in England than in the US – probably because the prospect of shifting liability for ‘costs’ is likelier under the broader English definition of the term to outweigh the disadvantages of making a settlement offer. Similarly, within the US, Rule 68 appears to be used more widely in disputes covered by one-way pro-plaintiff fee-shifting statutes such as Title VII, since the Supreme Court held in *Marek v. Chesny* that attorneys’ fees shifted under such statutes are to be considered ‘costs’ under Rule 68.

Because of the relatively infrequent use of Rule 68 in US courts, a number of American critics have in recent years supported its expansion – either by extending its coverage to attorneys’ fees generally, or by making the procedure available to plaintiffs as well as defendants. (It should be

---

² Under the US Internal Revenue Code, Sec. 7430, if a taxpayer makes an offer to the Internal Revenue Service to settle a claim for taxes allegedly owed, and if the Internal Revenue Service rejects the offer and later obtains a judgment against the taxpayer that is not greater than the taxpayer’s offer, the taxpayer may be entitled to recover reasonable litigation costs, including attorneys’ fees.
noted, however, that providing the procedure to plaintiffs is meaningless to the extent that they are already entitled to collect costs when they prevail; in such circumstances, the opportunity to make an offer of judgment can only advantage defendants.) The U.S. Republican Party’s recent ‘Contract with America’, for example, would have established just such a generalized offer-of-judgment rule in federal diversity cases. Such proposals have commonly been supported by the claim that they will reduce expenditures on litigation by encouraging parties to make more reasonable settlement offers and to accept such offers when they are made. Their proponents have also argued that it is fairer to charge the costs of trial to the party who, by refusing a reasonable settlement, causes those costs to be incurred.

In general, the economic literature on offers of judgment is substantially less developed than that on pure indemnification, and many interesting questions remain to be fully investigated, including the effect of the procedure on strategic behavior in negotiations. The place to begin any analysis of the offer of judgment, however, is with the observation that it is essentially an option to convert disputes over damages into disputes over liability. To see this, compare two cases: one in which it is clear that the defendant has acted negligently but unclear whether the plaintiff’s injuries are 1000 or 3000 (with the two possibilities being equally likely), and a second in which it is clear that damages are 4000, but an even gamble whether the defendant is liable at all. In both cases, expected damages are 2000, but absent an offer-of-judgment procedure, the plaintiff’s position is stronger in the former. She is certain to prevail at trial and to recover some fraction of her costs, even if it is only court fees. In the latter case, she runs the risk of paying both her costs and a portion of the defendant’s. Under Rule 68 or a similar procedure, however, the defendant can convert the former dispute into a partial settlement of 1000 combined with a dispute over whether the defendant is liable for an additional 2000. In this converted dispute, the defendant stands an even chance of avoiding liability for the plaintiff’s costs and of recovering his own. This improves his expected position to what it would be in the case of pure liability, at the plaintiff’s expense.

The example illustrates two lessons. First, a rule authorizing defendants but not plaintiffs to make offers of judgment redistributes wealth from plaintiffs to defendants in disputes that are entirely or partly over damages, as both Priest (1982) and Miller (1986) have suggested. Second, such offers have no effect in disputes that are purely over liability. If the only possible trial outcomes are verdicts of zero or 4000, for instance, there is no advantage to the defendant in making a Rule 68 offer of less than the full 4000. If he offers a lesser amount, he will be liable for costs in
the event of a plaintiff’s verdict and certain to receive costs in the event of a defendant’s verdict – just as he would if he made no offer at all. Similarly, a less-than-full offer does not affect the possible payoffs for the plaintiff. The defendant could of course offer to settle for the full 4000, but the plaintiff should be happy to accept such an offer whether or not Rule 68 is in force.

The offer-of-judgment procedure, accordingly, cannot affect whether an offer is made or accepted in such cases.

With these points taken as caveats, the effects of offer-of-judgment rules are roughly analogous to those of indemnification in general. The possibility that costs will be shifted following a settlement offer both raises the stakes of the case and lowers the perceived marginal cost of legal expenditure, thus increasing incentives to expend resources at trial. The effect is less than under the pure English rule, however, since only post-offer expenditures are liable to be shifted. Similarly, the opportunity to make an offer of judgment increases expected payoffs for plaintiffs who expect to win large awards at trial, and lowers expected payouts of defendants who expect awards to be low, emboldening such parties to pursue litigation.

The effect of offers of judgment on the settlement decision depends, like the effect of indemnification generally, on the parties’ attitudes toward risk and on the model of settlement that one thinks appropriate. Under the Landes-Posner-Gould optimism model, offers of judgment tend to lower the chances of settlement between risk-neutral parties, since, as Priest (1982) suggests and Miller (1986) and Chung (1996) confirm, such offers lower the reservation values of optimistic defendants more than they do those of optimistic plaintiffs, thus reducing the potential settlement range. Offers of judgment also increase the risk of litigation, though not as much as pure indemnification does; this encourages risk-averse parties to settle, but risk-preferring parties to litigate. Anderson (1994), who extends the optimism model to include the possibility of bargaining stalemate, reaches similar results. Within Bayesian models of settlement, the outcome appears more complicated. Cooter, Marks and Mnookin (1982) conjecture that an offer-of-judgment rule, by effectively taxing hard offers and subsidizing soft ones, should encourage settlement. Spier (1994), however, in a model in which defendants make offers to plaintiffs with private information, finds that the procedure leads to more settlement than the pure American rule in cases where the plaintiff’s private information relates solely to the size of the award, less settlement than the American rule in cases where the plaintiff’s private information relates solely to the probability of liability, and an ambiguous effect in other cases. She also demonstrates a similar result using a mechanism-design model that, instead of specifying any particular bargaining process, assumes that the parties use a Pareto-efficient trading mechanism in the style of Myerson and Satterthwaite (1983). As
with pure indemnification, accordingly, the effect of offers of judgment on settlement probabilities appears to depend on the sources of the underlying dispute.

7.2 Fee Shifting Conditioned on the Margin of Victory

In the United States, a variety of statutory and judicially created rules allow courts to award partial or full indemnification in lawsuits in which the losing party’s case is deemed after the fact to be of sufficiently low merit. Such rules include the common-law torts of barratry, abuse of process, and malicious prosecution, the traditional authority of courts of equity to exercise their discretion in the interests of justice, the sanctions for discovery abuse provided by Federal Rule of Civil Procedure 37, and the (just amended) provisions of the Internal Revenue Code requiring the government to pay a taxpayer’s reasonable litigation costs upon a court finding that the government’s position in a tax dispute was substantially unjustified. Similarly, as Pfennigstorf (1984) reports, indemnification awards in most other Western legal systems are likely to be more generous in cases where the loser’s legal or factual position appears weak.

The possibility of tying indemnification to the merits of the losing case has attracted increased attention in recent years, in part as a response to the growth of litigation practice under US Federal Rule of Civil Procedure 11. This rule requires persons filing court papers to warrant that their filings are well grounded, and authorizes courts to impose monetary sanctions on parties whose filings are found to be frivolous, harassing, or made for purposes of delay. Limiting fee shifting to cases of particularly low merit has seemed to many commentators an attractive compromise between the English and American rules, since it protects clearly deserving litigants without imposing unnecessary risk on those who bring colorable claims in good faith.

As Bebchuk and Chang (1996) have pointed out, the effect of policies such as Rule 11 is to condition fee shifting on the winner’s margin of victory; those who win in a rout receive indemnification, while those who win narrowly do not. They confirm the conventional wisdom in a formal model, showing that such policies, if designed properly, can do a better job than either the English or the American rule at encouraging meritorious suits (and by analogy, defenses) and discouraging frivolous ones. The reason is that such policies make use of the parties’ private ex ante information regarding the merits of the case. A party who loses by a large margin is less likely to have believed ex ante that her case had merit; conversely, one who wins by a large margin is less likely to have believed that her case lacked merit. While the optimal fee-shifting rule depends on the distribution of judicial and litigant error, it is possible by altering the
threshold for fee shifting to regulate the proportion of potential claims and defenses that are actually brought into the system. Because its effects are zero-sum, however, margin-based fee shifting is still less efficient than policies that decouple one side’s payments from the other side’s recovery (see generally Polinsky and Rubinfeld 1993). Policies that tax or subsidize individual parties based on the *ex post* quality of their case may be best of all; they may also, of course, be the most difficult to administer.

While other incentive effects of margin-based fee shifting have not been formally explored, it appears likely that it has analogous consequences to indemnification generally, though in lesser degree. These consequences recapitulate the discussion in earlier sections of this chapter and can be surveyed in brief. First, to the extent that such policies succeed in encouraging meritorious claims and defenses and discouraging frivolous ones, they will tend to improve incentives for primary substantive behavior. The complications described in Section 6 above, however, remain to be analyzed. It is possible, for instance, that the prospect of shifting litigation costs to the other side following a commanding victory will induce excessive caretaking *ex ante*, though the benefits of doing so are less than under the pure English rule.

Second, margin-based fee shifting will both raise the stakes of litigation and decrease its expected marginal cost, inducing the parties to intensify their efforts at trial. Schmalbeck and Myers (1986) argue that this effect will be relatively minor, since in a truly frivolous case there is little that the parties can do to change the outcome. Their argument is open to question, however, as the substantial amount of litigation effort under Rule 11 illustrates (see, for example, Kobayashi and Parker 1993, who discuss the incentive effects of recent amendments intended to reduce such ‘satellite’ litigation). While margin-based indemnification has little effect on cases that are clearly contestable or clearly frivolous, in many disputes the colorability of the losing case is less obvious. Parties in such intermediate cases, thus, will have an incentive to increase their expenditures in order to influence the size of the margin of victory.

Third, margin-based fee shifting will decrease the likelihood of settlement to the extent that the parties have a difference of opinion regarding the chances of indemnification. Optimistic parties will exaggerate the likelihood that they will win by a large margin and underestimate the likelihood that they will lose by a large margin. The prospect of indemnification will cause such parties to toughen their settlement demands, reducing the range for settlement. Since pessimistic and unbiased parties will have an incentive to settle in any event, the net consequence will be an increase in trials. Similarly, given private information regarding the probability of a one-sided outcome, margin-based fee shifting will increase the variance of
the parties’ reservation settlement values, encouraging tougher bargaining
and hence fewer settlements. The increased risk of trial, however, works to
counteract such effects for risk-averse litigants.

In sum, however, the case for at least some margin-based fee shifting
appears stronger than the case for indemnification generally, on grounds
of both fairness and efficiency. Parties who lose lawsuits decisively are
probably more deserving of sanction than those who lose barely. The
social value of litigation is probably higher in close cases, whether mea-
ured by the public benefits of legal precedent or by more libertarian
considerations. And margin-based fee shifting seems to do a better job at
providing improved incentives for primary behavior, and has lower costs
in terms of incentives for increased expenditure at trial. In light of the
relatively limited theoretical and empirical work on this particular topic,
however, these conclusions must be regarded as tentative.

8. Empirical Evidence on the Effects of Fee Shifting
Given the complexity and ambiguity of the aforementioned considera-
tions, it would plainly be desirable to have some hard empirical evidence
to bring to the policy debate. Unfortunately, such evidence is sparse. What
evidence exists on the effects of fee shifting falls into three categories: labora-
tory research on bargaining behavior by experimental subjects, numerical
simulations of litigation behavior based on empirically obtained
parameters, and econometric evidence primarily drawn from a single
policy experiment: Florida’s experience with the English rule in medical
gainpractice cases from 1980 through 1985.

8.1 Laboratory Experiments
Coursey and Stanley (1988) tested the effects of fee shifting within an
experimental setting they designed to simulate the process of bargain-
ing under threat of trial. They divided their subjects (students at the
University of Wyoming) into pairs and instructed them to attempt to
divide between themselves a number of tokens that were subsequently con-
vertible into cash. If time expired before the subjects reached agreement,
the tokens were divided through a random drawing, intended to represent
an uncertain court award. To simulate rational expectations, the subjects
were presented with the probability distribution of awards arising from
the random drawing, and to simulate the costs of trial, the subjects were
collectively fined an amount equaling 40 percent of the total value of the
tokens in the event the drawing had to take place.

The experimenters conducted negotiations using three different cost allo-
cation rules. Under the simplest procedure, the fine was divided between
the two parties equally, in an intended simulation of the American rule.
A second group of subjects negotiated under a rule whereby the fine was paid entirely by the party who received the smaller portion of the tokens in a random drawing; this was intended to simulate the English rule. Yet a third group negotiated under an offer-of-judgment rule intended to simulate Rule 68: the plaintiff paid the entire fine if the draw awarded her an amount less than or equal to the defendant’s last proposal; and the fine was otherwise split equally.

The result of this experiment was that subjects settled more frequently under the English than under the American rule. Under Rule 68 settlement was likeliest of all; and in addition, the plaintiff was much more likely to be the accepting party. The authors also found that settlements were more favorable to the defendant under Rule 68 than under the English rule, consistent with the theoretical predictions outlined in the previous section. The American rule was most favorable of all for the defendant, though the authors ascribed this result to the behavior of one especially risk-averse individual. These findings are consistent with theoretical models that predict increased settlement on the basis of simple risk aversion. Because the experimental design ruled out the possibility of optimism or private information, however, its results cannot be extrapolated to situations in which such phenomena, which could cause the English rule to reduce settlement, are present.

In a separate series of survey experiments, Rowe and various co-authors studied the effects of cost allocation rules on lawyers’ and students’ responses to a variety of bargaining situations presented by hypothetical tort and civil rights cases. While this experimental design suffered from the weakness that the subjects were not provided with any direct financial incentives, the more realistic nature of the problems and the subjects’ professional status and experience provided at least some motivation to bargain seriously. The results of the experiments, however, were mixed. Rowe and Vidmar (1988) found that there was little difference between the American rule and a modified Rule 68 (enhanced to cover attorneys’ fees and to allow plaintiffs as well as defendants to make offers) on law students’ willingness to accept offers of settlement, although they did find an effect on the size of counteroffers, as well as a difference in plaintiff acceptance rates, between modified Rule 68 and a one-way pro-plaintiff rule. Anderson and Rowe (1996) replicated this experiment with practicing lawyers, also examining the subjects’ behavior under an alternative fee-shifting rule in which the maker of a rejected offer had to pay the rejecting party’s subsequent reasonable attorneys’ fees. They found that while modified Rule 68 did not appreciably affect plaintiffs’ minimum asks relative to the American rule, it did raise the maximum amounts that defendants were willing to offer. Finally, Rowe and Anderson (1996)
considered the effects of a modified Rule 68 on hypothetical bargaining in civil rights cases otherwise governed by a pro-plaintiff rule. They found that replacing this one-way rule with an enhanced Rule 68 significantly lowered plaintiffs’ minimum asks, as well as the gap between plaintiff asks and defendant offers. In all, these results suggest that fee shifting has its strongest effect when it is one-sided and when the favored side is risk averse or liquidity constrained.

More recently, Main and Park (2000) find that shifting between the English and American rules has no effect on the frequency of pretrial settlements, but that the English rule leads to greater settlement amounts when the probability of the plaintiff prevailing is relatively large. Their experimental data is derived from repeatedly and randomly pairing 14 undergraduate students to play a five-minute negotiating game in which bids were transmitted back and forth by intermediaries. The subjects were modestly compensated. Each pairing of undergraduates played several times under different rules, and one play was selected at random for purposes of determining payoffs. This experimental design allowed repeated play and learning without creating wealth effects. (However, it also diminished the impact of the monetary reward on each play.)

Main and Park (2002) present experimental data on the effect of a defendant-offer-of-judgment rule when such a rule is applied against the backdrop of the English rule for costs. Thus, in those circumstances in which the offer-of-judgment rule does not operate — as when the defendant does not make an offer or the plaintiff wins more than the offer — the losing side pays all costs. Main and Park find that an offer-of-judgment rule has no effect on the frequency of settlement but lowers the settlement amount. The experimental design in Main and Park (2002) is similar to that in Main and Park (2000). Thirty-eight university students were repeatedly and randomly paired. The researchers modestly compensated their subjects based on the rule-affected outcome of a three-minute computer-interactive negotiation game. Each pairing of subjects played several times under different rules, and one play was selected at random for purposes of determining payoffs.

8.2 Simulations

A number of authors have attempted to estimate the quantitative effects of fee shifting by numerically simulating the behavior of theoretical models. Katz’s (1987) approach is illustrative. He developed algebraic formulas, based on a linear approximation to the standard model of litigation expenditure, that relate the difference in expenditure per case between the English and American rules to two empirical parameters: the ratio of total expenditure to the stakes of the case, and the elasticity with which parties...
increase their expenditures in response to higher stakes. Using empirical estimates of these parameters taken from the University of Wisconsin’s Civil Litigation Research Project (Trubek et al. 1983), he calculated the likely effects of switching from the American to the English rule, concluding that such a switch would increase expenditures per case in the neighborhood of 125 percent.

Such a large increase in cost per case, however, could be expected to lead to a reduction in the number of cases or to increased settlement. In an attempt to measure this anticipated reduction, Hause (1989) extended the Landes-Gould optimism model to allow for variable expenditure, and calculated its numerical behavior for a range of possible parameter values. He concluded that the increased costs per case under the English rule were sufficient to outweigh any effects of optimism, resulting on balance in an increased frequency of settlement. Hersch (1990) recalculated Hause’s simulations under the assumption that trial expenditure is determined in a conjectural-variations rather than a Nash equilibrium. He found that both settlement and costs per case rose, though by a lesser amount than Hause had estimated. The parameter values Hause and Hersch used, however, were not based on any empirical data. Donohue (1991a) recalculated Hause’s simulations using what he argued were more plausible parameter estimates, and concluded that the English rule would increase trials on balance.

Hylton (1993a, 1993b) used numerical simulation to estimate the effects of fee shifting on primary behavior. He concludes that while litigation is more frequent under the English than the American rule, levels of substantive compliance under the two rules would be similar. Best of all, according to his calculations, is one-way pro-plaintiff fee shifting, which leads to the highest level of compliance and least amount of litigation. These conclusions, however, depend both on the functional form used in his simulations, and on the specific assumptions of his theoretical model.

Watanabe (2006), which is described below, combines simulation techniques with econometric estimation in order to study the effects of litigation reform on settlement timing.

8.3 Econometric Evidence of Actual Disputes

Schwab and Eisenberg (1988) report on a 1976 statute that established one-way pro-plaintiff fee shifting in federal constitutional tort cases (that is, cases in which the federal government is sued for violating the plaintiff’s constitutional rights). They find some evidence that the statute was followed by a decline in plaintiff success rates at trial and by an increase in trials relative to other federal civil actions, but little evidence of any increase in the number of lawsuits filed. These results suggest that the statute had its primary effect
in encouraging plaintiffs to bargain more aggressively in settlement negotiations, consistent with Rowe et al.'s survey experiments, as well as with the theoretical predictions of the optimism and Bayesian settlement models discussed in Section 5 above. Because of the relatively low magnitude of their quantitative estimates, however, and because their observations were muddied by the fact that some courts shifted fees on a discretionary basis before the statute was passed, the authors present their findings as tentative.

Fournier and Zuehlke (1989) develop an econometric model of settlement behavior in which the plaintiff's settlement demand and the probability of settlement are jointly determined. Using nonlinear methods that correct for data censoring (that is, the fact that the amount of the demand is only observed when settlement takes place), they estimate their model using data from a nationwide survey of civil federal filings between 1979 and 1981. The data classify disputes according to the type of legal claim at issue (for example, tort, copyright, antitrust) and include information regarding the alleged damages and the number of litigants in each case, the mean and variance of trial awards in litigated cases within each subject-matter classification, and separately prepared estimates of government litigation costs in each case classification, which the authors argue serve as a reasonable proxy for private litigation costs. The coefficients of the resulting equations suggest, not surprisingly, that settlement demands are positively correlated with mean trial awards and alleged damages, and negatively correlated with litigation costs. More interestingly, they also suggest that settlement is more likely in cases and categories with high alleged damages, high mean and variance of trial awards, multiple parties, and low (!) litigation costs. These latter results are at odds with the theoretical predictions of most models of settlement, although the negative relationship between settlement and potential trial awards is consistent with a hypothesis of risk aversion.

Because a fraction of the sampled filings were subject to the English rule, Fournier and Zuehlke were able to estimate its effects as well. They find that fee shifting is negatively correlated with both the probability of settlement and the size of the settlement demand, although the latter effect is not statistically significant. This finding lends some support to the predictions of both the optimism and Bayesian models. The filings in their sample governed by the English rule, however, were few in number and concentrated in a few specialized areas, so this evidence cannot be regarded as especially strong. It is possible that this correlation reflects differences in the types of disputes covered by the rule, rather than effects of indemnification. The authors' unusual findings regarding the other determinants of settlement also call this result into question.

More instructive is Hughes and Snyder's research on Florida's
experiment with the English rule in medical malpractice cases. The Florida statute, passed in 1980 with the support of the state medical association, also provided an offer-of-judgment procedure and exempted insolvent parties from the obligation to pay indemnification; it was repealed in 1985 with the support of its original proponents following a series of expensive and well-publicized plaintiff verdicts. In Snyder and Hughes (1990), the authors use a bivariate probit procedure to analyze insurance company data on closed claims filed before, during, and after the period in which the rule was in effect, and estimate the effects of indemnification on plaintiffs’ decisions to drop claims, settlement, and defendants’ expenditure on lawyers. Their findings lend support to several of the theoretical predictions outlined in previous sections of this chapter. Specifically, they find that in cases governed by the English rule, (1) a significantly higher percentage of claims were dropped at an early stage of the litigation, consistent with the proposition that fee shifting encourages risk-averse and low probability plaintiffs to exit the system; (2) defendants spent significantly more per case, in amounts consistent with Katz’s simulations, in both settled and in litigated cases; (3) holding other case characteristics constant, the likelihood of litigation increased, consistent with the optimism model. Because dropped cases tended disproportionately to have characteristics that would have made them likelier to go to trial had they remained in the system, however, the authors conclude that the English rule decreased the frequency of litigation on balance. Because of greater expenditure per case, however, total expenditures on litigation still increased.

In a subsequent article analyzing the same data set (Hughes and Snyder 1995), the authors find that the English rule was associated with an increased frequency of plaintiff success rates at trial, increased jury awards, and larger out-of-court settlements. These increases were significant not just statistically but in absolute terms; for instance, the average judgment in litigated cases increased from $25,190 in cases governed by the American rule to $69,390 in cases governed by the English rule. These results appear to be driven by the case selection effects detailed in the first article. The authors suggest that their results vindicate the proposition that indemnification improves the quality of claims brought, although they admit the possibility of an alternative explanation that low-damage cases are merely being driven away by the higher costs of litigation. They conclude that fee shifting, contrary to the assertions of some legal practitioners, is not necessarily an antiplaintiff policy. Rather, it benefits plaintiffs with high-quality or high-damage claims at the expense of those with low-quality or low-damage claims, and possibly at the expense of defendants. These conclusions, if valid, would explain why the financial advantages expected by the Florida statute’s original proponents did not appear to materialize;
they would also suggest that the statute improved the deterrent effect of civil liability. Whether such an improvement would be worth the increased litigation expenditures it occasioned, however, and whether it would be replicated in other areas of law with different substantive and procedural characteristics from medical malpractice, remain open questions.

Two recent empirical studies emphasize the impact of cost-shifting rules on the dynamics of litigation, and in particular, the timing of settlement. Watanabe (2006) also analyzes Florida data, but from a time period following repeal of that state’s experiment with the English rule. Watanabe studies the effect of the English rule on the timing of settlement and the costs of delayed resolution. He finds that the English rule results in longer and costlier disputes because, with more at stake, parties have a greater incentive to wait longer for new information regarding the merits of the plaintiff’s case before making the decision to settle.

Watanabe uses a methodology that combines econometric estimation and numerical simulation. His analysis proceeds in three steps.

First, he constructs a theoretical model of settlement dynamics. Potential litigants meet to discuss settlement after injury has been suffered, but before a suit has been filed. During this prefilling phase of settlement negotiations, the plaintiff may file suit, thus causing bargaining to enter a second, litigation phase. The pre-litigation phase is time-bound by the statute of limitations; the litigation phase is time-bound by the scheduled court date. The parties begin negotiations with different beliefs about the plaintiff’s chance of prevailing at trial. They identically discount future costs and benefits, and they agree, and are certain regarding, what the plaintiff would recover should she prevail at trial. In each period of each phase of negotiation: (1) new information may arrive upon which the parties update their beliefs regarding the plaintiff’s chance of winning at trial, (2) one (randomly selected) party makes a settlement offer, (3) the other decides whether or not to accept the offer, and (4) in the pre-litigation phase, the plaintiff decides whether or not to accelerate the process by filing suit. Plaintiffs have contingent fee arrangements with their attorneys under which the attorneys pay all costs. Thus, Watanabe’s model is a kind of sequential bargaining model with non-common priors and exogenous learning.

Second, Watanabe estimates his model. To do so, Watanabe uses a data set on medical malpractice insurance claims in Florida for serious injuries or death that were closed between 1985 and 1999. The data set contains detailed information on the time, mode, cost, and terms of settlement, the occurrence and time of filing, and litigant characteristics. Using numerical methods (combined with functional form assumptions), Watanable determines the conditional probability distribution of observable variables that
would result from each possible array of unobserved parameter values. He then estimates such unobserved parameter values by choosing the array of such values that assigns the maximum likelihood to the configuration of observable variables that are actually observed in the data.

Third, Watanabe conducts policy experiments on his so-calibrated model. Of particular relevance to the topic at hand, Watanabe finds that moving to a loser-pays-all allocation of legal fees would delay settlement and increase legal costs. Under the rule modification that Watanabe studies, the plaintiff herself, as opposed to her attorney, pays defendant’s legal costs if the plaintiff loses. For each party, the English rule causes a greater divergence in each party’s payoffs as between prevailing and losing. Watanabe describes the intuition for additional delay this way:

[The] trade-off between [the] potential for learning new information and [the] extra legal cost due to delay determines the equilibrium of the model. [The] value of learning increases with loser-pay-all legal fee allocation because [the] difference in payoff from winning and losing judgment is increased. Hence, settlement timing delays and legal costs increase.

Yoon and Baker (2006) obtain quite different results. They use insurance data on the disposition (including settlement) of filed cases in New Jersey to study the effect of New Jersey’s expansion of its bilateral offer-of-judgment rule to allow the categorical recovery of unlimited attorneys’ fees. (The unilateral federal rule allows recovery of attorneys’ fees only in certain situations.)

Yoon and Baker obtained their data from an insurance company (whose name they are unable to disclose). The cases in Yoon and Baker’s sample consist of suits by non-policyholders against the insurance company for injuries allegedly caused by policyholders. Most suits in the database arise from automobile and homeowners’ insurance policies.

Yoon and Baker employ a “difference-in-differences” approach. Their treatment group consists of New Jersey litigants involved in suits before and after the offer-of-judgment rule was strengthened. Their control group consists of similarly situated litigants in surrounding states that did not experience a similar rule change. Yoon and Baker examine whether differences in relevant variables across the two time periods differ across their treatment and control groups. Importantly, Yoon and Baker argue that the New Jersey rule change that they study – because it was judicial and not legislative – was likely exogenous and not itself a reflection of underlying, unobservable differences between the treatment and control groups.

Yoon and Baker find that the New Jersey rule change reduced time to resolution by roughly 7 percent and reduced the insurance company’s attorneys’ fees by roughly 20 percent. They conclude that “a substantial
cost-shifting mechanism would be an effective means of increasing the efficacy of offer-of-judgment rules.”

Yoon and Baker hypothesize that their results are driven mainly by the effect of the rule change on settlement timing. They suggest that expanding the range of shiftable costs under the bilateral offer of judgment rule caused parties to reach settlement more quickly (though not more often, according to their data), and that this reduction in the duration of litigation, in turn, was the dominant force acting on attorneys’ fees. Why did the rule change reduce time to settlement? Yoon and Baker postulate that, among many conflicting forces, the dominant effects of expanded cost shifting were these: (1) parties were induced to make more generous settlement offers, and (2) parties regarded trial as a less favorable prospect because the increased stakes implied a harmful escalation in spending.

In contrast, Watanabe (2006), discussed above, finds that moving to a loser-pays-all allocation of legal fees delays settlement and thereby increases legal costs. Several factors might explain the difference in the implications for settlement timing as between Watanabe (2006) and Yoon and Baker (2006). One possible factor, of course, is that the rule changes considered in each study are not identical: the expanded cost shifting in Yoon and Baker, but not in Watanabe, is conditional on the proposal and rejection of a settlement offer. However, it is unclear from existing theory what impact this difference might have on settlement timing.

Another possible factor concerns differences in case composition. While Watanabe looks at medical malpractice cases, Yoon and Baker consider cases arising from automobile accidents and property damage. In Watanabe (2006), cases take longer to settle with a cost-shifting rule because, the stakes of the case being greater, parties are willing to wait longer for new information revealed in pretrial process. It is possible that this wait-to-learn dynamic is more pronounced in medical malpractice cases than in auto accident and property damage cases. Plausibly, the information flow in medical malpractice cases is both more substantial and more extenuated – given the more prominent role of scientific and medical expertise.

9. Conclusion
All in all, despite the substantial scholarly and popular attention that the question of indemnity for legal fees has attracted, the number of robust conclusions that can be drawn regarding its consequences are few. Fee shifting does appear to increase legal expenditures per case, in some cases significantly. It also encourages parties with poorly grounded legal claims to settle or to avoid litigating them in the first place, and has a similar effect
on litigants who are averse to risk, regardless of the merits of their cases. Aside from these generalizations, most of the other propositions commonly asserted about fee shifting can be neither verified nor rejected. It is unclear whether fee shifting increases the likelihood of settlement, whether it decreases total expenditures on litigation or total payouts by defendants, or whether on balance it improves incentives for primary behavior. It is even unclear whether fee shifting makes it easier for parties with small meritorious claims to obtain compensation, in light of the increased costs per case that it induces. In this regard, the relative lack of systematic empirical investigation of these questions is particularly lamentable.

In light of this state of affairs, one is tempted to conclude that the amount of scholarly attention directed to this topic exceeds its actual social importance. The continued popular and political interest in fee-shifting rules, however, makes this conclusion problematic. While some support for fee shifting arises from its relative simplicity and its status as the international majority rule, much of its continued appeal undoubtedly stems from its association with deeply held notions of corrective justice – and specifically, from the idea that a party who is determined ex post to be in the right should be made financially whole. Counterarguments based on economic efficiency, or indeed on any ex ante perspective, can never entirely rebut this simple yet powerful intuition.

Whether the English rule is more just than the American rule, or whether its greater fairness justifies its incentive properties, cannot be settled by lawyers or economists alone. The citizenry as a whole must decide whether the principle of full compensation for victorious litigants outweighs the procedural values of providing citizens with an open forum for grievances and an opportunity to be heard, the uncertainty imposed on those who cannot predict the outcome of court decisions, and the political implications of regulating legal fees through a system of bureaucratic oversight rather than through private contract between attorney and client. Moreover, as Prichard (1988) and Hylton (1996) have observed, rules of cost allocation feed back through the selection of cases to influence the development of other areas of substantive and procedural law. Rules that encourage parties to raise relatively innovative claims and defenses help to break down precedent, while rules that penalize risk-taking and novel arguments help to preserve traditional formal categories. Given the pervasive influence of ostensibly procedural rules on substantive outcomes, it may not be possible to separate the policy of fee shifting from deeper questions of what the law should be.
304 Procedural law and economics

Bibliography


Procedural law and economics


Other References

Cases
*Alyeska Pipeline Service Co. v. Wilderness Society*, 421 US 240 (1975)