Contingent fees for personal injury litigation

Patricia Munch Danzon*

This study presents a theoretical analysis of contingent fee and hourly wage contracts for legal services. In contrast to previous analyses, it concludes that, with risk neutral plaintiffs and attorneys, the contingent fee induces the amount of attorney's effort that would be chosen by a fully informed plaintiff who was paying an attorney by the hour. Both the expected gross recovery and the expected attorney's fee are the same under a contingent fee as they are under an hourly fee system. For the risk averse plaintiff, expected utility is unambiguously higher with a contingent fee. Empirically based estimates show that regulation of contingent fees may have significant effects on the number of suits and the size of awards.

1. Introduction

■ Contingent fees are the dominant form of payment for plaintiff attorneys in personal injury litigation in the United States. The typical contingent fee is a predetermined fraction of the award, paid if and only if the plaintiff wins the case. By contrast, defense attorneys are typically paid an hourly wage for time spent, regardless of the outcome of the case.

There is a long tradition of hostility toward contingent fees. They are prohibited in England and Canada and have been singled out for regulation in the United States.[1] The rationale for this traditional hostility is summarized in McKinnon (1964): (1) giving the attorney the right to finance litigation allegedly promotes "nuisance" suits with little legal merit; (2) contingent fees are said to be excessive; and (3) the attorney's stake in the claim creates a conflict of interest with the client which impedes settlement—the attorney is allegedly more prone to gamble.

The traditional view, that contingent fees stimulate "excessive" litigation, differs sharply from the conclusion reached by previous economic analysis. Schwartz and Mitchell (1970) conclude that, relative to the benchmark of an hourly wage contract with a fully informed, risk-neutral client, the contingent fee results in fewer hours per case and consequently lower gross recoveries and lower fees. Accepting this conclusion, Clermont and Currivan (1978) propose replacing the pure contingent fee with a contingent hourly plus percentage fee and Reder (1978) advocates either outright sale of claims by plaintiffs

* The Hoover Institution of War, Revolution, and Peace.

This research was funded in part by Contract No. 600-76-1050 between the Health Care Financing Administration, U.S. Department of Health, Education, and Welfare, and The Rand Corporation. Support was also provided by the Hoover Institution. The views expressed are not necessarily shared by the research sponsors. I would like to acknowledge helpful comments from John Cogan, Howard Marvel, Bridger Mitchell, Charles E. Phelps, W. Craig Riddell, participants in the Law and Economics Seminar at Stanford University, Alvin Klevorick, and two anonymous referees.

[1] Since 1975, 18 states have enacted limits on contingent fees for medical malpractice cases, but hourly fees are unrestricted. The Federal Tort Claims Act limits contingent fees to a flat 25% on claims against the federal government. New York and New Jersey have sliding scale limits, whereby the allowed percentage decreases with the size of recovery.

to attorneys or public provision of a certifier to reconcile the conflicting interests of plaintiff and attorney.

This article reexamines these contradictory beliefs about the effects of a contingent fee relative to an hourly wage contract. Section 2 analyzes the effect of these two polar types of contract on attorney effort per case, and hence the plaintiff's expected gross recovery, and the division of the gross recovery between the net award to the plaintiff and the attorney's fee, assuming risk neutrality. Section 3 introduces risk aversion. Section 4 discusses the effects of contingent fees on the number of claims filed. Section 5 provides rough estimates of the effects of regulation of contingent fees. Section 6 considers the broader issue of the optimal form of contract for legal services, for both plaintiff and defense, and shows why the conclusions of the literature on principal and agent relationships (Harris and Raviv, 1978; Stiglitz, 1975; Shavell, 1979) are not directly applicable to the market for legal services. The final section summarizes the results and draws limited policy conclusions. The analysis is set in the context of personal injury litigation, but applies to other contexts where an agent is paid a percentage commission, contingent on output.

2. Hourly wage and contingent fee contracts with risk neutrality

■ This section presents models of the hourly wage contract (Model A) and two alternative formulations of the contingent fee contract: the Schwartz-Mitchell formulation (Model B) and a reinterpretation (Model C) based on an alternative specification of competition for cases. Risk neutrality of both client and attorney is assumed. The decision to hire an attorney is taken as given.

The following notation is used:

p = probability of a positive recovery by the plaintiff;
A = gross recovery conditional on a positive recovery;[2]
w = wage rate of the attorney if paid by the hour;
L = input of plaintiff attorney hours on the case;
H = input of defense attorney hours;
α = attorney's contingent fee percentage;
$p(L, H)A(L, H)$ = judicial production function, giving the expected award as a function of L and H.

For notational simplicity, this function will be denoted pA.

I adopt the Cournot-Nash assumption of noncooperation between plaintiff and defense, but assume that each side anticipates a payoff function that incorporates the equilibrium response of the other. Thus, in equilibrium each party optimizes against the optimal investment of the other.[3] Denote the optimized inputs $\hat{L}, \hat{H}$. Positive but diminishing returns to plaintiff's attorney's input are assumed:

$$\frac{dp}{dL} \equiv \frac{\partial p}{\partial L} + \frac{\partial p}{\partial \hat{H}}\frac{\partial \hat{H}}{\partial L} > 0; \qquad \frac{dA}{dL} \equiv \frac{\partial A}{\partial L} + \frac{\partial A}{\partial \hat{H}}\frac{\partial \hat{H}}{\partial L} > 0; \qquad \frac{d^2p}{dL^2}, \frac{d^2A}{dL^2}, \frac{\partial^2 pA}{\partial L^2} < 0.$$

□ **Model A: hourly wage contract with risk neutrality.** The plaintiff is assumed to have costless knowledge of the payoff function, $pA(L, \hat{H})$, and the attorney's input, L.

---

[2] "Recovery" and "award" are used interchangeably to refer to court awards and out-of-court settlements. In the risk-neutral model, $A(L, H)$ may be interpreted as the expected value of a probability distribution of awards. With risk aversion, $A(L, H)$ must be interpreted as deterministic.

[3] The noncooperation assumption may be more plausible in litigation than in the traditional duopoly context because in the former the facts and identity of the principals, if not the agents, change with each case.

Decision rule. The plaintiff selects L to maximize his expected net recovery:

$$\underset{L}{\text{Max}}\ \phi = pA - wL, \tag{1}$$

which yields:

$$\frac{d(pA)}{dL} = w. \tag{2}$$

Equation (2) implies that the informed client maximizes his expected net recovery by investing to the point where the attorney's gross marginal product is equal to his hourly wage rate.

□ **Model B: Schwartz-Mitchell (S-M) contingent fee contract with risk neutrality.** The plaintiff is assumed to be ignorant of the payoff function, $pA(L, \hat{H})$, and of L.[4] The unconstrained attorney can therefore choose L to maximize his expected net profit, given that his fee is some fraction, α, of the gross recovery, and his time has an opportunity cost, w, equal to the wage he could earn on other cases.

Decision rule. The attorney selects L to maximize his expected net profit:

$$\underset{L}{\text{Max}}\ \phi = \alpha pA - wL, \tag{3}$$

which yields:

$$\alpha \frac{d(pA)}{dL} = w. \tag{4}$$

Equation (4) implies that the attorney equates *his share* of the gross marginal product to the opportunity cost of his time. Comparing equations (2) and (4), since $\alpha < 1$, equilibrium occurs at a larger gross marginal product. Hence there are fewer hours per case with a contingent fee than with an hourly wage. The plaintiff's probability of winning and the plaintiff's gross recovery are therefore lower.

The S-M conclusions with respect to the effect of the contingent fee on the *net* recovery of the plaintiff and the fee of the attorney depend on specific assumptions about competition in the market for legal services. S-M assume that attorneys compete for cases by bidding down α until hourly earnings on contingent fee cases are reduced to w. Taken alone, this assumption implies that both the plaintiff's net recovery and the attorney's fee are lower with a contingent fee.[5]

The S-M conclusion, that the contingent fee contract fails to maximize the net value of the plaintiff's case because of a suboptimal input of attorney time, is analogous to the traditional view that share contracts in general are inefficient because they induce suboptimal factor inputs. Cheung (1968) has shown that actual share-cropping contracts stipulate factor inputs as well as output shares, thereby circumventing private incentives to invest less than the mutually optimal level.

In the case of contingent fees for legal services, the plaintiff's ignorance of the production function precludes solving the problem of distorted incentives by detailed *ex ante* specification of the attorney's input. The typical contract stipulates only the attorney's share of output and sometimes the payment of filing fees, etc. It is shown in Model *C*

[4] Schwartz and Mitchell (1970) do not consider the role of the defense.

[5] S-M modify this conclusion by introducing a general equilibrium adjustment in the market for legal services. Assuming a fixed number of personal injury and other cases, and a fixed supply of attorney hours, S-M conclude that introducing a contingent fee for personal injury cases unambiguously reduces the attorney's fee (since L and hence w fall) and may reduce the plaintiff's net recovery, depending on whether the reduction in gross recovery (due to lower L) is greater or less than the reduction in the lawyer's fee. If w falls negligibly, because demand for the attorney's time is elastic, the initial conclusion holds—the plaintiff's net recovery is lower with a contingent fee.

below, however, that competition *for* cases and monitoring of outcomes can provide a perfect substitute for fully informed and costless monitoring of inputs.

□ **Model C: reinterpretation of contingent fee contract with risk neutrality.** All assumptions of the S-M model are retained with one exception. Both models require that potential plaintiffs canvass attorneys for the best offer. But whereas S-M assume attorneys bid only on the fee percentage, α, I assume that attorneys compete in terms of the net recovery they offer a client, $pA(1 - \alpha)$. Bidding is in terms of the probability of winning and expected gross recovery as well as the fee percentage.[6] To maximize his chances of getting the case, each attorney bids a p, A, and α that maximize the expected net recovery to the plaintiff, subject to the condition that the attorney covers the opportunity cost of his time. Although the bid is with respect to expected outcome, this implies a commitment of L. It is further assumed that competitive monitoring of outcomes forces the attorney to deliver the implicitly promised effort. The realism of these assumptions is discussed below.

Decision rule. The attorney selects L and α to maximize the expected net recovery of the client, subject to the constraint that the attorney cover the opportunity cost of his time:

$$\underset{L,\alpha,\mu}{\text{Max}}\ \phi = pA(1 - \alpha) + \mu(\alpha pA - wL), \tag{5}$$

where μ is a Lagrange multiplier.

Differentiating with respect to L, α, and μ, respectively, gives

$$(1 - \alpha)\frac{d(pA)}{dL} + \mu\left(\alpha\frac{d(pA)}{dL} - w\right) = 0, \tag{6}$$

$$-pA + \mu pA = 0, \tag{7}$$

$$\alpha pA - wL = 0. \tag{8}$$

From equation (7), $\mu = 1$. Substituting this result in equation (6) yields

$$\frac{d(pA)}{dL} = w. \tag{6a}$$

Equation (6a) defines the attorney's optimum input of hours, $\hat{L}$, which in turn determines the expected award, $\hat{p}\hat{A}$, that the attorney will bid to try to get the case. The optimum $\hat{L}$ equates the opportunity cost of time, w, with the gross marginal product. This is identical to the equilibrium condition under an hourly wage contract, when a perfectly informed client determines attorney input to maximize the net value of his claim (equation (2)).

Although L and pA are the same under the contingent fee contract and the hourly wage contract, the division of the gross recovery into attorney's fee and client's net recovery is different under the two contracts. Equation (8) shows that the *expected* fee, αpA, is equal to the fee that would have been paid under an hourly wage contract, since w and L are invariant. But the *actual* fee on cases won, αA, exceeds the cost of time by a multiple that is the inverse of the optimized *ex ante* probability of winning, $\hat{p}$. On medical malpractice cases with attorney representation, roughly 60% of the claims obtain some positive recovery (Danzon, 1980). Assuming this is correctly anticipated, contingent fees on average will appear to overcompensate attorneys 66% for their time spent on cases won.[7]

---

[6] This type of bidding occurs in real estate markets. The agent suggests a listing price as well as a contingent commission percentage.

[7] $(.6)^{-1} = 1.66$. It is misleading to view contingent fees on cases won as compensating attorneys for time spent on cases lost. With competition, an attorney can only charge a client for the expected hours on *his* case, not for hours spent on other cases he has lost.

Thus, if $p < 1$, the competitive contingent fee exceeds the value of time actually spent because of *ex ante* uncertainty, even if the attorney is risk neutral.[8] The plaintiff's realized net recovery on cases won is therefore lower with a contingent fee than with an hourly wage contract.

A further implication of this model is that the attorney's share, α, provides a measure of the elasticity of output (expected recovery) with respect to effort, $N_{pA,L}$. Multiplying equation (6a) by L/pA and substituting in equation (8) yield:

$$\alpha = \frac{wL}{pA} = \frac{d(pA)}{dL}\frac{L}{pA} = N_{pA,L}. \tag{9}$$

These conclusions differ from those reached by S-M because of different assumptions about the nature of competition. Both models assume that competitive bidding for cases reduces contingent fees to competitive levels.[9] They differ as to whether competition is based on the client's net recovery, $pA(1 - \alpha)$, or is confined to the attorney's share, α. The former is more plausible, because without an estimate of pA, the client cannot evaluate the dollar implications of alternative bids. But granted bidding in terms of net recovery, some mechanism to enforce delivery of the amount bid is clearly key to the invariance conclusion. In the case of contingent fees for real estate agents, the listed price, which is the analogue of the expected recovery, is public information, as is the final sales price. Competitive bidding on list prices and public scrutiny of discrepancies between list and final sale prices should enforce maximization of net return for the client.

In the case of personal injury litigation, however, the attorney's bid price is known only to the client and possibly to a referring attorney. Enforcement depends on concern for future referral business from both sources. The referring attorney is often paid a fee, sometimes some fraction of the recovery. Such a system would, if it were common, provide the necessary enforcement mechanism. There is no systematic evidence of the extent of the referral market (Dietz, Baird, and Berul, 1973; Curran, 1976).[10] Presumably it is more prevalent where the value of technical expertise fosters specialization among attorneys, such as medical malpractice or product liability. Thus, monitoring by attorneys may substitute for monitoring by clients in areas where repeat or referral business by clients is rare and expertise is relatively deficient.

An alternative way to evaluate the models is to consider the plausibility of their implications. First, consider the effect of a change in w. The S-M model implies that equilibrium is restored entirely by a change in α, whereas the alternative model predicts an equilibrating change in $\hat{L}$, but not necessarily any change in α. Crude evidence is inconsistent with the S-M model. Between 1950 and 1980 the numbers of lawyers per capita increased roughly 70% and the relative earnings of lawyers showed substantial fluctuations (Pashigian, 1977), but contingent fee percentages were remarkably stable. The 1957 mean fee percentage for personal injury cases in New York was identical to the 1971 mean fee for a national sample of malpractice cases.

Second, the S-M model implies a unitary elasticity of output with respect to effort, $N_{pA,L}$ since

$$\alpha \frac{d(pA)}{dL} = w = \alpha \frac{pA}{L}.$$

Since α is typically .33, this implies that the attorney's gross marginal product is three

[8] With imperfect competition, the attorney input and gross and net recovery are lower than with a competitive market for legal services. But the outcome is still invariant to the type of contract, provided an attorney's degree of monopoly power is the same, whether he sells his time on an hourly or contingent fee basis.

[9] Assuming equal hourly earnings on contingent fee and hourly wage cases is plausible, given attorney mobility between the two markets, and is supported by the empirical evidence in Dietz, Baird, and Berul (1973).

[10] Concern for reputation mitigates incentives for suboptimal investment, but it does not ensure optimal effort on each case.

times his wage rate. Therefore, the plaintiff would be better off buying legal defense insurance and paying an attorney by the hour, if insurance were available with less than a 200% loading charge. But legal defense insurance has not flourished in the United States as it has in Europe, in part because of the contingent fee alternative in the United States (Pfennigstorf, 1975). Thus, whatever the load factor on legal defense insurance (and it is surely less than 200%), the effective load factor on contingent fee business, due to imperfect monitoring of attorneys, is apparently less.

Although this evidence suggests that the S-M model is inappropriate, it obviously does not prove the perfect monitoring required for the alternative model. Hereafter I use the alternative model to illustrate the polar case of perfect monitoring.

3. Hourly wage and contingent fee contracts in the presence of risk aversion

■ When there is risk aversion, the conclusion that the same outcome results with hourly wages as with contingent fees does not hold. To demonstrate why, it is convenient to assume that L can be specialized to produce either p or A separately, so that the production function can be written $p(L_1, H)A(L_2, H)$, where L_1 denotes plaintiff attorney hours assigned to increasing p and L_2 denotes plaintiff attorney hours assigned to increasing A.[11]

With risk aversion, expected utility replaces wealth as the maximand. Let $U(V)$ denote the plaintiff's utility function at his initial wealth, and let subscripts 1 and 0 denote states in which the plaintiff wins and loses, respectively. Initially, assume that the attorney is risk neutral.

□ **Model D: hourly wage contract with plaintiff risk aversion.**

Decision rule. The plaintiff selects L_1 and L_2 to maximize his expected utility:

$$\underset{L_1,L_2}{\text{Max}}\ \phi = pU[V + A - w(L_1 + L_2)] + (1 - p)U[V - w(L_1 + L_2)]. \tag{10}$$

Maximization of equation (10) with respect to L_1 and L_2 yields:

$$p\frac{\partial A}{\partial L_1}\frac{U_1'}{\bar{U}'} = w \tag{11}$$

$$\frac{\partial p}{\partial L_2}\frac{(U_1 - U_0)}{\bar{U}'} = w, \tag{12}$$

where $\bar{U}' = pU_1' + (1 - p)U_0'$ = expected marginal utility.

□ **Model E: contingent fee contract with plaintiff risk aversion.** Competitive bidding for cases implies maximization of the plaintiff's expected utility, subject to the constraint that the attorney cover the opportunity cost of his time.

Decision rule. The attorney selects L_1, L_2, and α to:

$$\underset{L_1,L_2,\mu,\alpha}{\text{Max}}\ \phi = pU[V + A(1 - \alpha)] + (1 - p)U[V] + \mu[\alpha pA - w(L_1 + L_2)]. \tag{13}$$

Maximization with respect to α implies $\mu = U_1'$. Substituting this result in the first-order conditions for L_1 and L_2 yields:

$$\phi_{L_1} = p\frac{\partial A}{\partial L_1} = w \tag{14}$$

[11] It is easily verified that the invariance conclusion holds in the risk neutral case with this production function.

$$\phi_{L_2} = \frac{\partial p}{\partial L_2}\left[\frac{U_1 - U_0}{\bar{U}'_1} + \alpha A\right] = w. \tag{15}$$

A comparison of Model *D* and Model *E* shows that, in contrast to the risk-neutral plaintiff, the risk-averse plaintiff is affected by the type of contract. Equations (14) and (11) imply that for a *given p,* the net recovery is still maximized with a contingent fee, but not with an hourly wage.[12] But, the effect of plaintiff risk aversion on L_2 and hence on p cannot be signed unambiguously. It depends on the parameters of the production and utility functions.[13]

Although it cannot be shown that the risk-averse plaintiff's expected or realized recovery will necessarily be higher with a contingent fee than with an hourly fee, his expected utility is unambiguously higher with the contingent fee. Consider any L that the plaintiff would choose with an hourly wage contract. The attorney could offer a contingent fee contract with the same implicit L and hence the same expected gross and net recovery, but with an α which narrows the range of possible wealth for the plaintiff and thus increases the plaintiff's expected utility.[14]

The effect of attorney risk aversion on L and hence on pA under a contingent fee contract also depends on the parameters of the production and utility functions. It is clear, however, that for any L, the expected fee must exceed the hourly wage costs ($\alpha pA > wL$) if the attorney is risk averse, whereas if the attorney prefers risk, then $\alpha pA < wL$. The attorney's risk tolerance therefore tends to increase the plaintiff's net recovery from a given gross recovery.

4. Number of suits

■ So far I have considered the investment of effort, given the decision to file suit. This section discusses the effect of risk attitude and type of contract on the decision to file.

□ **Hourly wage.** The risk-neutral plaintiff will hire an attorney on an hourly fee if the expected payoff is positive:

$$pA - wL - C > 0, \tag{16}$$

where A is defined net of any settlement offer obtained without representation, and C is the plaintiff's time and other costs of filing suit.

The risk-averse plaintiff requires that the utility of the gamble exceed the utility of his certain wealth:

$$pU[V + A - wL - C] + (1 - p)U[V - wL - C] \geq U[V], \tag{17}$$

where V is now defined to include any offer obtained without an attorney. Substituting an empirically based estimate of 2.0 for λ in a constant relative risk aversion utility function yields:[15]

$$\frac{V_1}{V} \geq \frac{pV_0 + (1 - p)V_1}{V_0}, \tag{18}$$

where V_1 is the plaintiff's wealth if he wins and V_0 is his wealth if he loses.

---

[12] With risk aversion, $\frac{U'_1}{\bar{U}'} < 1$.

[13] Derivations of the results for the quadratic and constant relative risk aversion utility functions are available from the author.

[14] The α which offers the plaintiff an equal expected net recovery also satisfies the constraint in (13):

$$V + pA - wL = V + pA - \alpha pA \text{ implies } \alpha pA = wL.$$

[15] Brown and Deaton (1972) survey several empirical studies which report values of λ of approximately 2.

In 1976 the average malpractice award was $29,456 (NAIC, 1977). Assuming a one-third fee and $C = 0$, this implies an average net recovery of roughly $20,000.[16] A plaintiff with average risk aversion but with assets of less than $20,000 would be unwilling to hire an attorney on an hourly basis to file a malpractice suit of average value, if $p < .66$. In fact, plaintiffs won in only 41% of the cases. The deterrence to filing is greater: the lower p, the lower V, the greater the aversion to risk, and the higher w or, equivalently, the plaintiff's borrowing costs.

□ **Contingent fee.** With a contingent fee, the risk-neutral plaintiff would file if $pA(1 - \alpha) > C$, whereas for the risk-averse plaintiff, the necessary condition is:

$$pU[V + A(1 - \alpha) - C] + (1 - p)U[V - C] \geq U[V]. \tag{19}$$

Thus, even with a contingent fee, the plaintiff's risk aversion reduces the number of suits filed if the plaintiff's costs are significant.

For the attorney, the necessary condition for accepting a case is that the expected utility of the contingent fee is at least equal to his certain income alternative:

$$pZ(Y + \alpha A) + (1 - p)Z(Y) \geq Z(Y + wL), \tag{20}$$

where $Z(Y)$ denotes the attorney's utility of initial wealth. Thus, the limited truth in the common allegation that the contingent fee stirs up suits is that a risk-preferring attorney would accept a case that would be rejected by a risk-neutral attorney and would charge an expected fee less than the cost of his time. Ignoring risk preference and assuming attorney risk neutrality is the norm, equation (20) reduces to:

$$\alpha pA - wL \geq 0, \tag{20a}$$

which may be compared with equation (16), the necessary condition for a fully informed risk-neutral plaintiff to hire an attorney with an hourly fee. Let $C = \delta pA$. Then, if $\alpha < 1 - \delta$, some suits that would be brought by a fully informed, risk-neutral client will be rejected by a risk-neutral attorney with a contingent fee.

5. Constraints on contingent fees

■ The purpose of this section is to show the effects of regulatory and minimum time constraints on the fee percentage. Contrary to the common belief that fee controls simply redistribute income from attorney to client, regulation of α will induce a reduction in L and hence in p and A, unless unconstrained fees yield rents at the margin. The magnitude of these effects depends on whether the regulations are binding and on the elasticity of pA with respect to L. To determine how constraining the ceilings are, we first present a theoretical discussion and empirical evidence on the determinants of unregulated fees and then provide rough estimates of the effects of regulation.

□ **Determinants of α.** Thus far I have considered the production function for a specific case and specific attorney. In general, the production function and therefore the equilibrium α depend on shift parameters for such factors as the stakes of the case, the evidence for the plaintiff, the quality of the attorney, and the stage of disposition (settlement or trial).

If some minimum time input, $\bar{L}$, is required to handle any case, equation (5) must be solved subject to the additional constraints: $L \geq \bar{L}$ and $\alpha \leq 1$. If the constraints are inconsistent, because $w\bar{L}/pA(\bar{L}, \hat{H})$ implies $\alpha > 1$, the case will not be brought. If

[16] This overstates the expected net recovery of a "typical" claimant to the extent that the mean award exceeds the median, and for the risk-averse plaintiff with an hourly wage, the optimized A is less than this actual A realized with a contingent fee. Offsetting this upward bias, $\frac{1}{3}A$ is an upper bound for wL.

$\alpha \leq 1$ is nonbinding, but $\bar{L}$ exceeds the optimum when there are no constraints, $\hat{L}$, then α must exceed $\hat{\alpha}$, the value in the unconstrained case, if the increment in time produces less than a commensurate increase in pA. Taking the total differential of constraint (8), the elasticity of α with respect to L is

$$N_{\alpha,L} = \frac{d \ln \alpha}{d \ln L} = 1 - \frac{d \ln pA}{d \ln L} = 1 - N_{pA,L}.$$

If $N_{pA,L} = .33$, then $N_{\alpha,L} = .66$. It follows that if the optimum L subject to the constraint is twice the optimum without the constraint, α must increase from .33 to .55. Thus, minimum time constraints may effectively bar cases with low stakes or weak evidence (low $pA(\bar{L}, \hat{H})$). If they are brought, such cases will tend to have high α.

The limited data available on contingent fees provide some evidence on how the characteristics of the case affect α. A 1971 survey of contingent fees on medical malpractice cases shows that for respondents using a fixed percentage, the mode was 33⅓%, the mean and median were 36%, with a range of 25 to 50%. For those using a sliding scale, modal percentages were 33⅓% before trial and 40% through trial and appeals, but the range ran from 20% before trial to 50% through appeals. The increasing percentage for trial implies that the positive effect on α of higher L and lower p dominates the negative effect of higher A at trial.[17] The only evidence of the relation between α and A (a proxy for the stakes) is from a study of contingent fees for 3,000 personal injury cases in 1957 in New York City. The average fee ranged from 41% on cases under $1,000 to 29% on cases over $25,000, with an overall mean of 36% (Franklin, Chanin, and Mark, 1961).[18] This suggests that $\hat{L}$ increases less than in proportion to the stakes, possibly because of binding fixed costs on small cases.

□ **Effects of regulation of contingent fees.** The effect of regulation depends on the extent to which it is binding. The Federal Tort Claims Act ceiling of 25% is likely to be binding in most cases. The California medical malpractice sliding scale yields an average percentage of less than one-third on cases over $140,000.[19] The maximum allowed on an award of $1 million would be 14%, roughly one-third of the typical free market fee for a case taken to trial.

To obtain a crude estimate of the effect of binding controls, the condition of competitive equilibrium, $\alpha = wL/pA$, can be used to solve for the reduction in L, and hence the reduction in pA, necessary to accommodate an imposed reduction in α:

$$N_{pA,\alpha} = N_{pA,L}N_{L,\alpha} = N_{pA,L}[1 - N_{pA,L}]^{-1}.$$

If $N_{pA,L} = .33$, then $N_{pA,\alpha} = .5$. Thus, if α is cut by 50%, the equilibrium adjustment of L implies that pA will fall by 25%.

---

[17] If the attorney were sure of obtaining a settlement, S, equal to the expected court award, then the ratio of the fee percentage for trial, α_t, relative to that for settlement, α_s, is a measure of incremental time input required in trial relative to settlement:

$$\alpha_s S = \alpha_s pA = wL_s \qquad \alpha_t pA = wL_t \qquad \therefore \frac{\alpha_t}{\alpha_s} = \frac{L_t}{L_s} \approx \frac{.4}{.33} = 1.2.$$

If settlement is uncertain, or $S < pA$, or the attorney prefers risk, α_t/α_s understates L_t/L_s.

[18] In this sample, the fee percentage does not vary systematically by stage of disposition for a given size of recovery.

[19] The California sliding scale is 40% on the first $50,000, 33% on the next $60,000, 25% on the next $100,000, and 10% on any amount over $200,000 (California State Assembly Bill AB1, 1976). The average award in California in 1976 was $31,508. Unless the limits are revised, inflation will bring an increasing number of cases within the binding limits.

This rough calculation implies that regulation of contingent fees may significantly reduce the plaintiff's probability of winning and the plaintiff's gross recovery. Sharply decreasing sliding scales will affect cases with large stakes. Ceilings that make no distinction by stage of disposition will discourage litigation to verdict and appeals. Flat ceilings are most likely to be binding on cases with low stakes or weak evidence for the plaintiff, and on high quality lawyers. The reduction in expected payoff will reduce the number of claims filed. There may also be an indirect, dynamic effect on claim frequency and size of awards if the proplaintiff trend in rules of compensable damages and liability is slowed, as a result of reducing the incentive to litigate to verdict. Positive settlement offers will be made less frequently and for lower amounts.

There is some evidence to support these predictions. We estimate that fee ceilings for medical malpractice cases reduced average settlement size in 1976 by 9%, increased the proportion of cases dropped without payment from 43 to 48%, and reduced the proportion of cases litigated to verdict from 6.1 to 4.6% (Danzon and Lillard, 1981).[20] This suggests that unconstrained fees do not convey rents at the margin and that controls reduce not only fees but also compensation to plaintiffs.

6. Optimal forms of contract

■ I have shown that a contingent fee based on output dominates an hourly wage for a risk-averse plaintiff who hires a risk-neutral attorney, and I have argued that market selection will tend to eliminate risk-averse attorneys from the contingent fee business. I do not attempt a rigorous proof that the contingent fee is superior to all alternatives, but in this section I offer some preliminary thoughts on alternative contracts proposed by others and on contingent fees for the defense.

There is a growing literature on the role of incentives, risk, and information in determining optimum employment contracts. For example, Harris and Raviv (1978) consider a model where output depends on an agent's effort and some random factor, θ. They show that (1) if the probability distribution of θ and the production function are known and θ is observable *ex post,* then the optimal contract is independent of the agent's effort and depends solely on θ; (2) if θ is not observable, but the agent is risk neutral, then the optimal agent's fee depends only on output. Applying this model to legal services, even if all components of θ were observable *ex post,* a contract which depends only on θ is unlikely to be optimal because the plaintiff is ignorant of the production function and probability distribution of θ, *ex ante.* If attorneys are risk neutral, then the optimal contract would require outright sale of the claim, which is infeasible for reasons discussed below.

Stiglitz (1975) considers the optimal choice of piece and time rates in a similar context. He shows that if the employee is risk neutral, the optimal contract pays each worker a piece rate equal to his marginal product. The predominance of the contingent fee, which is a pure piece rate, is not a simple case of the Stiglitz model, because the fee is a percentage determined *ex ante* rather than the marginal product determined *ex post.* Because the client cannot distinguish the attorney's marginal product from the effect of random factors, a Stiglitz piece rate is not feasible.

Shavell (1979) shows that if a risk-averse principal employs a risk-neutral agent, the optimal agent's share is 1, that is, outright sale of the claim to the agent. In practice α never exceeds .5. The obstacles to the attorney's acquiring more than a 50% share are both legal and pragmatic. The common law doctrines of maintenance and champerty restrict an attorney's right to acquire a financial interest in a case.[21] Even in the absence

[20] These are counterfactual estimates, comparing actual outcomes with predicted values, had fee ceilings not been adopted in states which in fact adopted them.

[21] McKinnon (1964) cites evidence that courts would not enforce a fee in excess of 50%.

of legal restriction, however, outright sale is unlikely to be optimal since the value of a personal injury claim depends not only on the attorney's effort but also on the behavior of the plaintiff. Outright sale eliminates the plaintiff's incentives to maximize the value of the claim. At the limit, the profit-maximizing plaintiff could capture twice the value of the claim by selling it first to his attorney for its expected value and then selling his services to the defense to save them that expected value.[22] Clearly, the plaintiff's moral hazard is minimized by the hourly wage contract, maximized by outright sale, and reduced but not eliminated by the sharing arrangement of the contingent fee.

Designing a form of contract that provides efficient marginal incentives when output depends on the effort of both principal and agent is a complex problem not addressed here. The fact that plaintiffs overwhelmingly select contingent over hourly fees suggests that the costs, in terms of plaintiff moral hazard, are more than offset by the gains. I have explicitly considered only risk shifting. Clearly, the plaintiff's ignorance of $pA(L, H)$, monitoring costs, and higher borrowing costs for the plaintiff than the attorney all enhance the superiority of the contingent over the hourly fee.

Finally, Clermont and Currivan (1978) propose a contingent hourly wage plus some fraction of the recovery. This ignores the plaintiff's costs of monitoring L and, as proposed, has no mechanism to ensure competitive returns to attorneys.

The factors that lead to the dominance of the contingent fee for the plaintiff are mitigated when the defendant is an insurance company. Risk neutrality and comparable borrowing costs of principal and agent may be assumed; repeat business permits learning about the production function and reduces monitoring costs. But these differences do not apply when the defendant is an individual. This suggests that a crucial factor favoring hourly fees for defense attorneys may be the difficulty of monitoring the output of the defense attorney. The defense attorney's output is the reduction in payoff relative to what it might have been without his effort, which is unobservable. Thus the difficulty of measuring output precludes an output-based fee on the defense side.[23]

7. Conclusions

■ This article has shown that, given certain assumptions about the nature of competition, the contingent fee system induces the amount of attorney effort that would be chosen by a fully informed, risk-neutral plaintiff who was paying an attorney by the hour. With risk neutrality, attorney effort, expected outcome, and expected fee are identical under the contingent fee and hourly wage contracts. The actual fee realized on cases won is higher with the contingent fee, even if the attorney is risk neutral. Although the competitive requirements for the invariance conclusion may not be fully met in the market for legal services, there is evidence that the traditional model severely overpredicts the extent of suboptimal attorney input. With plaintiff risk aversion, the plaintiff's expected utility is higher with a contingent fee. But the probability of winning and the gross recovery may be higher or lower with a contingent fee depending on the parameters of the production and utility functions. Rough estimates suggest that regulation which reduces the contingent fee percentage by 50% will lead to a 25% reduction in the expected gross recovery.

If the benchmark of optimal expenditure on litigation is that which would be chosen by fully informed, risk-neutral plaintiffs, then regulation or prohibition of contingent fees

[22] Such a scenario is not totally fanciful. The possibility of profit from changing sides is demonstrated by so-called "Mary Carter" agreements, whereby one of several defendants settles secretly and then works for the plaintiff against the other defendants in court. The secret settlement provides that the final payment by the settling defendant is inversely related to the award obtained against the remaining defendants.

[23] The argument, that contingent fees prevail on the plaintiff's side because payment of an hourly wage could not be enforced in the event of no recovery, is surely spurious. Payment could be extracted before performance, as is common for most fee-for-service professional services.

will, if effective, result in suboptimal investment in pursuing claims. A rigorous analysis of whether this benchmark and hence the volume of litigation induced by contingent fees is too high, given the overall social objectives of the tort system, is beyond the scope of this article. Ordover (1978, 1979) has shown that when postaccident information about negligence is imperfect and plaintiffs face positive costs of bringing suit, the private incentive to litigate is insufficient. Essentially, the private calculus ignores the social benefit of enforcing tort standards. But this conclusion neglects costs, such as litigation costs of the defense and costs of operating the courts, imposed by the plaintiff. Only if these noninternalized costs exceed the noninternalized benefits does the contingent fee impose excessive costs on society.

References

BROWN, A. AND DEATON, A. "Surveys in Applied Economics: Models of Consumer Behavior." *Economic Journal,* Vol. 82 (December 1972).

CHEUNG, S.N.S. "Private Property Rights and Sharecropping." *Journal of Political Economy* (1968).

CLERMONT, K.M. AND CURRIVAN, J.D. "Improving on the Contingent Fee." *Cornell Law Review,* Vol. 63, No. 4 (April 1978).

CURRAN, W.J. "How Lawyers Handle Medical Malpractice Cases." National Center for Health Services Research, U.S. DHEW Publication No. (HRA) 76-3152.

DANZON, P.M. *The Disposition of Medical Malpractice Claims.* Santa Monica: The Rand Corporation, R-2622-HCFA, July 1980.

——— AND LILLARD, L.A. "Settlement out of Court: The Disposition of Medical Malpractice Claims." Hoover Working Paper E-82-5. Forthcoming in *Journal of Legal Studies.*

DIETZ, S., BAIRD, B.C., AND BERUL, L. "The Medical Malpractice Legal System" in Appendix to the *Report of the Secretary's Commission on Medical Malpractice,* Washington, D.C.: Department of Health, Education, and Welfare, 1973, pp. 73–89.

FRANKLIN, M.A., CHANIN, R.H., AND MARK, I. "Accidents, Money and the Law: A Study of the Economics of Personal Injury Litigation." *Columbia Law Review,* Vol. 61, No. 1 (January 1961).

HARRIS, M. AND RAVIV, A. "Some Results on Incentive Contracts." *American Economic Review,* Vol. 68, No. 1 (March 1978).

MCKINNON, F.B. *Contingent Fees for Legal Services.* Chicago: Aldine Publishing Company, 1964.

NATIONAL ASSOCIATION OF INSURANCE COMMISSIONS. *Malpractice Claims,* Vol. 1, No. 4 (May 1977).

ORDOVER, J. "Costly Litigation in the Model of Single Activity Accidents." *Journal of Legal Studies,* Vol. 7, No. 2 (1978).

———. "On the Consequences of Costly Litigation in the Model of Single Activity Accidents: Some New Results." *Journal of Legal Studies,* Vol. 10, No. 2 (1981).

PASHIGIAN, P. "The Market for Lawyers: Determinants of the Demand for and Supply of Lawyers." *Journal of Law and Economics,* Vol. 20, No. 1 (April 1977).

PFENNIGSTORF, W. *Legal Expense Insurance.* Chicago: American Bar Foundation, 1975.

REDER, M.W. "Contingent Fees in Litigation with Special Reference to Medical Malpractice" in S. Rottenberg, ed., *The Economics of Medical Malpractice,* Washington, D.C.: American Enterprise Institute, 1978.

SCHWARTZ, M.L. AND MITCHELL, D.J.B. "An Economic Analysis of the Contingent Fee in Personal Injury Litigation." *Stanford Law Review,* Vol. 22 (1970).

SHAVELL, S. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell Journal of Economics,* Vol. 55, No. 1 (Spring 1979).

STIGLITZ, J.E. "Incentives, Risk and Information." *Bell Journal of Economics,* Vol. 6, No. 2 (Autumn 1975).

WILSON, R.B. "On the Theory of Syndicates." *Econometrica* (January 1968).