ISSN 1045-6333

CONTINGENT FEES AND AGENCY COSTS

Bruce L. Hay

Discussion Paper No. 162

6/95
Revised 7/95

Harvard Law School
Cambridge, MA 02138

The Program in Law and Economics is supported by a grant from the John M. Olin Foundation.

CONTINGENT FEES AND AGENCY COSTS

Bruce L. Hay*

Abstract

This paper examines the operation of ordinary linear contingent fees in a model of litigation in which the recovery on a claim is a function of the lawyer's efforts. The object of the paper is to analyze the linear fee that maximizes the client's welfare in the presence of attorney moral hazard. The paper identifies the optimal fee as the one that minimizes two agency costs: underinvestment in the claim, and attorney rents. The paper also attempts to identify how the optimal linear fee varies with different case characteristics.

---

*Harvard Law School. My thanks to Scott Altman, Lucian Bebchuk, Victor Brudney, Christine Jolls, Louis Kaplow, Avery Katz, Bob Mnookin, David Rosenberg, Steve Salop, Steve Shavell, and workshop participants at Georgetown, Harvard, Yale, and the 1995 ALEA Annual Meeting (Berkeley) for their comments.

CONTINGENT FEES AND AGENCY COSTS

Bruce L. Hay*

A basic fixture in American litigation is the fee arrangement that gives the plaintiff's lawyer a set fraction -- say, one-third -- of the amount recovered from the defendant. It is apparently the most common method of financing many types of lawsuit brought by individuals.[1] As a result, it has attracted the attention of many observers and regulators of the legal profession in recent years. Despite the arrangement's pervasiveness and the attention it has received, however, a central problem concerning this simple fee arrangement remains relatively unexplored.

The problem is identifying the right contingent fee for a given case. Consider the following hypothetical case: a plaintiff, injured in an accident, hires a lawyer to sue for damages, agreeing to pay the lawyer fraction F of any recovery. The lawyer invests A dollars' worth of time and effort in the case, and the court awards the plaintiff B dollars in damages, from which the lawyer subtracts fraction F as her fee.[2] Is F the right fraction? What makes it better, or worse, than some other fraction G? What, in other words, characterizes the optimal contingent percentage fee for the plaintiff's case?

In addressing this problem, I have two main objectives. The first is to identify the formal

---

*Harvard Law School. My thanks to Scott Altman, Lucian Bebchuk, Victor Brudney, Christine Jolls, Louis Kaplow, Avery Katz, Bob Mnookin, David Rosenberg, Steve Salop, Steve Shavell, and workshop participants at Georgetown, Harvard, Yale, and the 1995 ALEA Annual Meeting (Berkeley) for their comments.

[1]See, for example, Deborah R. Hensler et al., Compensation for Accidental Injuries in the United States 135 (RAND Institute for Civil Justice 1991).

[2]For clarity of reference, the female pronoun is used to designate the lawyer, while the male pronoun is used to designate the client.

properties of the optimal percentage fee, defined in terms of maximizing the client's net expected recovery in litigation. The second is to determine how the optimal fee varies with different case characteristics. To examine these matters, I use a simple principal-agent model of litigation.

The model focuses on the class of cases -- presumably numerous -- in which the success of the plaintiff's claim depends on the quality and quantity of effort his lawyer invests in the litigation. In this set of cases, the lawyer's fee has a dual impact on the plaintiff's recovery. On the one hand, it obviously determines *the division, between lawyer and client, of the gains (or losses) from the litigation*; taking the proceeds from the litigation as given, the fee determines how the lawyer and client split these proceeds. On the other hand, the fee also determines *the magnitude of the gains (or losses) from the litigation*: the proceeds from the litigation are not given, but are a function of lawyer effort; and the lawyer's investment of effort in the case depends on the reward she anticipates getting from the client. Identifying the optimal percentage fee requires, at bottom, understanding the relation between these two effects.

This problem has received surprisingly little systematic attention. Previous work has shown that percentage fees generally make the lawyer an imperfect agent of the client, and has demonstrated the theoretical superiority of alternatives to percentage fees.[3] But given the

---

[3]The first paper to analyze the diverging interests of lawyer and client under the linear fee was Murray L. Schwartz & Daniel J. B. Mitchell, An Economic Analysis of the Contingent Fee in Personal-Injury Litigation, 22 Stan. L. Rev. 1125 (1970), though analogous sharing arrangements had previously been analyzed in other context (such as sharecropping, in which the landowner takes a fixed fraction of the farmer's output).

The principal alternatives to the simple linear contingent fee that have been proposed and/or used are the following:

(1) Complex nonlinear contingent fees, in which the lawyer's percentage depends on the amount of work done or the amount recovered. See, for example, Calif. Code § 6146 (requiring the percentage collected to vary with the amount recovered); Geoffrey P. Miller, Some Agency Problems in Settlement, 16 J. Legal Stud. 189, 201 (1987) (describing contracts in which the lawyer's percentage depends on the stage of litigation in which the suit is resolved); Kevin M. Clermont & John D. Currivan, Improving on the Contingent Fee, 63 Cornell L. Rev. 529 (1978) (proposing contingent hourly fees combined with percentage bonuses).

(2) Hybrid contingent fee in which the lawyer both takes a fraction of the recovery *and* makes a fixed side payment

pervasiveness of percentage fees in practice, it is worth trying to identify the features of the optimal percentage fee, without worrying about the additional (important) issue of what type of fee might do a better job.[4] Throughout my analysis, therefore, I will focus exclusively on percentage fees in which the attorney receives a simple linear fraction of the award.

My topic should be clarified in one other respect. My objective is to analyze the fee that is *privately* optimal, not *socially* optimal. I focus on the fee that maximizes the client's welfare in litigation (subject to the constraint that the lawyer will act to maximize *her own* welfare). But litigation generates external costs and benefits, including administrative costs and deterrence of certain activities. There is no reason, in general, to suppose that the fee that maximizes the client's net expected recovery is also the fee that generates optimal deterrence from a social standpoint.[5] But since minimizing principal-agent conflicts between lawyer and client is often taken to be an important goal , I focus on that problem without worrying about the complex problem of social optimality.

The article is organized as follows. Section I describes the central features of the model, which are (1) positive but diminishing returns to the claim from attorney effort, and (2) attorney moral hazard. Section II uses the model to identify the formal properties of the optimal

---

to the client. This approach encompasses the possibility of having the lawyer simply purchase the claim outright from the client (thereby acquiring a 100 percent share of the claim). See, for example, Marc J. Shukaitis, A Market in Personal Injury Tort Claims, 16 J. Legal Stud. 329 (1987) (arguing for allowing plaintiffs to sell claims to lawyer); Daniel L. Rubinfeld & Suzanne Scotchmer, Contingent Fees for Attorneys: An Economic Analysis, 24 RAND J. Econ. 343 (1993) (presenting a more general model of hybrid contingent fees).

[4]This is especially true because some of the proposed alternatives are not feasible under existing law. Hybrid arrangements (see note __ above) in which the client a fixed amount, regardless of whether the claim succeeds, probably run afoul of prohibitions on champerty and maintenance. See Shukaitis, supra note ___.

[5]See Steven Shavell, The Social versus the Private Incentive to Bring Suit in a Costly Legal System, 11 J. Legal Stud. 333 (1982); Louis Kaplow, The Value of Accuracy in Adjudication , 23 J. Legal Stud. 307, 338-42 (1994).

percentage fee. Section III introduces a more specific model of the attorney's production function in order to derive some comparative-static results on how the optimal fee varies with differences in the production function. Section IV examines some of the consequences of the analysis for different types of fee regulation.

I. THE MODEL

A. Assumptions

The basic model we will analyze rests on three central assumptions.

1. Claim Value and Lawyer Effort

Our focus is on cases in which the expected value of a claim depends, in large part, on the efforts of the plaintiff's attorney -- the quantity and quality of work she puts into investigation, discovery, and so forth.[6] We will assume that additional effort generally produces positive but diminishing returns. Formally, let

x = Level of effort invested in a claim; $x \geq 0$;

$w(x)$ = Resulting expected value of the claim; $w' > 0$, $w'' < 0$.

The attorney's efforts may affect the probability of recovery, the amount of recovery, or both. We assume that some uncertainty attends the outcome of a case, so that w represents the expected value of a range of possible outcomes. In addition, we assume there are no fixed costs associated with the litigation, meaning that the net return to lawyer investment is positive for all levels of

[6]Obviously this can only be true up to a point. Many factors bearing on the claim's strength are beyond the lawyer's power to control during the litigation, such as the content of available evidence, existing legal materials, and the lawyer's own level of skill and experience.

investment.[7]

2. Attorney Control

We also assume the lawyer (rather than the client) has effective control of the litigation. Informational barriers are assumed to prevent the client from monitoring, or specifying contractually, the the quality of the attorney's performance except in the broadest terms. For purposes of the model, this means the lawyer is free to select the level of effort she puts into the case.

3. Cases Go to Trial

To keep the analysis tractable, we will ignore the complications associated with the choice between trial and settlement.[8] Instead, we will simply assume that all cases go to trial. The assumption is not so unreasonable as it may sound. While it is true that most cases settle, it is plausible to suppose that in most cases the settlement amount reflects the expected value the claim *would* have if it went to trial. As a result, the optimal percentage fee in a world in which cases settle is not very different from the optimal percentage fee in a world in which all cases go to trial.

B. The Client's Problem

In this setting, the interaction between the plaintiff and his lawyer essentially takes the form of a game depicted in Figure 1. Initially, the client hires the lawyer to work under a contingent fee that awards the the lawyer a fixed fraction of the recovery. The model will will use the following notation for the fee:

---

[7]More precisely, let if $\overline{x}$ be some positive value of x, and let $\ddot{x} = \overline{x}/2$. To say that there are no fixed costs means that, if $w(\overline{x})$-$\overline{x}$ is positive, so is $w(\ddot{x}) - \ddot{x}$.

[8]The difficulty of introducing that choice is that it may generate two lawyer production functions, one affecting the settlement value of the claim and one affecting its trial value, which would have to be analyzed simultaneously.

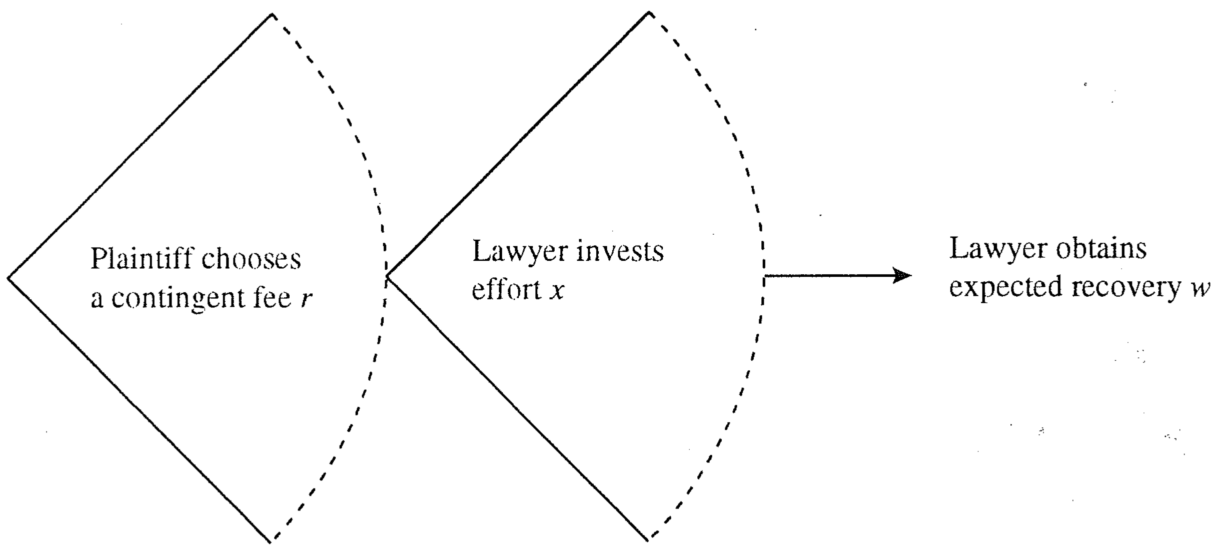

FIGURE 1. — Sequence of events in the model.

r = Fraction of the recovery given to the lawyer pursuant to the contingent fee arrangement; $0 < r < 1$.

After entering into the contract, the lawyer chooses a level of effort x to invest in the case, yielding an expected recovery of w. The client's expected payoff from the game is then given by

$$(1-r)w. \qquad \textbf{(1)}$$

In analyzing the game, our objective is to identify the fee that maximizes this expression.[9] What makes this problem difficult is that the lawyer's fee affects both terms in the expression; it determines not only the client's relative share $(1-r)$ of the recovery, but also the expected value (w) of the recovery. This is because the choice of effort is left up to the lawyer, whose objective is to maximize *her own* payoff, which is given by

$$rw - x. \qquad \textbf{(2)}$$

Our task, therefore, is to find the value of r that maximizes expression **(1)**, subject to the constraint that the lawyer will choose the effort level that maximizes expression **(2)**.[10]

It is worth emphasizing here that we are interested in identifying the *best* percentage fee for the client in this game, rather than in predicting the *actual* fee that the lawyer and client will agree to. Another way of putting this is that we want to see what value of r would emerge in a perfectly competitive market -- including full information and free entry -- in which lawyers and clients can contract only over the value of r.[11] We put to one side imperfections (such as

---

[9] We assume, therefore, that the client is risk-neutral. The same assumption is made about the lawyer.

[10] An additional constraint is that expression **(2)** must be positive, or the lawyer will not take the case. As we will see, this constraint is not binding (provided that the lawyer is free to choose her level of effort and that there are fixed costs associated with the litigation), since rw-x is then positive for *any* value of r.

[11] Of course, in a truly perfect competitive market, which included costless enforcement of contracts, the contract between lawyer and client would not be so limited; it would would either use a different compensation device, or would (enforceably) specify the amount of effort the lawyer had to exert.

asymmetric information) that might in practice lead the client to agree to a suboptimal fee arrangement. Our question, in effect, is what fee an omniscient social planner would have the client give the lawyer in the first stage of the game, if the planner's objective were to maximize the client's net expected recovery.

To my knowledge, this game has not been analyzed in previous work. There is of course a well-developed theoretical literature on principal-agent contracts in general,[12] as well as an extensive literature on contingent fees for legal services.[13] So far as simple linear percentages are concerned, however, both bodies of work are confined to (a) showing that such fee arrangements do not properly align the agent's incentives with those of the principal (the lawyer tends to underinvest in the litigation), and (b) identifying superior alternatives to simple linear percentages. So far as I am aware, no one has examined what the optimal linear percentage is, taking as given that such an arrangement must be used (and that the client has no way of controlling the lawyer's input).[14]

---

[12]See, for example, Sanford J. Grossman & Oliver D. Hart, An Analysis of the Principal-Agent Problem, 51 Econometrica 7 (1983); Steven Shavell, Risk Sharing and Incentives in the Principal and Agent Relationship, 10 Bell J. Econ. 55 (1979); Paul Milgrom & John Roberts, Economics, Organization & Management ch. 7 (1992).

[13]See, for example, the works cited in note ___ above.

[14]Patricia Danzon has analyzed the optimal percentage fee in contracts in which *the lawyer can commit herself (at the time the contract is entered into) to exert a specified level of effort* -- so that the contract between lawyer and client effectively contains two terms, the fee and the amount of effort. See P.M. Danzon, Contingent Fees for Personal Injury Litigation, 14 Bell J. Econ. 213, 216-17 (1983). In her model, the lawyer commits to a given level of effort by promising the client a given recovery; the commitment is enforced by reputational bonding. This assumes away a crucial aspect of the problem; namely, the task of creating appropriate *incentives* for the lawyer when there is no direct way (such as a contract specifying her effort level) of *controlling* her handling of the litigation. The problem of finding the optimal contingent fee is is most interesting and important when we assume -- as we do in what follows -- that the lawyer's choice of effort is a decision made at the margin, after the fee agreement is entered into.

A focus on marginal investment decisions also distinguishes the present paper from James D. Dana & Kathryn E. Spier, Expertise and Contingent Fees: The Role of Asymmetric Information in Attorney Compensation, 9 J. L. Econ. & Org. 349 (1993), which examines the use of contingent fees in settings where the lawyer has better information than the client about the quality of the client's case, where quality is independent of lawyer effort. In their model, there is only a single level of effort available to the lawyer, so inducing effort at the margin is not an issue.

II. THE AGENCY COST TRADEOFF

A. Two Sources of Agency Cost

Consider the maximum possible value the client's claim can have, net of litigation costs -- the total surplus generated by investing in the claim up to the point at which the marginal cost of litigation equals the marginal return to the claim. This surplus is given by

$$\hat{w} - \hat{x}, \tag{3}$$

where $\hat{x}$ denotes the ideal level of investment, and $\hat{w}$ denotes the resulting expected recovery on the claim. This is the amount, in expected terms, that the client would reap if he could represent himself as skilfully (and at the same cost) as can the lawyer.[15] In hiring a lawyer to represent him instead, the client wants to reap as much of this potential surplus as possible.

There are two basic sources of agency cost the client needs to worry about in selecting a percentage fee. The first concerns investment in the claim. As is well known, the lawyer who owns only a fractional interest in the claim will invest less in the claim than someone who owned the entire claim. The greater this shortfall in investment -- captured by the expression

$$(\hat{w}-\hat{x}) - (w-x) \tag{4}$$

-- the smaller the "pie" available for division between lawyer and client.

The second source of agency cost concerns the division of the pie itself. For the client to capture the full surplus generated by the litigation, the lawyer must take away, in expected terms, no more than her investment. (She may earn ex post rents to offset the risk of nonrecovery; but ex ante, she should earn only her opportunity cost, or else the client winds up with less than the

---

[15]An equivalent formulation is to assume that client and lawyer could, with zero transaction costs, write a contract obliging the lawyer to maximize the net value of the claim.

full surplus.) The lawyer's total expected return from the case is given by

$$rw - x. \tag{5}$$

This figure cannot, in equilibrium, be less than zero; but the more it exceeds zero, the smaller the client's share of the pie becomes.

The performance of a fee can be judged by its effect on the sum of these two agency costs -- underinvestment and lawyer rents. For simple algebra shows that this sum, given by

$$[(\hat{w}-\hat{x}) - (w-x)] + [rw - x],$$

is equivalent to the expression

$$(\hat{w}-\hat{x}) - (1-r)w,$$

which is simply the gap between the potential surplus in the case and the client's net expected recovery under the fee. The client's task, therefore, is to find the fee that minimizes the sum of these agency costs

The difficulty he faces is that using a linear percentage fee, he cannot minimize both at the same time. The two types of agency cost move, in effect, in opposite directions. On the one hand, *the lower the fee -- the further it is from 100% -- the greater the investment shortfall.* Differentiating **(4)** with respect to r and applying the chain rule gives

$$\frac{dx}{dr}(1 - w'). \tag{6}$$

But our assumptions that the return to effort is positive but diminishing and that the lawyer acts to maximize her own welfare ensure that $w' > 1$ and that $dx/dr > 0$.[16] It follows that **(6)** is negative,

[16]The lawyer chooses x to maximize rw-x. The first-order condition of the solution is that $rw' = 1$, which

meaning the investment shortfall grows as the fee drops. Intuitively, the smaller the attorney's stake, the less she invests in enlarging the pie.

On the other hand, *the greater the attorney's fee -- the further it is above zero -- the greater her rents.* Differentiating **(5)** with respect to r yields

$$(rw' - 1)\frac{dx}{dr} + w. \tag{7}$$

But our assumptions ensure that rw'-1 is nonnegative, implying that **(7)** is positive. Thus, the lawyer's rents, increase with the fee. The intuition here is that, because returns to effort are diminishing, inducing the lawyer to work more *at the margin* requires raising her fee; but then she is in effect overcompensated for her earlier investment (since, by assumption, she would have made that earlier investment even without the higher fee).

The client, then, faces an unavoidable tradeoff between minimizing lawyer rents and minimizing the investment shortfall. The fee that minimizes their sum is the fee that satisfies the following equality:

$$r = \frac{(w')^2}{(w')^2 + w(-w'')}. \tag{8}$$

As is clear from this expression, which we derive in the Appendix, the location of the optimal fee will vary from case to case, depending on the slope and magnitude of the lawyer's production function. Wherever the optimal fee is located, however, it will generate positive agency costs of

---

implies that $w > 1$. Differentiating the first-order condition with respect to r and applying the chain rule yields $rw''(dx/dr) + w' = 0$, which implies that $dx/dr > 0$.

both kinds.

B. Variants of the Basic Model

Let us use r^* to denote the optimal linear fee granted the assumptions in the basic model; and let x^* and w^* denote the corresponding investment level and expected judgment. The payoffs to lawyer and client under the optimal fee in the basic model are then as follows:

Plaintiff: $(1-r^*)w^*$
Lawyer: $r^*w^* - x^*$

Let us see how the analysis changes if we vary the assumptions underlying the basic model. I leave formal demonstrations to the Appendix; here it is enough to state the intuition.

1. External Constraints on Investment Decisions

The basic model assumes that the lawyer is free to invest as little effort in the case as she pleases, regardless of her agreement with the client. This assumption is not entirely plausible. There must be *some* positive level of effort to which she can enforceably commit herself -- some "floor" on the quality of her input in the case. (That floor might be supplied by professional regulation or by reputational concerns.) But only if that "floor" is relatively high -- fairly close to x^* -- does it make any difference to the analysis.

To see the point, suppose the lawyer must invest some minimum effort level $\overline{x}$ if she takes the case.[17] To get the lawyer to invest $\overline{x}$, the client can get away with paying the lawyer no rents, because the floor takes away the lawyer's power to make marginal investment decisions. However, to get the lawyer to invest *more* than $\overline{x}$, the client must pay a fee that fully compensates her at the margin. If, therefore, the floor ($\overline{x}$) is sufficiently below the optimal investment level

[17]If she invests less than $\overline{x}$ she will be disbarred, sued for malpractice, or something of the sort.

(x^*), the client will find it in his interest to pay the lawyer the same fee (and generating the same payoffs) he would in the basic model.

2. Fixed Costs

The basic model also assumes that there are no fixed costs associated with the lawyer's production function. That is, we have assumed that even at low levels of investment, the plaintiff's case is a paying proposition -- in the sense that the net return to the claim from the lawyer's investment is positive.[18] In some cases, it is plausible to instead to suppose that the lawyer must make a substantial investment to realize a net return to the claim -- meaning that if she invests less than that, her investment will be wasted.[19]

Let us suppose, then, that in pursuing the case the lawyer must incur some fixed cost F, or else the claim will yield a negative net return on her investment.[20] Then the optimal linear fee will be the same as in the basic model, but will give the plaintiff and the lawyer the following respective payoffs:

Plaintiff: $(1-r^*)w^*$
Lawyer: $r^*w^* - x^* - F$

Thus, the client's payoff is the same as in the simple model, but the lawyer's rents are smaller.

The intuition here is that, despite the presence of fixed costs, the lawyer's *marginal* investment decisions under a given fee will be no different than they would if there were no fixed

---

[18] See note ___ above.

[19] This might be true in, say, a complex product-liability or malpractice case. The case is not a paying proposition unless a lot is invested in it.

[20] We assume that the return to her variable investment, that is $w(x)$, is the same as in the basic model. Thus, if she has incurred F, the payoff to her investment of x is $w(x)$; if she has not incurred F, the payoff to her investment of x is 0. We also assume that F is small enough to satisfy the following inequality: $r^*w^* - x^* \geq F$.

costs. Provided, therefore, that r^* is large enough to cover her total expected investment, the presence of fixed costs makes no difference except to lower the lawyer's rents.

3. Settlement

The basic model also does not take into account the possibility that the case may settle without trial. But if the settlement amount reflects the expected judgment at trial, the prospect of settlement does not affect the plaintiff's choice of fee. Thus, suppose that cases settle for some fraction (or multiple) q of the expected judgment; we can assume the settlement occurs before the lawyer has invested anything in the case. So long as the value of q is independent of the fee, the optimal linear fee remains r^*. The respective payoffs to plaintiff and lawyer under the optimal fee are then

Plaintiff: $(1- r^*)qw^*$
Lawyer: r^*qw^*

Observe that if q is less than one, the client does worse here than in the basic model. Note, however, that even if q is a lot less than one, the lawyer's rents are likely to be greater than they are in the basic model.

C. Agency Costs and the Market for Legal Services

Under either the basic model or any of the above modifications, then, agency costs of both types will be present under the optimal. Their precise magnitude for a given case (with production function $w(x)$) will vary, depending on which of the assumptions just discussed we embrace. But under practically any account, the optimal linear fee will result in "too little" being recovered from the defendant (less than $\hat{w}$, the surplus-maximizing recovery), and in "too much" being paid to the lawyer (more than x, the opportunity cost of her actual investment).

It is important to emphasize that this problem has nothing to do with the competitiveness of the market for legal services. It might be asked why -- if the market were perfectly competitive -- either type of agency cost could exist in equilibrium. The existence of an investment shortfall -- unrealized surplus from the claim -- means, in effect, that a resource is being underexploited. Why doesn't market competition lead to the most efficient deployment of the resource? Similarly, why would not lawyers' rents be eliminated through a process of competitive bidding, so that clients had to pay no more than the lawyers' opportunity costs?

The simple answer is that *restricting the fee arrangement to a simple percentage of the award* prevents the type of price competition that would (in the absence of that restriction) eliminate both types of inefficiency. If, for example, our model allowed the lawyer to make a "side payment" to the client in exchange for a share of the claim, both agency costs might disappear under the optimal fee.[21] But this is because the subject matter of the contract would be different from the one we are considering. In a market in which lawyer and client can contract only over the value of r, the client cannot avoid trading off lawyer rents against investment shortfall. This is true even if that market for legal services is (but for that restriction) perfectly competitive.

Two corollaries to this conclusion are worth highlighting. First, if -- for whatever reason -- attorney compensation takes the form of a simple percentage of the award, then levels of defendant liability are lower, and levels of plaintiff compensation a lot lower, than they would be if there were "ideal" price competition (with side payments and so forth) among lawyers. Defendants pay less because less is invested in the claim; plaintiffs get a lot less *both* because less

[21] The optimal fee arrangement might consist of the lawyer in effect purchasing the claim from the client; such an arrangement would in principle both encourage full investment in the claim and strip away the lawyer's rents. (The purchase would have to be structured in such a way as to ensure client participation in the litigation; perhaps the lawyer would only purchase 90 percent of the claim.)

is invested in the claim *and* because they have to pay rents to the lawyer.[22] (Whether this is bad from a social standpoint depends on unspecified details of the liability system.)[23]

Second, the prospect of earning ex ante rents from contingent-fee representation should attract entry by lawyers until there is in effect an oversupply of lawyers -- each spending only part of her time actually representing clients, and the rest of her time trying to get clients. In the resulting competitive equilibrium, the rents earned from representing clients would be dissipated by the time and effort spent looking for new cases and clients. (This effort presumably represents a pure deadweight loss.)

III. FEE ARRANGEMENTS AND CASE CHARACTERISTICS

What qualities of a case determine the value of the optimal linear fee, and the resulting payoffs to lawyer and client? In order to investigate this issue, we must give a more explicit formal representation of the properties of different production functions. To investigate this issue, I use a more precisely specified model of the lawyer's production function. (To keep the analysis tractable, I restore the three assumptions of the basic model.)

Part of my purpose here is to see how the optimal fee arrangement varies with different case properties. For example: As the stakes in a case rise, it might be asked what happens to the optimal fee percentage (does it go up or down?); to the lawyer's rents; and to the client's share of the potential surplus. I am also interested, however, in seeing what *equilibrium values* these

---

[22]In a world of perfect price competition, the lawyer would make the surplus-maximizing investment $\hat{x}$ and collect no rents; the defendant would pay $\hat{w}$, and the client would take home $\hat{w} - \hat{x}$ (in expected terms). In contrast, if the lawyer works under r^*, thedefendant pays w^* (which is less than $\hat{w}$). If the plaintiff got to keep the full surplus of the lawyer's efforts, he would take home $w^* - x^*$ (which is less than $\hat{w} - \hat{x}$); but since he does not get to keep the full surplus, he gets even less than that.

[23]Such as the deterrent value of liability, the administrative costs of running the legal system, and so on.

different features of the optimal arrangement may have. Proposals for fee regulation often rest on the assertion that lawyers collect too much and clients collect too little. It may be of some interest to see, in some simulated cases, what lawyers and clients would earn under the optimal fee.

A. Determinants of the Optimal Fee Arrangement

On an intuitive level, the principal determinants of the optimal fee must be the costs of litigating the case; the likelihood of success in the case; and the likely amount of recovery in the event of success. That intuition, however, is obviously incomplete, because it ignores the endogenous character of the costs and rewards of litigation. It is the lawyer's production function curve -- rather than the point she chooses on the curve -- that is (in the model) exogenous to the fee she works under. We need to ask, therefore, what characteristics of the lawyer's production function determine the optimal fee in a given case.

In order to investigate this issue, let us give a more explicit formal representation of the properties of different production functions. Consider the set of hypothetical cases in which the lawyer's production function has the following structure:

$$w(x) = \alpha(1\text{-}e^{-\beta x}), \tag{9}$$

where

α	=	An exogenous parameter representing the maximum potential amount, in expected terms, at stake in the case ($\alpha > 0$); and
β	=	An exogenous parameter affecting the marginal productivity of the lawyer's efforts ($\beta > 0$).

The natural logarithm base e is used primarily for computational ease; the qualitative conclusions of our analysis would hold if we used a different exponential base.

I have selected this functional form to focus on, for two reasons. First, it is fair to speculate that in most cases, the expected recovery approaches a maximum possible value -- a sort of ceiling -- which the lawyer's efforts cannot push it beyond. It would be unrealistic to suppose that the expected recovery has no limit as a function of lawyer effort. It is reasonable to assume, however, that the ceiling varies from case to case -- depending on the nature of the injury, the strength of the evidence, and so forth. The parameter α is designed to capture this feature; the greater the value of α, the higher the ceiling. (See Figure 2.) We can think of it α as reflecting, in a rough way, the amount potentially at stake in a case.[24]

[FIGURE 2]

Second, the ease with which the lawyer can "approach" the ceiling may also vary from case to case. In some instances, it may be fairly easy to bring the expected recovery up to a point close to its ceiling; one might expect this in relatively simple cases involving, say, only a few witnesses and comparatively settled legal rules. In such cases, a relatively small investment takes the case near to the ceiling, and further investment has only a tiny payoff. In more complex cases, involving scientific evidence and contested legal standards, it may be more costly to bring the case's value up to a point near its ceiling; after a relatively small investment has been made, the payoff to additional investment is still substantial. The parameter β is designed to reflect this variation; the higher the value of β, the less costly it is for the lawyer to bring the case's value up a

---

[24]Thus, suppose the lawyer's efforts can raise the *likelihood* of recovery (that is, the chance of winning on liability) to about 50 percent, but not above that; and suppose that her efforts can bring the expected *amount* of the defendant's payment (conditional on the defendant's being found liable) up to about $200,000, but not above that. Then we would say the value of α -- the ceiling on the expected recovery -- is $100,000. Accordingly, to say that two cases have different values of α may reflect differences in the likelihood of winning, differences in the size of the recovery in the event of victory, or both.

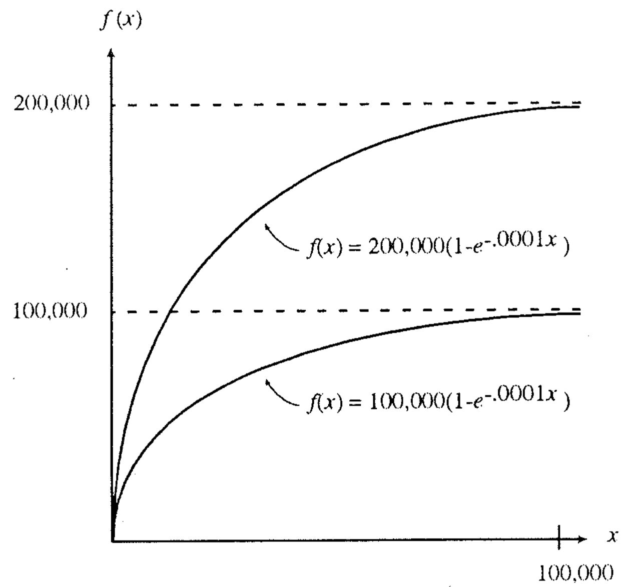

FIGURE 2 . — The effect of varying α between \$100,000 and \$200,000, where the value of β is fixed at .0001. (Drawn to approximately 50% vertical scale.)

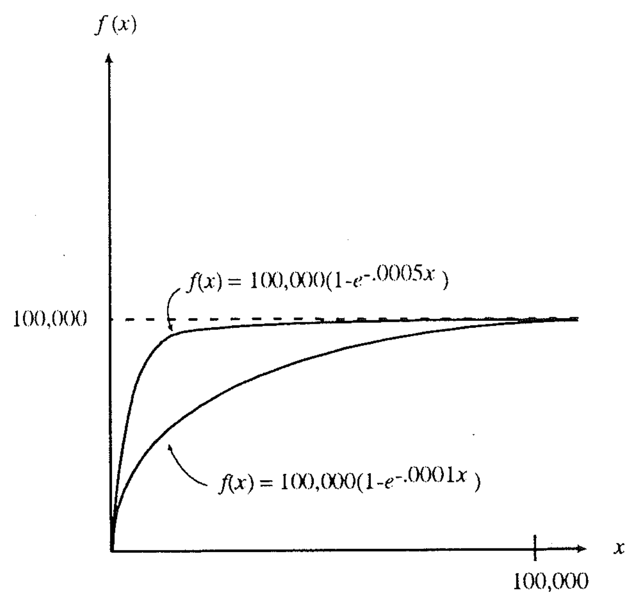

FIGURE 3 . — The effect of varying β, where the value of α is fixed at \$100,000. (Drawn to approximately 50% vertical scale.)

given fraction of the distance to the ceiling.[25] (See Figure 3.)

[FIGURE 3]

This description is meant to highlight the theoretical distinction between the two parameters. In principle it is possible, on the one hand, to have a case in which the *ceiling is high* but in which it is *not costly to cover a substantial fraction (say, 90 percent) of the distance* to the ceiling. It is also possible, on the other hand, to have a case in which the ceiling is low, but in which it is very costly to cover that same fraction of the distance toward the ceiling. Conceptually, the *height* of the ceiling bears no necessary relation to the *ease of moving toward* the ceiling.

But whether the two parameters are so separate in practice is another matter. In many instances it is plausible to suppose that α and β are not independent but rather inversely correlated with each other -- in other words, that cases with high ceilings tend to have lower values of β than cases with low ceilings. The reason for this is that, generally speaking, litigation frequently becomes more cumbersome as the stakes go up. The defendant, having more at stake, may spend more to resist the plaintiff's claim;[26] cases with a lot of money on the line often involve numerous parties;[27] the procedures courts use to resolve a case often become more complex as the stakes go up.[28] Thus, while we should bear in mind the conceptual distinction between α and β in the

---

[25]Note that variation in the value of β(as in the value of α) may reflect differences in lawyers' quality rather than anything intriinsic to the case.

[26]See Richard A. Posner, Economic Analysis of Law 568 (3d ed. 1992).

[27]Even if not named by the plaintiff, they may be brought in by the defendant seeking to shift liability to others.

[28]Big cases are more prone to wind up in federal court, are less likely to be sent to court-annexed arbitration, and so on.

analysis to follow, we should also remember the possible (inverse) correlation between them.

B. Case Parameters and Optimal Fee Arrangements

Using expression **(8)**, we find that the optimal fee is simply given by

$$r^* = \frac{1}{\sqrt{\alpha\beta}} . \qquad \textbf{(10)}$$

Our objective is to see what this fee arrangement yields for lawyer and client. Since the mathematics involved is simple but space-consuming, I omit it here.[29]

Table 1 summarizes the effect of parameter changes on different features of the fee arrangement. The first two columns of the table treat the parameters as independent; the second two columns treat them as inversely correlated.

TABLE 1: EFFECT OF PARAMETER VARIATION ON DIFFERENT FEATURES OF FEE ARRANGEMENT

	One Parameter Held Constant		Product $\alpha\beta$ Held Constant	
	α increased	β increased	α increased	β increased
Fee Percentage	-	-	n.c.	n.c.
Client's Recovery:				
Amount	+	+/-	+	-
Fraction of $\hat{w}$-$\hat{x}$	+	+	n.c.	n.c.
Lawyer's Profits:				
Amount	+	+/-	+	-
Rate of return	+	+	n.c.	n.c.

NOTATION: "+" means an increase for the feature in question; "-" means a decrease; "+/-" means there may be movement in either direction; "n.c." means there is no change.

[29]It is presented in Bruce L. Hay, Contingent Fees and Agency Costs, Discussion Paper No. 162, Harvard Law School Program in Law and Economics (June 1995).

1. <u>The Optimal Fee</u>

If we hold one parameter constant, then an increase in the other parameter should lead to a *decrease* in the percentage given to the lawyer. The intuition here is fairly straightforward. An increase in α raises the marginal return to the lawyer's effort, meaning in essence that to get her to invest a given amount she need only be paid a rather small fraction of the recovery. An increase in β implies that the production function curve levels off quickly,[29] meaning there is not much payoff from investing a lot in the claim.[30] Both of these effects favor giving the lawyer a comparatively small fee. Thus, all else equal *the higher the ceiling on the expected recovery, or the less costly it is for the lawyer to approach the ceiling, the lower the optimal fee percentage*.

In contrast, as Table 1 indicates, cases with the same value of $\alpha\beta$ should have the same fee. Thus, *cases in which the ceiling is high but in which it is costly for the lawyer to move upward* should involve the same fee as *cases in which the ceiling is low but in which it is easy for the lawyer to move upward.* If, therefore, the parameters are inversely correlated in the way we speculated above, there might be relatively little variation in the optimal fee; cases with radically different ceilings will nonetheless might have similar optimal fees.

2. <u>Lawyer Profits</u>

---

[29]Figure 4 on page __ gives an example. At low levels of investment, the high-β curve is very steep, meaning that the lawyer is willing to work for a relatively low fee; and at high levels of investment, the high-βcurve is very shallow, meaning that there is not much return to the claim from investing additional effort.

[30]The client does best by having the lawyer invest comparatively low effort, investing only so long as the curve remains steep. An analogous way of seeing the point is as follows: the low-β curve in Figure 4 is comparatively shallow at the beginning (relative to the other curve), so the lawyer needs a higher fee to make the initial investment; and the curve is comparatively steep later on (again, relative to the other curve), meaning the client wants the lawyer to keep investing. Both factors favor giving the lawyer a higher fee in the low-β case.

We examine two measures of the lawyer's profits under the optimal fee. One is the absolute amount of her rents; this is given by the difference between her share of the expected recovery and her total investment; the other is the rate of return -- or percentage profits -- the lawyer earns on her investment.[31] As Table 1 indicates, the lawyer's profits do not necessarily move in same direction as the optimal fee. Consider the two first columns of the table: an increase in one parameter (holding the other constant) generally leads to an *increase* in the lawyer's profits, even though it leads to a decrease in her fee percentage.

The reason for this counterintuitive result is that the change in parameters affects the lawyer's equilibrium investment in the case, and the resulting expected recovery. These changes can easily more than offset the reduction in the fee percentage. For example, as α increases, so does the marginal return to investment. As a result, the amount the lawyer collects thus goes up, even though her fee percentage goes down: in effect, the pie expands faster than her share shrinks. Now consider what happens when β goes up: the lawyer can, with a relatively small investment, produce a big recovery, meaning her level of investment goes down. Her fee percentage also goes down, but not as quickly. The net effect is that her profits may go up.

3. Client Recoveries

As Table 1 indicates, the client's recovery -- measured either in absolute terms or as a fraction of the potential surplus -- generally increase as either parameter increases. One interesting implication of this result that under the optimal fee, *client recoveries generally move in the same direction as lawyer profits; the more the lawyer nets from the case, the more the*

[31] Absolute rents are given by $r^*w^*-x^*$; the rate of return is given by $(r^*w^*-x^*)/x^*$. The latter measure may be more useful for comparing cases with different ceilings.

client nets from the case. Another interesting implication is that all else equal, *the higher the ceiling in the case, the less acute the agency problems stemming from the use of a percentage fee:* as the value of α increases, so does the fraction the client gets of the potential surplus from his claim.

C. Optimal Arrangements in Sample Cases

To this point the analysis has focused on the qualitative relations between case parameters and fee arrangements, without giving any sense of the magnitudes involved. I now present some numerical simulations of a few sample cases, in order to see what the optimal fee arrangements look like in quantitative terms. These cases are purely hypothetical, involving seemingly reasonable assumptions concerning the range of possible parameters in actual cases.

The simulations will use three values of α. The first value is $20,000; this corresponds to cases in which the stakes are relatively low, for example a minor personal-injury suit. The second value of α is $500,000, which corresponds to tort cases with much higher stakes. These two figures appear to span much of the range of possible values of α, at least for the domain of contingent-fee suits brought by individuals.[32] As a third value of α, I have chosen the intermediate value of $100,000, which corresponds to cases with moderately high stakes.

To select some appropriate values of β, we proceed more indirectly. Since it is not

TABLE 2: SAMPLE CASE PARAMETERS

[32] See Deborah R. Hensler et al., Trends in Tort Litigation: The Story Behind the Statistics 12-24 (RAND Institute for Civil Justice 1987); Mark A. Peterson, Civil Juries in the 1980s: Trends in Jury Trials and Verdicts in California and Cook County, Illinois (RAND Institute for Civil Justice 1987).

Parameters		Surplus-Maximizing Investment ($\hat{x}$)	Resulting Expected Recovery ($\hat{w}$)
α	β		
20,000	1	$10	$19,999
"	.1	$76	$19,990
"	.01	$530	$19,900
"	.001	$2,996	$19,000
"	.0001	0	0
"	.00001	0	0
"	.000001	0	0
100,000	1	$12	$99,999
"	.1	$92	$99,990
"	.01	$691	$99,900
"	.001	$4,605	$99,000
"	.0001	$23,026	$90,000
"	.00001	0	0
"	.000001	0	0
500,000	1	$13	$499,999
"	.1	$108	$499,990
"	.01	$852	$499,900
"	.001	$6,215	$499,000
"	.0001	$39,120	$490,000
"	.00001	$160,944	$400,000
"	.000001	0	0

intuitively obvious even what order of magnitude β should have, let us examine what happens when we match different values of α and β. In particular, for the values of α we have selected, let us look at what the potential surplus is in cases with varying values of β. Table 2 provides this information for a sample set of values of β from .000001 to 1.

This table enables us to weed out all but a comparatively narrow range of β values. At one extreme, for values of .000001 or less, the potential surplus is zero, so there is no point in asking about the optimal fee arrangement. At the other extreme, when we use values of .01 or greater, the potential surplus seems unrealistically high for most cases: the slope of the

TABLE 3: OPTIMAL FEES IN SAMPLE CASES

β	α = 20,000	α = 100,000	α = 500,000
.001	r^* = 22%	r^* = 10%	r^* = 4%
.0005	32%	14%	6%
.0001	71%	32%	14%
.00005	---	45%	20%
.00001	---	---	45%

production function is so steep that only a minuscule investment by the lawyer brings the claim's value very near the ceiling.[33] Cases in which this happens may exist, but it seems reasonable to assume that in most cases the payoff to lawyer effort is somewhat less spectacular.

Accordingly, we will restrict our attention to values of β that fall between these two extremes. We will use five arbitrarily-chosen values of β in this range: .00001, .00005, .0001, .0005, and .001. This gives us a total of fifteen possible combinations of α and β. For three of these combinations, the potential surplus is zero;[34] this leaves a total of twelve cases in our sample.

1. The Optimal Fee

Table 3 indicates the optimal fee for each of the sample cases. Two points are worth emphasizing about these figures. One is the considerable variation in the optimal fee for different cases; thefee that maximizes the client's welfare can fee can range from well below the stereotypical 33 percent to well above it. A second point is that the optimal fee drops as the value of either parameter rises. Notice, however, that as one moves from the top left-hand region of the

[33]For example, suppose the ceiling is $100,000. As the Table shows, when β = .1, an investment of $92 by the lawyer bringsthe claim up from zero to $99,900!

[34]If α= 20,000, then the case is not worth anything unless β > .0005. If α=100,000, the case is not worth anything unlessβ > .00001.

TABLE 4: LAWYER PROFITS IN SAMPLE CASES

Parameters			
α	β	Amount	Rate of Return
20,000	.001	$1,974	130%
"	.0005	$2,022	88%
"	.0001	$677	20%
"	.00005	---	---
"	.00001	---	---
100,000	.001	$6,697	290%
"	.0005	$8,230	210%
"	.0001	$10,110	88%
"	.00005	$8,627	54%
"	.00001	---	---
500,000	.001	$18,253	587%
"	.0005	$24,101	437%
"	.0001	$41,151	210%
"	.00005	$47,811	149%
"	.00001	$43,135	54%

Table to the bottom right-hand region, the optimal fee remains relatively constant -- reflecting our earlier observation that cases with similar values of the product $\alpha\beta$ have similar optimal fees.

2. Lawyer Profits

Table 4 indicates the lawyer's profits under the optimal fee in each case. Observe that her rate of return may be enormously high -- the optimal fee may, *in expected terms*, easily give her a several fold return on her investment. Notice, in addition, that her profits rise sharply with α, and also generally rise with β.

3. Client Recoveries

Finally, Table 5 contains the client's recovery under the optimal fee. As the model predicted, agency problems under the contingent fee grow less acute with increases in either α or β. The client can capture over 90 percent of the potential surplus when the ceiling on the

TABLE 5: CLIENT RECOVERIES IN SAMPLE CASES

Parameters		Amount	Fraction of Potential Surplus
α	β		
20,000	.001	$12,056	75%
"	.0005	$9,351	70%
"	.0001	$1,716	56%
"	.00005	---	---
"	.00001	---	---
100,000	.001	$81,000	85%
"	.0005	$73,716	82%
"	.0001	$46,754	70%
"	.00005	$30,557	64%
"	.00001	---	---
500,000	.001	$456,279	92%
"	.0005	$438,754	90%
"	.0001	$368,579	82%
"	.00005	$320,000	77%
"	.00001	$152,786	63%

recovery is high, but can get at most 75 percent when the ceiling is low.

IV. SOME APPLICATIONS

A. Predicting Market Behavior

The model may be of some use in predicting patterns in the use of linear contingent fees in the market. Regarding the percentages agreed to by lawyer and client, the model would predict that contingent fee percentages would generally drop as the value of α or β rises, unless there is , unless there is an offsetting increase in the other parameter. Regarding lawyer profits, it would predict an effective hourly rate (varying with case parameters) well above what the lawyer would for noncontingent fee work. (This effect would presumably be coupled with an oversupply of contingent-fee lawyers, meaning that one might expect such lawyers to devote a substantial portion of their time searching for clients ["ambulance chasing"] rather than actually working on cases.) Regarding client recoveries, it would predict that linear contingent fees would perform best -- give the client the largest fraction of the potential surplus -- in high-stakes cases; accordingly, we might expect to see greater use of alternative fee arrangements in low-stakes cases than in high-stakes cases.

Unfortunately, the available data is insufficient to test, in even a rudimentary way, most of these conjectures. There is, however, a small amount of published data concerning the fee percentages collected by lawyers in different types of tort case. This data, compiled by the RAND Institute for Civil Justicein its studies of the tort system, sheds some light on how contingent fee percentages vary in different contexts. Two pieces of information are useful for our purposes.

The first concerns the average attorneys fee paid in different categories of claims. Table 6 provides this information for five categories of tort claim, arranged according to injury source. The first is a catchall category covering all types of personal injury claim combined.[35] The remainder consist of more specialized categories: auto accident injury claims,[36] medical malpractice claims,[37] tort claims arising out of large-scale domestic aviation disasters,[38] and claims

[35]The data appearing here is drawn from James S. Kakalik & Nicholas M. Pace, Costs and Compensation Paid in Tort Litigation, RAND Publication R-3391 (1986), at 40-41, 111-14; and from Deborah R. Hensler et al., Compensation for Accidental Injuries in the United States, RAND Publication R-3999-HHS/ICJ (1991), at 136.

[36]The data appearing here is drawn from Kakalik & Pace, supra note ___, at 40, 114.

[37]The data here is drawn from id. at 41.

[38]The data here is drawn from James S. Kakalik et al., Costs and Compensation Paid in Aviation Accident Litigation, RAND Publication R-3421-ICJ (1988) (hereinafter referred to as "Aviation Accident Litigation"), at .

TABLE 6: AVERAGE PLAINTIFF ATTORNEYS FEES AND RECOVERIES IN FIVE TORT CLAIM CATEGORIES

Claim Type	Average attorneys fee
All Torts Combined	
CLRP Data (1978) [a]	28%
RAND Data (1988) [b]	29%
Specific Categories	
Auto Accidents (1978) [c]	29%
Medical Malpractice (1971-73) [d]	32%
Asbestos Injury (1980-82) [e]	34%
Aviation Accidents (1977-82) [f]	17%

[a] All tort cases (654) closed in 1978 in twelve jurisdictions. Source: RAND Publication R-3391-ICJ (1986) (reporting University of Wisconsin Civil Litigation Research Project data).

[b] All claims (371) for nonfatal accidents brought during 1988 by individuals in a random sample of roughly 18,000 households. (Excludes claims for illnesses and latent injuries caused by toxic exposures.) Source: RAND Publication R-3999-HHS/ICJ (1991).

[c] All auto accident cases (362) closed in 1978 in twelve jurisdictions. Source: RAND Publication R-3391-ICJ (1986) (reporting University of Wisconsin Civil Litigation Research Project data).

[d] Based on data collected from 400 malpractice lawyers in the early 1970s. Source: RAND Publication R-3391-ICJ (1986) (reporting HEW data).

[e] 513 randomly-drawn cases closed during the years 1980-82. Source: RAND Publication R-3132-ICJ (1984).

[f] All claims (933) arising from large-scale domestic aviation accidents during 1976-82. Source: RAND Publication R-3421-ICJ (1988).

for injuries related to asbestos exposure.[39]

The figures indicate the estimated average attorneys fee paid by the client, exclusive of expenses.[40] Several caveats are important here: It appears that in the *majority* of these cases, a

[39]See James S. Kakalik et al., Variation in Asbestos Litigation Compensation and Expenses, RAND Publication R-3132-ICJ (1984) (hereinafter referred to as "Asbestos Litigation"), at 89.

[40]Expenses refer to reimbursement of the attorney for filing fees and the like. The Hensler data, supra note ___, concerns fees alone, independent of expenses. The remaining sources. The remaining data sources do not separate fees and expenses, but indicate the average proportion of fees to expenses in the amount paid to the attorney for auto accident cases and all torts combined. See Kakalik & Pace, supra note ___, at 115; I have used these figures to

simple linear contingent fee was used; but this is not true of *all* of the cases.[41] Further, the data are fromdifferent years, and in some instances were acquired through different sampling methods. Notice also that these are *average* figures, grouping together very disparate claims; each figure may mask substantial variation within each category. (For example, the 29 percent figure for auto accident cases does not imply that everyplaintiff signed up for a 29 percent fee.)[42] For these reasons, the data should be taken as merely suggestive of the differences between the categories.

The figures show considerable variation in the percentages paid to attorneys. The average fee in auto accident cases is about the same as the average for all cases combined. In contrast, the average medical malpractice case fee is about 14 percent higher than than the combined figure, and the average asbestos exposure case fee is about 25 percent higher than the combined figure. Finally, plaintiffs in aviation cases pay their lawyers, on average, a figure 40 percent *below* the combined figure.

The variations among these categories seem roughly congruent with the model's predictions. Aviation cases involve high potential damages;[43] liability (and the extent of the plaintiff's injury) is generally clear and frequently uncontested;[44] and few defendants are involved

estimate the average fees. (For malpractice and asbestos cases, I have used the average proportion for all torts combined, which is 9:1.)

[41] See Deborah R. Hensler et al., Compensation for Accidental Injuries in the United States 135-36 (RAND Institute for Civil Justice 1991). Most of the remainder involve more complex nonlinear contingent fees. Examples of nonlinear fees include those in which the percentage collected by the lawyer depends on the amount of time spent on the case; on the stage of litigation in which the case is resolved; or on the amount recovered. Unfortunately, the RAND data do not distinguish the cases in which linear and nonlinear fees were used.

[42] Indeed, our model would predict great variation within each category -- for example, all else equal, we would expect a lower fee in an auto case involving a fatality than in one involving minor injuries.

[43] Virtually all are wrongful death actions; victims often healthy and high wage earners.

[44] Defendants often agree not to contest liability in exchange for an agreement not to seek punitive damages. Kakalik. Since the injury is almost always a fatality, there is nothing to fight about except the extent of the plaintiff's

in the litigation.[45] Hence the value of α and β are both relatively high, so we would expect fee percentages to be relatively low.[46]

Now consider malpractice and asbestos cases. On the one hand, the α parameter does not seem to explain the higher-than-average fees encountered in these categories. Damages are often high in these categories (though typically lower than in aviation cases),[47] though the availability of evidence supporting the plaintiff's case may vary a great deal.[48] The value of α in these categories is almost certainly generally lower than in aviation cases. But there is no clear reason to suppose that that the average malpractice and asbestos case has a lower ceiling than the average for all cases combined.

On the other hand, approaching the ceiling is probably more difficult in these categories than in most cases. Defendants in medical malpractice and asbestos exposure cases typically vigorously contest liability,[49] and litigating the liability issue involves extensive use of experts. As

---

losses.

[45]See Kakalik et al., Aviation Accident Litigation, at 12, 20, 69. Kakalik and his coauthors emphasize several of these factors in explaining why contingent fees in aviation cases are lower than in other types of tort litigation. See *id.* at 57; Kakalik et al., Tort Litigation, at 41.

[46]The fact that liability is clear and damages high means the ceiling is high; the fact that defendant expenditures on liability are small and few defendants are involved means the ceiling is relatively easy to approach.

[47]Asbestos cases often involve disabling or fatal diseases, but the victims are almost always smokers and are typically relatively advanced in age. See Kakalik et al., Variation in Asbestos Compensation and Expenses, at 22, 34. Of malpractice cases, less than half typically involve permanent injuries or death. See Patricia M. Danzon, Medical Malpractice 22-23 (1985).

[48]Plaintiffs in asbestos cases must present proof of a causal link between an exposure and an injury that may not surface for several decades. In malpractice cases, the defendants are often the only witnesses to the events underlying the suit.

[49]Doctors are worried about the reputational effects of adverse outcomes in malpractice cases. See Kenneth Syverud, The Duty to Settle, __ Va. L. Rev. __, __ (1991). Asbestos defendants are worried about preclusive effects of adverse judgments and also have an incentive to discourage future claims. See ALI Reporters' Study, Enterprise Responsibility for Personal Injury, Vol. II, at 400 (1991).

a result, β is relatively low in both categories, perhaps explaining the relatively high fees in both. (Asbestos cases also involve many defendants and raise difficult proof-of-causation problems, further lowering the value of β in this category -- which may explain why fees are higher here than in malpractice cases.)[50]

A second useful piece of information in the RAND studies concerns the fees charged in the aviation accident category. This is the only category for which the average figures indicated in Table 6 have been broken down. Figure 4 indicates the range of percentages collected by attorneys from awards in this category.[51] As the Figure shows, there is a fair degree of variation in the percentages collected; some were below 5 percent, others above 40 percent.

[FIGURE 4]

Finally, Table 7 shows how the average attorney's fee varies with the economic losses sustained by the plaintiff.[52] Using variation in economic losses as a proxy for variation in α,[53] the pattern in Table 7 roughly tracks the model's prediction, at least in the direction the change takes. Within the category of aviation accident litigation, we would not -- owing to the special features of aviation cases -- expect the value of β to change much as the stakes in the case increase from,

---

[50]The average asbestos case has 16 defendants. See Kakalik et al., Costs and Compensation Paid in Tort Litigation, at 41. On the difficulty of establishing causation, see id. at 38. The RAND authors emphasize these factors in attempting to explain the relatively high fees in asbestos cases.

[51]Unfortunately, the percentages here include expense other than the attorney's fee. The RAND data are insufficient to enable us to draw the figure for attorneys' fees alone. Since expenses generally run to about 10 percent of the amount collected by the attorney, the figure is suggestive. (The RAND studies give no reason to suppose the shape of the figure is due to variations in the amount of expenses.)

[52]These fee figures are adjusted to remove the estimated 10 percent devoted to expenses other than attorney fees. The methodology underlying the eocnomic loss figures is described in Elizabelth M. King & James P. Smith, Economic Loss and Compensation in Aviation Accidents, RAND Publication R-3551-ICJ (1988).

[53]King & Smith, supra, found that the amount of compensation goes up with the plaintiff's economic losses.

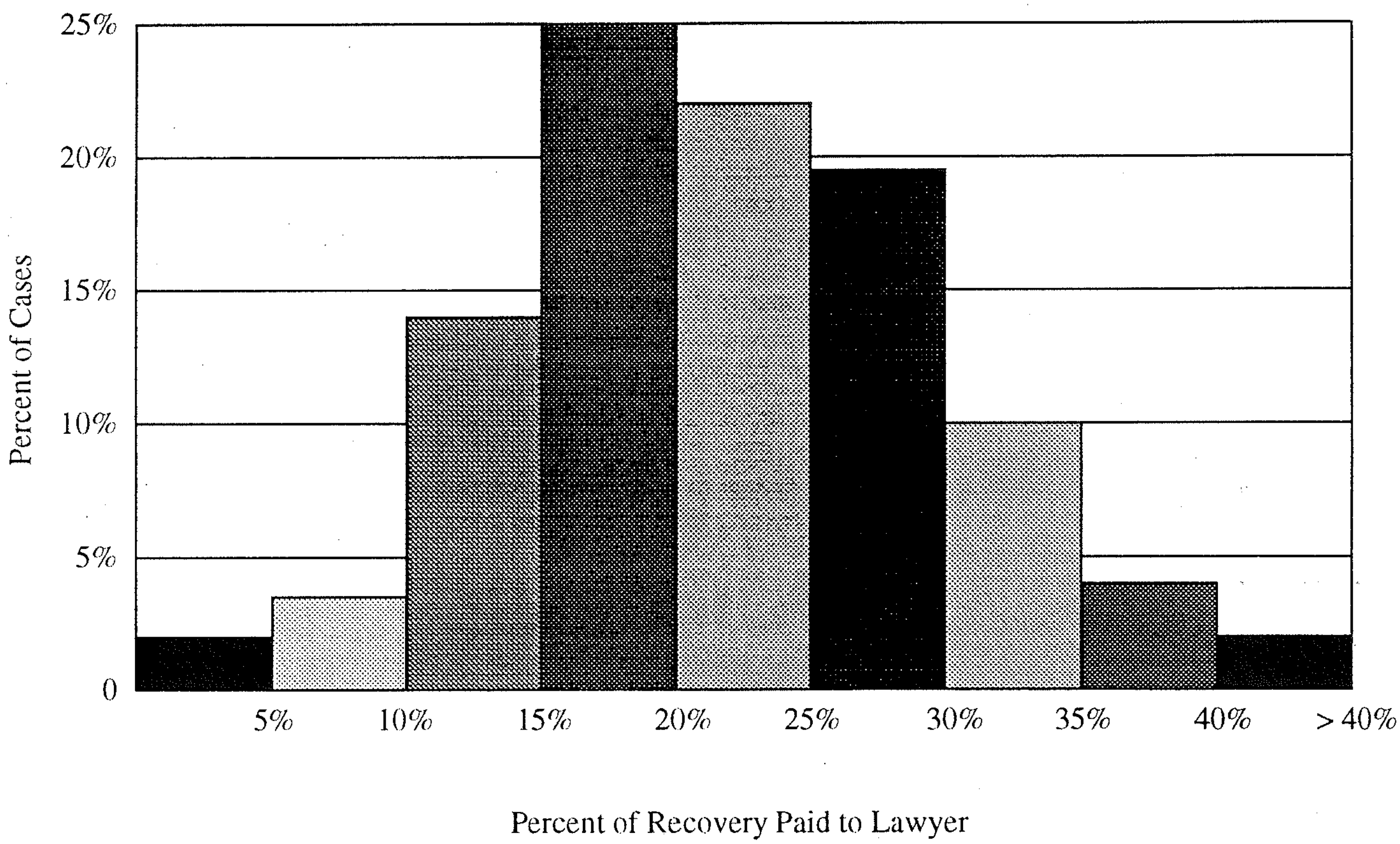

FIGURE 4. — Aviation cases. Note this includes expenses, which are typically 10 percent of the amount paid to the lawyer. Reproduced from James S. Kakalik et al., Costs and compensation Paid in Aviation Accident Litigation, Rand Publication R-3421, at 54 (1988).

TABLE 7: ATTORNEYS FEES IN AVIATION ACCIDENT CASES

Plaintiff's Economic Losses	Average Attorneys Fee
100,000-500,000	23%
500,000-1,000,000	22%
> 1,000,000	20%

Source: RAND Publication R-3421-ICJ (1988).

say, \$100,000 to \$1 million.[54] We would thus expect the fee to drop as the stakes increase, which it does by about 15 percent.

B. Consequences for Fee Regulation

Many states put an upper limit on the percentage a contingent fee contract can specify. Since the model does not attempt to describe the workings of the contingent fee market (it asks what linear fees lawyers *should* get, not what they *do* get), it has nothing to say about whether, or where, such regulation is justified. It does, however, offer some prescriptions about the *manner* in which such regulation is undertaken. (These points also apply to situations -- such as class actions -- in which courts must set the lawyer's fee directly because there is no client to sign a contract with.)

The first concerns the proper basis or rationale for fee regulation (assuming the objective is to maximize net client recoveries). Proponents of regulation typically focus on the profits lawyers earn under contingent fees. Bigger lawyer profits, the argument goes, mean less of the recovery is left over for the client; the aim of regulation is too prevent the lawyer from cutting too deeply into the client's share of the recovery. To support their argument, proponents of

[54]The features I have in mind here are the paucity of disputed issues and the small number of defendants (often only one).

regulation typically cite instances in which a lawyer working under a contingent fee earn an amount equal to many times the normal hourly rate for their services.[55]

Our analysis raises some cautionary signals about this rationale for regulation. I make no claim about whether lawyers in fact tend to earn excessive profits in the market; perhaps they do.[56] But if the objective of fee regulation is to maximize the welfare of the client, the model suggests that a single-minded focus on reducing lawyer profits is unwise. For as we have seen, increasing lawyer profits is often associated, at leat in principle, associated with greater net recoveries for the client. Our simulations showed that under quite plausible assumptions the optimal fee may in equilibrium give the lawyer a substantial multiple of her opportunity cost -- so that her effective hourly earnings are, in expected terms, several times the noncontingent hourly fee she could command.

It is also worth emphasizing that the profits earned by the lawyer under the optimal fee are unconnected to compensating the lawyer for the risk that she will lose the case and recover no fee. Critics of existing contingent fee practices often target fee arrangements that enable the lawyer to earn substantial profits in cases where there is no risk of losing the case.[57] But the optimal fee will always have this effect when returns to effort are positive but diminishing at the margin. The reason the lawyer is "overpaid" has to do not with *risk* but with *incentives*. When

---

[55]Critics have cited instances in which lawyers apparently have earned virtually riskless effecitve hourly rates of $30,000. See L. Brickman et al., Rethinking Contingency Fees 22 (1994).

[56]Studies conducted by researchers at the University of Wisconsin provide some data on the hourly earnings of contingent fee lawyers; see, for example, Herbert Kritzer, The Justice Broker: Lawyers and Ordinary Litigation 138-41 (1990). But there is no data that would enable us to say what the optimal margin in the case in question. For that reason, I do not examine the Wiconsin data here.

[57]See, for example, Lester Brickman, Contingent Fees without Contingencyies: *Hamlet* without the Prince of Denmark?, 37 UCLA L. Rev. 29 (1989); John F. Grady, Some Ethical Questions about Percentage Fees, 2 Litigation 20 (Summer 1976).

percentage fees are used, paying the lawyer rents in excess of her opportunity cost is an unavoidable byproduct of giving her the right incentives at the margin. Thus, the profit margins we saw in our simulations would obtain under the optimal fee even if litigation were riskless.

A second general lesson of the model is that to the extent they are used, fee caps should generally be designed to discriminate among cases based on their values of α and β. An across-the-board cap of X percent for all cases civil cases[58] is probably not the best approach, unless it is thought that α and β are perfectly inversely correlated. It is just as much a mistake to differentiate cases purely on the basis of the value of α -- which is in effect achieved by rules that call for the lawyer's share to go down as the amount recovered goes up[59] -- unless it is thought that α and β are completely uncorrelated. A preferable method is probably to fashion caps for particular substantive areas -- as some states have done, for example, for medical malpractice cases.[60]

Finally, a third implication of the model is that regulatory efforts (to the extent justified by some market failure) to find alternatives to the simple percentage fee should focus on low-stakes (low α) cases. Even under the best possible linear fee, clients in low-stakes wind up with a relatively small (compared to high-stakes cases) share of the potential surplus from their claim. Clients in these cases would benefit most from substitution of alternative fee structures that lessened the agency problems inherent in the contingent fee arrangement.

V. CONCLUSION

Both the predictive and prescriptive applications of the model are limited in an important

---

[58]See, for example, Mich Gen. Ct. R. 928.2; Okla. Stat. Ann. tit. 5 § 7 (West 1993).

[59]See, for example, Conn. Gen. Stat. § 52-251c; N.J. Ct. R. 1:21-7.

[60]See, for example, Calif. Code § 6146; Iowa Code § 147.138 (1993). This is not to say, of course, that the caps states have in fact selected are the correct ones.

respect. We have essentially taken a mechanism-design approach to our problem, asking what fee maximizes the client's welfare. We have not modelled the process by which lawyer and client in fact settle on a given fee. Developing an explicit model of the market for contingent fees, and determining the equilibrium fees that are likely to emerge in the market, are a major prerequisite to understanding and evaluating the contingent fee system. I examine this issue in separate research.

APPENDIX

Optimal Linear Fees

Optimal Linear Fee in the Basic Case

Here we derive the optimal fee for the case in which there are no fixed costs (meaning $w(x) > 0$ for all x) and in which the lawyer is free to choose whatever level of effort she wishes. In this case, the client wants to choose r to maximize

$$(1-r)w. \qquad \textbf{(A1)}$$

The lawyer's incentive compatibility constraint (differentiating with respect to x) is $rw'-1=0$. Solving for r and substituting the result into **(A1)** yields

$$w - \frac{w}{w'}. \qquad \textbf{(A2)}$$

The client's taskis to find the level of effort x that maximizes this expression. Differentiating with respect to x and setting the result equal to zero implies that **(A2)** attains a maximum when

$$w' = \frac{(w')^2 + w(-w'')}{(w')^2} \,. \tag{A3}$$

But since $r = 1/w'$, we find that **(A2)** attains a maximum when

$$r = \frac{(w')^2}{(w')^2 + w(-w'')} \,. \tag{A4}$$

External Constraints on the Lawyer's Investment

Suppose the lawyer is required to invest some minimum effort level $\overline{x}$ if she takes the case. The lawyer will invest $\overline{x}$ provide she is given a fee $\overline{r}$ satisfying

$$\overline{r} = \frac{\overline{x}}{\overline{w}}, \tag{A5}$$

where $\overline{w}$ is the expected judgment produced by the minimum effort level. However, the lawyer will not invest any more than $\overline{x}$ unless she is given a fee satisfying $r = 1/w'$. By assumption, r^* is the fee that maximizes $(1-r)w$ subject to that constraint. Thus, the client's choice boils down to r^* or $\overline{r}$.

The client's return from using r^* is $(1-r^*)w^*$; his return from using $\overline{r}$ is $\overline{w}-\overline{x}$. Subtracting the latter from the former yields

$$(1-r^*)[w^*- \overline{w}] - (r^*-\overline{r})\overline{w}; \tag{A6}$$

substituting in **(A5)** and rearranging terms, we find that **(A6)** is positive (meaning r^* is preferable to $\bar{r}$) iff

$$(1-r^*)[w^*-\bar{w}] \;>\; r^* - \frac{\bar{x}}{\bar{w}}. \tag{A7}$$

As $\bar{x}\rightarrow 0$, the limit of the left-hand side is $(1\text{-}r^*)w^*$, while the limit of the r.h.s. is r^*. It follows that a value of $\bar{x}$ can be chosen (one sufficiently close to zero) for which **(A7)** is satisfied. As $\bar{x}\rightarrow x^*$, the limit of the l.h.s. is zero, while the limit of the r.h.s. is some positive number.[61] It follows that a positive value of $\bar{x}$ can be chosen (one sufficiently close to x^*) for which **(A7)** is not satisfied.

Fixed Costs

Assume that there is some fixed cost F that the lawyer must incur to realize any return on the claim. Formally, this generates a new production function $v(x)$, where $v(x) = w(x\text{-}F)$. The client wants to choose r to maximize $(1\text{-}r)v(x)$. Now, the lawyer, if she has invested the fixed cost F, will choose x to maximize $rv(x+F) = rw(x)$. and the client's return will be $(1\text{-}r)w(x)$. By assumption, r^* maximizes $(1\text{-}r)w(x)$. Thus, conditional on the lawyer's willingness to incur the fixed cost F, we know that r^* is the optimal fee.

The question, then, is whether the lawyer is willing to incur F. The lawyer will be willing to incur that fixed cost provided that her net expected return from the case is nonnegative, that is, if

---

[61]The r.h.s. can be rewritten as $(r^*\bar{w}\text{-}\ \bar{x})/\bar{w}$. Since r^* gives the lawyer positive rents, we know $r^*w^*\text{-}x^*$ is positive. It follows that the above fraction approaches a positive number as $\bar{x}$ approaches x^*.

$$rw - (x+F) \geq 0 . \tag{A8}$$

If **(A8)** holds at r^*, therefore, r^* is optimal for the client. Suppose, on the other hand, that **(A8)** does not hold at r^*. To get the lawyer to invest anything in the case, the client must raise the fee at least to the point at which **(A8)** is an equality.[62] However, the client does not want to raise the fee *past* that point, since by assumption for fees greater than r^*, the value of $(1-r)w$ is decreasing in r.

Settlement

Assume that cases settle for some multiple q of the expected judgment (q is assumed to be some positive constant). Suppose we want to compare r^* to some other fee percentage $\tilde{r}$. By definition,

$$(1- r^*)w^* \quad > \quad (1-\tilde{r})\tilde{w}, \tag{A9}$$

where w^* represents the expected judgment if the lawyer works under r^*, and $\tilde{w}$ represents the expected judgment if the lawyer works under $\tilde{r}$. Expression **(A9)** implies that

$$q \times (1- r^*)w^* \quad > \quad q \times (1-\tilde{r})\tilde{w},$$

or, rearranging terms, that

$$(1- r^*)qw^* \quad > \quad (1-\tilde{r})q\tilde{w}.$$

This establishes that if r^* is preferable to $\tilde{r}$ when the case is certain to go to trial, it remains so given that the case will settle. This is true for any value of $\tilde{r}$.

---

[62]Differentiatingthe left-hand side of (A5) with respect to r yields $w + (dx/dr)[rw'-1]$, which is positive. In some cases, of course, F may be so large that (A5) does not hold even at $r=1$.

Simulated Cases

In what follows we derive the general results presented in Section III, for a lawyer production function defined as

$$w(x) = \alpha(1-e^{-\beta x}).$$

Optimal Fee Arrangements in Equilibrium

Plugging the relevant terms into (A5) and simplifying yields

$$r^* = \frac{1}{\sqrt{\alpha\beta}} . \tag{A6}$$

Solving the lawyer's maximization problem (the first-order condition being $r\alpha\beta e^{-\beta x} - 1 = 0$) yields the following results:

$$
\begin{aligned}
x &= \frac{\ln(r\alpha\beta)}{\beta} \\
w &= \alpha\left(1 - \frac{1}{r\alpha\beta}\right) \\
x^* &= \frac{\ln\sqrt{\alpha\beta}}{\beta} \\
w^* &= \alpha - \sqrt{\frac{\alpha}{\beta}} \\
\hat{x} &= \frac{\ln(\alpha\beta)}{\beta} \\
\hat{w} &= \alpha - \frac{1}{\beta}\,.
\end{aligned}
\tag{A7}
$$

The lawyer's net recovery and rate of return, respectively, are given by

$$
r^* w^* - x^* = \frac{\sqrt{\alpha\beta} - (\ln\sqrt{\alpha\beta} + 1)}{\beta} \tag{A8}
$$

$$
\frac{r^* w^* - x^*}{x^*} = \frac{\sqrt{\alpha\beta} - 1 - \ln\sqrt{\alpha\beta}}{\ln\sqrt{\alpha\beta}} \tag{A9}
$$

Differentiating **(A8)** with respect to α and β, respectively, yields

$$\frac{\sqrt{\alpha\beta} - 1}{2\alpha\beta^2}$$

$$\frac{\alpha(1 + 2\ln\sqrt{\alpha\beta} - \sqrt{\alpha\beta})}{\beta},$$

The first is always positive[63]; the latter may or may not be, depending on the value of $\alpha\beta$. Differentiating **(A9)** with respect to α and β, respectively, gives

$$\frac{\sqrt{\alpha\beta}(\ln\sqrt{\alpha\beta} - 1) + 1}{2\alpha(\ln\sqrt{\alpha\beta})^2}$$

$$\frac{\sqrt{\alpha\beta}(\ln\sqrt{\alpha\beta} - 1) + 1}{2\beta(\ln\sqrt{\alpha\beta})^2}.$$

Both of these are are always positive.

Turning to the client, his net recovery under the optimal fee is given by

$$(1 - r^*)w^* = \frac{(\sqrt{\alpha\beta} - 1)^2}{\beta}; \tag{A10}$$

the fraction of the potential surplus he captures is thus

$$\frac{(1 - r^*)w^*}{\hat{w} - \hat{x}} = \frac{(\sqrt{\alpha\beta} - 1)^2}{\alpha\beta - \ln(\alpha\beta) - 1}. \tag{A11}$$

[63] Recall that $\alpha\beta > 1$.

Differentiating **(A10)** with respect to α and β, respectively, gives

$$\frac{\sqrt{\alpha\beta} - 1}{\sqrt{\alpha\beta}}$$

$$\frac{(\sqrt{\alpha\beta} - 1)[\alpha(1 - \beta) + \sqrt{\alpha\beta})]}{\beta^2 \sqrt{\alpha\beta}}$$

The first is always postive; the latter is definitely positive when $\beta < 1$, and may or may not be positive otherwise. Differentiating **(A11)** with respect to α gives

$$\frac{\beta\left[\ln(\alpha\beta)\left(\frac{1}{\sqrt{\alpha\beta}} - 1\right) + \sqrt{\alpha\beta} + \frac{1}{\alpha\beta} - 1\right]}{(\alpha\beta - \ln(\alpha\beta) - 1)^2},$$

which is positive for all $\alpha\beta > 1$.