Stanford | Institute for Economic Policy Research (SIEPR)

Subrogation and the Theory of Insurance When Suits Can Be Brought for Losses Suffered

A. Mitchell Polinsky
Stanford Law School and NBER

Steven Shavell
Harvard Law School and NBER

March, 2018

Working Paper No. 17-007

Journal of Law, Economics, & Organization, forthcoming.

Last revised: March 2018

Subrogation and the Theory of Insurance When Suits Can Be Brought for Losses Suffered

A. Mitchell Polinsky[*]
Stanford Law School and National Bureau of Economic Research

Steven Shavell[**]
Harvard Law School and National Bureau of Economic Research

Abstract: The theory of insurance is considered here when an insured individual may be able to sue another party for the losses that the insured suffered—and thus when an insured has a potential source of compensation in addition to insurance coverage. Insurance policies reflect this possibility through so-called subrogation provisions that give insurers the right to step into the shoes of insureds and to bring suits against injurers. In a basic case, the optimal subrogation provisions involve full retention by the insurer of the proceeds from a successful suit and the pursuit of all positive expected value suits. This eliminates litigation risks for insureds and results in lower premiums—financed by the litigation income of insurers, including from suits that insureds would not otherwise have brought. Moreover, optimal subrogation provisions are characterized in the presence of moral hazard, administrative costs, and non-monetary losses and it is demonstrated that optimal provisions entail sharing litigation proceeds with insureds in the first two cases but not when losses are non-monetary.

Key words: insurance; subrogation; accidents; torts; litigation; moral hazard; administrative costs; non-monetary losses

---

[*] Josephine Scott Crocker Professor of Law and Economics, Stanford Law School, Stanford, CA 94305; email: polinsky@stanford.edu.

[**] Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School, Cambridge, MA 02138; email: shavell@law.harvard.edu.

We are grateful to editor Andrew Daughety, three referees, Kenneth Abraham, Scott Baker, Tom Baker, Mark Geistfeld, Daniel Gottlieb, Scott Harrington, Oliver Hart, Louis Kaplow, Saul Levmore, Stephen Mildenhall, Ariel Porat, David Rosenberg, Ted Sichelman, Kathryn Spier, and Ralph Winter for comments; James Fox, Anastasios Moraitis, and Paul Riskind for research assistance; and the John M. Olin Program in Law and Economics at Stanford Law School and the John M. Olin Center for Law, Economics, and Business at Harvard Law School for research support.

1. Introduction

In this article we consider the theory of insurance when insured individuals may be able to sue other parties for the losses they have suffered—and thus have a potential source of compensation in addition to insurance. Such situations are ubiquitous: an insured driver whose car is damaged in an accident may be able to sue another driver who caused the accident for repair costs; an insured homeowner whose house burns down due to a gas leak may be able to sue the utility company for the value of his residence; or a person covered by health insurance who slips and falls in a store may be able to sue the proprietor for medical expenses.

Insurance policies reflect the reality that it is often possible for an insured to sue an injurer for losses sustained. Namely, insurance policies not only promise to compensate insureds for their losses, they also usually include what are known as subrogation provisions that accord the insurer the right to step into the shoes of an insured and to sue a party who caused the insured's losses.[1] Suppose that driver A owns collision insurance coverage on his car and that driver B negligently causes an accident that damages A's car. Driver A would receive payment from his insurer for repair costs after the accident and then, under the subrogation terms of his insurance policy, driver A's insurer could sue driver B. Such a subrogated insurer typically would be authorized to retain much or all of the proceeds from suit[2] and also would bear much or all of the litigation costs.[3]

[1] For a general description of subrogation, see Jerry and Richmond (2012, pp. 648-54). See also Abraham and Schwarcz (2015: 266-68), Baker and Logue (2017: 297-300), and Harrington and Niehaus (2004: 195). Subrogation provisions may be explicit in insurance contracts, supplied by statute, or legally presumed to apply; see Jerry and Richmond (2012: 651-54). In some instances, subrogation effectively comes about when insureds themselves bring suit and then reimburse insurers for the coverage that they had received; see 16 *Couch on Insurance* §226:1, 3, 4, 41 (2017). On the origins of subrogation, which date from antiquity, see Marasinghe (1976a, b).

[2] On the practices and legal doctrines governing the amount that subrogated insurers retain from judgments, see generally 16 *Couch on Insurance* §223 (2017). The subrogation income that insurers collect can be significant;

In the United States, subrogation provisions are a common feature of property insurance, liability insurance, health and medical insurance, and disability insurance policies.[4] Subrogation provisions may also apply to governmentally-provided insurance, notably, to workers compensation, Medicare, and Medicaid programs.[5] The prevalence of subrogation provisions is similar in other countries as well.[6]

Our object here is to demonstrate that subrogation provisions are a fundamental feature of optimal insurance contracts—those that maximize the expected utility of insureds subject to the constraint that the insurer covers its expected expenses—and also to study the specific character of optimal subrogation provisions. To this end, we build on a model of subrogation introduced by Shavell (1987: 255-56) in which a risk-averse individual who confronts the risk of an accident may purchase insurance and, if an accident occurs, may have his insurer sue the injurer on his behalf. Suit is assumed to be successful only with a probability, in which event the injurer must pay damages equal to harm. Unlike Shavell, however, we suppose that litigation is costly.

In Section 2, we show that insurance policies with subrogation provisions are superior to pure insurance policies—policies that pay coverage after an accident occurs but that do not address the possibility of suit. In essence, there are two reasons for this conclusion. First, as identified by Shavell, the insured avoids the risk of litigation by subrogating his legal claims to

for example, a study cited in Perry et al. (2015) found that for automobile insurers such income amounted to 27% of their paid losses for collisions.

[3] The insurer would usually be responsible for the entirety of litigation costs when the insurer retains the full judgment but only for a proportional share of the costs when the insurer keeps less than the full judgment; see 16 *Couch on Insurance* §223: 113, 119 (2017).

[4] See, for example, Dobbyn and French (2016: 425-28) on property, liability, health and medical insurance, and Bleed (2001: 734) and Dobbyn and French (2016: 46) and on disability insurance.

[5] See, for example, Dobbyn and French (2016: 429-30) on workers compensation and Jerry and Richmond (2012: 652) on Medicare and Medicaid.

[6] See, for example, *Insurance Day* (2009), reviewing subrogation worldwide, including in England, Germany, France, Russia, Spain, Switzerland, Brazil, India, Singapore, China, and Australia.

the insurer. In exchange for forgoing his claims, the insured obtains a lower insurance premium because the insurer receives income from suit. Second, by subrogating his claims to the insurer, the insured obtains a reduction in the premium because the insurer will bring some positive expected value suits that the insured would not otherwise have pursued. Specifically, if the insured did not have a subrogation provision, he would not have pursued some positive expected value suits because of his risk aversion. The reduction in the premium due to the insurer bringing such suits did not arise in Shavell because he assumed that litigation is costless.[7]

In Section 3, we study the nature of optimal insurance policies with subrogation provisions in the presence of moral hazard, administrative costs, and non-monetary losses. Each of these factors leads to the optimality of less than complete insurance coverage. The shortfall of insurance coverage raises the question whether the insured should receive a share of the litigation proceeds obtained by a subrogated insurer in order to better compensate him. This question is the subject of debate in both legal literature and judicial opinions.[8] We address it with respect to each of the three factors just noted.

In the case of moral hazard, we assume that insureds can reduce the probability of one kind of accident by the exercise of unobservable care (for example, homeowners can reduce the risk of fire by storing flammables away from heat sources), whereas insureds play no role in the occurrence of a different kind of accident that is caused by other parties (home fires can arise due to the negligence of contractors). Optimal insurance coverage will be less than complete for the familiar reason that some exposure to risk will provide an incentive for insureds to take care to reduce accident risk—here the risk of the first kind of accident. However, we demonstrate that

[7] If litigation is costless, a risk-averse individual would bring all suits because he would not suffer any loss if a suit is unsuccessful.

[8] See, for example, Jerry and Richmond (2012: 657-60).

the optimal subrogation provision will award the insured a positive share of the damages from a successful suit. The reason is that the receipt of a share of damages by the insured will help to offset the portion of the loss not covered by insurance. Significantly, this implicit form of additional coverage will not contribute to moral hazard. That is because, for a suit to be successful, the accident must be found by a court to have been caused by another party (a contractor rather than a homeowner). Moreover, we show that it is desirable for the insurer to pursue some negative expected value suits because the benefit of the implicit additional coverage from the associated subrogation payments exceeds the expected net loss from bringing the suits.

With regard to administrative costs, we suppose that insurers bear expenses (notably, due to checking the veracity of claims) that rise with the level of coverage. Insurance premiums will then include a loading above their actuarially fair level, and optimal insurance coverage will be less than full, as is well-known. We show that the optimal subrogation provision will provide the insured a positive share of damages from a successful suit. The receipt of such damages by the insured will beneficially reduce the portion of the loss not covered by insurance, as in the case of moral hazard, and this additional coverage will be effectively free of administrative costs.[9] It also will be in the insured's interest for some negative expected value suits to be pursued by the insurer for essentially the reason noted in the case of moral hazard.

In the case of a non-monetary loss (such as the death of a spouse), it is well understood that optimal insurance coverage would not reflect such a loss (presuming that it does not affect the marginal utility of wealth). Optimal coverage would be restricted to monetary losses. Courts, however, award damages for both types of losses. It follows that under the optimal subrogation provision, the insurer would retain all of the damages from suit, including the damages for the

[9] The administrative costs at issue are negligible, as they are only those associated with transferring funds to the insured. (This statement is not inconsistent with the fact that litigation costs are distinctly positive.) See Section 3.B.

non-monetary loss. The insured would of course benefit from the insurer's receipt of damages for the non-monetary loss through a reduction in the insurance premium. Additionally, under the optimal subrogation provision, the insurer would bring suit if and only if the expected return is positive because, unlike in the moral hazard and administrative cost cases, the insured would purchase ideal insurance coverage.

We conclude in Section 4 with several observations about issues that are not addressed in our analysis, including litigation efficiencies enjoyed by insurers as a factor favoring subrogation, undesirable legal restrictions on subrogation, and the possible sale of claims as a substitute for subrogation.

Before proceeding, we note that little attention has been paid to subrogation in the economic literature on insurance[10] and we are aware of only three publications that formally address the topic: Shavell (1987: 255-56), Sykes (2001), and Gomez and Penalva (2015). As discussed above, Shavell recognized that subrogation is desirable because it shifts litigation risk from the insured to the insurer, but because he assumed that litigation is costless, he did not observe that subrogation is also desirable because it leads to the bringing of all positive expected value suits. Sykes examined the optimal form of subrogation provisions allowing for moral hazard, administrative costs, and non-monetary losses; he did not, however, consider why subrogation provisions are generally desirable. Sykes found that the insured should not receive any subrogation proceeds in the cases of moral hazard and administrative costs, and all the subrogation proceeds in the case of non-monetary losses—the opposite of our conclusions in all three cases. The reason for these differences is that Sykes made assumptions that lack apparent

[10] Notably, none of the thirty-seven articles in Dionne (2013)—*Handbook of Insurance*—considers subrogation.

economic justification.[11] Gomez and Penalva focus on identifying the effects of subrogation on the deterrence of accidents rather than, as here, on the form of the optimal insurance contract with subrogation.[12]

2. **Basic Analysis**

We consider here the standard model of insurance for accident risks, but supplemented by the ability of insured individuals to bring suit for harm that they suffer. In particular, we assume that identical risk-averse individuals face a chance of suffering harm from an accident; that they purchase insurance policies that maximize their expected utility; and that if an accident occurs, they will be able to sue a potentially liable party for recompense, where suit would involve a litigation cost and result in success only with a probability.[13]

Let

y = initial wealth of an individual;

$U(\cdot)$ = utility of an individual from wealth, where $U'(\cdot) > 0$ and $U''(\cdot) < 0$;

p = probability of an accident; $p \in (0, 1)$;

h = harm if an accident occurs; $h > 0$;

k = cost of bringing a suit; $k \geq 0$; and

q = probability of winning a suit, resulting in an award of h; $q \in [0, 1]$.

---

[11] See notes 23, 26, 28, and 30 below.

[12] However, in the course of their analysis, they duplicated Shavell's result. See also notes 18 and 40 below. Of additional interest is Viscusi's (1989) empirical investigation of subrogation, as well as legal literature on the subject (which often reflects economic views), including Kimball and Davis (1962), Abraham (1986: 153-55), Reinker and Rosenberg (2007), and articles cited in 16 *Couch on Insurance* Part XI (2017).

[13] We abstract in the model from the possibility of settlement because we want litigation to involve risk—otherwise we cannot study how subrogation alleviates litigation risk. We could have considered a more complicated model involving the possibility of both settlement and trial (notably, a model involving asymmetric information in settlement bargaining), and thus involving litigation risk. But we do not believe that such a model would lead to significant additional insight.

For simplicity, we assume that there is a single value of each of the above variables (and thus that these values are known to all parties).

A *pure insurance policy* is defined by a premium that the insured pays at the outset and a coverage amount that the insured receives immediately after an accident occurs but before a suit might be brought by the insured. This assumption about the time at which coverage is received reflects two aspects of reality: that a person will frequently have consumption needs that he must satisfy shortly after an accident (like replacing a car that he requires for transportation); and that litigation outcomes (or settlements) usually occur significantly later.[14]

An *insurance policy with a subrogation provision* adds to the pure insurance policy an instruction to the insurer whether to bring a suit on behalf of the insured after an accident occurs,[15] states that the insurer will bear litigation costs k if a suit is brought,[16] and specifies the share (perhaps zero) that the insurer will pay to the insured from the award h if a suit is successful.

Let

π = insurance premium;

c = insurance coverage; $c \geq 0$;

[14] In the course of analyzing the model presented here, we also investigated a model in which consumption occurs at two times—directly following an accident and after the outcome of litigation. In the latter model, it was optimal for the insured to receive compensation immediately following an accident. However, this two-consumption-period model did not provide much understanding beyond that gained from the model that we study here in which insurance payments are assumed to be made just after an accident.

[15] Because our purpose is to derive the terms of the insurance contract that maximizes the expected utility of insureds and it is possible for the contract to include instructions about when to sue, we study such instructions. In practice, however, we would not expect insurance contracts to have explicit instructions about when the insurer should bring a suit due to problems of asymmetry of information (see Section 4).

[16] This assumption is made for simplicity; it would be straightforward to show that it is a feature of the optimal contract in the basic analysis.

ϕ = instruction whether the insurer sues under a subrogation provision; if $\phi = 0$, suit is not brought; if $\phi = 1$, suit is brought; and

s = insured's share of the award h under a subrogation provision if a suit is successful; $s \in [0, h]$.

Thus, a pure insurance policy is described by a premium π and a coverage level c; and a policy with a subrogation provision is described additionally by an instruction ϕ about suit, and the insured's share s of an award.

By the optimal insurance policy we mean the policy (among either pure policies or policies with a subrogation provision) that maximizes the expected utility of the insured subject to the constraint that the premium equals the expected expenses of the insurer.

Under a pure insurance policy, the premium constraint is

$$\pi = pc. \tag{1}$$

The expected utility of the insured given c will depend on whether he would bring a suit if an accident occurred. If he would not, his expected utility would be

$$EU_N(c) = (1-p)U(y-\pi) + pU(y-\pi-h+c). \tag{2}$$

If the insured would bring a suit, his expected utility would be

$$EU_S(c) = (1-p)U(y-\pi) + p[qU(y-\pi+c-k) + (1-q)U(y-\pi-h+c-k)], \tag{3}$$

where h does not appear in the first term in brackets because the insured receives h from a successful suit. Therefore, given any c, the insured will bring a suit after an accident occurs if and only if (3) exceeds (2), and his expected utility will be $max[EU_N(c), EU_S(c)]$.[17] The optimal pure insurance policy therefore solves the problem

$$max\{max[EU_N(c), EU_S(c)]\} \text{over } c. \tag{4}$$

[17] We assume for concreteness that if the insured is indifferent between suing and not suing, he will not sue, and we will adopt similar conventions below without comment.

Let c^* denote the optimal level of coverage. We now show

Proposition 1. Under the optimal pure insurance policy,

(a) if suits are not optimal to bring, optimal coverage is full, $c^* = h$, and each insured thus has certain wealth of $y - ph$;

(b) if suits are optimal to bring, optimal coverage c^* is in $(k, h + k)$, where c^* is determined by

$$U'(y - pc) = qU'(y - pc + c - k) + (1 - q)U'(y - pc - h + c - k), \tag{5}$$

and each insured has expected utility of $EU_S = (1 - p)U(y - pc^*) + p[qU(y - pc^* + c^* - k) + (1 - q)U(y - pc^* - h + c^* - k)]$; and

(c) suits are optimal to bring if and only if their expected value exceeds a positive threshold: given any q and h, there exists a t satisfying $0 < t < qh$ such that if $qh - k > t$, suit is optimal to bring; and if $qh - k \leq t$, suit is not optimal to bring.

The proofs of this and subsequent propositions are contained in the Appendix.

Notes. (i) In part (a) the problem for the insured is the standard insurance problem in which optimal coverage is well-known to be full.

(ii) The intuition underlying (b) is that, if an accident occurs, the insured will bear k in litigation costs, suggesting that he would want to insure for at least that amount; and since he will also bear the risk of losing the suit, he might want to insure for as much as h more. It is plausible, however, that c^* would be significantly less than $h + k$, and, indeed, less than h. The reason is that c would often result in excessive total compensation, for if the suit is successful, the insured collects damages of h in addition to c. This point is reinforced by the fact that the probability q that suit would be won will tend to be high, since the hypothesis is that suit is optimal to bring.[18]

[18] Gomez and Penalva (2015) found that $c^* < h$ in their model, in which litigation costs were assumed to be zero; their result is a special case of our conclusion when $k = 0$.

(iii) Part (c) follows from the fact that the insured bears litigation risk, rendering some positive expected value suits not worth bringing.

Next consider the optimal insurance policy with a subrogation provision. In this case, if the instruction is $\phi = 0$, not to sue after an accident, the premium constraint will be (1), expected utility will be (2), and by Proposition 1(a), optimal coverage will be $c^* = h$ and utility will be $U(y - ph)$. If, however, $\phi = 1$, to sue after an accident, the premium constraint will be

$$\pi = pc - p(q(h - s) - k) \tag{6}$$

because the insurer will receive income of $h - s$ if it wins a suit and will incur the litigation cost k if there is an accident. Hence, the expected utility of the insured will be

$$EU_S(c, s) = (1 - p)U(y - \pi) + p[(1 - q)U(y - \pi - h + c) + qU(y - \pi - h + c + s)]. \tag{7}$$

The optimal insurance policy with a subrogation provision solves the problem

$$max[U(y - ph), max\ EU_S(c, s) \text{ over } c \text{ and } s], \tag{8}$$

and we have

Proposition 2. Under the optimal insurance policy with a subrogation provision,

(a) if suits are not optimal to bring, optimal coverage is full, $c^* = h$, and each insured has certain wealth of $y - ph$;

(b) if suits are optimal to bring, optimal coverage is full, $c^* = h$, the optimal subrogation share is zero, $s^* = 0$, and each insured has certain wealth of $y - ph + p(qh - k)$;

(c) suits are optimal to bring if and only if they have positive expected value, $qh - k > 0$, in which case insureds are better off by the positive amount $p(qh - k)$ than if suits are not brought; and

(d) the insured's expected utility is strictly higher than under the optimal pure insurance policy if suits have positive expected value (and otherwise is equivalent).

Notes. (i) The explanation of part (a) is the same as that above.

(ii) Part (b) holds because it is desirable for the insured to be fully protected against both accident risk and litigation risk, and this is accomplished when coverage is full and the subrogation share is zero.

(iii) Part (c) follows from (b). Because insureds are fully protected against all risks, they will be made better off from suit if and only if suit results in a lower premium. That will be true when suit has positive expected value.

(iv) Part (d) follows from observing that, under the optimal pure insurance policy, one of two inferior outcomes will occur when suits have positive expected value: either suits will not be brought, in which case wealth will be certain but lower than that under the optimal insurance policy with a subrogation provision because the insurer does not earn the additional expected income $p(qh - k)$ from bringing suit; or suits will be brought, but insureds will bear accident risk (because c^* is generally different from h) as well as litigation risk.

3. Moral Hazard, Administrative Costs, and Non-Monetary Losses

We now consider optimal subrogation provisions given three common factors that alter the nature of optimal insurance coverage: moral hazard, administrative costs, and non–monetary losses. Importantly, as noted in the introduction, each of these factors leads optimal insurance coverage to be less than sufficient to make the insured whole. We show that in the cases of moral hazard and administrative costs, the optimal subrogation provision provides for the insured to obtain a positive portion of the proceeds from successful litigation, but not in the case of non-monetary losses.

A. **Moral Hazard**

We here consider a problem of moral hazard, in which the incentive of an insured to take care to reduce the risk of an accident is dulled by his receipt of insurance payments. Specifically, we assume that the insured's level of care is not observable by the insurer and that it is non-monetary. Let

x = level of care of an insured; $x \geq 0$; and

$d(x)$ = disutility of care; $d(0) = 0$; $d'(0) = 0$; $d'(x) > 0$ for $x > 0$; $d''(x) > 0$.

We assume that that there are two mutually exclusive types of accident. The first type is caused by the insured, and its probability can be lowered by his exercise of care. For example, as we said in the introduction, a home owner can reduce the risk of fires by storing flammables away from the furnace. Let

$p_1(x)$ = probability of a type 1 accident, caused by an insured; $0 < p_1(x) < 1$; $p_1'(x) < 0$; $p_1''(x) > 0$.

The second type of accident is caused by another party, such as a contractor whose faulty installation of wiring leads to a house fire. We assume that the probability of this type of accident is exogenous.[19] Let

p_2 = probability of a type 2 accident, caused by another party; $p_2 > 0$.

Thus, the total probability of an accident is

$$p(x) = p_1(x) + p_2. \tag{9}$$

If an accident occurs, the insurer is assumed not to have any information about its type, so that the instruction ϕ whether to sue cannot be conditioned on its type. If a suit is brought,

[19] It will be seen that under this assumption, no moral hazard will be created by making a subrogation payment to the insured when an accident is of type 2. The significance of the assumption is only that the moral hazard associated with type 2 accidents is less than that associated with type 1 accidents; the moral hazard need not be zero.

however, the court is assumed to be able to determine which type of accident occurred and to award a judgment to the insured if and only if the accident was of the second type, caused by another party.[20]

Suppose that the instruction is not to sue after an accident. Then the insured's expected utility is

$$EU_N = (1 - p(x))U(y - \pi) + p(x)U(y - \pi - h + c) - d(x), \tag{10}$$

where neither c nor π can depend on x because the insurer is unable to observe it. The optimal insurance policy maximizes EU_N over c and π subject to

$$\pi = p(x)c, \tag{11}$$

and

$$\text{the insured maximizes } EU_N \text{ over } x, \text{ given } c \text{ and } \pi. \tag{12}$$

The role of moral hazard in this problem is obvious from (12) when $c = h$. Then the insured's expected utility is $EU_N = U(y - \pi) - d(x)$, in which case the insured would choose $x = 0$. If, however, c is less than h, then the insured's exposure to risk will be seen to lead him to choose $x > 0$. Let x_N be the x chosen by the insured under the optimal insurance policy assuming that suit is not brought.

Now assume that $\phi = 1$, that suit will be brought if an accident occurs. Then the insured's expected utility is

[20] We presume that the court does not observe the insured's level of care, and thus that the insurer cannot obtain such information from the court in order to combat moral hazard. The motivation for this assumption is that in reality there are many dimensions of care that a court would not be likely to consider. For instance, suppose that a homeowner sues a contractor because the contractor's negligent installation of wiring caused a fire that destroyed his kitchen. The evidence in the legal proceeding relevant to the homeowner's claim would concern the wiring in the kitchen. If the homeowner had also been negligent with respect to something he did bearing on fire risk elsewhere in the house—say leaving oily rags near the furnace in the basement—that fact would ordinarily not be relevant to the claim at issue, in which case the insurer would not learn about the home owner's negligence from the suit against the contractor.

$$EU_S = (1 - p(x))U(y - \pi) + p(x)[(1 - q)U(y - \pi - h + c) + qU(y - \pi - h + c + s)] - d(x). \tag{13}$$

The optimal insurance policy maximizes EU_S over c, s, and π subject to the constraints

$$\pi = p(x)c - p(x)(q(h - s) - k), \tag{14}$$

and

the insured maximizes EU_S over x, given c, s, and π. (15)

Observe that q, the probability of winning a suit conditional on an accident having occurred, is $p_2/(p_1(x) + p_2)$, because our assumption is that a suit is won if and only if the accident turns out to have been caused by another party. Because q depends on x we will sometimes write it as $q(x)$. Note that $p(x)q(x)$, the unconditional probability that a judgment will be obtained, is $(p_1(x) + p_2)[p_2/(p_1(x) + p_2)] = p_2$. The explanation is that all accidents caused by other parties will result in liability, whereas no accidents caused by insureds will result in liability. It follows that *the level of care cannot affect the number of judgments* because x influences only the probability of accidents caused by insureds. Relatedly, because $p(x)q = p_2$, $p(x)(1 - q) = p(x) - p_2 = p_1(x)$; that is, the probability of having an accident and then losing at trial is the probability of having an accident that was caused by the insured.

Hence, (13) becomes

$$EU_S = (1 - p(x))U(y - \pi) + p_1(x)U(y - \pi - h + c) + p_2U(y - \pi - h + c + s) - d(x), \tag{16}$$

and (14) becomes

$$\pi = p(x)(c + k) - p_2(h - s). \tag{17}$$

Consequently, the choice of care of insureds affects their premiums only through the coverage amount and litigation costs, not through the return to insurers from bringing suits.

Given (16), the constraint (15) can be shown to be equivalent to the first-order condition

$$-p_1'(x)[U(y-\pi) - U(y-\pi-h+c)] = d'(x); \tag{18}$$

that is, the marginal benefit from reducing the risk of type 1 accidents (and thus bearing $h - c$) equals the marginal cost of care. Note that the subrogation payment s does not appear in (18) and thus does not directly affect the insured's choice of care; this is because s is obtained only if there is a judgment, and the probability p_2 of a judgment is not affected by his care.

We can now describe the optimal insurance policy.

Proposition 3. Suppose that moral hazard is associated with the provision of insurance coverage (as described). Then under the optimal insurance policy with a subrogation provision,

(a) if suits are not optimal to bring, optimal coverage is incomplete, $c^* < h$;

(b) if suits are optimal to bring, optimal coverage is incomplete, $c^* < h$, and the optimal subrogation payment is positive, $s^* > 0$; and

(c) suits are optimal to bring whenever they would have positive or zero expected value if suits were not brought, $q(x_N)h - k \geq 0$, and some negative expected value suits are also optimal to pursue.

Notes. (i) When suits are not optimal to bring, the insurance policy becomes an example of the standard insurance policy involving moral hazard, in which $c^* < h$ is optimal so as to induce the insured to take some positive level of care.[21]

(ii) When suits are optimal to bring, the result that $c^* < h$ is essentially due to the reason just given. The explanation for $s^* > 0$ is as follows. First, since $c^* < h$, the marginal utility of wealth after an accident and receipt of c^* is higher than in the no-accident state, meaning that arranging for a positive s is similar to arranging for a beneficial increase in insurance coverage. Second, the possible countervailing argument that receipt of a payment s from the insurer would

[21] See, for example, the synthesis by Winter (2013: 208-12).

dilute the incentives of the insured to take care is not applicable: as we observed (after (18)), the insured's receipt of s is independent of his choice of care x because s is obtained only when an accident is caused by another party.[22] Consequently, raising s from zero reduces the risk-bearing created by $c^* < h$ without contributing to moral hazard.[23]

(iii) That suits are optimal to pursue when they have positive expected value is based on the following logic. Given any contemplated level of coverage c, premiums can be lowered if such suits are brought, and the insured need not bear any additional risk since $s = 0$ could be chosen. This suggests that insureds must be better off, but the effect of a lower premium on the optimally chosen level of care needs to be taken into account to complete this argument. That suits with zero or small negative expected value are optimal to bring is explained by the point made in note (ii) that, since $c^* < h$, risk-bearing is reduced by bringing suit because that permits $s^* > 0$.

(iv) We have implicitly assumed that the insurance contract would not be renegotiated, but it would be in the insured's and insurer's ex post interest to do so: since $s^* > 0$, the insured bears risk, and because the insurer is risk-neutral, they would both benefit if the insured were to receive a certain amount before a suit is brought (for example, $qs - \varepsilon$ for a small ε). However, it is not in the ex ante interest of the parties for this to occur; such a renegotiated contract would contribute to moral hazard, for the payment to the insured would be made whether or not he

[22] Actually, arranging for $s > 0$ turns out to *increase* care. The reason is that when $s > 0$, the premium rises, and an increase in the premium will elevate care (because the premium increase is equivalent to a reduction in the insured's wealth). See the demonstration that $x'(s) > 0$ in step (v) of the proof of part (b).

[23] Sykes (2001: 393-94) supposed that the problem of moral hazard is identical whether an accident results in a successful suit or does not. If that were true, s^* would equal zero. But it is implausible that the fact that a suit is successful would provide no information to an insurer about the role of the insured in an accident (in our case, about whether the accident was caused by the insured or by another party).

caused the loss.[24] Consequently, the insurer has an interest in suppressing renegotiation and may be able to do so by prohibiting employees from changing stated policy terms. Moreover, renegotiation may not occur because of the cost of bargaining or may fail due to asymmetric information. To the extent that renegotiation does not occur or fails for the preceding reasons, our rationale for s^* to be positive would apply.

B. Administrative Costs

We suppose here that insurers bear administrative costs that rise with the level of coverage. Notably, the cost to an insurer of checking the validity of a claim would tend to increase with the magnitude of coverage.[25] We assume for simplicity that such administrative costs rise in proportion to coverage. Let

λ = the administrative cost per dollar of coverage; $\lambda > 0$,

where λ will also be referred to as the *loading factor*.

If the instruction in the policy is not to sue after an accident, the premium constraint is

$$\pi = p(1 + \lambda)c. \tag{19}$$

In this case, expected utility given c is $EU_N(c)$ as stated in (2), where π is determined by (19).

When the instruction is to bring suit if an accident occurs, the premium constraint is

$$\pi = p(1 + \lambda)c - p(q(h - s) - k). \tag{20}$$

[24] The potential for renegotiation to which we refer is a version of that studied in Fudenberg and Tirole (1990) and related literature on renegotiation and moral hazard. In that literature, possible renegotiation occurs after the agent has chosen his unobservable level of care and before the outcome that depends on his care choice—loss or no loss in the context of insurance—eventuates. Here, the renegotiation occurs after the possible loss—but before a suit is resolved.

[25] For a discussion of this and other reasons for a loading on coverage, see, for example, Harrington and Niehaus (2004: 180-81). To elaborate on the reason mentioned in the text, consider a model with a z of levels of harm h distributed with positive density $f(h)$. Suppose that the cost to the insurer to verify that an insured has suffered at least harm z is $k(z)$, where $k'(z) > 0$. If the insured's coverage is c, the insurer will have to verify the full harm for $h \leq c$ and only harm up to c for $h > c$. Hence, the insurer's expected claim verification costs are

$$\int_0^c k(h)f(h)dh + k(c)[1 - F(c)],$$

which are increasing in coverage c.

Note that although the loading factor applies to insurance coverage c, we assume that it does not apply to the subrogation payment s to the insured. The justification for this assumption is twofold. First, if a suit is won, the court will have validated the insured's claim concerning the harm that he suffered. Second, the insurer would bear essentially no additional resource cost in making the payment of s—the cost of mailing a check or of wiring funds is negligible.[26] The expected utility of the insured will be $EU_S(c, s)$ as given in (7), with π determined by (20).

The optimal insurance policy with a subrogation provision solves the problem

$$max[max\ EU_N(c) \text{ over } c, max\ EU_S(c, s) \text{ over } c \text{ and } s], \tag{21}$$

leading to the following result.

Proposition 4. Suppose that there is a positive administrative cost λc associated with the provision of insurance coverage c. Then under the optimal insurance policy with a subrogation provision,

(a) if suits are not optimal to bring, optimal coverage is incomplete, $c^* < h$;

(b) if suits are optimal to bring, optimal coverage is incomplete, $c^* < h$, the optimal subrogation payment is positive, $s^* > 0$, and $c^* + s^* < h$; and

(c) suits are optimal to bring whenever they have positive or zero expected value, $qh - k \geq 0$, and some negative expected value suits are also optimal to pursue.

Notes. (i) When suits are not optimal to bring, c^* provides less than full coverage for a well-understood reason: were coverage full, a person would be locally risk-neutral, but because of the loading factor, reducing coverage a small amount would raise his expected wealth.[27]

[26] Sykes (2001: 392) overlooked this point and assumed that a positive loading factor is associated with making a subrogation payment to an insured.

[27] See, for example, Schlesinger (2013: 170).

(ii) When suits are optimal to bring, c^* is less than complete not only due to the loading factor, but also due to the possibility that a suit would be won and that a positive s^* would be obtained, reducing the need for coverage.

(iii) The reasons that $s^* > 0$ are essentially these: on one hand, since $c^* < h$, the marginal utility of wealth after an accident and receipt of c^* is higher than in the no-accident state; and on the other hand, the implicit premium that the insured pays to obtain s^* is actuarially fair because, as we observed above, there is no administrative cost associated with the payment of s^* to the insured.[28] In other words, an individual's election of $s^* > 0$ is similar to his purchasing insurance on actuarially fair terms in the event that his suit is won, when his usual insurance coverage was incomplete due to the loading factor.

(iv) However, choosing a positive s^* is not identical to purchasing insurance coverage, for the implicit premium paid for s^* (namely, pqs^*) is borne in each contingency faced by the insured, including when he has an accident and loses his subsequent suit. In that contingency, his wealth is lowest, so the increase in the premium is relatively costly in utility terms. This observation explains why $c^* + s^* < h$ rather than equals h.

(v) The explanation of the suit decision parallels that in the case of moral hazard. Bringing a positive expected value suit beneficially lowers the premium (and the insured need not bear any risk since $s = 0$ can be chosen), and bringing a suit with zero or small negative expected value is desirable since that permits $s^* > 0$ (with a probability), thereby partially offsetting incomplete coverage.

(vi) As in the case of moral hazard, we have implicitly assumed that the insurance contract would not be renegotiated. Again, it would be in the interest of the insured and insurer

[28] In contrast, Sykes (2001: 391-92) concluded that s^* is zero. This is because he assumed that the same loading factor λ is associated with a payment s to an insured as with insurance coverage c. But, for the reason provided above, we believe that the proper assumption is that there is no loading cost associated with a positive s.

to do so: because $s^* > 0$, the parties would have an incentive to arrange for the insured to obtain a certain payment (less than qs). Such an agreement would not, however, alter the thrust of our point, which is that the optimal subrogation contract results in the insured receiving a portion of any financial benefit obtained by the insurer from the bringing of a suit.

C. <u>Non-Monetary Losses</u>

Now suppose that the losses from an accident include both a monetary component and a non-monetary component, such as when the injured party both bears medical expenses and suffers pain. Let

u = loss of utility due to the non-monetary component of harm; $u > 0$; and

a = court award for the non-monetary loss; $a > 0$.

If a person is involved in an accident and his level of wealth is z, assume that his utility will be $U(z) - u$. Accordingly, the occurrence of the non-monetary loss will not affect his marginal utility of wealth.[29] In all other respects, the model analyzed in this subsection is that employed in Section 2, including that the monetary loss is h and that the court will award h for this component of loss.

If suit will not be brought after an accident, the premium will be (1) and the insured's expected utility given c will be

$$EU_N(c) = (1 - p)U(y - \pi) + p[U(y - \pi - h + c) - u]. \tag{22}$$

If a suit will be brought after an accident, the premium will be

$$\pi = pc - p(q(h + a - s) - k) \tag{23}$$

because the insurer's expected profit given that an accident occurs is $q(h + a - s) - k$. Hence, the insured's expected utility given c and s will be

[29] This assumption seems most realistic, though cases in which non-monetary losses affect the marginal utility of wealth are considered in the literature on insurance; see originally Arrow (1974) and Cook and Graham (1977).

$$EU_S(c, s) = (1 - p)U(y - \pi) \tag{24}$$

$$+ p[(1 - q)U(y - \pi - h + c) + qU(y - \pi - h + c + s) - u].$$

The optimal insurance policy with a subrogation provision solves

$$max[max\ EU_N(c) \text{ over } c,\ max\ EU_S(c, s) \text{ over } c \text{ and } s], \tag{25}$$

and is as follows.

Proposition 5. Suppose that there are both monetary and non-monetary losses when an accident occurs. Then under the optimal insurance policy with a subrogation provision,

(a) if suits are not optimal to bring, optimal coverage equals the monetary loss alone, $c^* = h$;

(b) if suits are optimal to bring, optimal coverage equals the monetary loss alone, $c^* = h$, and the optimal subrogation share is zero, $s^* = 0$; and

(c) suits are optimal to bring if and only if they have positive expected value, $q(h + a) - k > 0$.

Notes: (i) The result in part (a) is the well-understood point that optimal insurance coverage for a non-monetary loss is zero when such a loss does not affect the marginal utility of wealth (Arrow 1974; Cook and Graham 1977).

(ii) The result in part (b) is due to two factors: on one hand, optimal insurance coverage is limited to monetary losses *h* as just observed; on the other hand, the insured can forgo collecting a positive subrogation share but still benefit from the court award *a* for the non-monetary loss through a lower premium.[30]

(iii) The result in part (c) is due to the point that the insurer's income from suits, and thus the insured's reduction in his premium, will be maximized if and only if all positive expected

[30] Sykes (2001: 389) found that it is optimal for the insured to retain the entire award *a* for non-monetary losses. That is because he assumed that it is impossible for the insured—who brings suit himself—to transfer the award *a* to the insurer. There is no apparent economic justification for this assumption.

value suits are pursued. Unlike in the cases of moral hazard and administrative costs, there is no motive to bring suits with small negative expected values in order to improve insurance coverage through a positive subrogation payment since the insured will already have been fully covered for his monetary loss.

4. Discussion

We conclude with observations about several issues that we did not consider above.

Litigation efficiencies enjoyed by insurers. An additional factor favoring subrogation is that insurers can probably more cheaply prosecute cases on behalf of their insureds because of economies of scale in litigation and informational advantages. Although large law firms might also benefit from economies of scale, insurers arguably would still enjoy lower costs because, in order to fulfill their insurance functions, they would already have obtained information relevant to determining the magnitude of losses and the behavior of insureds.

Cooperation of insureds in the litigation process. When suit is desirable under the terms of subrogation provisions, the cooperation of insureds—notably, supplying information needed in the litigation process and giving testimony—will be important to achieving a successful outcome. This suggests not only that optimal subrogation provisions would include clauses requiring such cooperation—which they do in fact[31]—but also that it would be beneficial for the subrogation share paid to insureds to be positive to better motivate them to cooperate.

Asymmetric information and delegation of the decision to litigate. Although we assumed that both insureds and insurers know the probability of winning at trial q and litigation costs k, in reality insureds especially will tend to have imperfect information about these variables. Thus, instructions in the insurance policy concerning litigation cannot practically depend on q and k. If

[31] See Abraham and Schwarcz (2015: 208) for a typical example.

as a result the insured delegates the litigation decision to the insurer, the insurer may or may not bring suit when the insured would wish. In the basic case in Section 2 and in the case of non-monetary losses, the insured's and the insurer's interests will be aligned. In those two cases the insured desires that the subrogation payment s is zero and that suits are brought if and only if they have positive expected value. Clearly, if s is zero—meaning that the insurer keeps the entire proceeds from litigation—the insurer will decide to bring suits exactly when they have positive expected value. However, in the cases of moral hazard and loading, there will be a conflict of interest between the insured and the insurer. For then the insured desires that s is positive and that suits are brought if they have positive expected value or a small negative expected value. But if s is positive, the insurer will not capture the entire judgment and therefore will not bring some positive expected value suits and of course will not bring any negative expected value suits. In the light of such conflicts of interest between insureds and insurers, it is not surprising that subrogated insurers face general duties not to unduly compromise the well-being of insureds in litigation.[32]

Legal restrictions on subrogation. Subrogation provisions may be constrained by law,[33] and such constraints may adversely affect the welfare of insureds.[34] Notably, the law bars subrogated insurers from collecting awards for pain and suffering.[35] This limitation reduces the expected utility of insureds because, as we discussed in Section 3.C, insureds would prefer not to receive any compensation for pure losses in utility; they would prefer instead to obtain lower

[32] See Jerry and Richmond (2012: 156-57).

[33] See Jerry and Richmond (2012: 651-52) and Dobbyn and French (2016: 432-34).

[34] On this general theme see Reinker and Rosenberg (2007).

[35] Courts generally permit subrogated insurers to collect damages only for types of losses that were insured; see, for example, 16 *Couch on Insurance* §223: 85 (2017). Thus, because insurance coverage against pain and suffering is essentially non-existent, subrogated insurers would be barred from retaining pain and suffering damages.

premiums that subrogation of pain and suffering awards would generate. Another example is that the law precludes subrogation of awards for wrongful death even though insureds would presumably wish their life insurers to have subrogation rights for the general reasons discussed in Section 2.[36]

The "made whole" rule of subrogation. According to an important default rule of the law on subrogation, an insured who has not been fully compensated for his loss by his insurance coverage should be given a share of litigation proceeds sufficient to make him whole.[37] Only after the insured is made whole is the insurer permitted to retain any litigation proceeds. Our results that under moral hazard and administrative costs it is optimal for insureds to receive a positive portion of damages, when their insurance coverage is less than full, point in the direction of the made-whole rule; but insureds will not necessarily be made whole (and definitely will not be in the case of administrative costs, where $c^* + s^* < h$). However, our conclusion that damages for non-monetary losses should be retained by the insurer implies that the made whole rule would be undesirable to apply in that case.

Selling claims as an alternative to subrogation. It might be thought that many of the virtues of subrogation could be obtained by insureds if they purchased pure insurance policies and sold their claims to other parties. In that way, insureds would obtain value for their claims without bearing litigation risk. However, the sale of claims is not observed in fact. The explanation appears to lie in several factors. First, legal barriers to such sales exist.[38] Second,

---

[36] See Jerry and Richmond (2012: 653-54) on the prohibition against subrogation for life insurers. Life insurance may be viewed as insurance for an insured's beneficiary (say a spouse) against the cessation of the insured's stream of income. As such, the lesson from Section 2 is that subrogation for life insurers would benefit insureds by avoiding overinsurance against loss of the insured's stream of income—resulting in a lower premium—and reducing litigation risk.

[37] On this "made whole" rule, see, for example, Jerry and Richmond (2012: 654-57).

[38] See, for example, Sebok (2011: 74-121) and 6 *American Jurisprudence* 2d Assignments §§53, 55 (2018).

purchasers of claims would be at a cost disadvantage relative to insurers, as mentioned in the first comment above. Third, risk-bearing would not be eliminated through the sale of claims because the revenue obtained by insureds would vary depending on the expected value of their claims.[39]

Subrogation versus insurance contingent on the outcome of litigation. One might believe that many of the benefits of subrogation could be achieved under an insurance policy that provides coverage only after a suit is resolved. In the setting of Section 2, consider a policy that pays an amount equal to litigation costs plus harm if a suit is lost and for just litigation costs if a suit won. Under this policy, the final levels of wealth of the insured would be the same as those under optimal subrogation. However, such an arrangement would not in fact be desirable for insureds because, as we explained in Section 2, insureds generally would want to receive compensation immediately after an accident because of their consumption needs at that time.

Consumer ignorance of subrogation terms. It is plausible that when individuals purchase insurance, they often do not consider the possibility that they might be able to bring a suit for losses that they might suffer, and they are unaware of the concept of insurance subrogation (and thus of its presence or absence in their policy). Given consumer ignorance of this nature, competitive pressures would still be likely to lead to subrogation, for its adoption by insurers would permit them to lower premiums, and individuals would presumably notice and find lower premiums attractive. However, we would not necessarily expect the specific terms of subrogation provisions to be in insureds' interests. Notably, insurers might wish to retain as much of the proceeds from suit as possible, whereas we have shown that insureds should obtain a share of the proceeds in the cases of moral hazard and administrative costs.

[39] In theory, this problem could be cured by having the insured sell his right to bring suits before any accident has occurred. That kind of sale, however, would still suffer from an informational disadvantage relative to subrogation.

Subrogation and social welfare. Although our focus has been on subrogation as a feature of insurance contracts that maximize the expected utility of insureds, subrogation also has broader effects that influence social welfare. Notably, subrogation results in a greater volume of suit against potentially liable parties (compare Proposition 1(c) to Proposition 2(c)), and thus increases the deterrence of undesirable acts as well as the litigation costs that society incurs. An analysis of the effects of subrogation on social well-being would take such factors into account.[40]

[40] Gomez and Penalva (2015) undertake such an analysis in comparing a subrogation regime, a regime barring subrogation, and a regime in which insurance benefits are subtracted from damage payments. See also the remarks in Abraham (1986: 154-55) and Shavell (1987: 235-40).

Appendix

Proof of Proposition 1: (a) If suits are not optimal to bring, then from (2) and $\pi = pc$, the insured's problem is to maximize $(1-p)U(y-pc) + pU(y-pc-h+c)$ over c. Jensen's inequality gives us

$$(1-p)U(y-pc) + pU(y-pc-h+c) \tag{A1}$$
$$\leq U((1-p)(y-pc) + p(y-pc-h+c)) = U(y-ph),$$

where the inequality is strict if $y - pc \neq y - pc - h + c$, or if $c \neq h$. And since if $c = h$, expected utility is $U(y-ph)$, (A1) implies that $c = h$ is optimal.

(b) Using (3) and $\pi = pc$, we obtain

$$EU_S{}'(c) = -p(1-p)U'(y-pc) + p(1-p)[qU'(y-pc+c-k) \tag{A2}$$
$$+ (1-q)U'(y-pc-h+c-k)],$$

implying that

$$EU_S{}'(0) = p(1-p)\{[qU'(y-k) + (1-q)U'(y-h-k)] - U'(y)\} > 0, \tag{A3}$$

so that $c^* > 0$. Hence, c^* must satisfy

$$EU_S{}'(c) = -p(1-p)U'(y-pc) + p(1-p)[qU'(y-pc+c-k) \tag{A4}$$
$$+ (1-q)U'(y-pc-h+c-k)] = 0,$$

which yields (5). Note also that

$$EU_S{}''(c) = p^2(1-p)U''(y-pc) + p(1-p)^2[qU''(y-pc+c-k) \tag{A5}$$
$$+ (1-q)U''(y-pc-h+c-k)] < 0,$$

which is the second-order condition for a maximum of (3), so that (5) identifies a (unique) maximum. When (5) holds, it must be that

$$y - pc^* - h + c^* - k < y - pc^* < y - pc^* + c^* - k. \tag{A6}$$

The first inequality here implies that $c^* < h + k$ and the second that $c^* > k$, so that c^* is in $(k, h + k)$.

(c) We first demonstrate that there is a $\tilde{k}$ satisfying $0 < \tilde{k} < qh$ such that for $k < \tilde{k}$ suit is preferred, at $k = \tilde{k}$ suit and no suit are equally desirable (and by our convention, suit will not occur), and for $k > \tilde{k}$ no suit is preferred. Let $EU_S(k)$ be expected utility when c is chosen optimally given k. At $k = 0$, the insured will be better off if he brings suit than not, for he might win the suit. Hence, $EU_S(0) > U(y - ph)$. At $k = qh$, the insured will be worse off if he brings suit, for

$$EU_S(qh) = (1 - p)U(y - pc) + p[qU(y - pc + c - qh) + (1 - q)U(y - pc - h + c - qh)] \qquad \text{(A7)}$$
$$< (1 - p)U(y - pc) + pU(y - pc + c - h) < U(y - ph).$$

The first inequality follows from the use of Jensen's inequality on the term in brackets, and the second inequality also follows from Jensen's inequality. Since $EU_S(k)$ is continuous and decreasing in k, there must be a unique $\tilde{k}$ satisfying $0 < \tilde{k} < qh$ at which $EU_S(\tilde{k}) = U(y - ph)$ and thus having the claimed properties.

The threshold t in (c) equals $qh - \tilde{k}$. For if $qh - k > qh - \tilde{k}$, or equivalently if $k < \tilde{k}$, suit is optimal to bring; and similarly, if $qh - k \leq qh - \tilde{k}$, or if $k \geq \tilde{k}$, suit is not optimal to bring. Also, because $\tilde{k} = qh - t$ and $0 < \tilde{k} < qh$, we have $0 < qh - t < qh$, or $0 < t < qh$. □

Proof of Proposition 2: (a) The proof of this part is identical to that of Proposition 1(a).

(b) By Jensen's inequality, we have

$$EU_S(c, s) = (1 - p)U(y - \pi) + p[(1 - q)U(y - \pi - h + c) + qU(y - \pi - h + c + s)] \qquad \text{(A8)}$$
$$\leq U((1 - p)(y - \pi) + p[(1 - q)(y - \pi - h + c) + q(y - \pi - h + c + s)])$$
$$= U(y - \pi + p(c - h) + pqs).$$

Using (6) we have that $y - \pi + p(c - h) + pqs = y - ph + p(qh - k)$. Hence,

$$EU_S(c, s) \leq U(y - ph + p(qh - k)), \tag{A9}$$

where the inequality is strict unless $y - \pi = y - \pi - h + c = y - \pi - h + c + s$, which is to say, unless $c = h$ and $s = 0$. Hence, we know that $c^* = h$ and $s^* = 0$.

(c) If the insurer is instructed not to bring a suit, the insured's wealth is $y - ph$, whereas if the insurer is instructed to bring a suit, the insured's wealth is $y - ph + p(qh - k)$. Consequently, it will be optimal to instruct the insurer to bring a suit if and only if $qh - k > 0$, in which case the insured's wealth will rise by $p(qh - k)$.

(d) If $qh - k \leq 0$, then by Propositions 1(a) and 1(c) and 2(a) and 2(c) the insured's expected utility is the same under the optimal pure insurance policy and the optimal insurance policy with a subrogation provision. If $qh - k > 0$, then by Proposition 2(b) and 2(c), a suit will be brought under the optimal insurance policy with a subrogation provision and utility will be $U(y - ph + p(qh - k))$. There are two possible outcomes under a pure insurance policy, depending on whether $qh - k$ exceeds the threshold level t. If $qh - k \leq t$, then by Proposition 1(a) and 1(c), a suit will not be brought under a pure insurance policy and utility will be $U(y - ph) < U(y - ph + p(qh - k))$. If $qh - k > t$, then by Proposition 1(b) and 1(c), a suit will be brought under a pure insurance policy and expected utility will be

$$\begin{aligned} U_S &= (1 - p)U(y - pc^*) + p[qU(y - pc^* + c^* - k) + (1 - q)U(y - pc^* - h + c^* - k)] \\ &< U(y - ph + p(qh - k)), \end{aligned} \tag{A10}$$

where the inequality follows from Jensen's inequality, which must apply strictly because $y - pc^* = y - pc^* + c^* - k = y - pc^* - h + c^* - k$ cannot hold. Hence, again, the insured is worse off than under the optimal insurance policy with a subrogation provision. □

Proof of Proposition 3: (a) We demonstrate that $c^* < h$ when $\phi = 0$ in several steps.

(i) c^* cannot exceed h: From (10), we have

$$EU_N'(x) = -p_1'(x)[U(y-\pi) - U(y-\pi-h+c)] - d'(x), \tag{A11}$$

since $p'(x) = p_1'(x)$. If $c \geq h$, (A11) is negative for $x > 0$; thus, $x = 0$ would be chosen by the insured. Hence, using (10) and (11), we know that for any $c > h$,

$$EU_N = (1-p(0))U(y-p(0)c) + p(0)U(y-p(0)c-h+c) - d(0), \tag{A12}$$

and for $c = h$,

$$EU_N = U(y-p(0)h) - d(0). \tag{A13}$$

Jensen's inequality implies that (A13) exceeds (A12), so $c > h$ cannot be optimal.

(ii) For any c in $[0, h]$ and any π, the insured's choice of x is uniquely determined by the first-order condition $EU_N'(x) = 0$: If $c < h$, (A11) implies that $EU_N'(0) > 0$, so that $x > 0$ must be optimal. Hence, $EU_N'(x) = 0$ must hold, that is

$$-p_1'(x)[U(y-\pi) - U(y-\pi-h+c)] = d'(x). \tag{A14}$$

If $c = h$, we noted in step (i) that $x = 0$ is optimal; and when this is the case, (A14) also holds. Thus, we know that for any c in $[0, h]$ and π, the chosen x must satisfy (A14). We also note that the second-order condition for a local maximum is satisfied.[41] To see that the solution to (A14) is unique, rewrite it as

$$[U(y-\pi) - U(y-\pi-h+c)] = -d'(x)/p_1'(x). \tag{A15}$$

Since $-d'(x)/p_1'(x)$ is strictly increasing in x, the solution to (A14) must be unique.[42]

(iii) For any c in $[0, h]$, the insured's choice of x is uniquely determined by the first-order condition (A14) and by (11): In other words, the claim is that this x is uniquely determined by the condition[43]

---

[41] $EU_N''(x) = -p_1''(x)[U(y-\pi) - U(y-\pi-h+c)] - d''(x) < 0$ for any $c \leq h$.

[42] In other words, we have justified the use of the first-order approach in describing the choice x of the insured given his policy. On this issue, see, for example, Rogerson (1985).

$$-p_1'(x)[U(y - p(x)c) - U(y - p(x)c - h + c)] = d'(x). \tag{A16}$$

We know from step (ii) and (11) that (A16) must hold at any x that is chosen given c. To show that the x solving (A16) is unique, note first that if $c = h$, then $x = 0$ is obviously the unique solution to (A16). Now assume that $c < h$ and suppose to the contrary that (A16) holds for some $x_1 < x_2$. Then

$$U(y - p(x_1)c) - U(y - p(x_1)c - h + c) > U(y - p(x_2)c) - U(y - p(x_2)c - h + c), \tag{A17}$$

since $p(x_1) > p(x_2)$ and U is concave. We can rewrite (A16) as

$$U(y - p(x)c) - U(y - p(x)c - h + c) = -d'(x)/p_1'(x). \tag{A18}$$

Using (A18), (A17) implies that

$$-d'(x_1)/p_1'(x_1) > -d'(x_2)/p_1'(x_2). \tag{A19}$$

Yet (A19) cannot hold because, as we observed above, $-d'(x)/p_1'(x)$ is strictly increasing in x, a contradiction. Let us denote the x that solves (A16) given c by $x(c)$.

(iv) $c^* < h$: We have shown that the problem of maximizing (10) over c subject to (11) and (12) is equivalent to maximizing

$$EU_N(c) = (1 - p(x(c)))U(y - p(x(c))c) + p(x(c))U(y - p(x(c))c - h + c) - d(x(c)) \tag{A20}$$

over $c \in [0, h]$. Using (A16), we have

$$\begin{aligned} EU_N'(c) = {} & -[p'(x(c))x'(c)c + p(x(c))][(1 - p(x(c)))U'(y - p(x(c))c) \\ & + p(x(c))U'(y - p(x(c))c - h + c)] + p(x(c))U'(y - p(x(c))c - h + c). \end{aligned} \tag{A21}$$

Hence,

$$EU_N'(h) = -p'(0)x'(h)hU'(y - p(0)h). \tag{A22}$$

[43] Note that the uniqueness of the solution to (A16) does not follow from (ii). Suppose that for some c, there exist $x_1 < x_2$ that each satisfy (A16). That would not be inconsistent with (ii), since π would be $p(x_1)c$ for x_1 and $p(x_2)c$, a different value, for x_2.

To prove that $c^* < h$, we show that $EU_N'(h) < 0$, which will hold if $x'(h) < 0$. To demonstrate the latter, we implicitly differentiate (A16) with respect to c and solve for $x'(c)$. The sign of $x'(c)$ must equal the sign of[44]

$$\partial\{-p'(x)[U(y - p(x)c) - U(y - p(x)c - h + c)] - d'(x)\}/\partial c, \tag{A23}$$

which is

$$p'(x)[p(x)U'(y - p(x)c) + (1 - p(x))U'(y - p(x)c - h + c)]. \tag{A24}$$

At $c = h$, (A24) is $p'(x)U'(y - p(x)h) < 0$. Thus, $c^* < h$.

(b) We demonstrate that $c^* < h$ and $s^* > 0$ when $\phi = 1$ in several steps.

(i) c^* cannot exceed h: Using (16) we have

$$EU_S'(x) = -p_1'(x)[U(y - \pi) - U(y - \pi - h + c)] - d'(x). \tag{A25}$$

If $c \geq h$, the term in brackets is less than or equal to zero, so that $EU_S'(x) < 0$ for $x > 0$. Hence, the insured would choose $x = 0$ for $c \geq h$. Thus, for any $c > h$,

$$\begin{aligned} EU_S(c, s) = {} & (1 - p(0))U(y - \pi(c, s)) + p_1(0)U(y - \pi(c, s) - h + c) \\ & + p_2U(y - \pi(c, s) - h + c + s); \end{aligned} \tag{A26}$$

and for $c = h$ and $s = 0$,

$$\begin{aligned} EU_S(h, 0) &= (1 - p(0))U(y - \pi(h, 0)) + p_1(0)U(y - \pi(h, 0)) + p_2U(y - \pi(h, 0)) \\ &= U(y - p(0)h). \end{aligned} \tag{A27}$$

But it can readily be verified that $U(y - p(0)h)) > EU_S(c, s)$ by Jensen's inequality. This means that a policy with $c^* > h$ could not be optimal.

(ii) For any c in $[0, h]$ and any s and π, the insured's choice of x is uniquely determined by the first-order condition (18), $EU_S'(x) = 0$: If $c < h$, (A25) implies that $EU_S'(0) > 0$, so that $x > 0$ is optimal and $EU_S'(x) = 0$ must hold. If $c = h$, we observed in step (i) that $x = 0$, and it is clear

[44] Equation (A16) is of the form $F(x, c) = 0$, where F is the first-order condition for the optimal choice of x by the insured. Implicitly differentiating, we obtain $F_x x'(c) + F_c = 0$, so that $x'(c) = -F_c/F_x$. We know that $F_x < 0$, for this is the second-order condition for x to have been chosen optimally. Hence, the sign of $x'(c)$ equals the sign of F_c.

that $EU_S'(0) = 0$ holds. We also note that if x satisfies (18), the second-order condition for a local maximum is satisfied. That the solution is unique follows from the argument made in step (ii) of the proof of part (a).

(iii) For any c in $[0, h]$, the insured's choice of x is uniquely determined by the first-order condition (18) and by (14): That is, x is uniquely determined by

$$-p_1'(x)[U(y - (p(x)(c + k) - p_2(h - s))) - U(y - (p(x)(c + k) - p_2(h - s)) - h + c)] \tag{A28}$$
$$= d'(x).$$

This claim follows from an argument essentially identical to that made in step (iii) of the proof of part (a). We continue to let $x(c)$ denote the insurer's choice of x given c (suppressing s in the notation).

(iv) $c^* < h$: Having shown in step (i) that $c^* \leq h$, we can prove that $c^* < h$ if we demonstrate that $c^* = h$ is not possible. To this end, we first observe that if $c^* = h$, it must be that $s^* = 0$. In particular, if $c = h$, we know from step (i) that $x = 0$. Hence, if $c = h$ and $s > 0$, we have

$$EU_S(h, s) = (1 - p(0))U(y - \pi(h, s)) + p_1(0)U(y - \pi(h, s)) \tag{A29}$$
$$+ p_2U(y - \pi(h, s) + s),$$

whereas from (A27) we have $EU_S(h, 0) = U(y - p(0)h))$. Since $U(y - p(0)h)) > EU_S(h, s)$ by Jensen's inequality, $c = h$ and $s > 0$ cannot be optimal.

Finally, we show that the conclusion that $c^* = h$ and $s^* = 0$ leads to a contradiction—that the insured would be better off if $c < h$ when $s = 0$. We have

$$EU_S(c, 0) = (1 - p(x(c)))U(y - \pi(c, 0)) + p(x(c))U(y - \pi(c, 0) - h + c) - d(x(c)), \tag{A30}$$

so that, using (18),

$$dEU_S(c, 0)/dc = –(1 – p(x(c))(p'(x(c))x'(c)(c + k) + p(x(c)))U'(y – \pi(c, 0)) \quad \text{(A31)}$$

$$+ (1 – p(x(c)) – p'(x(c))x'(c)(c + k))p(x(c))U'(y – \pi(c, 0) – h + c).$$

Hence,

$$dEU_S(h, 0)/dc = –p'(x(h))x'(h)(h + k)U'(y – \pi(h, 0)). \quad \text{(A32)}$$

Because $x'(h) < 0$,[45] $dEU_S(h, 0)/dc < 0$, meaning that c^* must be less than h.

(v) $s^* > 0$: We now hold c constant at $c^* < h$ and show that if $s = 0$, expected utility can be raised by increasing s, implying that $s^* > 0$. To this end, let us treat x as a function of s, determined implicitly by (18) and (17), and π as a function of s determined by (17). Hence, we may write EU_S as $EU_S(x(s), \pi(s), s)$. Thus, our object is to demonstrate that $dEU_S(x(0), \pi(0), 0)/ds > 0$. Using (16) and (18), we have

$$dEU_S(x(s), \pi(s), s)/ds = –\pi'(s)\{(1 – p(x(s)))U'(y – \pi(s)) + \quad \text{(A33)}$$

$$p_1(x(s))U'(y – \pi(s) – h + c^*) + p_2U'(y – \pi(s) – h + c^* + s)\}$$

$$+ p_2U'(y – \pi(s) – h + c^* + s).$$

Observe that $\pi'(s) = p'(x(s))x'(s)(c^* + k) + p_2$. Hence, at $s = 0$, (A33) becomes

$$–p'(x(0))x'(0)(c^* + k)[(1 – p(x(0)))U'(y – \pi(0)) + p(x(0))U'(y – \pi(0) – h + c^*)] \quad \text{(A34)}$$

$$+ p_2(1 – p(x(0)))[U'(y – \pi(0) – h + c^*) – U'(y – \pi(0))].$$

Since $c^* < h$, the second term in (A34) is positive. Thus, a sufficient condition for (A34) to be positive is that the first term is positive, which will be true if $x'(0) > 0$. To see that this holds, substitute (17) into (18) and implicitly differentiate the resulting expression with respect to s and solve for $x'(s)$. The sign of $x'(s)$ must equal the sign of[46]

[45] The sign of $x'(c)$ equals the sign of the partial derivative with respect to c of $–p_1'(x)[U(y – \pi) – U(y – \pi – h + c)] – d'(x)$, using the logic explained in note 44. Since $\pi = p(x)(c + k) – p_2h$, the partial derivative with respect to c is $–p_1'(x)p(x)[U'(y – \pi – h + c) – U'(y – \pi)] + p_1'(x)U'(y – \pi – h + c)$. At $c = h$, this becomes $p_1'(x)U'(y – \pi) < 0$.

[46] See note 44 above.

$$\partial\{-p'(x)[U(y - p(x)(c^* + k) + p_2(h - s)) \tag{A35}$$

$$- U(y - p(x)(c^* + k) + p_2(h - s) - h + c^*)] - d'(x)\}/\partial s,$$

which is

$$p'(x)p_2[U'(y - p(x)(c + k) + p_2(h - s)) - U'(y - p(x)(c + k) + p_2(h - s) - h + c)] > 0, \tag{A36}$$

where the inequality follows from observing that the bracketed term is negative because $c^* < h$. Thus, $x'(0) > 0$ and $s^* > 0$.

(c) Again, we prove the claim in steps.

(i) If $q(x_N)h - k > 0$, suits are optimal to bring: Suppose to the contrary that suits are not optimal to bring and let $c_N{}^*$ be the optimal c in this case. We will show that there is an insurance policy under which suits are brought and insureds are better off, which will be a contradiction. In particular, hold c fixed at $c_N{}^*$ and let

$$s = (q(x_N)h - k)/q(x_N). \tag{A37}$$

Then $s > 0$ by our hypothesis and, using (14),

$$\pi(c_N{}^*, s) = p(x_N)c_N{}^* - p(x_N)(q(x_N)(h - s) - k) = p(x_N)c_N{}^*. \tag{A38}$$

Because c and π have not changed, it is clear from (18) and (A14) that x_N will continue to be chosen by the insured. The insured must be better off under this policy because he has the same coverage, pays the same premium, chooses the same level of care, but obtains $s > 0$ whenever he wins a suit, which occurs with probability $p_2 > 0$. Hence, it must be optimal for suit to be brought.

(ii) If $q(x_N)h - k = 0$, suits are optimal to bring: Suppose to the contrary that suits are not optimal to bring. We will show that there is an insurance policy under which suits are brought and insureds are better off. Again, hold c fixed at $c_N{}^*$ and let $s = 0$. Then $\pi(c_N{}^*, s) = \pi(c_N{}^*, 0) = p(x_N)c_N{}^*$. Because c and π have not changed, x_N will still be chosen. Hence, the insured is just as well off as he was when suits were not brought. However, by part (b) of the proposition, $s^* > 0$. Thus, the policy with c_N^* and $s = 0$ cannot be optimal, meaning that there exists a policy involving suit that is superior, and thus better than the optimal policy involving no suit.

(iii) Some suits for which $q(x_N)h - k < 0$ are optimal to bring: Let k_o be such that $q(x_N)h - k_o = 0$. We claim that for all k in $(k_o, k_o + \varepsilon)$, for a sufficiently small positive ε, suits are optimal to bring. This will prove the claim since for such k, $q(x_N)h - k < 0$. Let $EU_S(c^*, s^*, k)$ denote expected utility when suit is brought and c and s are chosen optimally given k. Also, observe that $EU_N(c_N^*)$ does not depend on k. Now in (ii) we showed that $EU_S(c^*, s^*, k_o) > EU_N(c_N^*)$. And since $EU_S(c^*, s^*, k)$ is continuous in k, we must have $EU_S(c^*, s^*, k) > EU_N(c_N^*)$ for all k within ε of k_o for a sufficiently small positive ε. Hence, suits are optimal to bring for k in $(k_o, k_o + \varepsilon)$. □

Proof of Proposition 4: (a) If $c^* = 0$, then $c^* < h$. If $c^* > 0$, it is determined by the first-order condition from (2), making use of (19),

$$p(1 + \lambda)[(1 - p)U'(y - \pi) + pU'(y - \pi - h + c)] = pU'(y - \pi - h + c). \qquad \text{(A39)}$$

Since $\lambda > 0$, (A39) implies that $(1 - p)U'(y - \pi) + pU'(y - \pi - h + c) < U'(y - \pi - h + c)$, or that $U'(y - \pi) < U'(y - \pi - h + c)$, meaning that $c^* < h$.

(b) We first show that $s^* > 0$. Suppose otherwise, that $s^* = 0$. Then the proof of (a) implies that $c^* < h$; for even though π is determined by (20) rather than (19), (A39) still holds.

We next demonstrate that $c^* < h$ and $s^* = 0$ leads to a contradiction—which will imply that $s^* > 0$. Observe that

$$EU_S'(s) = -pq(1 - p)U'(y - \pi) + pq[(1 - pq)U'(y - \pi - h + c + s) - (1 - q)pU'(y - \pi - h + c)], \qquad \text{(A40)}$$

so that

$$EU_S'(0) = pq(1 - p)[U'(y - \pi - h + c) - U'(y - \pi)] > 0, \qquad \text{(A41)}$$

where the inequality follows from $c^* < h$. Since $EU_S'(0) > 0$, s^* cannot be 0, a contradiction.

Last, we show that $c^* + s^* < h$, which will also imply that $c^* < h$. Now since $s^* > 0$, $EU_S'(s) = 0$ holds, which from (A40) is

$$(1 - p)U'(y - \pi) + p[qU'(y - \pi - h + c + s) + (1 - q)U'(y - \pi - h + c)] \quad \text{(A42)}$$

$$= U'(y - \pi - h + c + s).$$

If $c^* + s^* < h$ is not true, then $c + s \geq h$, implying that $y - \pi - h + c + s \geq y - \pi$. Also, since $s > 0$, $y - \pi - h + c + s > y - \pi - h + c$. Thus, the wealth argument on the right-hand side of (A42) is greater than or equal to each of the three wealth arguments on the left-hand side, and strictly greater than $y - \pi - h + c$. Hence, $U'(y - \pi - h + c + s)$ is less than or equal to each of the marginal utilities on the left-hand side, and strictly less than $U'(y - \pi - h + c)$. Furthermore, the probability weights on the three marginal utilities on the left-hand side add to 1. Accordingly, the left-hand side of (A42) is less than the right-hand side, a contradiction. It follows that $c^* + s^* < h$.

(c) We first show that suits that have positive expected value will be brought. Suppose otherwise, that a suit with $qh - k > 0$ is not brought and coverage c^* is purchased. Expected utility will be

$$EU_N(c^*) = (1 - p)U(y - \pi) + pU(y - \pi - h + c^*), \quad \text{(A43)}$$

where $\pi = p(1 + \lambda)c^*$. But this expected utility can be improved upon with an insurance policy in which suit is brought, coverage is maintained at c^*, and $s = 0$. Then expected utility will be

$$EU_S(c^*, 0) = (1 - p)U(y - \hat{\pi}) + pU(y - \hat{\pi} - h + c^*), \quad \text{(A44)}$$

where $\hat{\pi} = p(1 + \lambda)c^* - p(qh - k)$. Since $qh - k > 0$, $\hat{\pi} < \pi$ and hence (A44) exceeds (A43), contradicting the optimality of not bringing suit.

Next consider a suit that has zero expected value. If such a suit is not brought, expected utility will be (A43), and again consider a policy in which suit is brought, coverage remains at c^*, and $s = 0$. Then (A44) will apply, with $\hat{\pi} = p(1 + \lambda)c^* - p(qh - k) = p(1 + \lambda)c^* = \pi$. Hence, $EU_S(c^*, 0) = EU_N(c^*)$. But by (b), we know that if a suit is brought, $s^* > 0$. Accordingly, $EU_S(c^{**}, s^*) > EU_S(c^*, 0)$, where c^{**} is the optimal c when suits are brought. Thus, $EU_S(c^{**}, s^*) > EU_N(c^*)$, contradicting the optimality of not bringing suit.

That it is optimal to bring some negative expected value suits follows from the preceding paragraph. Because it is strictly optimal to bring suit when $qh - k = 0$, by continuity, it must also be strictly optimal to bring suit if litigation costs are $k + \varepsilon$ for small positive ε. □

Proof Proposition 5: (a) We want to show that

$$EU_N(h) = U(y - ph) - pu > (1 - p)U(y - pc) + pU(y - pc - h + c) - pu = EU_N(c) \tag{A45}$$

for $c \neq h$. This is equivalent to $U(y - ph) > (1 - p)U(y - pc) + pU(y - pc - h + c)$ for $c \neq h$, which follows from Jensen's inequality.

(b) We want to show that

$$\begin{aligned} EU_S(h, 0) &= U(y - ph + p(q(h + a) - k)) - pu \\ &> (1 - p)U(y - \pi) + p[(1 - q)U(y - \pi - h + c) \\ &\quad + qU(y - \pi - h + c + s) - u] = EU_S(c, s) \end{aligned} \tag{A46}$$

for any $c \neq h$ or any $s > 0$. Equivalently, we want to show that

$$\begin{aligned} &U(y - ph + p(q(h + a) - k)) > \\ &(1 - p)U(y - \pi) + p[(1 - q)U(y - \pi - h + c) + qU(y - \pi - h + c + s)] \end{aligned} \tag{A47}$$

for any $c \neq h$ or any $s > 0$. Now expected wealth on the right-hand side of this inequality is

$$(1-p)(y-\pi) + p[(1-q)(y-\pi-h+c) + q(y-\pi-h+c+s)] = \tag{A48}$$

$$y - \pi - p(h-c) + pqs = y - [pc - p(q(h+a-s)-k)] - p(h-c) + pqs$$

$$= y - ph + p(q(h+a) - k,$$

which is the wealth on the left-hand side. Hence Jensen's inequality demonstrates that (A47) holds, and strictly because the levels of wealth $y-\pi$, $y-\pi-h+c$, and $y-\pi-h+c+s$ are not all equal if $c \neq h$ or if $s > 0$.

(c) We know from the preceding steps that $EU_N(c^*) = U(y-ph) - pu$ and that $EU_S(c^*, s^*) = U(y-ph+p(q(h+a)-k)) - pu$. Thus, suit will be desirable if and only if

$$U(y-ph+p(q(h+a)-k)) > U(y-ph), \tag{A49}$$

which is to say if and only if $p(q(h+a)-k) > 0$ or, equivalently, $q(h+a)-k > 0$. □

References

Abraham, Kenneth S. 1986. *Distributing Risk***:** *Insurance, Legal Theory, and Public Policy*. New Haven: Yale University Press.

Abraham, Kenneth S., and Daniel Schwarcz. 2015. *Insurance Law and Regulation*. *Cases and Materials*. 6th ed. St. Paul: Foundation Press.

Arrow, Kenneth J. 1974. "Optimal Insurance and Generalized Deductibles," 1974 *Scandinavian Actuarial Journal* 1-42.

Baker, Tom, and Kyle D. Logue. 2017. *Insurance Law and Policy: Cases and Materials*. 4th ed. Frederick, MD: Wolters Kluwer.

Bleed, Lisa N. 2001. "Enforcing Subrogation Provisions as 'Appropriate Equitable Relief' Under ERISA Section 502(A)(3)," 35 *University of San Francisco Law Review* 727-50.

Cook, Philip J., and Daniel A. Graham. 1977. "The Demand for Insurance and Protection: The Case of Irreplaceable Commodities," 91 *Quarterly Journal of Economics* 143-56.

Dionne, Georges, ed. 2013. *Handbook of Insurance.* 2nd ed. New York: Springer.

Dobbyn, John F., and Christopher C. French. 2016. *Insurance Law in a Nutshell*. 5th ed. St. Paul: West Academic.

Fudenberg, Drew, and Jean Tirole. 1990. "Moral Hazard and Renegotiation in Agency Contracts," 58 *Econometrica* 1279–1319.

Gomez, Fernando, and Jose Penalva. 2015. "Tort Reform and the Theory of Coordinating Tort and Insurance," 43 *International Review of Law and Economics* 83-97.

Harrington, Scott E., and Gregory R. Niehaus. 2004. *Risk Management and Insurance*. 2nd ed. New York: McGraw-Hill.

Insurance Day. 2009. "International Comparative Review of Subrogation Law," September.

Jerry II, Robert H., and Douglas R. Richmond. 2012. *Understanding Insurance Law*. 5th ed. New Providence, NJ: Lexis-Nexis.

Kimball, Spencer L., and Don A. Davis. 1962. "The Extension of Insurance Subrogation," 60 *Michigan Law Review* 841-72.

Marasinghe, M. L. 1976a. "An Historical Introduction to the Doctrine of Subrogation: The Early History of the Doctrine I," 10 *Valparaiso University Law Review* 45-65.

————. 1976b. "An Historical Introduction to the Doctrine of Subrogation: The Early History of the Doctrine II," 10 *Valparaiso University Law Review* 276-99.

Perry, Bill, David Rosenberg, and Ianika Tzankova. 2015. "Money Machine: Subrogation Strategy is Evolving in the US." *Insurance Day*, April 23.

Reinker, Kenneth S., and David Rosenberg. 2007. "Unlimited Subrogation: Improving Medical Malpractice Liability," 36 *Journal of Legal Studies* S261-89 (June).

Rogerson, William P. 1985. "The First-Order Approach to Principal-Agent Problems," 53 *Econometrica* 1357-67.

Schlesinger, Harris. 2013. "The Theory of Insurance Demand," in G. Dionne, ed., *Handbook of Insurance*. 2nd ed. New York: Springer.

Sebok, Anthony J. 2011. "The Inauthentic Claim," 64 *Vanderbilt Law Review* 61-139.

Shavell, Steven. 1987. *Economic Analysis of Accident Law*. Cambridge: Harvard University Press.

Sykes, Alan. 2001. "Subrogation and Insolvency," 30 *Journal of Legal Studies* 383-99.

Viscusi, W. Kip. 1989. "The Interaction between Product Liability and Workers' Compensation as Ex Post Remedies for Workplace Injuries," 5 *Journal of Law, Economics, & Organization* 185-210.

Winter, Ralph A. 2013. "Optimal Insurance Contracts Under Moral Hazard," in G. Dionne, ed., *Handbook of Insurance*. 2nd ed. New York: Springer.