© The Journal of Risk and Insurance, 1996, Vol. 63, No. 2, 185–206.

Insurable Interest, Options to Convert, and Demand for Upper Limits in Optimum Property Insurance

Rod Garratt
John M. Marshall

Abstract

Upper limits in property insurance contracts can result directly from the consumer's demand for them. They are demanded because the consumer has options to convert or move out of damaged property rather than merely to restore it to its previous condition and occupy it. In the absence of transactions costs, finding the optimum upper limit is equivalent to finding the insurable interest. When real estate transactions are costly, binding upper limits are demanded, but they may be higher than the upper limits that an insurer should impose to ensure incentive compatibility and mitigate moral hazard. Moreover, the concept of insurable interest divides to become two distinct concepts: the upper limit of legitimate demand, and the upper limit that the prudent insurer would use.

Introduction

Soon after the 1991 fire in Oakland, California, the press reported that many of the houses that were destroyed were not fully covered by insurance. Later it was observed that the houses were being rebuilt as larger, grander, restyled structures. These facts seem paradoxical and mutually contradictory, and they raise two interesting questions in the theory of insurance. The first question is whether binding upper limits on coverage might be rationally demanded by consumers. If so, the underinsurance of the Oakland houses, which the press had characterized as inadvertent, might instead be deliberately chosen by the owners and even approximately optimizing for them.

The rationale for choosing an upper limit on coverage is found in the second question raised by the Oakland case: whether a damaged house is typically replaced by a replica of itself or by something different. Intuitively, the answer is something different, more modern, and more appropriate to current demand. Because the highest and best use of the site has changed in the years since the house

---

Rod Garratt is Assistant Professor and John M. Marshall is Professor, both in the Department of Economics, University of California, Santa Barbara.

The authors thank the referees, the Editor, Cheng-Zhong Qin, Michael Smith, and numerous participants in the 1994 Risk Theory Seminar for generous comments. They also thank the Academic Senate of the University of California, Santa Barbara, for research support.

was originally built, the owner chooses a value-maximizing replacement, not a replica. Moreover, in deciding to purchase insurance, the owner can anticipate the replacement decision and demand coverage that accords with it. The anticipated replacement defines the extent of economic loss and leads to the choice of insurance with a binding upper limit.

Options to convert and demands for upper limits are inseparably related not only in the specific case of Oakland but in more general cases. The relationship always depends upon the structure of transactions costs. For example, the property owner who is converting for resale into a liquid market—the first case described below—is different from the one who converts for her own private consumption—the second case below. More subtly, when the property owner who resells into a liquid market incurs fixed transactions costs—the third case below—he is sometimes motivated by these costs to modify his demand for insurance.

An unusual factor that enters the analysis of options to convert or to move is the indivisibility of the decision. Formal treatment of indivisibility in the analysis of insurance is new, but indivisibility itself is completely intuitive in this context. Some indivisibilities arise directly from the preferences of the consumer—the choice is between the currently occupied dwelling and the house most preferred of all the alternatives. In other situations, indivisibility is induced by fixed transactions costs. The homeowner faces search costs, realtors' fees, and moving expenses that are fixed more or less independently of the quality of the chosen home. These costs encourage him to remain in the current home, even it if becomes unsuitable, rather than move to another that is only slightly different. More broadly, the fixed transactions costs of moving explain why people do not typically move from one house to a slightly different alternative. The typical move is to a house that is substantially different from its predecessor, an indivisible decision.

MARKET VALUES: THE OAKLAND STORY

In the stylized facts about the Oakland situation, transactions costs do not play a central role. In order to focus on the effects of options to convert, the loading of insurance prices above fair value is assumed to be zero. To examine the situation, suppose that both houses and land of all qualities can be bought and sold freely at market prices that are available without search costs, and that owners can borrow and lend freely in financial markets. Property insurance is written in terms of damage, denoted by t, which is the cost to restore the damaged property, without converting it, to the value it had before the damaging event. Damage is not itself the economic loss, for economic loss is dependent upon the decision made about repairing, restoring, or converting the damaged property.

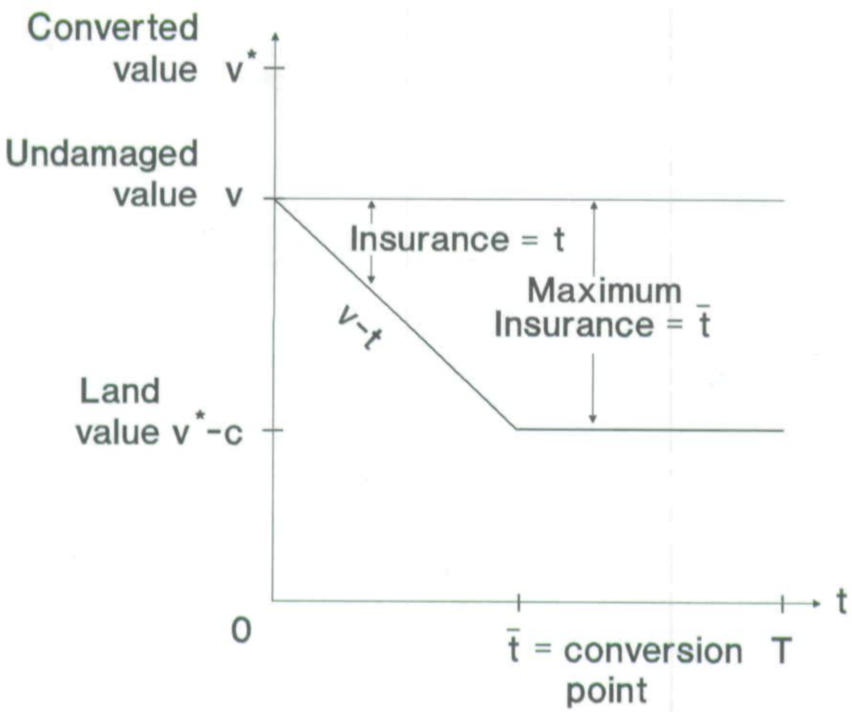
The conversion decision is illustrated in Figure 1. The value of the existing property is v. If the property is rebuilt, there is a unique best use for it, and the cost to clear the land and build that optimum improvement is c. The cost c is a constant that ignores any differences in salvage value that might attach to different levels of damage. The value of the converted property is v^*, and the value of the land (not cleared of the incumbent building) is $v^* - c$. Clearly, $v^* - c < v$, for otherwise the property would already be converted. In the event of damage, conversion takes place if the value of the land that is implied by the conversion option is above the value of the existing property, net of the cost of repairing the damage. That is, conversion is chosen if

$$v^* - c > v - t. \tag{1}$$

Rearranging this condition, the option to convert is exercised if the damage exceeds a threshold level $\bar{t} = v - (v^* - c)$.

The implications for insurance are immediate. Since conversion occurs whenever damage exceeds $\bar{t}$, loss of wealth by the owner is limited to that value. Thus, full insurance of wealth means purchasing insurance for damage up to $\bar{t}$, but not beyond. At fair prices, the property owner optimally chooses a binding upper limit equal to $\bar{t}$. Further insurance is not demanded because it is not worth the cost. Thus, the typical property owner demands an upper limit on coverage that is less than the greatest possible damage, and, when damage exceeds the upper limit, the old structure is replaced with a different, more valuable one.

Figure 1
The Option to Convert



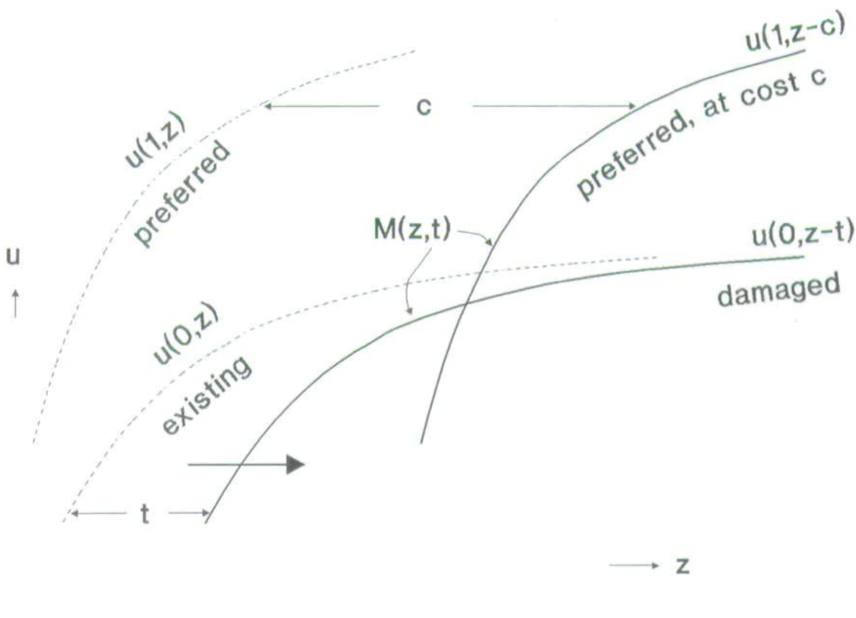
Real Estate Transactions Costs: The Ranch House Case

In the model of the previous section, the decision to convert is independent of the decision about where to live. Such separation prevails in the absence of transactions costs. In the cases analyzed below, transactions costs in real estate are present, and consequently the decision to convert interacts with the consumption decision.

To illustrate the demand for binding upper limits in the presence of transactions costs, consider the case of a house owned by a rancher and built on her spacious ranch. The building is not saleable separately, nor is there a market for the small parcel of land on which it stands. The building and land cannot be sold as a unit because no one but the rancher is interested in living there. In case of fire, there is no market in which the rancher can simply purchase another house. She must rebuild the house as it was or build an alternative one.

The situation of the rancher in the absence of insurance is illustrated in Figure 2. Utility is measured vertically, and consumption other than housing is measured horizontally. The nonhousing good on the horizontal axis is the numeraire and denoted z. The notation z also stands for wealth, measured in the numeraire. The rancher experiences utility u(0, z) from the existing house at consumption z. From the preferred house and the same consumption level, utility is u(1, z). However, there is a price c to be paid for the preferred house. Thus, starting with nonhousing wealth z, the rancher can consume only z – c in the preferred house, and the utility realized is u(1, z–c). The position of the latter curve is found by taking the curve u(1, z) and shifting it to the right by the amount c, as shown in Figure 2.

Figure 2
Ranch House Preferences



The single crossing illustrated in Figure 2 is a result of the natural supposition that housing is a normal good. Because housing is indivisible, the term normal has the special meaning developed for such goods by Cook and Graham (1977). Specifically, housing is a normal good because the reservation price that the rancher is willing to pay to move to the preferred house—the horizontal distance between the curves in Figure 2—is an increasing function of wealth. The implication is that, to the left of the crossing, the rancher experiences higher utility by remaining in the existing house, while, to the right of it, she prefers to move.

In this context, damage t is the cost of restoring the existing house to its previous *condition*. The idea of restoring the house to its previous *value* has no meaning because there is no market for the house. Damage shifts the utility function for the existing house to the right, just as the cost of building did for the preferred house. Utility as a function of wealth z, given damage t, is $u(0, z-t)$, as illustrated in Figure 2. As t increases, the existing-house curve $u(0, z-t)$ slides progressively to the right, while the preferred-house curve $u(1, z-c)$ stands still.

It is natural to view t as a state of the world and to define consumer preferences through a state-dependent utility function that depends on t. Let the variable $e(t) = 0$ or 1 indicate the house occupied by the rancher. Then the state-dependent utility function is

$$u(e, z-(1-e)t-ec), \tag{2}$$

and the problem of optimum insurance is to supply the proper amount of contingent wealth z in each damage state t.

The insurance contract of the rancher pays $I(t)$ in damage state t. The rancher's initial nonhousing wealth is w, the premium for insurance is P, and consumption while residing in the existing house is $w + I(t) - P$. In case of conversion, the preferred house is built, and consumption is $w + I(t) - P - c$. In these terms, utility becomes

$$u(e, w+I(t)-P-(1-e)t-ec). \tag{3}$$

Given that the probability density function of damage is $f(t)$, the problem of finding optimum insurance at fair prices is to

$$\underset{P,\, I(t),\, e(t)}{\text{Maximize}} \int_0^T u(e, w+I(t)-P-(1-e)t-ec)f(t)dt \tag{4}$$

$$\text{subject to } P = \int_0^T I(t)f(t)dt \tag{5}$$

$$I(0) = 0. \tag{6}$$

Fair prices are used here, as in the Oakland case, in order to isolate how the conversion option affects the optimum. Others have studied optimum insurance with loaded prices but without conversion options (for instance, Arrow, 1971; Raviv, 1979; and Gollier, 1987, 1992), and their contributions might be the starting point for some future analysis of conversion options with loaded prices.

The solution to the problem of equations (4) through (6) only partially resembles an insurance contract and is given by the following proposition.

Proposition 1

The optimum indemnity in the problem of equations (4) through (6) has the following form: For some values t_0 and t_1 satisfying $0 \leq t_0 \leq T$ and $t_0 < t_1$, the solution is

$$I(t) = t \text{ on } 0 \leq t \leq t_0 \tag{7}$$

$$I(t) = t_1 \text{ on } t_0 < t \leq T. \tag{8}$$

All proofs are given in the Appendix. The intuition behind the solution is interesting. The rancher has nonconcave, state-dependent utility functions represented, in each state t, by M(z, t), which is the upper envelope of the two utility functions in Figure 2. The problem of finding optimum insurance is then to

$$\underset{z(t)}{\text{Maximize}} \int_0^T M(z,t)f(t)dt \tag{9}$$

$$\text{subject to } w = \int_0^T z f(t)dt. \tag{10}$$

To see the solution, consider for each t the concave hull of M(z, t) which is denoted $M^*(z, t)$. Substituting the concave hulls in equation (9) creates a surrogate problem that is quite easy to solve. In particular, contingent wealths z(t) are adjusted until there is some level of marginal utility λ that is common to all states. Then optimum insurance satisfies, for all t

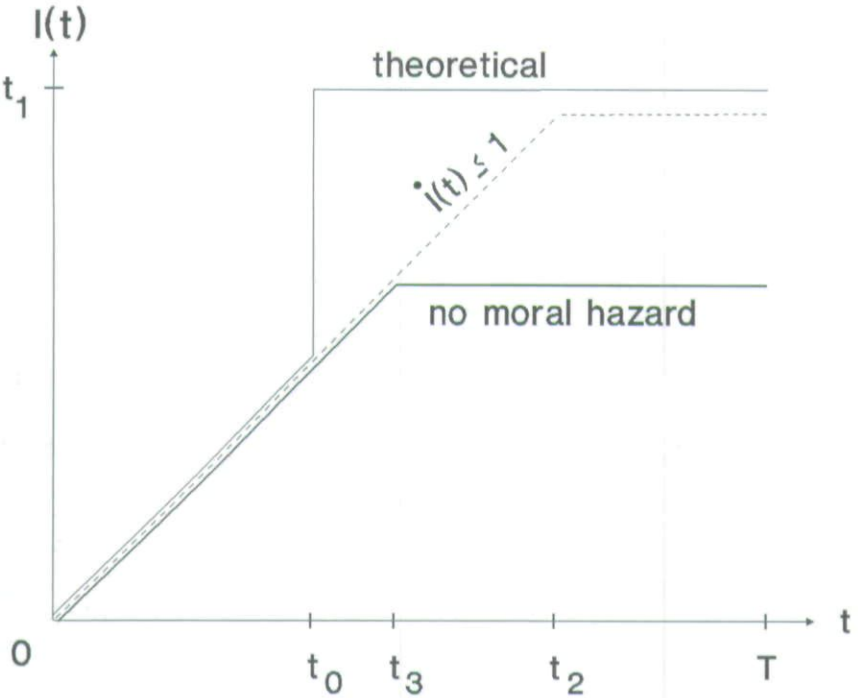
$$M_1^*(z(t),t) = \lambda. \tag{11}$$

By construction as a concave hull, each $M^*(z, t)$ has a linear portion over the nonconcavity of M(z, t). As t increases in Figure 2, the lower, existing-house curve moves from left to right, and therefore the linear portion of $M^*(z, t)$ becomes progressively steeper. For low values of t, the linear portion of $M^*(z, t)$ is rather flat, and the optimum marginal utility falls below the linear portion on the existing-house curve. As damage increases, the level t_0 is reached at which the slope of the linear portion just equals the critical marginal utility λ. For greater levels of damage, the optimum marginal utility always falls above the linear portion of $M^*(z, t)$ on the preferred-house curve.

The solution of the surrogate problem also solves the underlying, nonconcave problem. Why? The optimum of the surrogate is an upper bound on the optimum for the nonconcave problem. It is also feasible for the nonconcave problem, except possibly in the single state t_0, which can be ignored because t_0 occurs with probability zero. Hence, the solution of the surrogate is the solution of the original problem.

The solution described in Proposition 1 with t_0 strictly between 0 and T appears as the light line in Figure 3. The rancher stays in the existing house and demands full insurance for damage up to t_0. For damage barely beyond that point, she abandons the existing house and builds the preferred house while the indemnity jumps discontinuously upward. Because higher levels of damage to the abandoned house are now irrelevant, the optimum indemnity stays at the same value, which is the binding upper limit in the optimum contract. Full insurance of all damage is the special case of $t_0 = T$. It occurs when the rancher is so poor that she does not build the preferred house regardless of damage. A rancher with very high wealth builds the preferred house immediately, and in this case $t_0 = 0$.

Figure 3
Upper Limits and Moral Hazard



The optimum contract with t_0 strictly between 0 and T involves a gamble in wealth and therefore is not an enforceable insurance contract.[1] To ensure a legal contract, the problem must be solved again with an additional restriction that coverage cannot increase at a rate faster than one,

$$\dot{I}(t) \leq 1. \tag{12}$$

The constraint (12) is stronger than the requirement that indemnity cannot exceed damage, that is, $I(t) \leq t$. The latter restriction, which is familiar from Raviv (1979)

[1] Risk-loving in optimum insurance is not entirely new. For instance, Sinn (1982) shows that individuals who seek to purchase wealth or liability insurance may behave as risk lovers if society provides a guaranteed minimum wealth level.

and elsewhere, does not prevent upward jumps in the optimum coverage and thus does not have the needed effect.[2] With constraint (12), the optimum insurance contract is continuous, there is no gambling on risk, and there is a binding upper limit on coverage, as shown in the following proposition.

Proposition 2

The optimum indemnity in the problem of equations (4) through (6) with the added constraint (12) has a solution of the following form: There is a value t_2 satisfying $0 \leq t_2 \leq T$ such that the solution is

$$I(t) = t \text{ on } 0 \leq t \leq t_2 \tag{13}$$

$$I(t) = t_2 \text{ on } t_2 \leq t \leq T. \tag{14}$$

The contract derived in Proposition 2 is illustrated by the dashed line in Figure 3. The upper limit on coverage is not imposed by the insurer. It is freely demanded by the client under the restriction that indemnity cannot increase faster than damage.

Insurable Interest

Propositions 1 and 2 have some surprising implications for the concept of insurable interest. The concept is one among the many strands of insurance law that are collected under the title of the "principle of indemnity" (see Keeton, 1960, chap. 2). According to the principle, an insured can collect no more than the value to him or her of the destroyed property, and the needed value is the insurable interest. Thus, the sole owner of an ocean cargo valued at $1 million has an insurable interest of $1 million, a half-owner of a similar cargo has an insurable interest of $500,000, and so on.

The example is familiar, and it illustrates a duality in the idea of insurable interest. Within the concept of insurable interest are two distinct, economically relevant ideas: the legitimate demand for insurance by an honest client, and the prudent limitation set by an insurer who remains skeptical that the client might willfully destroy the property. Thus, the honest half-owner of the cargo legitimately demands $500,000 in coverage and is unwilling, at fair or worse-than-fair prices, to insure for a higher value. The prudent insurer doubts the client's intentions, is unwilling to create any ill-directed incentives, and insures only up to $500,000, its prudent limitation. In this and all of the standard examples of insurable interest, there is a happy coincidence between legitimate interest and prudent limitation.

The dual concepts underlying insurable interest are not always in close agreement. Propositions 1 and 2 show that, when property can be converted, the happy coincidence between legitimate demand and prudent limitation vanishes. A client with no intent to defraud can legitimately demand insurance like that in

---

[2] The gamble provided by the theoretical solution of Proposition 1 can be approximated by a contract that satisfies $I(t) \leq t$. The contract has a deductible for small losses, a segment in which insurance increases one-for-one with damage, and then an upward jump in coverage at the conversion point.

Proposition 2, but the insurance company cannot prudently supply such a contract to a suspected scoundrel. In the contract, which is illustrated by the dashed line in Figure 3, the rancher abandons the existing house and moves into the preferred house at a particular, critical level of damage, but her demand for insurance extends beyond that point and into a range where utility rises with increasing damage. In such situations, an unscrupulous client can benefit by causing or aggravating the damage. Thus, the upper limit in Proposition 2 represents the legitimate demand of an honest client but not the prudent limitation set by the insurer.

The prudent insurer imposes an upper limit that prevents the rise in utility. The restriction is

$$u(e(t), w + I(t) - P - (1-e(t))t - e(t)c) \text{ is nonincreasing in } t. \tag{15}$$

In the presence of this restriction, equation (12) is superseded. Thus, the problem of optimum demand for insurance with the prudent limitation set by the insurer is that of equations (4) through (6) with the additional restriction of equation (15). Its solution is called the no–moral hazard solution.

The solution to this new problem has the expected form of full insurance up to a binding upper limit, as shown in the following proposition.

Proposition 3

The optimum indemnity in the problem of equations (4) through (6) with the added constraint (15) has a solution of the following form: There is a value t_3 satisfying $0 \leq t_3 \leq T$ such that the solution is

$$I(t) = t \text{ on } 0 \leq t \leq t_3 \tag{16}$$

$$I(t) = t_3 \text{ on } t_3 \leq t \leq T. \tag{17}$$

The no–moral hazard solution is shown by the heavy line in Figure 3. The point where moral hazard starts is t_3, which is the prudent limitation set by the insurer. In order to quantify t_3, the insurer needs to know the insured's preferences and solve them for the conversion point, a task that involves interesting quantitative challenges. In practical cases, the upper limit chosen by the insurer is derived from judgment and rules of thumb. There is, in addition, a considerable temptation to increase coverage beyond the prudent limitation, which is below the legitimate demand of an honest client, and good underwriting may yield to that temptation for suitably selected cases.

Many Housing Choices

The presence of additional alternatives does not affect the analysis much. The best way to see this is to redraw Figure 2 with another curve representing a third housing choice. The existing house is the lowest and left-most curve, and it moves from left to right with increasing damage while the other curves stand still. Under the no–moral hazard contract, the homeowner switches to whichever alternative house is encountered first, and that occurs exactly at the binding upper limit in the policy. Under the theoretical contract, the discontinuous upward jump is to one of the two alternative houses. Since at most one of the alternative houses plays an active role, the presence of a second alternative is largely irrelevant, as are third, fourth, etc. alternatives.

FIXED TRANSACTIONS COSTS

An important type of transactions cost is the fixed cost borne by consumers in moving from one house to another. Fixed transactions costs can lead to upper limits—or closely related structures—in the optimum insurance contract.

Consider a homeowner in a liquid real estate market. His home and all of the alternatives he might consider can be readily bought or sold, and the home is optimal for the land so there is no conversion decision. The homeowner thinks of moving from his current house to a lower-cost, less-preferred one. In moving, however, he would bear fixed transactions costs such as search costs, realtors' fees, financing costs, and moving expenses. Because the transactions costs are all fixed costs, the homeowner does not move from the current residence to another that is marginally worse. He moves—if at all—in a discontinuous leap to a house that is much less preferred.

The situation is illustrated in Figure 4. Given consumption of the nonhousing good equal to z, utility in the current house is u(1, z), and utility in the less-preferred house is u(0, z). Moreover, there is a savings from moving down. Consumption of z in the current house corresponds to a consumption in the less-preferred house of z + s, where s is the saving net of transactions costs. The utility realized is u(0, z+s), a curve whose position is found by taking the curve u(0, z) and shifting it to the left by the amount s, as shown in Figure 4.

Damage t is the cost to restore the current house to its previous market value. Utility in the current house as a function of damage t is u(1, z–t), as illustrated in Figure 4. As t increases, the current-house curve u(1, z–t) slides progressively to the right while the curve for the less-preferred house u(0, z+s) stands still.

The homeowner has wealth w and pays a premium P to purchase an insurance contract that pays I(t). Consumption is w + I(t) – P – t in the current house and w + I(t) – P – t + s in the less-preferred house. Letting e = 1 be the current house and e = 0 the less-preferred house, utility is

$$u(e, w + I(t) - P - t + (1-e)s). \tag{18}$$

Given the probability density function of damage f(t), the insurance problem is to

$$\underset{P,\, I(t),\, e(t)}{\text{Maximize}} \int_0^T u(e, w+I(t)-P-t+(1-e)s)f(t)dt \tag{19}$$

$$\text{subject to } P = \int_0^T I(t)f(t)dt \tag{20}$$

$$I(0) = 0. \tag{21}$$

The optimum contract is given by the following proposition.

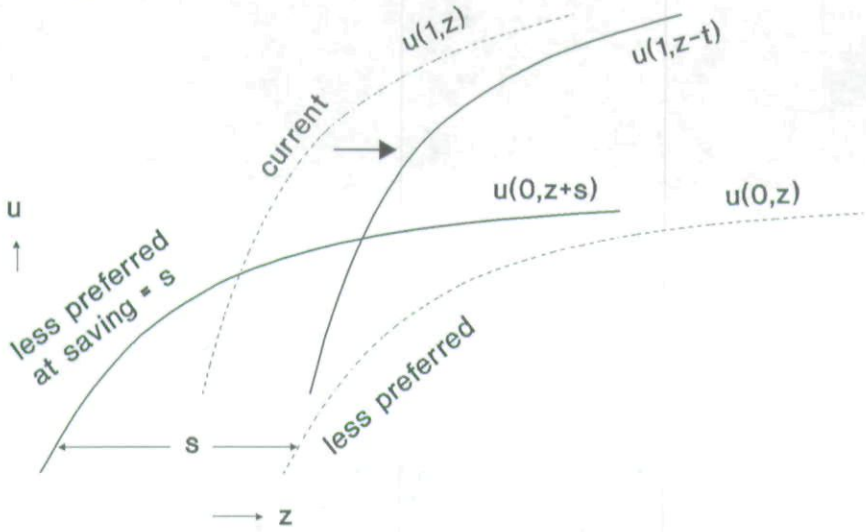
Proposition 4

The optimum indemnity in the problem of equations (19) through (21) has the following form: There is a value t_4 satisfying $0 \le t_4 \le T$ and a constant $k > 0$ such that

$$I(t) = t \text{ on } 0 \le t \le t_4 \tag{22}$$

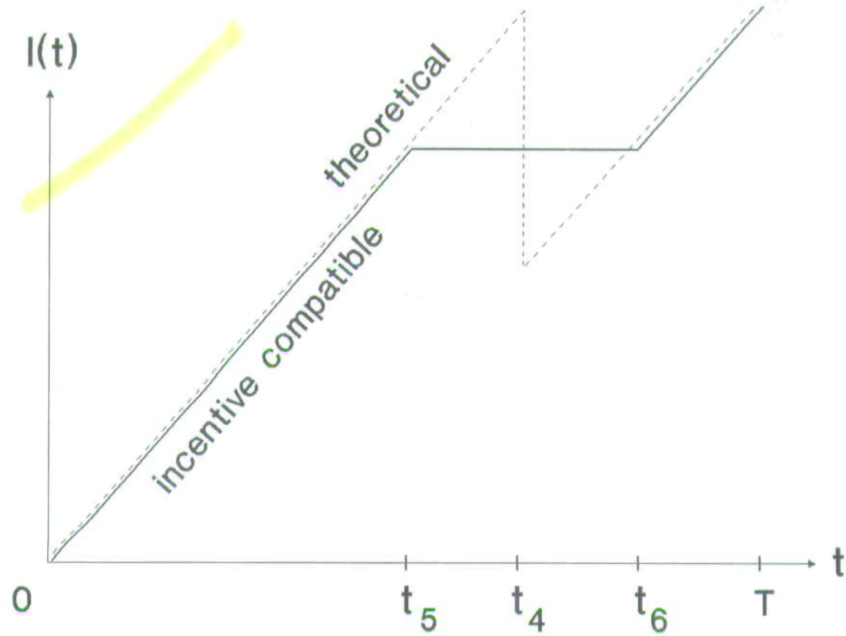
$$I(t) = t - k \text{ on } t_4 < t \le T. \tag{23}$$

Figure 4
Preferences with Fixed Transactions Costs



The interesting cases occur when t_4 is strictly between 0 and T. They are characterized by the sawtooth payment schedule illustrated by the dotted line in Figure 5. For damage up to the critical level t_4, the homeowner has full insurance and stays in the current house, but for damage above the critical level, he moves to the less-preferred house and accepts an indemnity that is less than the amount of damage. Such an insurance policy makes sense for a homeowner whose wealth is slightly less than that for which the current house would be an optimum choice. The fixed cost of a transaction prevents him from making a marginal adjustment, and therefore he remains in the current house even though it is a bit too expensive. Full insurance of all losses occurs when $t_4 = T$. Homeowners having high enough wealth choose such insurance and never exercise the option to move down. On the other hand, homeowners having low enough wealth immediately exercise the option to move down. In their case, $t_4 = 0$.

Figure 5
Incentive Compatibility



A homeowner whose optimum contract has the sawtooth payment schedule has an incentive to under-report large damages. In case of damage slightly larger than t_4, the homeowner reports damage slightly less than t_4, obtains a higher indemnity, and remains in the preferred, current house. Under-reporting of damage raises the expected value of indemnities and thus upsets the fair pricing of the contract. To avoid this problem, the insurer should impose a constraint on the payments so that the homeowner reports damage honestly or, equivalently, so that any false reporting is irrelevant. The needed constraint is that indemnity must be nondecreasing as reported damage increases,

$$\dot{I}(t) \geq 0. \tag{24}$$

The resulting solution is given by the following proposition.

Proposition 5

The optimum indemnity in the problem of equations (19) through (21) with the added constraint (24) has a solution of the following form: There are values t_5 and t_6 satisfying $0 \leq t_5 \leq t_6 \leq T$ such that

$$I(t) = t \text{ on } 0 \leq t \leq t_5 \tag{25}$$

$$I(t) = t_5 \text{ on } t_5 \leq t \leq t_6 \tag{26}$$

$$I(t) = t - (t_6 - t_5) \text{ on } t_6 \leq t \leq T. \tag{27}$$

The constraint that indemnity is nondecreasing in reported damage produces an optimum contract possessing features that are similar to those of insurance with a binding upper limit. Because of the constraint, the sawtooth dotted curve in Figure 5 is replaced by the monotone solid curve. The flat portion of the solid curve is, in effect, the upper limit demanded by the homeowner. Depending upon the exact parameters, the right-hand end of the flat spot may fall close to or equal to maximum total damage, and in those cases the optimum contract looks much like an ordinary insurance contract with a binding upper limit.

The contract in Proposition 5 reflects the observations made earlier about the dual nature of insurable interest. There is again a divergence between the legitimate demand of the homeowner and the prudent limitation imposed by the insurer, but here the direction of the divergence is reversed. The prudent limitation, which is equal to the maximum possible damage, is now greater than the homeowner's demand for insurance. Instead of demanding full insurance, the homeowner intends to move to a less-preferred house if damages are great. Thus, by his choice of an upper limit, he reveals the extent of his interest in the insured property. Another consumer owning a similar house and having higher wealth would have legitimate demand more nearly in accord with the prudent limitation of the insurer. But that homeowner, too, might have other options that define legitimate demand in conflict with the prudent limitation.

CONCLUSION

Much of the literature on optimum insurance is about insurance of wealth. The viewpoint abstracts from specifics and explores those features that are common to a large class of insurance problems. Upper limits in insurance are explained in that context by introducing regulatory constraints on policies (Raviv, 1979) or by including a bankruptcy option (Huberman, Mayers, and Smith, 1983). In contrast, the analysis given in this article maintains that insuring property damage is not the same thing as insuring wealth. Indeed, a binding upper limit on insured damage is necessary for full insurance of wealth. At least that is the case with zero transactions costs. In other cases, two of which are discussed above, the optima are not derived through the mediating concept of wealth but directly from maximization of expected utility. With or without transactions costs, the results flow from the fact that property can be traded or converted.

The analysis is an example of insurance theory that focuses not on wealth but on things that are insured. This branch of the theory studies the influence of the idiosyncratic features of each type of insurance, such as life, health, contract performance, liability, and casualty. Other examples of the insurance of things are the insurance of irreplaceable commodities (Cook and Graham, 1977) and analyses of health insurance through the use of state-dependent utility (e.g., Pauly, 1968; Zeckhauser, 1970; and Garratt and Marshall, 1994).

In all of the situations considered here, binding upper limits on coverage are optimum from the homeowners' point of view. Like the Oakland residents, they underinsure optimally and, in case of heavy damage, exercise their options.

APPENDIX

Proof of Proposition 1

In the problem of equations (4) through (6), substitute the notation

$$z(t) = w + I(t) - P. \tag{A1}$$

The constraint in equation (5) becomes

$$w = \int_0^T z(t)f(t)dt. \tag{A2}$$

With this notation, the problem of equations (4) through (6) can be written as a problem of optimum control:

$$\underset{e,z}{\text{Maximize}} \int_0^T u(e, z\text{-}(1\text{-}e)t\text{-}ec)f(t)dt. \tag{A3}$$

$$\dot{n} = -zf(t),\ n(0) = w;\ n(T) = 0 \tag{A4}$$

$$e(t) = 0 \text{ or } 1. \tag{A5}$$

The costate variable of n is v(t), and its equation of motion is

$$\dot{v} = 0 \tag{A6}$$

because n does not affect the value of the Hamiltonian

$$H(n,z,e,v,t) = u(e, z - (1-e)t - ec)f(t) + v(-zf(t)). \tag{A7}$$

Optimum controls are denoted by superscript bars. Given the solution value of the costate variable v, the necessary condition for optimality of the control z is that it should maximize $H(\bar{n}, z, \bar{e}, v, t)$. The condition is

$$u_2(\bar{e}, \bar{z}-(1-\bar{e})t-\bar{e}c) = v, \tag{A8}$$

showing that, for fixed $\bar{e}$, $\bar{z} - (1-\bar{e})t - \bar{e}c$ is a constant. Define z_0 and z_1 to satisfy

$$u_2(0,z_0) = v, \tag{A9}$$

and

$$u_2(1, z_1-c) = v, \tag{A10}$$

respectively. Define t_1 to satisfy

$$t_1 = z_1 - z_0, \tag{A11}$$

and define t_0 to satisfy

$$t_0 = z_1 - z_0 - \frac{u(1, z_1 - c) - u(0, z_0)}{v}. \tag{A12}$$

These values can be visualized in Figure 2. For maximizing the Hamiltonian, note that the isobars of $H = u - vz$ are lines with slope v in Figure 2, and that the controls $\bar{z}$ and $\bar{e}$ are chosen to reach the highest isobar.

From these observations and equations (A8) through (A12), the optimum necessarily has

$$\bar{z}(t) = z_0 + t \text{ and } \bar{e}(t) = 0 \text{ on } 0 \leq t \leq t_0, \tag{A13}$$

and

$$\bar{z}(t) = z_1 \text{ and } \bar{e}(t) = 1 \text{ on } t_0 < t \leq T. \tag{A14}$$

The solution in equations (A13) and (A14) satisfies a sufficient condition proposed by Arrow (1968) and established by Seierstad and Sydsaeter (1977) as their Theorem 3. The condition is that the maximized Hamiltonian

$$H^*(n, v, t) = \underset{e,z}{\text{Max}}\ H(n, z, e, v, t) \tag{A15}$$

should be a concave function of n. Here, concavity is satisfied trivially because $\partial H/\partial n = 0$.

Translate the results back to the notations I(t) and P by letting $I(t) = \bar{z}(t) - \bar{z}(0)$ and $P = w - \bar{z}(0)$. Q.E.D.

Proof of Proposition 2

Including constraint (12), the optimum control problem is to

$$\underset{e,a}{\text{Maximize}} \int_0^T u(e, z-(1-e)t-ec)f(t)dt \tag{A16}$$

$$\dot{n} = -zf(t),\ n(0) = w;\ n(T) = 0 \tag{A17}$$

$$\dot{z} = a,\ z(0) \text{ and } z(T) \text{ free} \tag{A18}$$

$$0 \leq a \leq 1 \tag{A19}$$

$$e(t) = 0 \text{ or } 1. \tag{A20}$$

Existence.[3] The first part of the proof is to establish the existence of an optimum control that attains the maximum implied in equation (A16). The values of the maximand are bounded above by the value found in Proposition 1. The problem to be surmounted is that the space of controls is not compact, and Young (1980) has shown that it becomes compact when it is augmented to include all "chattering" controls. Because of compactness, the solution is attained on the

[3] Because the Hamiltonian and the maximized Hamiltonian are nonconcave, sufficiency theorems, such as those of Seierstad and Sydsaeter (1977) or Kamien and Schwartz (1971), are not directly applicable.

augmented space. The next task is to show that it is an ordinary control, not a chattering one, and that is done in the second part of the proof, where the solution is denoted ($\bar{n}$, $\bar{z}$, $\bar{e}$, $\bar{a}$).

Properties. By the Pontryagin Theorem (1965, p. 22), there are costate variables $v_1(t)$ and $v_2(t)$ necessarily associated with the solution. Given the costate variables, the Hamiltonian of the problem is

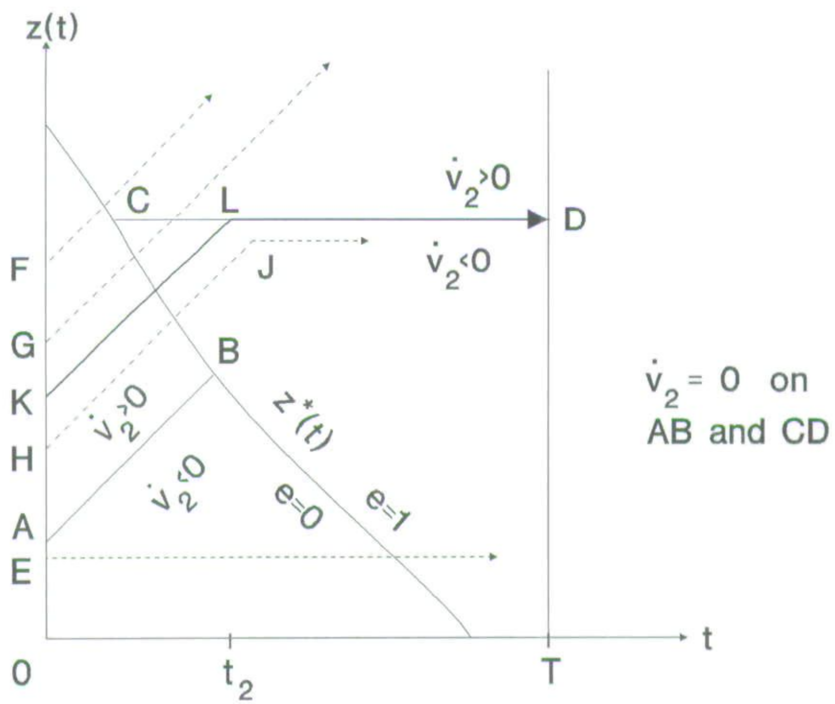
$$H(n,z,e,a,v_1,v_2,t) = u(e,z-(1-e)t-ec)f(t) - v_1 zf(t) + v_2 a. \tag{A21}$$

The necessary condition on control variable $\bar{e}$ is that it should maximize $H(\bar{n}, \bar{z}, e, \bar{a}, v_1, v_2, t)$, and it is achieved if $\bar{e}$ maximizes $u(e, \bar{z} - (1-e)t - ec)$. Define the boundary between the e = 0 area and the e = 1 area as $z^*(t)$ satisfying

$$u(0, z^*(t) - t) = u(1, z^*(t) - c). \tag{A22}$$

This locus is negatively sloped and illustrated in Figure A1.

Figure A1
Optimum Path KLD



Optimum choice of control variable a depends upon the value of the costate variable v_2. From equation (A21), the motion of v_2 is determined by

$$\dot{v}_2 = v_1(t) - u_2(\bar{e}, \bar{z} - (1-\bar{e})t - \bar{e}c). \tag{A23}$$

Also from equation (A21), it is clear that $\dot{v}_1(t) = -\partial H/\partial n = 0$. The constant function $v_1(t)$ is written as v_1. In Figure A1, the stable lines of v_2 are drawn. In the e = 0 area, $\dot{v}_2 = 0$ along the line AB, which satisfies

$$u_2(0, z-t) = v_1, \tag{A24}$$

and in the area where e = 1, the $\dot{v}_2 = 0$ locus CD satisfies

$$u_2(1, z-c) = v_1. \tag{A25}$$

Note that point B must lie southeast of point C, as drawn, because at point B the value of $u_2(e, z-(1-e)t - ec)$ jumps discontinuously upward as the control variable $\bar{e}$ switches from zero to unity. The jump is a consequence of normality of demand for housing. The areas where $\dot{v}_2 > 0$ and $\dot{v}_2 < 0$ are also shown in Figure A1.

A final necessary condition is that, because the endpoints of z are free, $v_2(0)$ and $v_2(T)$ are zero. For that reason, either $\dot{v}_2$ is constantly zero along the optimum path or else it must change sign. Thus, the optimum path cannot start at a point like E, for if it did, v_2 would immediately become negative, the control a would be forced to zero, the $\dot{v}_2 < 0$ region would never be escaped, and $v_2(T)$ could not be zero. Nor can the optimum path start at a point like F, for if it did, v_2 would immediately become positive, control a would be forced to unity, the $\dot{v}_2 > 0$ region would never be escaped, and $v_2(T)$ would not be zero. Similarly, a path starting at a point like G cannot be optimum because it passes into the zone of $\dot{v}_2 > 0$ while v_2 is still a nonnegative number. Again, $v_2(T)$ cannot be zero. Finally, a path starting at H cannot be optimum. Along this path, v_2 immediately becomes positive and grows up to its point of crossing with $z^*(t)$. From there, v_2 declines and reaches zero at point J. Thereafter, v_2 is negative and declining, implying $v_2(T) < 0$.

Thus, the optimum path must be like that starting at point K. Along the optimum path, v_2 immediately becomes positive and grows until it crosses $z^*(t)$, then declines and reaches zero at point L, remaining zero (because CD is the $\dot{v}_2 = 0$ locus) until point D is reached. Damage at point L is t_2. Translation from $\bar{z}$ to P and I is accomplished as in the proof of Proposition 1. Q.E.D.

Proof of Proposition 3

Invert the utility function to y(e, u) satisfying u = u(e, y(e,u)). Then the problem is written as

$$\underset{e,a}{\text{Maximize}} \int_0^T u(t)f(t)dt \tag{A26}$$

$$\dot{n} = -(y(e,u) + (1-e)t + ec)f(t),\ n(0) = w;\ n(T) = 0 \tag{A27}$$

$$\dot{u} = a,\ u(0) \text{ and } u(T) \text{ free} \tag{A28}$$

$$-m \le a \le 0 \tag{A29}$$

$$e(t) = 0 \text{ or } 1. \tag{A30}$$

As in the proof of Proposition 2, a solution to the problem exists, which is denoted ($\bar{n}, \bar{u}, \bar{e}, \bar{a}$), and the costate variables $v_1(t)$ and $v_2(t)$ are necessarily associated with the solution. Given the costate variables, the Hamiltonian of the problem is

$$H(n,u,e,a,v_1,v_2,t) = \tag{A31}$$

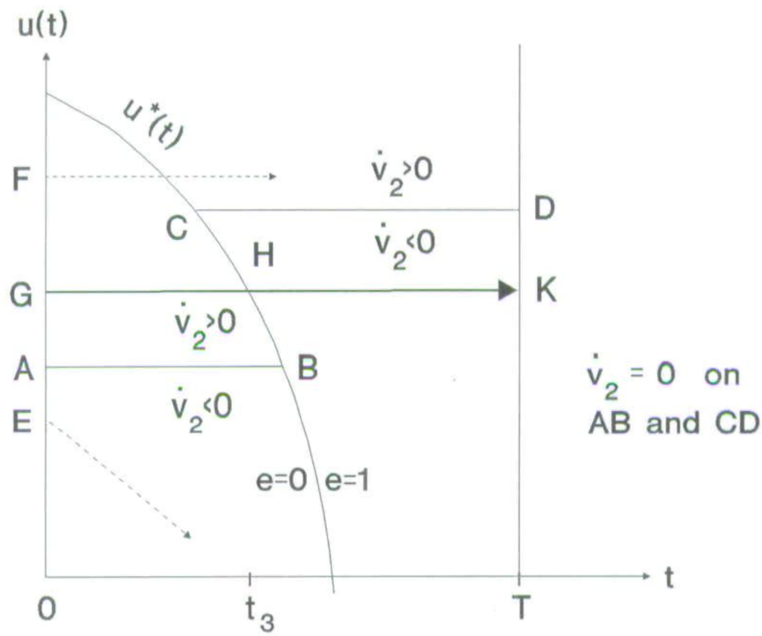
$$uf(t) - v_1(y(e,u) + (1-e)t + ec)f(t) + v_2 a.$$

The necessary condition on control variable e is that it maximizes $H(\bar{n}, \bar{u}, e, \bar{a}, v_1, v_2, t)$, and it is achieved if e minimizes $y(e, \bar{u}) + (1-e)t + ec$. Define the boundary between the $e = 0$ area and the $e = 1$ area as $u^*(t)$ satisfying

$$y(0, u^*(t)) + t = y(1, u^*(t)) + c. \tag{A32}$$

This curve is negatively sloped and illustrated in Figure A2.

Figure A2
Optimum Path GHK



From equation (A31), it is clear that $\dot{v}_1(t) = -\partial H/\partial n = 0$, and the constant $v_1(t)$ is denoted simply v_1. The motion of v_2 is

$$\dot{v}_2 = [1 - v_1 y_2(e,u)] f(t). \tag{A33}$$

Thus, $\dot{v}_2 = 0$ along a locus that satisfies

$$y_2(e,u) = \frac{1}{v_1}. \tag{A34}$$

In Figure A2, the stable line of v_2 is drawn. In the e = 0 area, it is a line segment AB, and in the e = 1 area it is CD. The height of AB is necessarily below that of CD because, given the normality of demand for the preferred house, marginal cost of utility $y_2(e, u)$ jumps abruptly downward when e changes from 0 to 1. The areas where $\dot{v}_2 > 0$ and $\dot{v}_2 < 0$ are also shown in Figure A2.

Because the endpoints of u(t) are free, $v_2(0) = v_2(T) = 0$, which means $\dot{v}_2$ must change sign along the optimum path. Thus, the optimum path cannot start at a point like E or F, for if it did, v_2 would be constantly either growing or declining over the whole of the necessarily level path from t = 0 to t = T.

The optimum path must be like that starting at point G. Along the optimum path, v_2 immediately becomes negative and falls until the path reaches point H, and then it begins to rise and reaches zero precisely at point K.

On the part of the optimum path from G to H,

$$\bar{z}(t) = y(0, u) + t \tag{A35}$$

is required to keep utility constant. From H to K, the requirement is

$$\bar{z}(t) = y(1, u) + c. \tag{A36}$$

Damage at point H is t_3. Since $v_2 < 0$ for all t, a = 0 for all t. Translation from $\bar{z}$ to P and I is accomplished as in the proof of Proposition 1. Q.E.D.

Proof of Proposition 4

Using the notation z(t) = w + I(t) – P, the problem of optimum control is to

$$\underset{e,z}{\text{Maximize}} \int_0^T u(e, z - t + (1 - e)s)f(t)dt \tag{A37}$$

$$\dot{n} = -zf(t),\ n(0) = w;\ n(T) = 0 \tag{A38}$$

$$e(t) = 0 \text{ or } 1. \tag{A39}$$

The solution is found as in the proof of Proposition 1. Given the constant value of the costate variable v, the necessary condition for optimality of the control z is

$$u_2(\bar{e}, \bar{z} - t + (1 - \bar{e})s) = v. \tag{A40}$$

Thus, for fixed $\bar{e}$, $\bar{z} - t + (1 - \bar{e})s$ is a constant. Define z_0 and z_1 to satisfy

$$u_2(0, z_0 + s) = v \tag{A41}$$

and

$$u_2(1, z_1) = v, \tag{A42}$$

respectively. Optimum $\bar{e}$ requires

$$u(0, z_0+s) - vz_0 = u(1, z_1) - vz_1. \tag{A43}$$

Finally, t_4 is determined from the budget equation

$$\int_0^{t_4}(z_1+t)f(t)dt + \int_{t_4}^{T}(z_0+t)f(t)dt = w. \tag{A44}$$

Given z_0, z_1, and t_4 that satisfy (A41) through (A44), the optimum necessarily has

$$\bar{z}(t) = z_1 + t \text{ and } \bar{e}(t) = 1 \text{ on } 0 \le t \le t_4 \tag{A45}$$

and

$$\bar{z}(t) = z_0 + t \text{ and } \bar{e}(t) = 0 \text{ on } t_4 < t \le T. \tag{A46}$$

Thus, $k = z_1 - z_0$ in the statement of the proposition.

Sufficiency follows as it did in the proof of Proposition 1 since u and $\dot{n}$ do not depend on n. Translation from $\bar{z}$ to P and I is accomplished as in the proof of Proposition 1. Q.E.D.

Proof of Proposition 5

Including constraint (24), the problem of optimum control is to

$$\underset{e,a}{\text{Maximize}}\int_0^T u(e, z\text{-}t+(1\text{-}e)s)f(t)dt \tag{A47}$$

$$\dot{n} = -zf(t),\ n(0) = w;\ n(T) = 0 \tag{A48}$$

$$\dot{z} = a,\ z(0) \text{ and } z(T) \text{ free} \tag{A49}$$

$$0 \le a \le 1 \tag{A50}$$

$$e(t) = 0 \text{ or } 1. \tag{A51}$$

As in the proof of Proposition 2, compactness is needed to ensure the existence of a solution. Compactness is achieved by augmenting the space of controls to include all chattering controls. Then the solution, which is denoted ($\bar{n}$, $\bar{z}$, $\bar{e}$, $\bar{a}$), is shown to be an ordinary control in the course of proving its other properties. Costate variables $v_1(t)$ and $v_2(t)$ are necessarily associated with the solution, and, given the costate variables, the Hamiltonian of the problem is

$$H(n,z,e,a,v_1,v_2,t) = u(e, z - t + (1-e)s)f(t) - v_1zf(t) + v_2a. \tag{A52}$$

The necessary condition on control variable e is that it maximizes $H(\bar{n}, \bar{z}, e, \bar{a}, v_1, v_2, t)$, and it is achieved if e maximizes $u(e, \bar{z} - t + (1-e)s)$. Define the boundary between the e = 0 area and the e = 1 area as $z^*(t)$ satisfying

$$u(0, z^*(t) - t + s) = u(1, z^*(t) - t). \tag{A53}$$

The curve is positively sloped and illustrated in Figure A3.

Optimum choice of control variable a depends upon the value of the costate variable v_2. From equation (A52), the motion of v_2 is

$$\dot{v}_2 = v_1(t) - u_2(e, z - t + (1-e)s). \tag{A54}$$

Also from equation (A52), it is clear that $\dot{v}_1(t) = -\partial H/\partial n = 0$. The constant function $v_1(t)$ is written as v_1. In Figure A3, the stable lines of v_2 are drawn. In the e = 0 area, $\dot{v}_2 = 0$ along the line AB, which satisfies

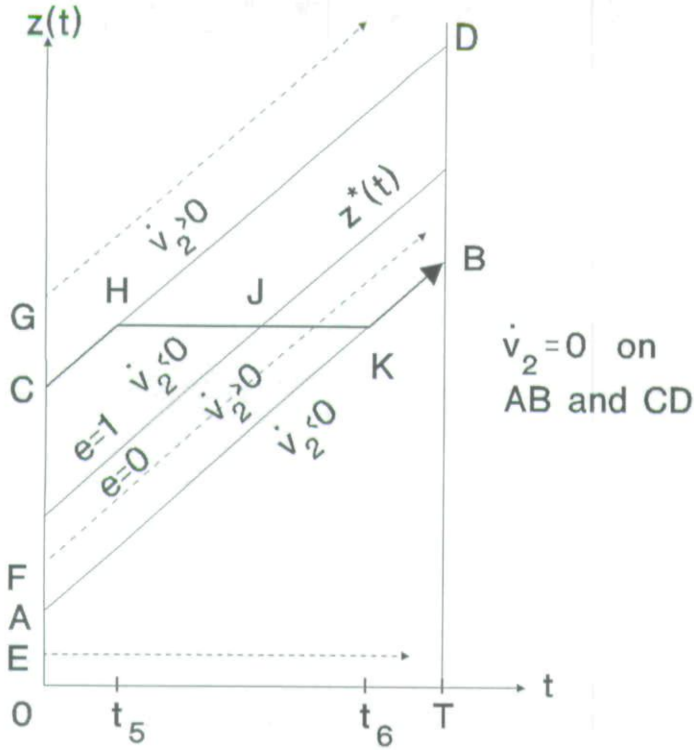
$$u_2(0, z-t+s) = v_{1,} \tag{A55}$$

and in the area where e = 1, the $\dot{v}_2 = 0$ locus CD satisfies

$$u(1, z-t) = v_{1.} \tag{A56}$$

The areas where $\dot{v}_2 > 0$ and $\dot{v}_2 < 0$ are also shown in Figure A3.

Figure A3
Optimum Path CHJKB



Because the endpoints of z(t) are free, $v_2(0)$ and $v_2(T)$ are zero. The optimum path cannot start at a point like E, for if it does, v_2 immediately becomes negative, the control a is forced to be zero, the $\dot{v}_2 < 0$ region is never escaped, and $v_2(T)$ cannot be zero. The optimum path cannot start at a point like F or G, for if it does,

v_2 immediately becomes positive, control a is forced to unity, the $\dot{v}_2 > 0$ region is never escaped, and $v_2(T)$ cannot be zero.

Thus, the optimum path must start at point C in Figure A3. The policy has full insurance up to H. Between H and J, a = 0 and $\dot{v}_2 < 0$. Between J and K, a = 0 and $\dot{v}_2 > 0$. At K, $v_2 = 0$ and a = 1 up to B. Damage is t_5 at point H and t_6 at point K. Translation from $\bar{z}$ to P and I is accomplished as in the proof of Proposition 1. Q.E.D

References

Arrow, Kenneth J., 1968, Applications of Control Theory to Economic Growth, in: G. B. Dantzig and A. F. Veinott, Jr., eds., *Mathematics and Decision Sciences* (Providence, R. I.: American Mathematical Society).

Arrow, Kenneth J., 1971, *Essays in the Theory of Risk Bearing* (Chicago: Markham).

Cook, Philip J. and Daniel A. Graham, 1977, The Demand for Insurance and Protection: The Case of Irreplaceable Commodities, *Quarterly Journal of Economics,* 91: 143–156.

Garratt, Rod and John M. Marshall, 1994, Triage, Efficacy of Medical Treatment, and Optimum Health Insurance, Working Paper in Economics 12–94, University of California, Santa Barbara.

Gollier, Christian, 1987, The Design of Optimal Insurance without the Nonnegativity Constraint on Claims, *Journal of Risk and Insurance,* 54: 312–324.

Gollier, Christian, 1992, Economic Theory of Risk Exchanges: A Review, in: Georges Dionne, ed., *Contributions to Insurance Economics* (Boston: Kluwer Academic Publishers).

Huberman, Gur, David Mayers, and Clifford W. Smith, Jr., 1983, Optimal Insurance Policy Indemnity Schedules, *Bell Journal of Economics,* 14: 415–426.

Kamien, Morton I. and Nancy Schwartz, 1971, Sufficient Conditions in Optimum Control Theory, *Journal of Economic Theory,* 3: 207–214.

Keeton, Robert E., 1960, *Basic Insurance Law* (St. Paul: West).

Pauly, Mark V., 1968, The Economics of Moral Hazard: Comment, *American Economic Review,* 58: 531–537.

Pontryagin, L. S., V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, 1965, *The Mathematical Theory of Optimal Processes* (New York: John Wiley).

Raviv, Artur, 1979, The Design of an Optimal Insurance Policy, *American Economic Review,* 69: 84–96.

Seierstad, Atle and Knut Sydsaeter, 1977, Sufficient Conditions in Optimal Control Theory, *International Economic Review,* 18: 367–391.

Sinn, Hans-Werner, 1982, Kinked Utility and the Demand for Human Wealth and Liability Insurance, *European Economic Review,* 17: 149–162.

Young, L. C., 1980, *Optimal Control Theory* (New York: Chelsea).

Zeckhauser, Richard, 1970, Medical Insurance: A Case Study of the Trade off Between Risk Spreading and Appropriate Incentives, *Journal of Economic Theory,* 2: 10–26.