Nitpicky Insurers and the Law of Contracts

(preliminary and incomplete)

Jean-Marc Bourgeon* Pierre Picard†

September 8, 2016

Abstract

Making insurance coverage conditional upon the circumstances of losses improves the risk sharing-incentive trade-off under moral hazard. If the insurer can offer complete contracts that specify the indemnity payment under all possible circumstances, then the optimal coverage depends both on the policyholder's financial loss and on a signal that reflects the likelihood of a high or low level of effort according to the circumstances. When circumstances cannot be exhaustively described in the insurance policy, this conditioning may go through legal disputes. We show how the law of contract should give some leeway to insurers to cut the indemnity in unfavorable circumstances in order to provide incentives to effort, while respecting legal principles based on the balance of probabilities.

Keywords: insurance, moral hazard, incomplete contracts.

JEL Classification Numbers: D82, D86, G22.

*INRA and Ecole Polytechnique (Department of Economics). Address: UMR Économie Publique, 16 rue Claude Bernard, 75231 Paris Cedex 05, France. E-mail: bourgeon@agroparistech.fr

†Ecole Polytechnique (Department of Economics). Address: Department of Economics, Ecole Polytechnique, 91128 Palaiseau Cedex, France. E-mail: pierre.picard@polytechnique.edu

1 Introduction

It is striking to observe how average citizens' perception of what is an insurance contract often differs from its usual definition by economists. The counterpart of the insurance premium as perceived by policyholders is generally nothing more than an imprecise expectation of what future indemnities will be, should a loss occur. This is particularly true for lines of risk where, rightly or wrongly, policyholders think that insurers have leeway in settling claims, and may nitpick on the indemnity payment if they have that possibility. On the other hand, economic analysis commonly relies on a crude description of what is an insurance policy, without room for ambiguity. An insurance contract is just characterized by the indemnity schedule that defines the insurance payment depending on the policyholder's loss, by the premium charged by the insurer, and sometimes by a policy dividend rule.[1] Defining an insurance policy in this way is suitable for analyzing risk-sharing and competitive interactions in a large variety of contexts, be they characterized by transaction costs, information asymmetry, parameter uncertainty, limits to risk pooling or other features.

Nevertheless, this approach imperfectly reflects actual insurance mechanisms. Frequently, the indemnity payment depends not only on the financial evaluation of the policyholder's loss, but also on the circumstances under which this loss occurred. This includes the causal mechanism that links the operative event with the damages, and the qualitative description of the loss itself. This often goes through exclusions that are stated in the small print of the contract. For instance, a corporate property policy may exclude the damages resulting from fire caused by an explosion, or from the transportation of hazardous materials. Similarly, a homeowner policy may exclude the theft that may result from a lack of minimum precaution, such as locking the door when leaving home.

Conditioning the insurance payment on the circumstances of the loss leads to a lower predictability of the coverage in the case of an accident. Hence, from a risk-sharing standpoint, there is no reason why this should be so. After all, people may inadvertently forget to lock their door! However, restricting coverage according to the circumstances of the loss may be worthwhile under moral hazard if circumstances are informative on the policyholder's effort, be it associated with risk prevention or loss reduction. This is a well-known result in incentive theory (in particular Holmström, 1979), but its implications for the design of insurance contract may not have been

[1] Of course, in practice, these ingredients of insurance contracts take many different forms, including, for instance, experience rating in automobile insurance or fee-for-service payment in health insurance.

sufficiently scrutinized. The first objective of this paper is to explore this issue.

To do so, we consider a setting in which the insurer receives a signal on the claim (an aggregated information about the circumstances of the loss) that is informative about the policyholder's behavior because some circumstances are more likely to occur than others, depending on the policyholder's effort. We begin by considering the simple case where the size of the potential loss is fixed, and we show that the insurer should pay full compensation or entirely deny the claim when the circumstances are very favorable or very unfavorable (i.e., when they most likely correspond to a high or low effort level, respectively), and he should provide partial coverage in more ambiguous cases. We then contemplate the case where the amount of the loss is uncertain. We show that the optimal indemnity schedule depends on whether the circumstances of the claim are a sufficient statistic of the policyholder's effort, i.e., whether they include all the relevant information to evaluate the probability distribution of losses. If so, then the optimal indemnity schedule takes a surprisingly simple form: the indemnity is calculated through a deductible that depends on the circumstances of the loss. The deductible vanishes or it fully cancels the compensation under the most favorable and most unfavorable circumstances, respectively, and it partially reduces the indemnity in intermediate cases. In essence, the retained loss only depends on the circumstances and not on the size of the claimed loss. This strongly differs from the standard result of optimal insurance under moral hazard where retained losses are a function of total losses, for instance under the form of a straight deductible contract or with a mixture of deductible and coinsurance. If the circumstances are not a sufficient statistic of the policyholder's effort, then the optimal contract should mix partial coverage at the margin and conditioning on the loss circumstances.

However, the circumstances that characterize a loss may be so diverse that it would be impracticable to describe them all in an insurance contract. This is not an issue in practice when the treatment of information is regulated by well formatted processes (for instances police and expert reports for car accidents) and all the possible circumstances are classified in a comprehensive way by insurance regulators, so that the interpretation of contracts precludes any ambiguity.[2] But the law of insurance contracts is not always a problem solver that leads to the straightforward settlement of the disputes arising between parties. When the diversity of possible circumstances is too large,

[2]This is generally the case for lines of business such as automobile and homeowner products, where many similar accidents occur each year, and the law of contracts has established a well-defined interpretation of circumstances with little room for ambiguity.

insurance law may not provide a unilateral answer to all possible disagreements between insured and insurer.

Considering the claim settlement stage, we may say that insurance policies are incomplete contracts when they do not specify the indemnity payment in an unambiguous way in all possible situations. As for any incomplete contract, once uncertainty on the state of nature is resolved, i.e., when an event triggers a claim for which the letter of the insurance contract does not provide a clear-cut answer, there is room for bargaining between insured and insurer. Disputes may be resolved by amicable settlement, or by resorting to an arbitrator or by going to court, but in some way or another, they are arbitrated by law. If well-established case law does not provide an immediate answer in an insured-insurer dispute, then the ultimate decision of judicial authorities will be based on fundamental (not case-specific) principles of the law of insurance contracts: duty of utmost good faith, prohibition of preexisting condition non-disclosure, cancellation of contract for negligence, invalidity of a claim because of misrepresentation, absence of evidence about the operative event, onus of the proof regarding the loss..., just to name a few.

The second part of the present paper analyzes this issue by immersing the moral hazard problem in an incomplete contracting setting, and by considering the interaction between insurer and insured after a claim is filed. Conditioning the indemnity on the circumstances of the loss can be indirectly reached through legal disputes, in which the insurer invokes a legal means (a non case-specific principle of insurance law) to fully or partially deny the claim, and courts may dismiss his allegation. On the basis of the (freely available) soft information on the loss circumstances included in the policyholder's claim, the insurer may either validate it or he may search for verifiable information through a costly audit in order to sustain a nitpicking strategy that would not be dismissed in court. We consider a setting where the standard of proof used by courts is the balance of probabilities, i.e., courts weigh up the evidence based on the circumstances of the loss and decides which was the most probable behavior of the policyholder, and ultimately whether the indemnity should be paid fully or not. Allowing insurers to allege possible misconduct of the policyholder is an indirect way to condition the insurance coverage on the circumstances of the loss, because these circumstances affect the final decision that will result from the application of the law of contracts.

In a previous paper related to the insurance fraud issue, (Bourgeon & Picard, 2014), we showed that cutting the indemnity according to the circumstances of the loss could be welfare improving, when such a nitpicking strategy acts as an incentive device for insurers to perform costly audits. We assumed that the law of insurance contracts restricts the contractually feasible indemnity cuts, but we didn't analyze the determinants of these legal stipulations. The present paper goes a step further by analyzing how the law of contracts may be an indirect way to condition insurance coverage on the circumstances of the loss, thereby improving the incentive-risk sharing trade-off that is inherent to insurance under moral hazard.

The rest of the paper is as follows. Section 2 introduces the model. Section 3 considers the case of complete insurance contracts, in which the contractual insurance indemnity is specified for all possible contingencies that may be at the origin of a claim. Section 4 raises the issue of incomplete insurance contract. Section 5 concludes. Proofs are in an appendix.

2 The model

Consider an insurance company providing coverage to a risk-averse individual (household or firm) against accidents that may result in a loss L. The behavior of the policyholder is indexed by $b \in \mathcal{B} = \{1, 2, \ldots, n\}$ with probability of accident π_b for behavior b and $\pi_1 < \pi_2 < \ldots < \pi_n$. Hence, $b = 1$ corresponds to a cautious behavior, with the lowest probability of accident and, in that case, we may say that the policyholder exerts effort. $b = 2, \ldots, n$ correspond to various types of misconducts, ranked in the order of increasing probability of accident.[3] The policyholder's von Neumann-Morgenstern utility function is written as $u(W_f) - d_b$ with $u' > 0, u'' < 0$, where W_f is her final wealth and d_b denotes the disutility of effort under behavior b, with $d_b < d_1$ for all $b \geq 2$. Any accident is characterized by its circumstances, i.e., by all the specificities of the operative event at the origin of the loss. The circumstances of an accident are denoted $\omega \in \Omega$, with $(\Omega, \mathcal{F}, \mathbb{P}_b)$ a probability space where the probability measure $\mathbb{P}_b$ depends on the policyholder's behavior.

3 Complete contracts

The insurer has an information system that reflects his greater or lesser capacity to verify the circumstances of the accident. It associates a real number $x \in [0, 1]$ to any possible circumstances $\omega \in \Omega$ in such a way that x satisfies the strict Monotone Likelihood Ratio Property (MLRP). In more formal terms, there exists $\varphi : \Omega \longrightarrow [0, 1]$ such that $x = \varphi(\omega)$ has density functions $g_b(x)$ for

[3] For instance, a car driver may exert a low level of effort because she does not adequately maintain her vehicle, or because her speed is not appropriate, or because she drives after drinking, or because of a mixture of these behaviours.

all $b \in \mathcal{B}$, and the likelihood ratios $g_{b+1}(x)/g_b(x)$ are increasing over $[0,1]$ for all $b = 1, \ldots, n-1$.[4] Hence, x is a signal of the policyholder's effort: in the case of an accident, it is more likely that the policyholder did not behave cautiously if x is large.[5] We denote $\phi_b(x) \equiv g_b(x)/g_1(x)$ for $b = 2, \ldots, n$, with $\phi_b'(x) > 0$.

An insurance contract specifies the indemnity $I(x)$ for all $x \in [0,1]$, and the premium P. Such contracts are said to be complete because the indemnity payment is conditioned on the circumstances of the accident through the insurer's information system. The policyholder's final wealth is $W_f = W - P$ if there is no accident, and $W_f = W - P - L + I(x)$ if there is an accident under circumstances ω such that $x = \varphi(\omega)$.

Under behavior b, the policyholder's expected utility is $Eu_b^* - d_b$ where

$$Eu_b^\star = (1 - \pi_b)u(W - P) + \pi_b \int_0^1 u(W - P - L + I(x))g_b(x)dx \tag{1}$$

The first and second terms in (1) correspond to the no-accident and accident states respectively. We assume that the parameters of the problem are such that it is optimal to induce the policyholder to behave cautiously, i.e., to abstain from any form of misconduct, which will be the case if the incentive constraints

$$Eu_1^\star - Eu_b^\star \geq d_1 - d_b \tag{2}$$

are satisfied for all $b \geq 2$. The insurance premium must at least cover the expected indemnity payments when the policyholder behaves cautiously, i.e.

$$P \geq \pi_1 \int_0^1 I(x)g_1(x)dx. \tag{3}$$

[4] We know from Milgrom (1981) that such an information system exists when circumstances $\omega \in \Omega$ are "comparable signals." Comparability means that, for any pair ω, ω' either ω is "more favorable than ω'" - in the sense that, for any non-degenerate prior on $b \in \mathcal{B}$, the posterior distribution after ω' dominates the posterior distribution after ω in the sense of strong FOSD, or ω' is more favorable than ω, or they are equivalent (i.e., they do not affect the beliefs whatever the prior). $x = \varphi(\omega)$ is distributed with density $g_b(x)$ if Ω is a subset of $\mathbb{R}^m$ and $\mathbb{P}_b$ can be represented by a density function on Ω. Function $\varphi(\omega)$ is not defined in a unique way. Different functions φ may exist, all of them satisfying strong MLRP. They correspond to different ways in which data on the accident circumstances can be used through a signal x. In what follows, we assume that $g_b(x) > 0$ for all $x \in (0,1)$ and all $b \in \mathcal{B}$, with c.d.f. $G_b(x)$. We have $G_b(x) < G_{b+1}(x)$ for all $x \in (0,1)$ because strong MLRP implies strong FOSD.

[5] More precisely, for any nondegenerate prior on b, an increase in x induces a FOSD shifts in the posterior probability distribution of b. Note that there exists $\varphi(.)$ such that x is a sufficient statistic for ω - see Milgrom (1981). This would correspond to a complete information system where x would gather all the information that could affect the posterior distribution of b.

Finally, we assume that overinsurance is ruled out, either for legal reasons or because the policyholder could deliberately create losses in order to pocket the insurance indemnity. Taking into account the non-negativity constraint on the indemnity, we should have

$$0 \leq I(x) \leq L \text{ for all } x \in [0,1]. \tag{4}$$

The optimal complete insurance contract maximizes $Eu_1^{\star}$ given by (1) with respect to P and $I(\cdot)$ subject to (2),(3) and (4). The optimal indemnity schedule characterized in Proposition 1. We know that conditioning the insurance indemnity on x is efficient - i.e., it allows the policyholder to reach a higher expected utility in comparison to the case where the policyholder receives a fixed indemnity in the case of an accident - because x is informative in the sense of Holmström (1979). The remainder of this section shows how this conditioning show be implemented.

Proposition 1 *Under moral hazard, the optimal complete indemnity schedule $I^{\star}(x)$ is continuous, with*

$$\begin{aligned} I^{\star}(x) &= L && \text{if } 0 \leq x < \underline{x}, \\ dI^{\star}(x)/dx &< 0 && \text{if } \underline{x} < x < \overline{x}, \\ I^{\star}(x) &= 0 && \text{if } \overline{x} \leq x \leq 1. \end{aligned}$$

and $0 \leq \underline{x} < \overline{x} \leq 1$. Furthermore, $\phi_b(x) \to 0$ for all $b \geq 2$ when $x \to 0$ is a sufficient condition for $\underline{x} > 0$ and $\phi_b(x) \to +\infty$ for all $b \geq 2$ when $x \to 1$ is a sufficient condition for $\overline{x} < 1$.

Proposition 1 is illustrated in Figure 1. It states that the optimal insurance policy provides full coverage, partial coverage or zero coverage, depending on the circumstances of the loss. The more favorable the circumstances (i.e., the lower x), the larger the indemnity. The higher and lower bound L and 0 are reached under the most favorable or worst possible circumstances, respectively, and there is partial coverage in the intermediary cases. In other words, the optimal trade-off between incentives and risk coverage is reached by conditioning the insurance compensation on the circumstances of the loss: the most favorable these circumstances, the larger the compensation. We may also write $I^{*}(x) = [1 - z^{\star}(x)]L$, with $z^{\star}(x) = 0$ if $x < \underline{x}$, $z^{\star}(x) \in (0,1)$ with $z^{*\prime}(x) > 0$ if $\underline{x} < x < \overline{x}$ and $z^{\star}(x) = 1$ if $x > \overline{x}$. Hence, the insurance contract provides full coverage when the circumstances of the loss are favorable enough, and there is a partial or total indemnity cut for less favorable circumstances.

Figure 1

The signal $x = \varphi(\omega)$ is not defined in a unique way since many functions φ lead to the strict MLRP property. In a very imprecise but intuitive way, we may anticipate that the more informative the signal, the stronger the dependence between the signal and the insurance coverage. To be more precise in what we mean by informativeness, let us consider a family of signals $x_\sigma \in [0,1]$ indexed by $\sigma \in \Sigma$, with densities $g_b(x,\sigma)$ for $b \in \{1,2,\ldots,n\}$ and likelihood ratios $\phi_b(x,\sigma) \equiv g_b(x,\sigma)/g_1(x,\sigma)$ such that $\partial\phi_b/\partial x > 0$ for all $b \geq 2$. Signal x_σ is generated by function $\varphi_\sigma : \Omega \longrightarrow [0,1]$ such that $x_\sigma = \varphi_\sigma(\omega)$. Parameter σ is used to measure the informativeness of the signal x_σ. Let $I^*(x,\sigma)$ be the optimal indemnity schedule when insurance coverage is conditioned on signal x_σ. Proposition 2 provides a sufficient condition for a change in the signal to induce a more direct dependence between the signal and the indemnity. It is illustrated in Figure 2.

Proposition 2 *Let* $\sigma_0, \sigma_1 \in \Sigma$. *Assume that* $x \longrightarrow \dfrac{\partial\phi_b(x,\sigma_1)/\partial x}{\partial\phi_b(x,\sigma_0)/\partial x}$ *is increasing for all* $b \in \{2,\ldots,n\}$. *Then, there exists* $x^\star \in (0,1)$ *such that*

$$\begin{aligned}
I^\star(x,\sigma_1) &\geq I^\star(x,\sigma_0) \quad \textit{if } 0 \leq x < x^\star, \\
I^\star(x,\sigma_1) &= I^\star(x,\sigma_0) \quad \textit{if } x = x^\star, \\
I^\star(x,\sigma_1) &\leq I^\star(x,\sigma_0) \quad \textit{if } x^\star \leq x \leq 1.
\end{aligned}$$

For the sake of illustration, consider the case where $G_1(x,\sigma_1) \equiv G_1(x,\sigma_0)$. Thus, the probability distributions of signals x_{σ_0} and x_{σ_1} over $[0,1]$ only differ when $b \geq 2$. In that case, proposition 2 implies

$$\begin{aligned}
E[I^\star(x,\sigma_1)|I^\star(x,\sigma_1) &> \overline{I}^\star] > E[I^\star(x,\sigma_0)|I^\star(x,\sigma_0) > \overline{I}^\star], \\
E[I^\star(x,\sigma_1)|I^\star(x,\sigma_1) &< \overline{I}^\star] < E[I^\star(x,\sigma_0)|I^\star(x,\sigma_0) < \overline{I}^\star],
\end{aligned}$$

where E is the expected value operator when x is distributed according to $G_1(x,\sigma_1) = G_1(x,\sigma_0)$ and $\overline{I}^* = I(x^\star,\sigma_1) = I(x^\star,\sigma_0)$. Thus, if the insurance indemnity $I^\star(x,\sigma_0)$ is larger (resp. lower) than $\overline{I}^\star$, then $I^\star(x,\sigma_1)$ will be even larger (resp. lower) than $I^\star(x,\sigma_0)$. In other words, conditioning the insurance payment on x_{σ_1} rather than x_{σ_0} exacerbates the variability of the insurance indemnity. This is a somewhat paradoxical conclusion when the policyholder's expected utility is higher with x_{σ_1} rather than x_{σ_0} since increasing the variability of the insurance payment is a priori detrimental for risk averse policyholders.+ Under the condition provided in proposition 2, it is nevertheless optimal to do so because of the effect of this change on the policyholder's incentives to exert effort.

Figure 2

In practice, conditioning insurance coverage on the circumstances of the loss frequently goes through exclusions. The insurer commits to pay an indemnity I in the case of an accident, except under well-defined circumstances. This amounts to restricting the indemnity schedule to

$$I(x) = \begin{cases} I \text{ if } x \notin \Upsilon \\ 0 \text{ if } x \in \Upsilon, \end{cases}$$

where $\Upsilon \subset [0,1]$ is the set of signals for which an exclusion applies.

Proposition 3 *The optimal exclusion-based insurance contract is such that $\Upsilon = [\hat{x}, 1]$ where $\hat{x} \in (0,1]$. Furthermore, $\phi_b(x) \to +\infty$ for all $b \geq 2$ when $x \to 1$ is a sufficient condition for $\hat{x} < 1$.*

Proposition 3 characterizes the optimal solution when the insurance contract just includes a fixed indemnity I and exclusions. It is illustrated in Figure 3. Unsurprisingly, an exclusion applies when x is larger than a threshold $\hat{x}$, that is to say, under the most unfavorable circumstances. If $\phi_b(x) \to +\infty$ for all misconducts b when $x \to 1$, then the less favorable circumstances signal that it is very likely that the policyholder misbehaved in some way, and thus an exclusion applies when x is close to 1.

Figure 3

Let us now turn to the case where accidents may be more or less severe. An accident still occurs with probability π_b, but we now assume that the loss is $\ell \in [0, \overline{\ell}]$. The insurance indemnity is $I(\ell, x)$ and $f_b(\ell, x)$ denotes the joint density of (ℓ, x) over $[0, \overline{\ell}] \times [0,1]$ when the policyholder's behavior is b. By a straightforward adaptation of the previous problem, the optimal complete insurance contract $\{P, I(\cdot)\}$ maximizes

$$Eu_1 = (1 - \pi_1)u(W - P) + \pi_1 \int_0^1 \int_0^{\overline{\ell}} u(W - P - \ell + I(\ell, x)) f_1(\ell, x) d\ell dx,$$

subject to

$$(\pi_b - \pi_1)u(W - P) - \int_0^1 \int_0^{\overline{\ell}} u(W - P - \ell + I(\ell, x))[\pi_b f_b(\ell, x) - \pi_1 f_1(\ell, x)] d\ell dx \tag{5}$$

$$\geq d_1 - d_b \quad \text{for all } b = 2, \ldots, n,$$

$$P \geq \int_0^1 \int_0^{\overline{\ell}} I(\ell, x) f_1(\ell, x) d\ell dx, \tag{6}$$

$$0 \leq I(\ell, x) \leq \ell \quad \text{for all } \ell, x \tag{7}$$

We know from Holmström (1979) that in standard insurance models with moral hazard where the contractual indemnity depends only on the loss, a straight deductible is optimal when the policyholder's conduct affects the probability of an accident but not the conditional distribution of losses, i.e., when the effort is targeted toward risk prevention. If the policyholder's effort mixes risk prevention and loss-reduction,[6] then the optimal coverage combines a deductible with partial coverage at the margin.[7] Holmström (1979) also shows that the indemnity should depend on any signal which is informative about the policyholder's effort. The circumstances of the loss are such a signal, and thus it is not surprising that the optimal insurance indemnity should be conditional on them.

There are two relevant cases to investigate depending on the content of information revealed by the circumstances of the accident. In the first case, the circumstances of an accident are a sufficient statistic for the policyholder's behavior: in that case, the density of ℓ conditionally on x does not depend on behavior b, and thus it may be written as

$$f(\ell|x) = \frac{f_b(\ell, x)}{g_b(x)},$$

where, as previously, $g_b(x)$ is the density of x under behavior b, assumed to satisfy MLRP. In other terms, the circumstances of the accident include all the relevant information to evaluate the probability distribution of losses.[8] We obtain the following result:

Proposition 4 *If the circumstances of a loss are a sufficient statistic for the policyholder's effort, then the optimal complete insurance contract is written as*

$$I^\star(\ell, x) = \max\{\ell - \Phi(x), 0\}.$$

where $\Phi(\cdot) : [0,1] \longrightarrow \mathbb{R}_+$ *is continuous, with* $\Phi(x) = 0$ *when* $0 \leq x \leq \underline{x}$, $\Phi'(x) > 0$ *when* $\underline{x} < x < 1$. *Furthermore,* $\lim_{x \longrightarrow 0} \phi_b(x) = 0$ *for all* $b \geq 2$ *is a sufficient condition for* $\underline{x} > 0$ *and* $\lim_{x \longrightarrow 1} \phi_b(x) = +\infty$ *for all* $b \geq 2$ *is a sufficient condition for* $\overline{x}(\ell) < 1$ *with* $\Phi(\overline{x}(\ell)) = \ell$.

Thus, when circumstances are a sufficient statistic for the effort level, the optimal complete insurance contract is characterized by a deductible $\Phi(x)$ that depends on the accident circumstances,

[6] Loss reduction efforts are called self-insurance by Ehrlich & Becker (1972).

[7] A positive deductible is always optimal when there is a mass of probability on the no-loss state. See Winter (2013).

[8] Put differently, we may say that knowing the policyholder's effort in addition to the circumstances of the accident would not change the probability distribution on losses.

with full coverage at the margin above the deductible. The deductible vanishes when the circumstances are the most favorable, i.e., when $\varphi(\omega) \leq \underline{x}$. In less favorable circumstances, the deductible is positive and the larger x, the larger the deductible. For given losses ℓ, the insurer may not pay any indemnity when circumstances are very unfavorable, i.e., when $x \geq \overline{x}(\ell)$.

The following result characterizes the optimal complete contract when the circumstances of the accident are not a sufficient statistic for the effort, and $f_b(\ell|x)$ satisfies strict MLRP, that is $\ell \longrightarrow f_b(\ell|x)/f_0(\ell|x)$ is increasing with ℓ, for all $x \in [0,1]$ and all $b \in \{2,\ldots,n\}$.

Proposition 5 *If the circumstances of an accident are not a sufficient statistic for the policyholder's effort, then the optimal complete insurance contract is written as*

$$I^\star(\ell,x) = \max\{\ell - \Phi(\ell,x), 0\},$$

where $\Phi(\cdot) : [0,\overline{\ell}] \times [0,1] \longrightarrow \mathbb{R}_+$ *is continuous, with* $\Phi'_\ell > 0, \Phi'_x > 0$ *when* $\Phi(\ell,x) > 0$. *Furthermore,* $\lim_{x \longrightarrow 0} f_b(\ell,x)/f_1(\ell,x) = 0$ *for all* $b \geq 2$ *is a sufficient condition for* $\Phi(\ell,x) = 0$ *when* x *is close to* 0, *and* $\lim_{x \longrightarrow 1} f_b(\ell,x)/f_1(\ell,x) = +\infty$ *for all* $b \geq 2$ *is a sufficient condition for* $\Phi(\ell,x) = \ell$ *when* x *is close to* 1.

Proposition 5 shows that there should be partial coverage at the margin when the loss is large enough and the circumstances are favorable enough for positive coverage to be optimal. Comparing Propositions 4 and 5, we conclude that the optimal contract should include complete or partial coverage at the margin, depending on whether or not the accidents' circumstances are a sufficient statistic for the policyholder's effort. As shown in Proposition 5, partial coverage at the margin reinforces incentives when circumstances do not include all the relevant information about effort. Otherwise, as shown in Proposition 4, the marginal coverage should be equal to one, and it is optimal to condition the deductible on the accident circumstances. In both cases, less favorable circumstances should lead to less generous indemnity. If circumstances signal a misconduct of the policyholder (i.e., x is close to 1), then cancelling the indemnity is optimal. Conversely, when circumstances signal that the policyholder has exerted the required effort level (i.e., x is close to 0), then full coverage of losses is optimal.

4 Incomplete contracts

Let us now consider the case where contractual insurance payments cannot be conditioned upon the circumstances of the loss. We assume that the insurer costlessly observes a signal about circumstances when he receives claims, but this is only soft information upon which insurance coverage cannot be contractually conditioned. The implementation of the contract requires information that can be verified by a third party, and gathering such a hard information goes through a costly audit of claims. In other words, when insurers decide to nitpick, they must have access to objective informations on which they can rely during the claim settlement process, particularly if it ends up in court.

Whether disputes are resolved by amicable settlement, by resorting to an arbitrator or by going to court, in some way or another they are resolved in compliance with the law of contracts. Hence, after privately observing a signal correlated with the circumstances of a loss, the insurer may decide to gather verifiable evidence on the behavior of the insured that will allow him to invoke the law of insurance contracts in order to justify a cut in the indemnity. The insurer is willing to perform an expensive audit process (checks and, possibly, testimonials, invoices, an inventory, etc. ...) if it reduces the expected indemnity payment by an amount larger than the audit cost.

In view of the evidences, the insurer may be allowed by law to wholly or partially cancel the contractual compensation. In concrete terms, the insurer may allege that the policyholder misbehaved in some way and he may invoke a legal means that justifies the cut in indemnity. For instance, in many property insurance settings, the law of contracts specifies that the policyholder has a duty of care, and the insurer is allowed by law to totally or partially reject the claim if he considers that this principle was disregarded, and if courts do not invalidate his decision. Table 1 provides examples of such a correspondence between alleged misconducts and legal means.

Let us follow a standard way in the analysis of conflicts arbitrated by law, which consists of assuming that judges decide by relying on the likelihood of the behavior alleged by each party. An insurer may allege that the policyholder misbehaved and thus, on the basis of the law of contracts, that the claim should be fully or partially denied. However, the insurer's allegations must be consistent with the empirical evidence provided by the circumstances of losses, for otherwise the judge would consider them as insurer's bad faith and they would be invalidated. In other words, insurance law may allow insurers to condition insurance payments on loss circumstances by

insurance lines	**misconducts**	**legal means**
health insurance	hiding smoker status	non-disclosure
life insurance	misrepresenting pre-existing heath conditions	misrepresentation
homeowner insurance	not taking maintenance measures if a major storm is announced	duty of care
property insurance	non revealing valuables that could be stolen	non-disclosure
property insurance	undervaluing an asset	misrepresentation
automobile insurance	driving on worn tires	duty of care
automobile insurance	driving under the influence of alcohol	reckless conduct
life insurance	Participating in an hazardous activity or sport	reckless conduct

Table 1: Example of misconducts and legal means

opening the door to legal indemnity cuts ("law completes contracts" as it is sometimes said), but its application is constrained by the approval of the judge on the basis of available information. This corresponds to the usual standard of proof for civil cases in the Common Law: judges are supposed to decide "on the balance of probabilities."[9]

To deal with these issues in a simple way, let us come back to our basic model where the size of the loss is given and denoted by L. When a claim is filed, the insurer privately observes a signal $s \in R$ defined by

$$s = x + \varepsilon,$$

still with $x = \varphi(\omega)$ and ε is a zero-mean random variable, with $cov(x, \varepsilon) = 0$. We assume that x can be verified by auditing the claim, which costs c to the insurer, and we denote $q(s) \in [0, 1]$ the audit probability when signal s is perceived.

[9] In such a setting, more serious allegations require more convincing evidence on the severity of the alleged misconduct. This is illustrated by Clarke (1997): "If the insurer defends with a contract exception, he must prove that exception. For example, a contract requirement that the insured "shall take all reasonable steps to safeguard the property insured" has been seen as an exception of negligence, so the claimant is not required to prove care as a condition precedent of cover, but it is for the insurer, if he wishes and can do so, to prove negligence by the insured. The onus on the insurer is to prove the exception on the balance of probabilities, but that onus will be heavier when the defence alleges fraud or wilful misconduct, such as arson, by the insured." In other words, the more severe the policyholder's misbehavior alleged by the insurer, the more demanding the evidence that is required to sustain his allegation.

When no audit is performed, the insurer pays I to the claimant. If x has been verified through an audit, then the insurance payment depends upon x through legal means that may be invoked by the insurer. For notational simplicity, we do not distinguish misconducts $b \in \{2, \ldots, n\}$ from the corresponding legal means that can be invoked by insurers and we suppose that insurance law specifies the insurer's leeway in the claim settlement process, i.e., to what extent a legal means allows him to cut or even cancel coverage. More precisely, the law of insurance contracts in the following is subsumed by the proportions $y_b \in [0, 1]$ of the claim that the insurer is allowed to cut for each behavior b.[10] When the insurer is allowed to fully cancel the claim, we have $y_b = 1$. We assume $y_1 = 0$ because the insurer can cut the indemnity only by alleging that the policyholder misbehaved in some way. We also postulate that law is constrained by a severity principle, according to which the severity of misconducts and the intensity of indemnity cuts are co-monotone, i.e., $y_b \leq y_{b+1}$ for all $b = 1, \ldots, n-1$.

Hence, if the insurer is in a position to invoke legal means b under circumstances ω such that $x = \varphi(\omega)$, then he may decide to cut the indemnity by a fraction $z(x)$ lower or equal to y_b, and the insurance payment is $[1 - z(x)]I$.

The previous model has also to be adapted to incorporate the possibility that policyholders may misbehave with some probability, so that there is actually an uncertainty on their behavior. To do so, let us assume that individuals may make error in the following sense: the policyholder who decides to make effort (i.e., to choose $b = 1$) misbehaves with probability $\eta \in (0, 1)$, and in that case she will choose $b = 1$ with probability $1 - \eta$ and each misconduct $b \in \{2, \ldots, n\}$ with probability $\eta/(n-1)$. A policyholder who decides to make effort is aware that she may inadvertently misbehave.[11]

[10] We consider deterministic legal stipulations although in practice the application of insurance law to concrete cases is often a matter of appreciation which is within the expertise of courts. The principle of good care is a good illustration of the incomplete delimitation of the insurer's leeway when it comes to validating or rejecting claims. As written by Clarke (1997): "Insurers sometimes speak as if the insured has a legal duty to act as a prudent uninsured. However, the insured is obliged to take care to prevent or avoid *any* insured loss at all only if the contract says so in very clear terms... That being so, the same should also be true and, arguably, is true of any duty to minimise the extent or effects of loss that has already occurred or started to occur, i.e. to prevent more loss. A line between the two is hard to draw."

[11] We assume a uniform distribution among the possible misbehaviors for analytical convenience. Other distributions would of course affect the informational contents of signals s and x and thus the conditional probabilites derived below. However, the reader should be readily convinced that other distributions would not change qualitatively our results.

Conditionally upon x, an audit is performed with probability $E[q(s)|x]$, and, in that case, the indemnity payment is $[1-z(x)]I$. Otherwise, the claimant receives the contractual indemnity I. Thus, the policyholder's expected utility when making effort is now written as

$$\begin{aligned} E\overline{u}_1^\star &= (1-\overline{\pi}_1)u(W-P)+\overline{\pi}_1\left\{\int_0^1 E[q(s)|x]u(w-P-L+[1-z(x)]I)\overline{g}_1(x)dx\right. \\ &\qquad \left. +u(w-P-L+I)\int_0^1(1-E[q(s)|x])\overline{g}_1(x)dx\right\} \end{aligned}$$

where

$$\begin{aligned} \overline{\pi}_1 &= (1-\eta)\pi_1+\frac{\eta}{n-1}\sum\nolimits_{h=2}^n \pi_h, \\ \overline{g}_1(x) &= (1-\eta)g_1(x)+\frac{\eta}{n-1}\sum\nolimits_{h=2}^n g_h(x) \text{ for all } x, \end{aligned}$$

and her disutility of effort is

$$\overline{d}_1=(1-\eta)d_1+\frac{\eta}{n-1}\sum\nolimits_{h=2}^n d_h.$$

Similarly, for all $b\geq 2$, the policyholder's expected utility is $Eu_b^\star-d_b$ where

$$\begin{aligned} Eu_b^\star &= (1-\pi_b)u(W-P)+\pi_b\left\{\int_0^1 E[q(s)|x]u(w-P-L+[1-z(x)]I)g_b(x)dx\right. \\ &\qquad \left. +u(w-P-L+I)\int_0^1(1-E[q(s)|x])g_b(x)dx\right\} \end{aligned}$$

Thus, the incentive constraints are now written as

$$E\overline{u}_1^\star-Eu_b^\star\geq \overline{d}_1-d_b \text{ for all } b\geq 2.$$

We have

$$\Pr(b=b_0|x)=\frac{g_{b_0}(x)\eta/(n-1)}{g_1(x)(1-\eta)+\sum_{b=2}^n g_b(x)\eta/(n-1)},$$

for each $b_0\in\{2,\ldots,n\}$, and

$$\Pr(b=1|x)=\frac{g_1(x)(1-\eta)}{g_1(x)(1-\eta)+\sum_{b=2}^n g_b(x)\eta/(n-1)}.$$

Given x, alleging misconduct $b_0\in\{2,\ldots,n\}$ is said to be "credible on the balance of probabilities" if it is more likely that the policyholder had misconduct b_0 or a worse misconduct $b\in\{b_0+1,\ldots,n\}$ than a better behavior $b\in\{1,\ldots,b_0-1\}$, i.e., if[12]

$$\sum\nolimits_{b=b_0}^n \Pr(b|x)\geq\sum\nolimits_{b=1}^{b_0-1}\Pr(b|x),$$

[12] Observe that this legal standard, also called "preponderance of evidence," is generally understood as implying a threshold degree of certainty just above 50%. Other legal standards, like "proof beyond a reasonable doubt" or "clear and convincing evidence," would correspond to different (larger) levels of certainty that could be easily dealt with in our setup without changing our results qualitatively.

or, equivalently if

$$\sum\nolimits_{b=b_0}^{n} \Pr(b|x) \geq \frac{1}{2}.$$

We denote $\hat{b}(x)$ the most serious misconduct that can be credibly alleged when signal x is perceived, i.e.,

$$\sum\nolimits_{b=\hat{b}(x)+1}^{n} \Pr(b|x) < \frac{1}{2} \leq \sum\nolimits_{b=\hat{b}(x)}^{n} \Pr(b|x),$$

with $\hat{b}(x) = 1$ if no misconduct $b_0 \in \{2, \dots, n\}$ is credible. The following lemma shows that larger x allow the insurer to credibly allege more serious misconducts.

Lemma 1 *Function $\hat{b}(\cdot) : [0, 1] \to \mathcal{B}$ is a non-decreasing step function.*

Lemma 1 allows us to write $\hat{b}(x) = b$ if $x_b \leq x < x_{b+1}$, for all $b \in \mathcal{B}$, with $x_1 = 0$ and x_b is given by

$$\sum\nolimits_{b'=b}^{n} \Pr(b'|x_b) = \frac{1}{2},$$

for all $b = 2, ..., m$, where $m \in \mathcal{B}$ is the most severe misconduct that can be confirmed by the judge, i.e., that is credible under the balance of probabilities when x is close to 1. The corresponding maximum indemnity cuts are given by $y_{\hat{b}(x)}$ as illustrated Figure 4 (where $m = 4$).

Figure 4

The insured-insurer-judge interaction is described by the following five-stage game:

- **Stage 1:** The individual takes out the insurance policy I, P. She chooses behavior $b \in \mathcal{B}$ and, should a loss occur, she files a claim. In that case, the insurer observes signal s.

- **Stage 2:** The insurer either directly validates the claim or triggers an audit. In that case, he incurs the audit cost c and he gets the verifiable information x.

- **Stage 3:** If an audit has been performed, the insurer either validates the claim or he alleges that the policyholder misbehaved according to scheme $b \in \{2, \dots, n\}$.

- **Stage 4:** The insured may decide to contest in court the insurer's allegation. The judge confirms the insurer's allegation if $b \leq \hat{b}(x)$, and he dismisses it otherwise.

- **Stage 5:** The indemnity paid to the claimant is I if the claim has been validated by the insurer or if the insurer's allegation b has been dismissed by the judge. Otherwise, the insurer pays an indemnity $(1 - z)I$, with $z \leq y_b$.

A subgame perfect equilibrium of this game is easily characterized. After observing s at stage 1, the insurer triggers an audit at stage 2 if $c \leq E[y_{\hat{b}(x)}|s]I$, since $b = \hat{b}(x)$ is the most severe allegation made at stage 3 that will not be dismissed by the judge at stage 4. Thus, an equilibrium audit strategy is defined by

$$q(s) = \begin{cases} 1 & \text{if } c \leq E[y_{\hat{b}(x)}|s]I, \\ 0 & \text{otherwise.} \end{cases} \tag{8}$$

It is easy to show that

Lemma 2 *The equilibrium audit strategy is a unit step function: $q(s) = 0$ if $s < s^\star$ and $q(s) = 1$ if $s \geq s^\star$, with $s^* \in \mathbb{R} \cup \{-\infty, +\infty\}$.*

Intuition is straightforward: claims should be audited when the signal s is bad enough to be considered as a red flag: it indicates that the circumstances of the loss are likely to be unfavorable (i.e., x is probably large). Cases where $s^* = \pm\infty$ correspond to corner solutions where claims are never (respect. always) audited because c is very large (respect. very low).

Under a competitive insurance market, the policyholder obtains the whole surplus of her relationship with the insurer. Hence, for an optimal insurance law, the insurance contract $\{P, I, z(\cdot)\}$, the strategy $q(\cdot)$, and the insurance law itself $\{y_b, b \in \mathcal{B}\}$ maximize $E\overline{u}_1^\star$ subject to

$$P \geq \pi_1 I \left(1 - \int_0^1 E[q(s)|x]z(x)\overline{g}_1(x)dx\right) + c\int_0^1 E[q(s)|x]\overline{g}_1(x)dx,$$

$$0 \leq I \leq L,$$

$$E\overline{u}_1^\star - Eu_b^\star \geq \overline{d}_1 - d_b \text{ for all } b \geq 2,$$

$$q(s) = \begin{cases} 1 & \text{if } s < s^*, \\ 0 & \text{if } s \geq s^*, \end{cases},$$

$$c = E[y_{\hat{b}(x)}|s^*]I$$

$$z(x) = y_{\hat{b}(x)} \text{ for all } x \in [0,1],$$

where the insurer's break-even constraint includes the expected audit cost c.

Proposition 6 *Assume $\sum_{b=2}^n \phi_b(x) \geq (1-\eta)/\eta(n-1)$ when x is close to 1 and that c is not prohibitively large. Then, an optimal insurance law is such that $y_b > 0$ in a non-empty subset of $\mathcal{B}$.*

Under the condition stated in Proposition 6, if an audit shows that the circumstances of the loss are the most unfavorable (i.e., when x is close to 1), then the insurer is in a position to credibly allege that the policyholder misbehaved in some way (i.e. $b \geq 2$). In other words, his allegation will not be dismissed by the judge. This allows him to cut indemnity at least under the most unfavorable circumstances, which is a desirable feature of the insurance contract, and this will be sustained by an equilibrium strategy where claims are verified with positive probability if the audit cost c is not too large.

Note however that the insurance law limits the indemnity cuts that the insurer is allowed to decide when we have $0 < y_b < 1$. In other words, when his allegation is not dismissed, the insurer may not be allowed to fully cancel the claim. In that sense, insurance law provides leeway to reduce indemnities when some misconduct is credibly alleged, but it does not give carte blanche to the insurer. Bad signal may also be emitted by good faith policyholders (i.e., who exerted the required level of effort) and this is why insurance law should also limit the insurers' opportunism by preventing them to exaggeratedly reduce the coverage.

5 Conclusion

The standard approach to optimal insurance under moral hazard states that insurers should incentivize policyholders by not paying any indemnity when the loss is smaller than a deductible, and by providing full or partial coverage at the margin above the deductible. However, when contracts are complete and the circumstances of an accident are informative of the policyholder's effort, the indemnity should depend at the same time on the financial value of the loss and on its circumstances. If the circumstances of the loss are a sufficient statistic of the policyholder's effort, the deductible only depends on the circumstances and not on the size of the claim. The less favorable the circumstances, the larger the deductible. At one extreme, it may fully cancel the indemnity (under the less favorable circumstances) or vanish at the other, when circumstances are particularly favorable. If the circumstances are not a sufficient statistic of the policyholder's effort, then the optimal contract adds partial coverage at the margin to conditioning on circumstances.

When the insurance policy does not specify the indemnity payment according to all the contingencies that may characterize the claim, insurance contracts are incomplete. Conditioning the indemnity on the circumstances of the loss is nevertheless desirable because it incentivizes the policyholders to exert a high level of effort. This conditioning can be indirectly reached through legal disputes. Legal principles that allow insurers to reduce compensations under unfavorable claim circumstances act as an incentive device, but their application is constrained by the judgment of courts who decide on the balance of probabilities and possibly by upper limits on acceptable indemnity cuts. Hence, framing the ability of the insurer to cut indemnities in the name of legal principles improves the efficiency of insurance contracting under moral hazard. We derived our results just by assuming that courts decide on the balance of probabilities. More realistic ways to model the bargain between insurer,insured and courts could be envisioned, but the same fundamental trade-off between conditioning the indemnity on circumstances to incentivize policyholders and limiting the opportunism of the insurers would remain, and thus similar results would emerge.

References

Bourgeon, J.-M. & Picard, P. (2014). Fraudulent claims and nitpicky insurers. *The American Economic Review*, 104(9), 2900–2917.

Clarke, M. (1997). *The Law of Insurance Contracts*. Informa. Lloyd's of London Press Ltd.

Ehrlich, I. & Becker, G. S. (1972). Market insurance, self-insurance, and self-protection. *Journal of political Economy*, 80(4), 623–648.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 10(1), 79–91.

Milgrom, P. R. (1981). Good news and bad news: Representation theorems and applications. *The Bell Journal of Economics*, 12(2), 380–391.

Winter, R. A. (2013). *Handbook of Insurance*, chapter Optimal Insurance Contracts Under Moral Hazard, (pp. 205–230). Springer New York: New York, NY.

Appendix

A Proof of Proposition 1

Let $\mu_b \geq 0$ and $\lambda \geq 0$ be Kuhn-Tucker multipliers associated with (2) and (3) respectively. Denoting $W(x) \equiv W - P - L + I(x)$, the first-order optimality conditions w.r.t. $I(x)$ and P are written as

$$u'(W(x))\left[1 - \sum_{b=2}^{n} \mu_b \left(\frac{\pi_b}{\pi_1}\phi_b(x) - 1\right)\right] \begin{cases} \leq \lambda & \text{if } \; I(x) = 0, \\ = \lambda & \text{if } \; 0 < I(x) < L, \\ \geq \lambda & \text{if } \; I(x) = L, \end{cases} \tag{9}$$

for all $x \in [0,1]$ and

$$\begin{aligned} \lambda \;\; = \;\; & (1-\pi_1)u'(W-P) + \pi_1 \int_0^1 u'(W(x))g_1(x)dx \\ & + \sum_{b=2}^{n} \mu_b \left[(\pi_b - \pi_1)u'(W-P) + \int_0^1 u'(W(x))[\pi_1 g_1(x) - \pi_b g_b(x)]dx\right] \end{aligned} \tag{10}$$

respectively. Let us first show that there is $b \in \{2, \ldots, n\}$ such that $\mu_b > 0$. Suppose $\mu_b = 0$ for all $b \geq 2$. Then (9) gives

$$\begin{aligned} \lambda \geq u'(W(x)) & \quad \text{if} \;\; I(x) = 0, \\ \lambda = u'(W(x)) & \quad \text{if} \;\; 0 < I(x) < L, \\ \lambda \leq u'(W(x)) & \quad \text{if} \;\; I(x) = L, \end{aligned}$$

As $u'' < 0$, it should be the case that $u'(W(x))$, and thus $I(x)$, is constant for all $x \in [0,1]$. Let $u'^{\star}$ and $I^{\star} > 0$ be these common values, with $u'^{\star} = \lambda$ if $I^{\star} < L$ and $u'^{\star} \geq \lambda$ if $I^{\star} = L$.[13] Suppose $I^{\star} < L$. Then (10) gives

$$\lambda = (1-\pi_1)u'(W-P) + \pi_1 \overline{u}'^{\star} = (1-\pi_1)u'(W-P) + \pi_1 \lambda$$

and thus $u'(W-P) = \lambda = u'(W-P-L+I^{\star})$, which contradicts $I^{\star} < L$. We thus have $I^{\star} = L$, and $Eu_b^{\star} = u(W-P)$ for all $b \geq 1$, which contradicts (2). Consequently, there exists $b \in \{2, \ldots, n\}$ such that $\mu_b > 0$. Using (9), the conditions $I(x) \leq L$ and $I(x) \geq 0$ can be written as

$$u'(W-P)\left[1 - \sum_{b=2}^{n} \mu_b \left(\frac{\pi_b}{\pi_1}\phi_b(\underline{x}) - 1\right)\right] \geq \lambda \;\; \text{if} \;\; \underline{x} \geq 0 \; (= \lambda \text{ if } \underline{x} > 0),$$

[13] We may assume that $I^{\star} > 0$, because we have restricted attention to the case where some degree of insurance coverage is optimal.

and

$$u'(W-P-L)\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\overline{x})-1\right)\right]\leq\lambda \ \text{ if } \ \overline{x}\leq 1 (=\lambda \text{ if } \overline{x}<1)$$

respectively. Hence, as $u''<0$, we get

$$\begin{aligned} u'(W-P)\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\underline{x})-1\right)\right] &\geq u'(W-P-L)\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\overline{x})-1\right)\right] \\ &> u'(W-P)\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\overline{x})-1\right)\right] \end{aligned}$$

and thus

$$1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\underline{x})-1\right)>1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\overline{x})-1\right).$$

which implies $\underline{x}<\overline{x}$ under $\phi_b'(x)>0$. Using (9) also gives

$$u'(W(x))\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(x)-1\right)\right]=\lambda \ \text{ if } \ \underline{x}<x<\overline{x},$$

and differentiating yields

$$I'(x)=\frac{u'(W(x))^2}{u''(W(x))}\sum_{b=2}^{n}\frac{\mu_b\pi_b}{\lambda\pi_1}\phi_b'(x)<0.$$

Assume $\phi_b(x)\to 0$ if $x\to 0$ for all $b\geq 2$, and suppose $\underline{x}=0$. We would have $u'(W(x))\longrightarrow \lambda/(1+\sum_{b=2}^{n}\mu_b)$ when $x\to 0$, and thus $u'(W(x))<\lambda$ and $I(x)=0$ for x close to 0, hence a contradiction with $I'(x)<0$. Thus $\underline{x}>0$ in that case. Assume $\phi_b(x)\to+\infty$ if $x\to 1$ for all $b\geq 2$, and suppose $\overline{x}=1$. We would have

$$1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(x)-1\right)<0$$

for x close to 1, which contradicts $\underline{x}<x<\overline{x}$. Thus $\overline{x}<1$ in that case.

B Proof of Proposition 2

Let $\sigma_1,\sigma_0\in\Sigma$ with $\sigma_1>\sigma_0$, and $X=\{x\in[0,1] \text{ such that } I(x,\sigma_1)=I(x,\sigma_0)\}\neq\emptyset$. Assume that $x\longrightarrow\dfrac{\partial\phi_b(x,\sigma_1)/\partial x}{\partial\phi_b(x,\sigma_0)/\partial x}$ is increasing for all $b\geq 2$. Let $W(x,\sigma)\equiv W-P-L+I(x,\sigma)$. For all x such that $0<I(x,\sigma)<1$, we have

$$u'(W(x,\sigma))\begin{cases}\leq A(x,\sigma) \text{ if } I(x,\sigma)=0,\\ = A(x,\sigma) \text{ if } 0<I(x,\sigma)<L,\\ \geq A(x,\sigma) \text{ if } I(x,\sigma)=L,\end{cases}$$

where

$$A(x,\sigma) = \frac{\lambda(\sigma)}{1 - \sum_{b=2}^{n} \mu_b(\sigma)[\phi_b(x,\sigma) - 1]},$$

and $\lambda(\sigma)$ and $\mu_b(\sigma)$ are Kuhn-Tucker multipliers when insurance indemnity is conditioned on x_σ. Thus, we have

$$I'_x(x,\sigma) = \frac{u'(W(x,\sigma))^2}{u''(W(x,\sigma))} \sum_{b=2}^{n} \frac{\pi_b \mu_b(\sigma)}{\pi_1 \lambda(\sigma)} \frac{\partial \phi_b(x,\sigma)}{\partial x}$$

Let $x^\star \in X$ with $I(x^\star,\sigma_1) = I(x^\star,\sigma_0) \in (0,L)$, which gives

$$\frac{I'_x(x^\star,\sigma_1)}{I'_x(x^*,\sigma_0)} = K(x^\star),$$

where

$$K(x) \equiv \frac{\sum\limits_{b=2}^{n} \pi_b \mu_b(\sigma_1) \partial \phi_b(x,\sigma_1)/\partial x}{\sum\limits_{b=2}^{n} \pi_b \mu_b(\sigma_0) \partial \phi_b(x,\sigma_0)/\partial x}$$

Under the assumption made in the Proposition, we have $K'(x) > 0$. Consider two possible cases.

Case 1: $I'_x(x^\star,\sigma_1) < I'_x(x^\star,\sigma_0)$, and since $I'_x(\cdot) < 0$, $K(x^\star) > 1$.

We have $K(x) > 1$ for all x larger than $x^\star$ and also $I(x,\sigma_1) < I(x,\sigma_0)$ if x is larger than $x^\star$ and close to $x^\star$. Suppose there exists $x^{**} \in X$, $x^{**} > x^\star$ such that $I(x^{**},\sigma_1) = I(x^{**},\sigma_0) > 0$ and $I(x,\sigma_1) < I(x,\sigma_0)$ for all $x \in (x^\star, x^{**})$. We would then have $I'_x(x^{**},\sigma_1) \geq I'_x(x^{**},\sigma_0)$, which gives

$$\frac{I'_x(x^{**},\sigma_1)}{I'_x(x^{**},\sigma_0)} = K(x^{**}) \leq 1,$$

which contradicts $K(x^\star) > 1$ and $x^{**} > x^\star$. Thus, we have $I(x,\sigma_1) \leq I(x,\sigma_0)$ for all $x \geq x^\star$, with a strong inequality if $I(x,\sigma_0) > 0$. The proof is easily adapted if $I(x^{**},\sigma_0) = 0$, with unchanged conclusion.

Symmetrically, we have $I(x,\sigma_1) > I(x,\sigma_0)$ for all x smaller than $x^\star$ and close to $x^\star$. Suppose there exists $x^{**} \in X$, $x^{**} < x^\star$such that $I(x^{**},\sigma_1) = I(x^{**},\sigma_0) \in (0,L)$ and $I(x,\sigma_1) > I(x,\sigma_0)$ for all $x \in (x^{**},x^\star)$. Then, we would have $I'_x(x^{**},\sigma_1) \geq I'_x(x^{**},\sigma_0)$, and thus

$$\frac{I'_x(x^{**},\sigma_1)}{I'_x(x^{**},\sigma_0)} = K(x^{**}) \leq 1.$$

Since $I(0,\sigma_0) = I(0,\sigma_0) = L$, there exists $x^{***} \in [0,x^{**})$ such that $I(x,\sigma_1) < I(x,\sigma_0)$ for all $x \in (x^{***},x^{**})$ and $I(x^{***},\sigma_1) = I(x^{***},\sigma_0)$, which gives $I'_x(x^{***},\sigma_1) \leq I'_x(x^{***},\sigma_0)$. If $I(x^{***},\sigma_0) < L$, we have $I'_x(x^{***},\sigma_1)/I'_x(x^{***},\sigma_0) \geq 1$, and thus

$$K(x^{***}) = \frac{I'_x(x^{***},\sigma_1)}{I'_x(x^{***},\sigma_0)} \geq 1.$$

Hence

$$K(x^{***}) \geq 1 > K(x^{**}),$$

which contradicts $K'(x) > 0$ and $x^{***} < x^{**}$. Thus $I(x, \sigma_1) \geq I(x, \sigma_0)$ for all $x \in [0, x^\star)$. This is obviously also true if $I(x^{***}, \sigma_0) = L$.

Case 2: $I'_x(x^\star, \sigma_1)/I'_x(x^\star, \sigma_0) < 1$, and thus $K(x^\star) < 1$.

In that case, there exists $x^{**} \in (0, x^\star)$ such that $I(x^{**}, \sigma_1) = I(x^{**}, \sigma_0)$ and $I(x, \sigma_1) < I(x, \sigma_0)$ if $x \in (x^{**}, x^\star)$. Suppose $I(x^{**}, \sigma_0) < L$. We have $I'_x(x^{**}, \sigma_1) \leq I'_x(x^{**}, \sigma_0)$ which gives $I'_x(x^{**}, \sigma_1)/I'_x(x^{**}, \sigma_0) \geq 1$. We would have

$$K(x^{**}) = \frac{I'_x(x^{**}, \sigma_1)}{I'_x(x^{**}, \sigma_0)} \geq 1,$$

which contradicts $K(x^\star) < 1$ and $x^{**} < x^\star$. The proof is easily adapted in the case $I(x^{**}, \sigma_0) = L$, with unchanged conclusion.

The proof can be straightforwardly extended to the cases $I(x^\star, \sigma_1) = I(x^\star, \sigma_0) \in \{0, L\}$.

C Proof of Proposition 3

Consider an exclusion-based optimal contract with indemnity $I^\star$ and premium P. Suppose that there exist x_0, x_1, x_2 such that $x_0 < x_1 < x_2$ and $[x_0, x_1] \subseteq \Upsilon$ and $(x_1, x_2] \nsubseteq \Upsilon$. We may assume $G_1(x_1) - G_1(x_0) = G_1(x_2) - G_1(x_1)$ w.l.o.g.. Change the indemnity schedule by substituting $(x_1, x_2]$ to $[x_0, x_1]$ in the exclusion area, and keep the indemnity $I^\star$ and the insurance premium P unchanged. Thus $Eu_1^\star$ is unchanged and the insurer's break-even constraint still holds. Let

$$\Delta u = u(w - P - L + I^\star) - u(w - P - L) > 0.$$

Denoting by $\Delta(Eu_1^\star - Eu_b^\star)$ the changes in the left-hand-side of the incentive constraints, we get

$$\begin{aligned}
\Delta(Eu_1^\star - Eu_b^\star) &= \Delta u \left(\int_{x_1}^{x_2} [\pi_b g_b(x) - \pi_1 g_1(x)] dx - \int_{x_0}^{x_1} [\pi_b g_b(x) - \pi_1 g_1(x)] dx \right) \\
&= \pi_b \Delta u \left(\int_{x_1}^{x_2} \left[\phi_b(x) - \frac{\pi_1}{\pi_b} \right] g_1(x) dx - \int_{x_0}^{x_1} \left[\phi_b(x) - \frac{\pi_1}{\pi_b} \right] g_1(x) dx \right) \\
&> \pi_b \Delta u \left[\phi_b(x_1) - \frac{\pi_1}{\pi_b} \right] \left(\int_{x_1}^{x_2} g_1(x) dx - \int_{x_0}^{x_1} g_1(x) dx \right) \\
&= 0,
\end{aligned}$$

for all $b = 2, \ldots, n$, where the inequality on the third line comes from $\phi'_b(\cdot) > 0$ and the last equality from $G_1(x_1) - G_1(x_0) = G_1(x_2) - G_1(x_1)$. Hence the incentive constraints are satisfied and not-binding after the change, which contradicts the optimality of the solution. Consequently, Υ is defined by a threshold $\hat{x}$ above which an exclusion applies.

For an exclusion-based insurance contract with threshold $\hat{x}$, we may write

$$\begin{aligned} Eu_b^\star &= (1-\pi_b)u(w-P) + \pi_b G_b(\hat{x})u(w-P-L+I) \\ &\quad + \pi_b[1-G_b(\hat{x})]u(w-P-L). \end{aligned}$$

The optimal contract maximizes $Eu_1^\star$ with respect to $I \in [0, L], P$ and $\hat{x} \in [0,1]$, subject to the incentive constraints $Eu_1^\star - Eu_b^\star \geq d_1 - d_b$ for all $b \geq 2$ and the insurer's break-even constraint $P \geq \pi_1 I G_b(\hat{x})$, with multipliers μ_b and λ, respectively. Similarly to the proof of Proposition 1, we have $\lambda > 0, \mu_b \geq 0$ and there exists $b \geq 2$ such that $\mu_b > 0$

We know that $\hat{x} > 0$ and $I > 0$ for otherwise the contract would not provide any coverage. Let $u(1) \equiv u(w-P-L+I)$ and $u(2) \equiv u(w-P-L)$. The first-order optimality condition for $\hat{x}$ is written as

$$[u(1)-u(2)]\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(\hat{x})-1\right)\right] \geq \lambda \ (= \lambda \text{ if } \hat{x} < 1),$$

which is incompatible with $\hat{x} = 1$ if $\phi_b(x) \to +\infty$ for all $b \geq 2$ when $x \to 1$.

D Proof of Proposition 4

When $0 < I(\ell, x) < \ell$, the first-order optimality condition is written as

$$u'(W-P-\ell+I(\ell,x))\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b f_b(\ell,x)}{\pi_1 f_1(\ell,x)}-1\right)\right] = \lambda.$$

Using $f_b(\ell, x)) = f(\ell|x)g_b(x)$ gives

$$u'(W-P-\ell+I(\ell,x))\left[1-\sum_{b=2}^{n}\mu_b\left(\frac{\pi_b}{\pi_1}\phi_b(x)-1\right)\right] = \lambda,$$

with $\lambda > 0$ and $\mu_b \geq 0$, with strong inequality for at least one b.

Using $u'' < 0$ and MLRP yields the characterization of the proposition.

The rest of the proof follows the same line as the proof of Proposition 1.

E Proof of Proposition 5

When $0 < I(\ell, x) < \ell$, the first-order optimality condition is written as

$$u'(W - P - \ell + I(\ell, x)) \left[1 - \sum_{b=2}^{n} \mu_b \left(\frac{\pi_b f_b(\ell, x)}{\pi_1 f_1(\ell, x)} - 1 \right) \right] = \lambda,$$

with $\lambda > 0$ and $\mu_b \geq 0$, with strong inequality for at least one b. Using $u'' < 0$ and MLRP yields the characterization of the proposition.

If $f_b(\ell, x) \equiv f_1(\ell, x)$ for all $b \geq 2$, then the optimal contract is a {straight deductible. Otherwise, using $u'' < 0$ and MLRP yields the characterization of the optimal indemnity schedule provided in the proposition. The rest of the proof follows the same line as the proof of Proposition 4.

F Proof of Lemma 1

Let

$$\begin{aligned} \Phi(b_0, x) &= \sum\nolimits_{b=b_0}^{n} \Pr(b|x) - \sum\nolimits_{b=1}^{b_0-1} \Pr(b|x) \\ &= \frac{\eta}{n-1} \left(\sum\nolimits_{b=b_0}^{n} g_b(x) - \sum\nolimits_{b=1}^{b_0-1} g_b(x) \right) - g_1(x)(1-\eta). \end{aligned}$$

for $b_0 \geq 2$. When $\hat{b}(x) \geq 2$, we have

$$\hat{b}(x) = \sup\{b_0 \in \mathcal{B} | \Phi(b_0, x) \geq 0\}.$$

Let $x' > x$. Using strict MLRP yields

$$\begin{aligned} g_b(x') &> g_b(x) \frac{g_{\hat{b}(x)}(x')}{g_{\hat{b}(x)}(x)} \quad \text{if} \quad b > \hat{b}(x), \\ g_b(x') &< g_b(x) \frac{g_{\hat{b}(x)}(x')}{g_{\hat{b}(x)}(x)} \quad \text{if} \quad b < \hat{b}(x). \end{aligned}$$

Hence, if $b_0 \geq 2, x' > x$, we have

$$\Phi(b_0, x') > \frac{g_{b_0}(x')}{g_{b_0}(x)} \Phi(b_0, x),$$

and thus $\Phi(b_0, x') > 0$ if $\Phi(b_0, x) \geq 0$. We deduce $\hat{b}(x') \geq \hat{b}(x)$ if $\hat{b}(x) \geq 2$ and $x' > x$, which implies that $\hat{b}(x)$ is non-decreasing in $[0, 1]$. It is thus a step function that takes its values in $\mathcal{B}$.

G Proof of Lemma 2

An increase in s shifts the conditional probability distribution of x in the sense of strong FOSD. Furthermore, $y_b \leq y_{b+1}$ for all $b \in \{1, ..., n-1\}$ and $\hat{b}(x)$ is non-decreasing from Lemma 1. Consequently, $y_{\hat{b}(x)}$ is non-decreasing with x and $E[y_{\hat{b}(x)}|s]$ is non-decreasing with s. Hence, $E[y_{\hat{b}(x)}|s]I \geq c$ implies $E[y_{\hat{b}(x)}|s']I \geq c$ if $s' > s$, which proves the Lemma.

H Proof of Proposition 6

Assume first $c = 0$. In that case, $q(s) = 1$ for all s. Let λ and $\mu_b, b \geq 2$, be Lagrange multipliers for the break-even constraint and for the type b incentive constraint, respectively. Let us restrict the set of feasible solutions to the case where $y_b = y \geq 0$ for all $b \geq 2$. Thus we have $z(x) = 0$ if $x < x_2$ and $z(x) = y$ if $x \geq x_2$. We denote $u(1)$ and $u'(1)$ (resp. $u(2)$ and $u'(2)$) the value of the utility function and of its derivative when $x < x_2$ (resp. when $x \geq x_2$). The first-order optimality conditions w.r.t. I and y are written as

$$\begin{aligned}
&\overline{\pi}_1 u'(1)\overline{G}_1(x_2) + \overline{\pi}_1 u'(2)(1-y)[1-\overline{G}_1(x_2)] - \lambda\overline{\pi}_1\{1 - y[1-\overline{G}_1(x_2)]\} \\
&-u'(1)\sum\nolimits_{b'=2}^{n} \mu_{b'}[\pi_{b'}G_{b'}(x_2) - \overline{\pi}_1\overline{G}_1(x_2)] \\
&-u'(2)(1-y)\sum\nolimits_{b'=2}^{n} \mu_{b'}\{\pi_{b'}[1-G_{b'}(x_2)] - \overline{\pi}_1[1-\overline{G}_1(x_2)]\} \\
&\geq 0, = 0 \text{ if } I < L.
\end{aligned} \tag{11}$$

and

$$\begin{aligned}
&\overline{\pi}_1 I[\lambda - u'(2)][1-\overline{G}_1(x_2)] \\
&+u'(2)I\sum\nolimits_{b'=2}^{n} \mu_{b'}\{\pi_{b'}[1-G_{b'}(x_2)] - \pi_1[1-\overline{G}_1(x_2)]\} \\
&\leq 0, = 0 \text{ if } y > 0,
\end{aligned} \tag{12}$$

respectively, where $\overline{G}_1(x)$ is the c.d.f. for density $\overline{g}_1(x)$.

Suppose that $y = 0$ at an optimal solution, which implies $I < L$. We then have $u(1) = u(2) \equiv u, u'(1) = u'(2) \equiv u'$ and (11) and (12) simplify to

$$\overline{\pi}_1(u' - \lambda) - u'\sum\nolimits_{b'=2}^{n} \mu_{b'}(\pi_{b'} - \overline{\pi}_1) = 0, \tag{13}$$

and

$$\pi_1(\lambda - u')[1-\overline{G}_1(x_2)] + u'\sum\nolimits_{b'=2}^{n} \mu_{b'}\{\pi_{b'}[1-G_{b'}(x_2)] - \overline{\pi}_1[1-\overline{G}_1(x_2)]\} \leq 0 \tag{14}$$

respectively. Using (13) to substitue for $\overline{\pi}_1(u' - \lambda)$ in (14) and simplifying give

$$\sum\nolimits_{b'=2}^{n} \mu_{b'} \pi_{b'} [\overline{G}_1(x_2) - G_{b'}(x_2)] \leq 0.$$

Since strong MLRP implies strong FOSD, we have $G_{b'}(x_2) < \overline{G}_1(x_2)$ for all $b' \geq 2$, hence a contradiction.

Thus, when $c = 0$, we have $z(x) > 0$ in a subset of $[0, 1]$ with positive measure. The optimal expected utility of the policyholder varies continuously with c, and thus the previous conclusion remains true when c is not too large.

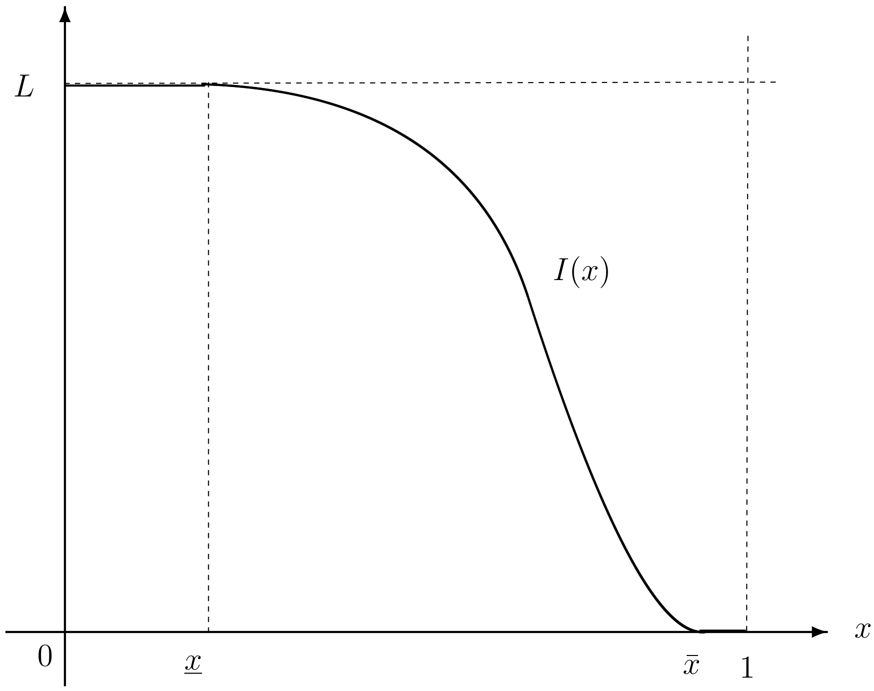

Figure 1: Indemnity schedule

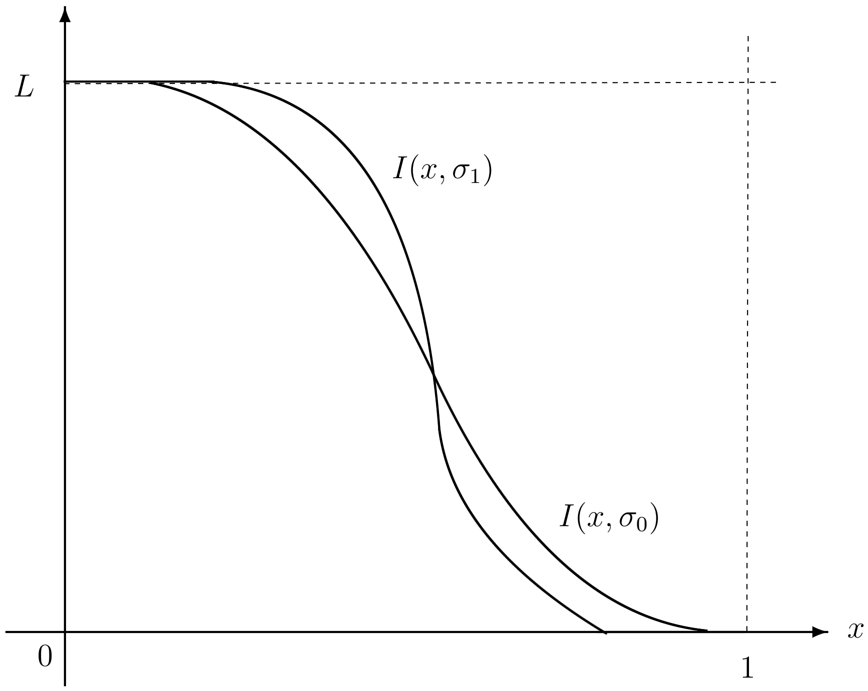

Figure 2: Informativeness of the signal and the indemnity schedule.

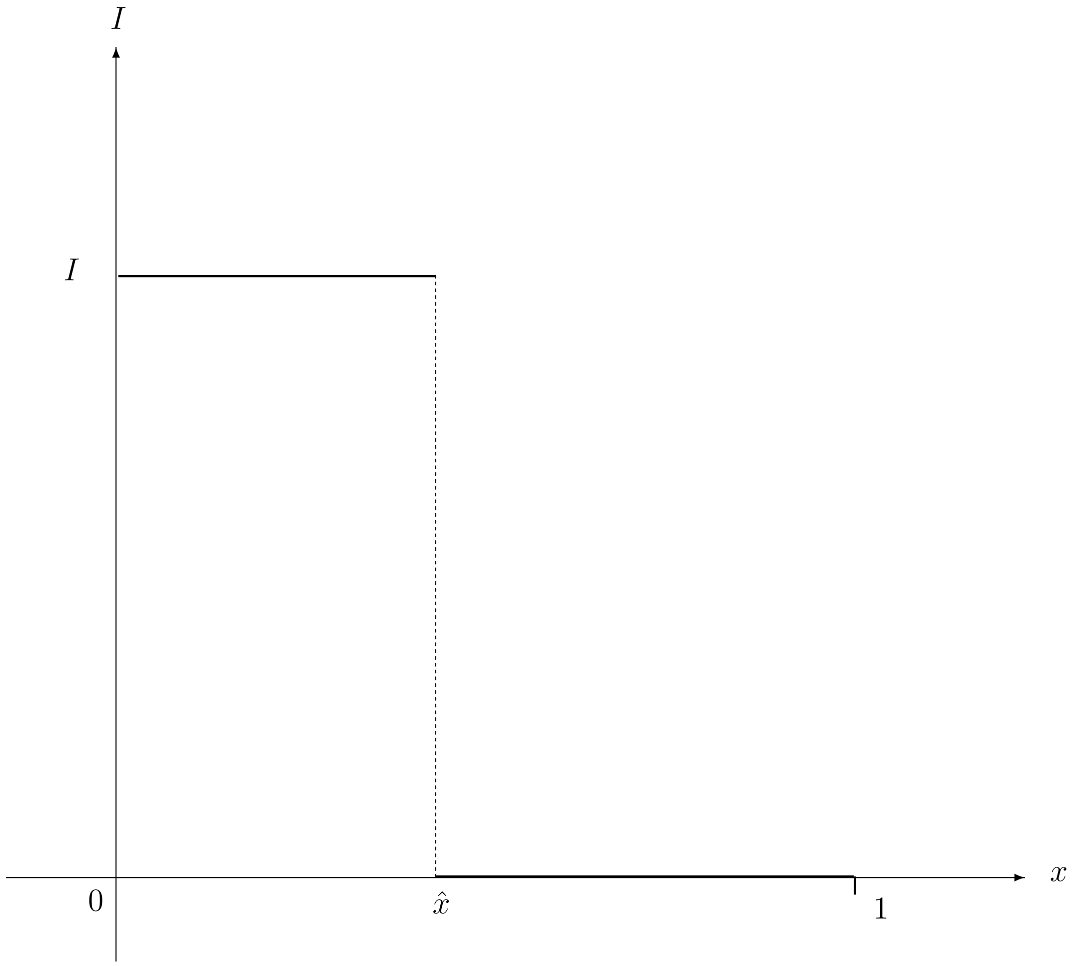

Figure 3: Exclusion-based insurance contract.

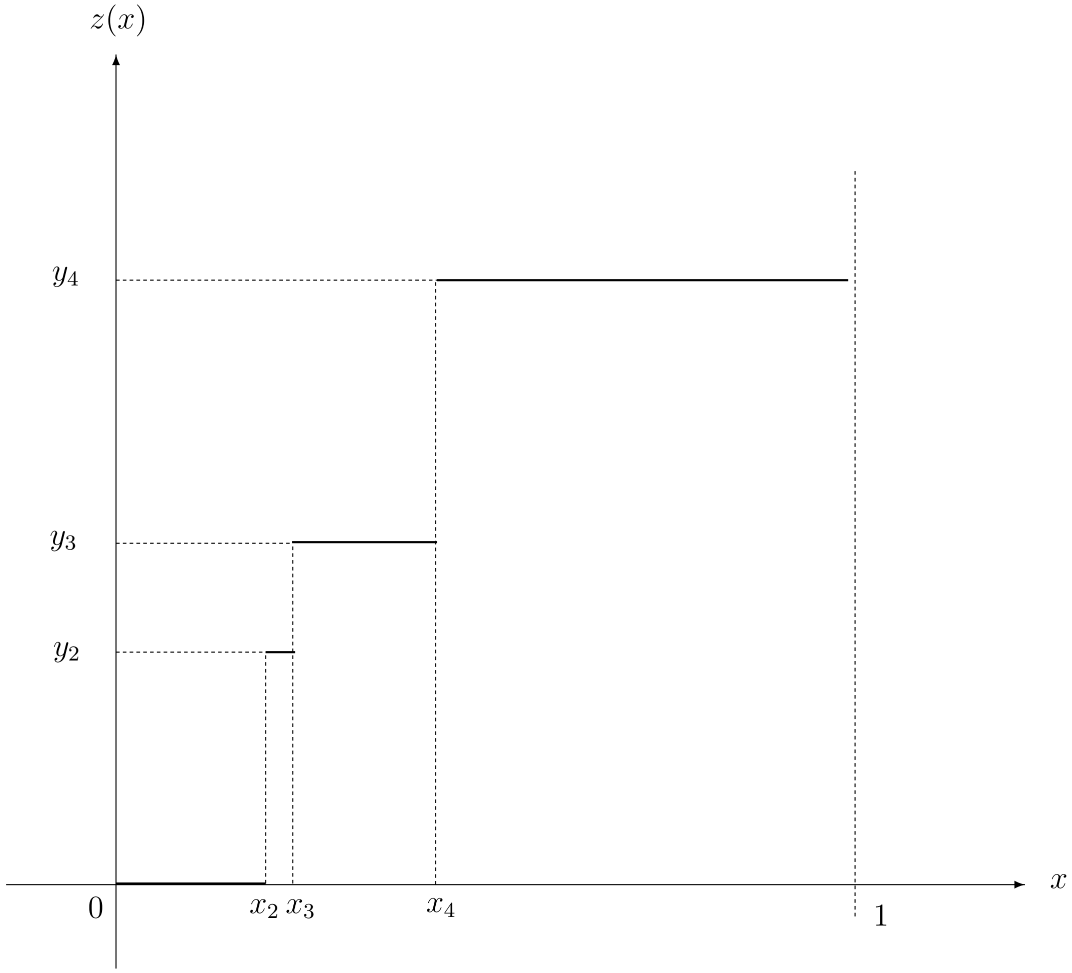

Figure 4: Optimal indemnity and the balance of probability.