Optimal Insurance with Adverse Selection[*]

Hector Chade[†]

Arizona State University
Department of Economics
Tempe, AZ 85287-3806

Edward Schlee[‡]

Arizona State University
Department of Economics
Tempe, AZ 85287-3806

This version: December 2007
(*Preliminary*)

Abstract

We solve the principal-agent problem of a monopolist insurer selling to an agent whose riskiness (probability of an accident) is private information, a problem introduced in Stiglitz (1977)'s seminal paper.

We first derive several important properties of optimal menus for an *arbitrary* type distribution. Besides some standard ones such as no overinsurance and no pooling at the highest type, we show that the optimal premium and indemnity are nonnegative for all types and that the principal always makes positive expected profit. More importantly, we provide a novel comparative static analysis of wealth effects, showing that the principal always prefers an agent facing a larger potential loss, as well as a poorer one if the agent's risk aversion decreases with wealth.

We then specialize to the case with a continuum of types distributed according to an smooth density function. We give simple sufficient conditions for complete sorting, optimal exclusion, and quantity discounts. The main result here is that, under two mild assumptions–the monotone likelihood ratio property for the density and decreasing absolute risk aversion for the agent–the optimal premium is 'S-shaped' in the amount of coverage, first concave, then convex.

We contrast these results with those of the standard monopoly pricing model with quasilinear preferences, and with the competitive insurance model.

---

[*] We have benefited from seminars at the University of Montreal, CEA-Universidad de Chile, University of Miami, and the 2007 Midwest Economic Theory Meetings at Michigan.

[†] Email: `hector.chade@asu.edu`

[‡] Email: `edward.schlee@asu.edu`

1 Introduction

Moral hazard and adverse selection are fundamental problems in insurance provision. The contracting literature has analyzed these problems extensively, both in competitive and monopolistic settings. For moral hazard, the workhorse has been the principal-agent model (including the case of monopoly provision).[1] For adverse selection, variants of the competitive model of Rothschild and Stiglitz (1976) dominate both the theoretical and the empirical literature–despite the lack of consensus on the appropriate model of competition under adverse selection. An exception is Stiglitz (1977), who analyzes a principal-agent version of the problem. He gives a clear graphical analysis of the two-type case, and derives a few properties of optimal insurance policies with a continuum of types. But to date little is known about the optimal monopoly insurance policy, even for the case of a finite number of types or continuum-type case with a continuous density.

We completely solve the problem of a monopolist insurer selling to a risk averse agent who is privately informed about the probability of suffering a loss of a given size. This problem does not fit the standard principal-agent model of a monopolist selling to a privately informed consumer (e.g., Maskin and Riley (1984)) for three reasons: (*i*) the agent's risk aversion implies that there are wealth effects (except for the CARA case); (*ii*) the agent's type enters the principal's objective function directly (common values); and (*iii*) the agent's reservation utility is type dependent. To our knowledge, no one has yet solved for a monopolist's optimal policy in a setting with all these features.

The paper is divided in two parts. In the first, we allow an arbitrary type distribution–requiring neither a finite support nor a continuous density function. Despite the generality, we extend *all* of the known results for the two-type case, and derive several additional properties of any optimal menu of insurance contracts.

Among the results that extend to the general case, we show that the optimal contract exhibits no pooling with the highest type, efficiency (full insurance) for the highest type, and that no type is overinsured. In addition, we show that the optimal premium, indemnity, and indemnity net of premium are nonnegative and co-monotone for all types, that the principal makes positive profit, and that the optimal policy is deterministic.

An important difference with the standard monopoly model (with quasilinear pref-

---

[1] Prescott and Townsend (1984) and Chiappori and Bernardo (2003) are exceptions.

erences) is that the agent's wealth matters. We determine how the agent's wealth and the loss size affects the principal's profit. Using monotone methods, we show that the principal *always* prefers an agent facing a larger loss, as well as a poorer one if the agent exhibits decreasing absolute risk aversion. Under moral hazard, by contrast, stronger assumptions are needed to determine how the agent's wealth affects the principal.

In the second part, we specialize to the case in which there is an interval of types distributed according to a smooth density. Since wealth effects preclude standard techniques used in the quasilinear case, we reformulate the problem to simplify application of optimal control arguments. In this context, we derive conditions for complete sorting of types, exclusion (or inclusion) of types, and for quantity discounts. For example, we show that a modification of the standard increasing hazard rate condition yields complete sorting.

More importantly, we provide a *general* result on the curvature of the optimal premium as a function of the amount of coverage: we show that if the density satisfies the monotone likelihood ratio property and the agent's risk aversion decreases with wealth, then the optimal premium is *S-shaped* in the amount of coverage, first concave, then convex. This curvature property stands in striking contrast with both the standard monopoly model and with competitive insurance models. In particular it rules out global quantity premia, an implication of (many) competitive insurance models (e.g., Rothschild and Stiglitz (1976)).

RELATED LITERATURE. The paper is closely related to three literatures. First, there is the insurance with adverse selection literature started by Rothschild and Stiglitz (1976) for competition, and Stiglitz (1977) for monopoly; each focuses on the two-type case.[2] We completely solve a more general problem than that described by Stiglitz in his seminal paper, and also compare some of the predictions with those from competitive adverse selection models. A large literature has tested implications of the *joint* hypothesis of adverse selection and (some version of) perfect competition in insurance. (Chiappori, Jullien, Salanié, and Salanié (2007) and Cawley and Philipson (1999) are recent examples.) Our results help isolate properties of insurance contracts due *solely* to adverse

---

[2] Biais, Martimort, and Rochet (2000), Section 4, consider a risk neutral monopoly 'market maker' who trades a risky asset with a risk averse investor. Although their model can be viewed as a monopoly insurance model, its properties are quite different from ours, as we explain at the end of Section 2.

selection. Second, there is the literature on principal-agent models with adverse selection, as illustrated by Mussa and Rosen (1978), Maskin and Riley (1984), Matthews and Moore (1987), Jullien (2000), Noldeke and Samuleson (2007), and Hellwig (2006). The already-mentioned special features of the insurance problem–wealth effects, common values, and type-dependent reservation utilities–are absent in Mussa and Rosen (1978), Maskin and Riley (1984), and Matthews and Moore (1987). Jullien (2000) emphasizes the third feature, and Noldeke and Samuleson (2007) allow for the second, but each restricts attention to quasilinear preferences (and to specific features of the solution). Hellwig (2006) derives the standard no-pooling and efficiency-at-the-top results in a general principal-agent problem with wealth effects, using a nontrivial extension of the Maximum Principle. We handle the three features of the problem under general assumptions about the type distribution using fairly elementary arguments. Finally, the paper is related to Thiele and Wambach (1999), who determine how wealth affects the principal's profit in the moral hazard case. In our adverse selection problem we are able to derive comparative static results under much weaker assumptions.

2 The Model

We model the optimal insurance contract design problem as a principal-agent problem with adverse selection. The agent has initial wealth of $w > 0$ and faces a potential loss of $\ell \in (0, w)$ with probability $\theta \in (0, 1)$. The agent's preferences are represented by a strictly increasing and strictly concave von Neumann-Morgenstern utility function $u(\cdot)$ on $\mathbb{R}_+$. The probability θ is private information to the agent.

The principal is risk neutral, with beliefs about the agent's type given by a cumulative distribution function $F(\cdot)$ with support $\Theta \subset (0, 1)$.[3] Let $\underline{\theta}$ and $\overline{\theta}$ be the smallest and largest elements of Θ; by assumption $0 < \underline{\theta} < \overline{\theta} < 1$. This formulation includes of course the discrete case with a finite number of types, the continuum of types case with an atomless density function, and mixtures of both.

The principal chooses, for each $\theta \in \Theta$, a contract (x, t) consisting of a premium t

---

[3] The suppport of a probability measure on the real line (endowed with the Borel σ-field) is the smallest closed set with probability one. Formally, $\Theta = \{\theta \in (0, 1) \mid F(\theta + \varepsilon) - F(\theta - \varepsilon) > 0 \text{ for all } \varepsilon > 0\}$. Although we assume $\Theta \subset (0, 1)$, in some examples we let the support be $[0, 1]$.

and an indemnity payment x in the event of a loss. The expected profit from a contract (x,t) purchased by a type-θ agent is $\pi(x,t,\theta)=t-\theta x$, and the ex-ante expected profit from a (measurable) menu of contracts $(x(\theta),t(\theta))_{\theta\in\Theta}$ is $\int_\Theta \pi(x(\theta),t(\theta),\theta)dF(\theta)$.

The expected utility of type-θ agent for a contract (x,t) is θ is $U(x,t,\theta)=\theta u(w-\ell+x-t)+(1-\theta)u(w-t)$. This function satisfies the following *single crossing property*: for any two *distinct* contracts (x',t') and (x,t) with $(x',t')\geq(x,t)$, and for $\theta'>\theta$, if $U(x',t',\theta)\geq U(x,t,\theta)$ then $U(x',t',\theta')>U(x,t,\theta')$. If $u(\cdot)$ is differentiable, this is equivalent to $-U_x(x,t,\theta)/U_t(x,t,\theta)$ strictly increasing in θ; i.e., indifference curves cross once, with higher types being willing to pay more for a marginal increase in insurance.

By the Revelation Principle, we can restrict attention to (measurable) menus $(x(\theta),t(\theta))_{\theta\in\Theta}$ such that the agent has incentives to participate and to announce her true type θ. Formally, the principal solves the following program:

$$\max_{x(\cdot),t(\cdot)}\int_\Theta \pi\left(x(\theta),t(\theta),\theta\right)dF(\theta)$$

subject to

$$U(x(\theta),t(\theta),\theta)\geq U(0,0,\theta) \qquad \forall\theta\in\Theta \tag{P}$$

$$U(x(\theta),t(\theta),\theta)\geq U(x(\theta'),t(\theta'),\theta) \qquad \forall\theta,\theta'\in\Theta. \tag{IC}$$

It is instructive to compare this problem with the standard monopoly problem (e.g., Maskin and Riley (1984)). In the standard model, the agent is assumed to have quasilinear utility, the agent's type does not directly enter the principal's payoff and the agent's reservation utility does not depend on its type. Here, the agent is risk averse, which implies nontrivial wealth effects (except for the case of CARA preferences); the agent's reservation utility is type dependent, since the expected utility of no insurance is different across types; and the agent's type enters the principal's payoff directly (i.e., it is a common values model). This last point directly leads to an important difference in the two settings: In the standard monopoly model, first-best profit is increasing in type (higher types demand more and the cost of selling to different types is the same); in the standard insurance model, first-best profit profit is not globally increasing in type (even though the demand for insurance at a fixed per-unit price is higher for higher types):

first-best profit equals the risk premium for that type, and the risk premium is concave the the probability of the loss, first increasing, then decreasing.

As mentioned (note 2), Biais, Martimort, and Rochet (2000), Section 4, consider a risk neutral monopoly ‘market maker’ who trades a risky asset to a risk averse investor with CARA utility with private information about both the return of the asset and his endowment; in particular, the investor knows the mean return of the asset, while the market maker just has a prior over the mean. Their risk neutral marker maker can be viewed as an insurer and the investor as an insuree. To make their model directly comparable with ours, suppose that the market maker knows the investor’s endowment. We can interpret the private information about the mean return as information about the *mean* loss in an insurance setting (rather than the *probability* of a loss). First-best profit is again the risk premium for a type. But since they assume CARA utility, first-best profit is *constant* in type: higher types simply have higher mean wealth, which does not affect the risk premium. This difference in the notion of a type leads to very different conditions for separation and curvature of optimal monopoly menus.

3 Arbitrary Distribution of Types: Main Results

3.1 Profitable Changes

Consider a contract that does not involve full insurance. Now change the indemnity in the direction of (but not beyond) full insurance and adjust the premium so that expected utility of a particular type of agent falls. Then the principal’s profit from that type increases. We repeatedly use this result to check for improvements to a given menu.

Lemma 1 (Profitable Changes) *Let $\theta \in \Theta$, and let $|x'' - \ell| < |x' - \ell|$ with $(x'' - \ell)(x' - \ell) \geq 0$. If $U(x'', t'', \theta) \leq U(x', t', \theta)$, then $\pi(x'', t'', \theta) > \pi(x', t', \theta)$.*

Proof. Fix $\theta \in \Theta$. Since the agent is strictly risk averse and $U(x', t', \theta) \geq U(x'', t'', \theta)$, it follows that $t'' - t' > \theta(x'' - x')$; for otherwise the consumption plan generated by the (x'', t'') would second-order stochastically dominate the plan generated by (x', t') and the agent would strictly prefer (x'', t'') to (x', t'). Thus, $t'' - \theta x'' > t' - \theta x'$. □

Intuitively, if a change from a given contract offers additional insurance and yet makes the agent worse off, it must be that the additional insurance is 'actuarially unfair.' But then the change must increase expected profit.

3.2 Properties of Optimal Menus

We now list several important properties of optimal menus. We emphasize that we derive these properties for an *arbitrary* type set $\Theta \subset (0, 1)$ and a *general* cumulative distribution function defined on it.

Theorem 1 (Properties of an Optimal Mechanism) *Any solution* $(x(\theta), t(\theta))_{\theta\in\Theta}$ *to the principal's problem satisfies the following properties:*

(i) (Monotonicity) $x(\theta)$, $t(\theta)$, *and* $x(\theta) - t(\theta)$ *are increasing in* θ*;*

(ii) (No Overinsurance) $x(\theta) \leq \ell$ *for almost all* θ*;*

(iii) (Nonnegativity) $x(\theta)$, $t(\theta)$, *and* $x(\theta) - t(\theta)$ *are nonnegative for almost all* θ*;*

(iv) (Participation) (P) *is binding for the lowest type* $\underline{\theta}$ *;*

(v) (Efficiency at the Top) *Without loss of generality,* $x(\overline{\theta}) = \ell$*;*

(vi) (No Pooling at the Top) *If* u *is differentiable, then* $x(\theta) < \ell$ *for almost all* $\theta < \overline{\theta}$*;*

(vii) (Profitability) *The principal's expected profit is positive.*

Proof. Appendix. □

The proof has many steps, but except for an occasional measure-theoretic detail, each of them is elementary. Result (i) follows immediately from incentive compatibility and the single-crossing property. We prove results (ii) and (iii) by contraposition: we use Lemma 1 to show that if a menu that does not satisfy either property, then there is another feasible menu that increases profit for a positive mass of agents. We prove (v) and (vi) similarly. For (iv), we show that if this property fails, the principal can feasibly reduce the utility of each type in each state by the same amount and hence increase expected profit. Finally, we prove (vii) by showing that there is a pooling contract that is accepted by a positive mass of high enough types with positive expected profit.

Stiglitz derives properties (i), (v), and (vi) for the two-type case. In independent work, Hellwig (2006) proves properties (v) and (vi) for an optimal tax problem with adverse selection, under the assumption that the type distribution is given by a density that may contain atoms. His proofs use first-order conditions from an optimal control problem (where much of the contribution is to extend the first order conditions to allow for discontinuous densities). Besides generalizing these results, Theorem 1 also shows that an optimal menu is nonnegative (property (*ii*)), and that there are always gains to trade (property (*vii*)). Moreover, as just mentioned we derive these properties for a *general* type distribution using fairly elementary arguments, repeatedly using Lemma 1 to find profitable deviations from a menu that fails one of these properties. As we discuss in Section 5, the arguments can be adapted to the standard monopoly pricing model with quasilinear utility (with or without common values). It would be interesting to know if similar arguments can be used in a more general principal-agent problem.

3.3 Comparative Statics of the Principal's Profit

The agent's risk aversion introduces wealth effects absent in the standard screening model with quasilinear utility: here, changing the loss amount or the agent's wealth changes the set of feasible menus. An important question how changes in the agent's wealth endowment affects the principal's profit: Does the principal prefer a richer or poorer agent; one facing a larger or smaller potential loss? In the first-best case (observable types), the answer is immediate: the demand for insurance goes up if an agent who exhibits decreasing absolute risk aversion (DARA) becomes poorer, or if the loss amount becomes larger; in either case profit increases. This intuition, however, is incomplete under adverse selection, since it ignores how the incentive compatibility and participation constraints change with wealth. If the constraint set grew as the agent becomes poorer (under DARA), then the conclusion would again be clear. Unfortunately, the constraint sets cannot be ordered by inclusion as we change the agent's wealth. Despite this difficulty, we answer these questions unambiguously.

Theorem 2 (Comparative Statics) *The principal's maximum profit is*

(i) increasing in the size of the loss ℓ;

(ii) increasing in the agent's risk aversion;

(iii) decreasing in w if the agent's preferences satisfy DARA.

The proof of each part is similar. First, we show that after each change — an increase in ℓ, risk aversion, or w (under DARA) — any menu that is optimal before the change continues to satisfy the *downward* incentive and participation constraints (although some upward incentive constraints may fail). We then show that if the principal simply lets each type choose its best contract from the original menu, then the principal's profit does not fall. Since the resulting menu will satisfy (IC) and (P), it follows that the principal's maximum expected profit cannot fall with the change.

Proof: Since (*iii*) follows from (*ii*), we just prove (*i*) and (*ii*). Fix a menu $(x(\theta), t(\theta))_{\theta\in\Theta}$ satisfying (IC) and (P), with $0 \leq x(\theta) \leq \ell$ for all $\theta \in \Theta$. (Any optimal menu satisfies these conditions almost everywhere.) Note that, by Lemma 1, $t(\theta) - \theta' x(\theta) \geq t(\theta') - \theta' x(\theta')$ if $\theta > \theta'$: i.e., profit increases if a type takes the contract offered to a higher type.

(*i*) THE PRINCIPAL PREFERS A LARGER LOSS SIZE. Let $\ell < \tilde{\ell} < w$, and let $\tilde{U}(x, t, \theta)$ ($U(x, t, \theta)$) be the expected utility of a type-θ agent for contract (x, t) when the loss is $\tilde{\ell}$ (ℓ). Fix $\theta' \in \Theta$ and let (χ, τ) be any nonnegative contract bounded above by $(x(\theta'), t(\theta'))$ and no better than $(x(\theta'), t(\theta'))$ for type θ' when the loss is ℓ:

(*a*) $(0, 0) \leq (\chi, \tau) \leq (x(\theta'), t(\theta'))$; and

(*b*) $U(\chi, \tau, \theta') \leq U(x(\theta'), t(\theta'), \theta')$.

Clearly, these inequalities are met if we set (χ, τ) equal to any *lower* type's contract in the menu $(x(\theta), t(\theta))_{\theta\in\Theta}$, or to the null contract $(0, 0)$.

We now show that $\tilde{U}(x(\theta'), t(\theta'), \theta') \geq \tilde{U}(\chi, \tau, \theta')$. If $(\chi, \tau) = (x(\theta'), t(\theta'))$, there is nothing to prove, so suppose that $(\chi, \tau) \neq (x(\theta'), t(\theta'))$, which, by (*a*) and (*b*), implies that $\chi - \tau < x(\theta') - t(\theta')$. To simplify notation, set $u_\ell(\theta') = u(w - \ell + x(\theta') - t(\theta'))$, $u_n(\theta') = u(w - t(\theta'))$, $u_n = u(w - \tau)$, $u_\ell = u(w - \ell + \chi - \tau)$, $\tilde{u}_\ell(\theta') = u(w - \tilde{\ell} + x(\theta') - t(\theta'))$, and $\tilde{u}_\ell = (w - \tilde{\ell} + \chi - \tau)$. Rewrite the inequality $U(x(\theta'), t(\theta'), \theta') \geq U(\chi, \tau, \theta')$ as

$$u_\ell(\theta') - u_\ell \geq \frac{1 - \theta'}{\theta'} \left(u_n - u_n(\theta')\right). \tag{1}$$

The strict concavity of $u(\cdot)$ and $\chi - \tau < x(\theta') - t(\theta')$ imply that $\tilde{u}_\ell(\theta') - \tilde{u}_\ell > u_\ell(\theta') - u_\ell$ and thus by (1) that $(1-\theta')\left(\tilde{u}_n(\theta') - \tilde{u}_n\right) + \theta'\left(\tilde{u}_\ell(\theta') - \tilde{u}_\ell\right) > 0$, or, equivalently,

$$\tilde{U}(x(\theta'), t(\theta'), \theta') > \tilde{U}(\chi, \tau, \theta'). \tag{2}$$

Since, for each $\theta' \in \Theta$, (2) holds for any $(\chi, \tau) \neq (x(\theta'), t(\theta'))$ satisfying (a) and (b), it follows that $(x(\theta), t(\theta))_{\theta \in \Theta}$ continues to satisfy all the *downward* incentive and participation constraints when the loss equals $\tilde{\ell}$.[4]

Let C be the closure of the set $\{(x(\theta), t(\theta)) \mid \theta \in \Theta\} \cup \{(0,0)\}$. Consider the problem of a type-θ agent choosing a contract in C to maximize expected utility. (A solution exists since C is compact.) Since $(x(\theta), t(\theta))_{\theta \in \Theta}$ satisfies (IC) when the loss is ℓ, the set C is completely ordered by the usual vector inequality $\geq$ on $\mathbb{R}^2$. It follows from (2) that any maximizer of $\tilde{U}(\cdot, \cdot, \theta)$ on C is greater than or equal to $(x(\theta), t(\theta))$.[5] Consider a menu defined by choosing for each $\theta \in \Theta$ any maximizer of $\tilde{U}(\cdot, \cdot, \theta)$. It satisfies (IC) and (P) when the loss is $\tilde{\ell}$, and by Lemma 1 has higher profit than $(x(\theta), t(\theta))_{\theta \in \Theta}$.

(*ii*) THE PRINCIPAL PREFERS A MORE RISK AVERSE AGENT. Let $v(\cdot)$ be more risk averse than $u(\cdot)$: i.e., $v(\cdot) = T(u(\cdot))$, for some strictly increasing and strictly concave function $T(\cdot)$. Denote by $V(x, t, \theta)$ the expected utility of a type-θ agent with von Neumann-Morgenstern utility function $v(\cdot)$.

As before, fix $\theta' \in \Theta$ and let (χ, τ) be any point satisfying (*a*) and (*b*) from the proof of (*i*). We will show that $V(x(\theta'), t(\theta'), \theta') \geq V(\chi, \tau, \theta')$.[6] As in (*i*), we can suppose that $\chi - \tau < x(\theta') - t(\theta')$. Let $u_\ell(\theta')$, $u_n(\theta')$, u_n and u_ℓ be defined as before and set

---

[4] If there were only a finite number of types, the proof would be complete at this point, since (IC) can be replaced by the downward incentive compatibility constraints and monotonicity in that case.

[5] In other words, as the loss increases, the maximizers of expected utility on C *strongly increase* in the sense of Shannon (1995), pp. 215-16. Equation (2) establishes that the expected utility function satisfies the strict single crossing property in (x, t) and ℓ, so the conclusion in this sentence also follows from her Theorem 4.

[6] This inequality follows from Theorem 1 in Jewitt (1987). We nonetheless include its simple proof.

$\Delta T_i = T(u_i(\theta')) - T(u_i)$ and $\Delta u_i = u_i(\theta') - u_i$, for $i = \ell, n$. Then

$$\begin{aligned} V(x(\theta'), t(\theta'), \theta') - V(\chi, \tau, \theta') &= \theta \Delta T_\ell + (1-\theta)\Delta T_n \\ &= \theta \frac{\Delta T_\ell}{\Delta u_\ell} \Delta u_\ell + (1-\theta) \frac{\Delta T_n}{\Delta u_n} \Delta u_n \\ &> \frac{\Delta T_n}{\Delta u_n} (\theta \Delta u_\ell + (1-\theta)\Delta u_n) \\ &\geq 0, \end{aligned}$$

where the first inequality follows from the strict concavity of $T(\cdot)$, and the second from the monotonicity of $T(\cdot)$ and $U(x(\theta'), t(\theta'), \theta) \geq U(\chi, \tau, \theta)$. Hence $(x(\theta), t(\theta))_{\theta \in \Theta}$ satisfies the downward incentive compatibility and participation constraints after the increase in risk aversion. As in the proof of (i), now let each type choose a best contract in C (the closure of the original menu in $\mathbb{R}^2$). Any such menu satisfies (IC) and (P) after the agent becomes more risk averse, and has at least as much profit as $(x(\theta), t(\theta))_{\theta \in \Theta}$. Hence, an increase in risk aversion cannot lower the principal's maximum expected profit. □

It is instructive to contrast Theorem 2 with the analysis of wealth effects in the principal-agent problem with moral hazard (see Thiele and Wambach (1999)). Under moral hazard, a decrease in agent's wealth loosens the incentive constraints, but tightens the participation constraint under DARA. *Stronger* assumptions on preferences are required to conclude that the principal prefers a poorer agent, as in the first best case. Under adverse selection, an decrease in agent's wealth loosens both the downward incentive and the participation constraints under *just* DARA, and we show that the potential tightening of the upward incentive constraints do not lower expected profit.

Note that, if profit and initial wealth or loss size were observable, then Theorem 2 would be testable: it predicts that profit is higher on menus offered to less wealthy agents or to those exposed to higher potential losses.

3.4 A Reformulation

Thus far we have restricted contracts to be deterministic: each type is offered a single premium-indemnity pair. We now justify this restriction. The simplest way to do so is to reformulate the problem by a standard change of variables: instead of choosing a menu of contracts, the principal chooses a menu of state-contingent utilities. We use this reformulation for the rest of the paper.

Given a menu $(x(\theta), t(\theta))_{\theta\in\Theta}$, define, for each $\theta \in \Theta$, $u(\theta)$ and $\Delta(\theta)$ as follows:

$$\begin{aligned} u(\theta) &= u(w - t(\theta)) && (3) \\ \Delta(\theta) &= u(w - t(\theta)) - u(w - \ell + x(\theta) - t(\theta)). && (4) \end{aligned}$$

A menu $(x(\theta), t(\theta))_{\theta\in\Theta}$ uniquely defines a menu $(u(\theta), \Delta(\theta))_{\theta\in\Theta}$. Conversely, given $(u(\theta), \Delta(\theta))_{\theta\in\Theta}$, we can recover $(x(\theta), t(\theta))_{\theta\in\Theta}$ as follows:

$$\begin{aligned} t(\theta) &= w - h(u(\theta)) && (5) \\ x(\theta) &= l - (h(u(\theta)) - h(u(\theta) - \Delta(\theta))), && (6) \end{aligned}$$

where $h = u^{-1}$. Theorem 1 (*i*), (*ii*), and (*iii*) imply that, for any optimal menu, $u(\cdot)$ and $\Delta(\cdot)$ are decreasing in θ, and that, for almost all $\theta \in \Theta$, $0 \leq \Delta(\theta) \leq \Delta_0 = u(w) - u(w - \ell)$.

Thus, an equivalent formulation of the principal's optimal contracting problem is

$$\max_{u(\cdot),\Delta(\cdot)} \int_{\Theta} [w - \theta\ell - (1-\theta)h(u(\theta)) - \theta h(u(\theta) - \Delta(\theta))]dF(\theta)$$

subject to

$$\begin{aligned} u(\theta) - \theta\Delta(\theta) &\geq u(\theta') - \theta\Delta(\theta') && \forall \theta, \theta' \in \Theta \\ u(\theta) - \theta\Delta(\theta) &\geq U(0, 0, \theta) && \forall \theta \in \Theta \\ \Delta(\theta) &\leq \Delta_0 && \forall \theta \in \Theta \\ \Delta(\theta) &\geq 0 && \forall \theta \in \Theta. \end{aligned}$$

In other words, we can think of a menu of contracts as specifying, for each type θ, an amount of utility $u(\theta)$ in the no loss state, and a decrease in utility $\Delta(\theta)$ in case of a loss. In this formulation, the constraints become *linear* in the screening variables, and the objective function is strictly concave in them. This immediately implies that stochastic mechanisms cannot improve upon deterministic ones.[7]

---

[7] Arnott and Stiglitz (1988), Proposition 10, proved a similar result for the two-type case. Our

Proposition 1 (Deterministic Mechanisms) *Any solution to the principal's problem involves a deterministic contract for almost all types.*

Proof. Suppose the principal offers each type θ in a set of positive probability a contract consisting of random variables $(\tilde{x}(\theta), \tilde{t}(\theta))$. In the reformulated problem this implies that the principal offers each type in that same set a contract consisting of random variables $(\tilde{u}(\theta), \tilde{\Delta}(\theta))$. Since the constraints are linear in these variables, any type-θ agent's constraints are satisfied if $\tilde{u}(\theta)$ and $\tilde{\Delta}(\theta)$ are replaced by their expected values. But profit increases from this change since the objective function is strictly concave. $\square$

4 Continuum of Types: The Density Case

We now specialize to the case in which $u(\cdot)$ is C^2, $\Theta = [\underline{\theta}, \overline{\theta}]$, and $F(\cdot)$ is C^1 with density $F'(\cdot) = f(\cdot)$ that is positive and differentiable on $(\underline{\theta}, \overline{\theta})$. Except for the possibility of a zero density at the endpoints, this distribution assumption is the most common one used in contracting problems with adverse selection. Yet, except for Stiglitz (1977), this case has been neglected by the literature on insurance under adverse selection. Under this assumption we can provide strong results for complete sorting of types, exclusion, and for the existence of 'quantity discounts.'

In an abuse of notation, we now let $U(\theta)$ be the expected utility of type $\theta \in \Theta$ for a menu satisfying (IC) and (P). We prove in the Appendix that if $f(\cdot)$ is continuous and positive in the interior of Θ, then $U(\cdot)$ is continuously differentiable. Hence any type's best choice from a feasible menu is unique, and, by the Envelope Theorem (Milgrom and Segal (2002), Theorem 3), we have $U'(\theta) = -\Delta(\theta)$.[8] With this result, we can formulate the principal's problem as a standard optimal control problem.

4.1 The Optimal Control Problem

From now on, let $\dot{U}(\theta)$ denote the derivative of U. Incentive compatibility requires $U(\cdot)$ to be convex, and therefore $\Delta(\cdot)$ to be nonincreasing. Moreover, Theorem 1 (*iv*) yields

---

reformulation allows us to extend this result to any number of types arbitrarily distributed.

[8] It is routine in the contracting literature to appeal to the Envelope Theorem before making any assumptions on the type distribution. But differentiability of the value function can easily fail if there are atoms or if the density function is not positive on the interior of the support.

$U(\underline{\theta}) = U(0,0,\underline{\theta})$. We can now write the principal's problem as the following optimal control problem with (continuous) control variable $\Delta(\cdot)$, (continuously differentiable) state variable $U(\cdot)$, and a free endpoint at $U(\overline{\theta})$:

$$\max_{U(\cdot),\Delta(\cdot)} \int_{\underline{\theta}}^{\overline{\theta}} [w - \theta\ell - (1-\theta)h(U(\theta) + \theta\Delta(\theta)) - \theta h(U(\theta) - (1-\theta)\Delta(\theta))] f(\theta) d\theta$$

subject to

$$\begin{aligned}
\Delta(\cdot) & \quad \text{nonincreasing} & & (7)\\
\Delta(\theta) & \geq 0 \quad & \forall\theta & \quad (8)\\
\Delta(\theta) & \leq \Delta_0 \quad & \forall\theta & \quad (9)\\
\dot{U}(\theta) & = -\Delta(\theta) \quad & \forall\theta & \quad (10)\\
U(\underline{\theta}) & = U(0,0,\underline{\theta}) & & (11)\\
U(\overline{\theta}) & \quad \text{free.} & & (12)
\end{aligned}$$

Consider the 'relaxed problem' that ignores (7)-(9), and let $\lambda(\cdot)$ be the costate variable of the problem.[9] The Hamiltonian is

$$H(U,\Delta,\lambda,\theta) = [w - \theta\ell - (1-\theta)h(U(\theta)+\theta\Delta(\theta)) - \theta h(U(\theta)-(1-\theta)\Delta(\theta))]f(\theta) - \lambda(\theta)\Delta(\theta),$$

and any solution to the relaxed problem must satisfy

$$\begin{aligned}
-\lambda(\theta) &= f(\theta)\theta(1-\theta)[h'(U(\theta)+\theta\Delta(\theta)) - h'(U(\theta)-(1-\theta)\Delta(\theta))] & (13)\\
\dot{\lambda}(\theta) &= f(\theta)[(1-\theta)h'(U(\theta)+\theta\Delta(\theta)) + \theta h'(U(\theta)-(1-\theta)\Delta(\theta))] & (14)\\
\lambda(\overline{\theta}) &= 0, & (15)
\end{aligned}$$

as well as (10) and (11).[10]

Note that $\lambda(\theta) \leq 0$, with strict inequality if $\Delta(\theta) > 0$, and $\dot{\lambda}(\theta) > 0$ for all θ.

---

[9] Because of wealth effects, the objective function is nonlinear in $U(\cdot)$, unlike the standard case with quasilinear preferences. Thus, we cannot appeal to the usual argument that applies Fubini's Theorem to reduce the problem to a single variable in $\Delta(\cdot)$.

[10] The Hamiltonian is strictly concave in (U,Δ), so these conditions are also sufficient for optimality.

Integrating (14) with respect to θ, using (15), and replacing the resulting expression in (13), yields the following optimality condition:

$$f(\theta)\theta(1-\theta)[h'(U(\theta)+\theta\Delta(\theta))-h'(U(\theta)-(1-\theta)\Delta(\theta))]=\int_{\theta}^{\overline{\theta}}a(s)f(s)ds, \qquad (16)$$

where $a(s)=(1-s)h'(U(s)+s\Delta(s))+sh'(U(s)-(1-s)\Delta(s))>0$. It follows from (16) that the omitted constraint (8) is satisfied for all types. Note that, in line with Theorem 1 (*ii*) and (*iv*), $\Delta(\overline{\theta})=0$ and $\Delta(\theta)>0$ for all $\theta<\overline{\theta}$. That is, $\overline{\theta}$ gets full insurance at the optimum and, to ensure incentive compatibility, all other types get partial insurance.[11] Equation (16) illustrates the standard *efficiency vs. information rent* trade-off of contracting problems with adverse selection: the left side is the marginal benefit of providing type θ with additional insurance, i.e., more efficiency, while the right side is the marginal cost of doing so since it leads to an increase in the information rent left to all higher types so as to ensure that incentive compatibility is satisfied.

The Implicit Function Theorem and (16) imply that $\Delta(\cdot)$ is C^1 on Θ (except possibly at the endpoints when f is zero there). Hence in the full program we can replace (7) with $\dot{\Delta}(\theta)\leq 0$ for all $\theta\in\Theta$. We next determine when the solution to the relaxed problem satisfies this constraint.

4.2 Complete Sorting

Differentiate (13) with respect to θ and use (10) to obtain, after some algebra,

$$\dot{\Delta}(\theta)=\frac{\lambda(\theta)[f'(\theta)\theta(1-\theta)+f(\theta)(1-2\theta)]-f(\theta)\theta(1-\theta)\dot{\lambda}(\theta)}{f(\theta)^2\theta^2(1-\theta)^2[\theta h''(U(\theta)+\theta\Delta(\theta))+(1-\theta)h''(U(\theta)-(1-\theta)\Delta(\theta))]}. \qquad (17)$$

Since $h''(\cdot)>0$, the denominator of (17) is positive, and the sign of $\dot{\Delta}(\theta)$ depends on the sign of the numerator. We now derive conditions for *complete sorting* of types at the optimal contract; i.e., for $\dot{\Delta}(\cdot)<0$ everywhere. Notice that, as already proved in

---

[11] This result is obvious if $f(\overline{\theta})>0$. And if $f(\overline{\theta})=0$, then $\Delta(\theta_n)$ tends to zero for any sequence θ_n in Θ tending to $\overline{\theta}$. To see this second point, divide both sides of (13) by $f(\theta_n)$ and use the Mean Value Theorem to write the right side as $\psi(\theta)(1-F(\theta_n))/f(\theta_n)$. The conclusion now follows since $\lim_{\theta\to\overline{\theta}} f(\theta)/(1-F(\theta))=\infty$ (Barlow, Marshall, and Proschan (1963), pp. 377-378).

Theorem 1, there is *no pooling of types at the top*; i.e., $\dot{\Delta}(\overline{\theta}) < 0$.[12]

Lemma 2 (Complete Sorting) *The optimal menu sorts all types who obtain some insurance completely if and only if* $f(\cdot)$ *satisfies, for every* θ,

$$\frac{f'(\theta)}{f(\theta)} \geq \frac{3\theta - 2 - b(\theta)}{\theta(1-\theta)}, \tag{18}$$

where $b(\theta) = h'(U(\theta) - (1-\theta)\Delta(\theta))/[h'(U(\theta) + \theta\Delta(\theta)) - h'(U(\theta) - (1-\theta)\Delta(\theta))]$.

Proof. Appendix. $\square$

The problem with condition (18) is that $b(\cdot)$ is endogenous. But it immediately yields $f'(\theta)/f(\theta) > (3\theta-2)/\theta(1-\theta)$ as a sufficient condition for complete sorting, a fact pointed out by Stiglitz (1977). To improve upon this condition, note that $b(\theta) \geq b^l$ for every θ, where $b^l = h'(u(w-\ell) - (\overline{\theta} - \underline{\theta})\Delta_0)/[h'(u(w) + (\overline{\theta} - \underline{\theta})\Delta_0) - h'(u(w-\ell) - (\overline{\theta} - \underline{\theta})\Delta_0)] > 0$. By Lemma 2, complete sorting follows if $f'(\theta)/f(\theta) > (3\theta - 2 - b^\ell)/\theta(1-\theta)$. Although it depends only on primitives, this condition is difficult to verify and does not even imply what we know from Theorem 1, that there is no pooling at the top.

The next result addresses these issues. Let $\rho(\theta) = \frac{f(\theta)}{1-F(\theta)}$ denote the *hazard rate* of the distribution. We use the following standard terminology: $f(\cdot)$ satisfies the monotone hazard rate condition (MHRC) if $\rho(\cdot)$ is increasing in θ; it satisfies the monotone likelihood ratio property (MLRP) if $f'(\cdot)/f(\cdot)$ is decreasing in θ. As is well-known (and easy to check), the MLRP implies the MHRC.

Theorem 3 (Complete Sorting: Sufficient Conditions) *The optimal menu completely sorts all types who obtain some insurance if*

(i) $\frac{\rho'(\theta)}{\rho(\theta)} > \frac{3\theta-1}{\theta(1-\theta)}$ *for all* $\theta \in [\underline{\theta}, \overline{\theta}]$*;*

(ii) $f(\cdot)$ *satisfies MLRP and either* $\overline{\theta} \leq 1/2$ *or* $f'(\cdot) \geq 0$*; or*

(iii) $f(\cdot)$ *is* C^1, $f'(\cdot)/f(\cdot)$ *is bounded on* Θ*, and* ℓ *is sufficiently small (how small depends on the primitives).*

Proof. The proofs of parts (*i*) and (*ii*) are in the Appendix. To prove part (*iii*), note that $\lim_{\ell \to 0} u(w-\ell) = u(w)$ and thus $\lim_{\ell \to 0} \Delta_0 = 0$. Consequently, $\lim_{\ell \to 0} b^l = \infty$.

---

[12] If $f(\overline{\theta}) > 0$, then (14) and (15) imply that $\dot{\Delta}(\overline{\theta}) < 0$. If $f(\overline{\theta}) = 0$, then $f'(\theta)/f(\theta) \to -\infty$ as $\theta \to \overline{\theta}$, so in either case $\limsup_{\theta \to \overline{\theta}} \dot{\Delta}(\theta) < 0$, implying $\dot{\Delta}(\theta) < 0$ for types near $\overline{\theta}$.

Since the ratio $f'(\cdot)/f(\cdot)$ is bounded, there exists a threshold for the loss, $\hat{\ell} > 0$, such that (18) is satisfied for all types if $\ell \in (0, \hat{\ell})$. $\square$

In the standard contracting model with quasilinear utility and private values, it is well-known that the MHRC implies complete sorting. Condition (i) is a modification of that familiar condition: it is weaker than the MHRC for $\theta < 1/3$, stronger otherwise. In particular, if $\bar{\theta} \leq 1/3$, then the MHRC also implies complete sorting in our model. One can show that MHRC is sufficient for an arbitrary interval of types if $u(\cdot)$ exhibits constant absolute risk aversion (CARA), or if we somehow ignore the presence of common values. The interaction of *both* effects, however, makes it difficult to confirm whether or not the MHRC suffices for complete sorting for an arbitrary interval of types. Part (*ii*) of Theorem 3 shows that the result obtains if we impose the MLRP and either restrict the support of θ or require the density to be nondecreasing. In particular, it is easy to confirm that (*ii*) is satisfied for the class of densities on $[0,1]$ given by $f(\theta) = (1+\alpha)\theta^{\alpha}$, $\alpha \geq 0$, a class which includes the uniform distribution.

Besides providing sufficient conditions using the hazard rate and likelihood ratio, Theorem 3 also shows that if $f(\cdot)$ is C^1 and the likelihood ratio is bounded, then there is a region of losses for which sorting is complete. In this case (7) can be omitted from the problem if we restrict attention to small losses.

4.3 Optimal Exclusion

It immediately follows from (7) and (9) that the set of types that receive some insurance at the optimum is an interval $[\theta_0, \overline{\theta}]$, with $\theta_0 \geq \underline{\theta}$. In this section we prove two results that shed light on the optimal value of θ_0. We first derive a sufficient condition for no type to be excluded from the optimal contract. We then provide a sufficient condition for a subset of low types to be excluded from the optimal contract.

Proposition 2 (No Exclusion) *Suppose that either Theorem 3* (*i*) *or* (*ii*) *holds. If* $f(\underline{\theta})$ *is sufficiently large, then no type is excluded from the optimal menu of contracts.*

Intuitively, Proposition 2 shows that no type is excluded if the presence of low types in the population (density) is significant enough.

Proposition 3 (Exclusion) *Suppose that either Theorem 3 (i) or (ii) holds. There exists a* $k > 0$ *that depends on primitives such that, if* $f(\tilde{\theta})/(1 - F(\tilde{\theta})) < k/\tilde{\theta}(1 - \tilde{\theta})$*, then all types* $\underline{\theta} \leq \theta \leq \tilde{\theta}$ *are excluded from the optimal menu.*

The sufficient condition for exclusion stated in Proposition 3 depends *only* on the primitives of the problem, i.e., $u(\cdot)$, w, ℓ, and $f(\cdot)$, and hence it can be readily checked. In particular, it suggests that a type is excluded if it is close to zero and if its presence in the population (density) is insignificant.

4.4 Curvature and Quantity Discounts

We now turn to the 'curvature' properties of the optimal menu, including quantity discounts. This question is important for several reasons. First, quantity *discounts* are a common form of price discrimination in practice, so it is natural to ask whether a monopolist would use them. Second, in competitive insurance models, quantity *premia* instead of discounts are the rule since equilibrium prices equal marginal cost; we would like to know if this prediction holds for a monopolist.

Let $(x(\theta), t(\theta))_{\theta \in \Theta}$ be an optimal menu. Since the coverage $x(\cdot)$ cannot increase unless the premium $t(\cdot)$ increases, there is an increasing function $T(\cdot)$ on $[x(\underline{\theta}), \ell]$ such that $t(\theta) = T(x(\theta))$ for all $\theta \in [\underline{\theta}, \overline{\theta}]$. We want to know when $T(x)/x$ is nonincreasing. A simpler question is to determine when $T(\cdot)$ is concave and we focus on this simpler property.

To begin, note that $T(\cdot)$ cannot be concave if a positive measure of agents pool at any (x, t) with $x > 0$. For suppose that $x(\theta_0) = x(\theta_1) = \tilde{x} > 0$, with $\theta_1 > \theta_0$. Setting $u_\ell = u(w - \ell + x(\theta) - t(\theta))$ and $u_n = u(w - t(\theta))$, we then have $\dot{T}(\tilde{x}^-) \leq \theta_0 u'_\ell/(\theta_0 u'_\ell + (1 - \theta_0)u'_n) < \theta_1 u'_\ell/(\theta_1 u'_\ell + (1 - \theta_1)u'_n) \leq \dot{T}(\tilde{x}^+)$, implying that $T(\cdot)$ cannot be concave. So we assume for the rest of this section that the optimal menu sorts types completely.

Since we are assuming complete sorting, $x(\cdot)$ is strictly increasing in θ, and thus its inverse, call it $z(\cdot)$, is well defined (i.e., $\theta = z(x)$). We can now represent the optimal mechanism as a *nonlinear premium schedule* $T(x) = t(z(x))$. We now calculate the second derivative of $T(\cdot)$, which we denote by $\ddot{T}(\cdot)$.

Lemma 3 (Curvature) *Let $T(\cdot)$ be an optimal nonlinear premium schedule that completely sorts types. Then $\ddot{T}(x(\theta)) < 0$ if and only if*

$$\frac{f'(\theta)}{f(\theta)} > \frac{3\theta - 2 + c(\theta)}{\theta(1-\theta)}, \tag{19}$$

where $c(\theta) = \theta u_n'' u_\ell'^2 / [\theta u_n'' u_\ell'^2 + (1-\theta) u_\ell'' u_n'^2]$.

Proof. Appendix. □

Since $c(\theta) < 1$ for every θ, it immediately follows from (19) that the optimal menu exhibits quantity discounts if $f'(\theta)/f(\theta) > (3\theta - 1)/\theta(1-\theta)$. For example, this condition holds if $f(\cdot)$ is uniform on $[\underline{\theta}, \overline{\theta}]$ with $\overline{\theta} < 1/3$.

We now characterize the curvature properties of the optimal menu, under the assumptions that the density satisfies the MLRP and that $u(\cdot)$ exhibits DARA.

Theorem 4 (S-shaped Property) *Let $T(\cdot)$ be an optimal nonlinear premium schedule that completely sorts types, and suppose that both the MLRP and DARA hold. Then*

(i) there is a coverage $\hat{x} \in [x(\underline{\theta}), \ell]$ such that T is concave below and convex above $\hat{x}$;

(ii) if f' takes on both positive and negative values, then T is concave on an interval of positive length if $\underline{\theta} < 1/3$ and convex on an interval of positive length if $\overline{\theta} > 2/3$.

Proof. (i) Denote the right side of (19) by $g(\theta)$. We first show that $c'(\theta) \geq 0$ implies that $g'(\theta) > 0$. We have

$$g'(\theta) = \frac{c'(\theta)}{\theta(1-\theta)} + \frac{3\theta^2 - 4\theta + 2 - c(\theta)(1-2\theta)}{\theta^2(1-\theta)^2}.$$

Since $c(\theta) \in (0,1)$, it follows that $3\theta^2 - 4\theta + 2 - c(\theta)(1-2\theta) > 3/2 > 0$. Therefore, $c'(\theta) \geq 0$ implies that $g'(\theta) > 0$.

We have $\dot{T}(x) = \theta u_\ell' / [(1-\theta) u_n' + \theta u_\ell']$ (see the proof of Lemma 3); rearranging yields $(1-\theta) u_n' / \theta u_\ell' = (1/\dot{T}) - 1 > 0$ and use the equality to rewrite $c(\theta)$ as

$$c(\theta) = \frac{1}{1 + \frac{r_\ell}{r_n}\left(\frac{1}{\dot{T}} - 1\right)}, \tag{20}$$

where r_i is the Arrow-Pratt risk aversion measure in state $i = \ell, n$. Hence

$$c'(\theta) = -\Omega \left[\frac{\partial \frac{r_\ell}{r_n}}{\partial \theta} \left(\frac{1}{\dot{T}} - 1 \right) - \frac{r_\ell}{r_n} \frac{\ddot{T}\dot{x}}{\dot{T}^2} \right], \tag{21}$$

where $\Omega = (1 + \frac{r_\ell}{r_n}(\frac{1}{\dot{T}} - 1))^{-2}$. Since the menu is increasing in θ, DARA implies that $\partial \frac{r_\ell}{r_n} / \partial \theta \leq 0$. By (21), if $\ddot{T}(x(\theta_0)) \geq 0$, then $c'(\theta_0) \geq 0$ and so $g'(\theta_0) > 0$. Thus, $g(\cdot)$ crosses the decreasing function $f'(\cdot)/f(\cdot)$ at most once from below, so there is an interval of types $(\hat{\theta}, \overline{\theta}]$ with $T(\cdot)$ is convex on the interval $\{x(\theta) | \theta \in (\hat{\theta}, \overline{\theta}]\}$ and concave otherwise. Setting $\hat{x} = x(\hat{\theta})$ completes the proof that T is S-shaped. (ii) follows from (19) and $c(\theta) \in (0, 1)$. $\square$

Example 1 (Uniform distribution, log utility) *Let the type distribution be uniform on* $[0, 1]$ *and utility be log. Combining Theorems 3 and 4, the optimal menu sorts types completely and is S-shaped. For log utility, it is easy to verify that the function* $c(\cdot)$ *in Lemma 3 is equal to the identity function. By (19),* T *is concave on* $[0, x(\frac{1}{2})]$ *and convex on* $[x(\frac{1}{2}), \ell]$*, and hence exhibits quantity discounts, at least for small coverage levels. If* f *is instead uniform on* $[0, \frac{1}{2}]$*, then* T *is globally concave, and hence exhibits quantity discounts globally.*

Theorem 4 is surprisingly sharp in light of the complications of common values and wealth affects–and the difficulty in coming up with *general* curvature results in principal-agent models with adverse selection. We impose only mild assumptions on the density (MLRP) and preferences (DARA), and thus the theorem covers a large class of cases.

That optimal menus exhibit quantity discounts over some range *distinguishes* the monopoly case from the competitive insurance model of Rothschild and Stiglitz (1976). In that model, the equilibrium premium for a type θ is $t(\theta) = \theta x(\theta)$, where $x(\theta)$ is the equilibrium coverage for type θ. Hence, there are always quantity premia, since $t(\theta)/x(\theta)$ and $x(\theta)$ increase in θ.[13] Theorem 4 implies that quantity premia are not an implication of adverse selection in insurance as such, but from the joint imposition of adverse selection and (some form of) perfect competition.

---

[13] Wilson (1993), pp. 382-84, presents an insurance example in which losses are normally distributed and the type is the mean of the loss distribution. He shows that under Ramsey pricing, the premium is also a convex function of the coverage level.

Remark 1 *Cawley and Philipson (1999) use the implication from Rothschild and Stiglitz (1976) that insurance premia should be convex as a function of the coverage amount to test for the presence of adverse selection in (term) life insurance. They regress the premium on a quadratic function of the coverage amount and find that the coefficient on the squared term is zero, and that the intercept is positive. They conclude that the estimated affine function–which implies quantity discounts–is evidence against adverse selection in life insurance. But by Theorem 4, such a function* could *be consistent with a monopoly insurer facing adverse selection.*[14]

5 Comparison with the Standard Monopoly Model

In their seminal paper on monopoly pricing under adverse selection, Maskin and Riley (1984) assume that type-θ buyer's preferences over quantity-payment pairs (x, t) are of the form $U(t, x, \theta) = v(x, \theta) + t$, where $v(x, \theta) = \int_0^x p(q, \theta)dq$, and $p(\cdot, \theta)$ is decreasing in q and $p(q, \cdot)$ is increasing in θ (p. 172). The cost of selling x units to any consumer is cx, where c is a positive constant. As already mentioned, theirs is a private-values model with no wealth effects on demand, in contrast to our insurance model with common values and wealth effects.

Unlike the analysis of Section 3, Maskin and Riley (1984) restrict attention either to a finite number of types or a continuum of types distributed according to an atomless density. A careful inspection of the proofs of Lemma 1 and Theorem 1 reveals that the arguments can be adapted to their model. That is, each property has an analogue in the standard monopoly model, and each can be derived for a *general* type distribution. For example, in that setting Lemma 1 says that a change in the quantity in the direction of a type's first-best quantity increases the monopolist's profit if the payment rises enough to make the type worse off.[15]

Although the analysis of exclusion and complete sorting is different in the two models,

---

[14] Rothschild and Stiglitz (1976) impose state-independent utility. For life insurance, state-*dependent* utility is surely a more appropriate assumption. And note that if the *only* private information is the dependence of utility on the state (and not the probability of death), then perfect competition generates a linear premium.

[15] Indeed, one can even introduce common values in the monopoly pricing model and adapt these results, so long as the first best quantity is increasing in types.

a far more important difference is the analysis of curvature and quantity discounts. For instance, in Maskin and Riley (1984), there are always quantity discounts for the highest types. That need not be true for a monopoly insurer.

To understand why we do not always get quantity discounts at the top, differentiate $t(\theta)/x(\theta)$ with respect to θ:

$$\frac{d(t(\theta)/x(\theta))}{d\theta} = \frac{\dot{x}(\theta)}{x(\theta)}\left(\frac{\dot{t}(\theta)}{\dot{x}(\theta)} - \frac{t(\theta)}{x(\theta)}\right) \geq -\frac{\dot{x}(\theta)}{x(\theta)^2}\pi(\theta),$$

with equality if and only if $\theta = \overline{\theta}$. The inequality comes from

$$\frac{\dot{t}(\theta)}{\dot{x}(\theta)} = \frac{\theta u'(w-\ell+x-t)}{\theta u'(w-\ell+x-t)+(1-\theta)u'(w-t)} \geq \theta$$

with equality if and only if $\theta = \overline{\theta}$. Hence, if $\dot{x}(\theta)\pi(\theta) < 0$, then revenue per unit, $t(\cdot)/x(\cdot)$, must be rising at θ. Moreover, in a small enough neighborhood including the highest type, $t(\cdot)/x(\cdot)$ is decreasing if and only if profit is positive for the highest type. In Maskin and Riley (1984), profit from every type is nonnegative, and is *always* positive for the highest type, since it has the highest willingness to pay and the marginal cost of supplying to any type is the same. As is evident from even the two-type case, profit from the highest type can easily be negative in an optimal insurance menu.[16] Indeed, if we assume CARA utility to get rid of wealth effects, we can illustrate dramatically how common values affects curvature.

Example 2 (Uniform distribution and CARA) *Let $u(z) = -e^{-rz}$, and rewrite $U(x,t,\theta)$ in the quasilinear form*

$$\tilde{U}(x,t,\theta) = \ln U(x,t,\theta) = w - t - (1/r)\ln(\theta e^{r(\ell-x)} + (1-\theta)).$$

The inverse demand is $\partial\tilde{U}(x,t,\theta)/\partial x = p(x,\theta) = \theta e^{r(\ell-x)}/[\theta e^{r(\ell-x)} + (1-\theta)]$. Maskin

---

[16] To illustrate this possibility in the continuum model, consider a sequence of distributions with common support on $[\underline{\theta},\overline{\theta}]$ converging weakly to a distribution that puts probability one on the midpoint $\hat{\theta} = (\overline{\theta}+\underline{\theta})/2$. In the limit, the optimal menu gives full insurance to type $\hat{\theta}$. If the agent is not too risk averse, profit from the highest type must eventually be negative, since the limiting menu pools all types in $[\hat{\theta},\overline{\theta}]$ at the first-best contract for $\hat{\theta}$–just as happens in the two-type case if the types are far enough apart or the agent is not too risk averse.

and Riley show that the optimal tariff in their private-values model is concave in a neighborhood of $x(\theta')$ if $\partial[p_1(1-F)/p_2 f]/\partial\theta > 0$ at θ' (and convex if the inequality is reversed), where $p_1 = \partial p(x,\theta)/\partial x$ and $p_2 = \partial p(x,\theta)/\partial\theta$. Applying their condition to the inverse demand $\partial\tilde{U}(x,t,\theta)/\partial x$, we get $\ddot{T}(x(\theta)) < 0$ if and only if

$$\frac{f'(\theta)}{f(\theta)} > \frac{1-2\theta}{\theta(1-\theta)} - \frac{f(\theta)}{1-F(\theta)}. \tag{22}$$

Now let $f(\cdot)$ be uniform on $[0,1]$. By (22), the optimal tariff for the private-values case is ***convex*** *for $\theta < 1/3$ and* ***concave*** *for $\theta > 1/3$. For our insurance model, however, the premium is* ***concave*** *for $\theta < 1/3$ and* ***convex*** *for $\theta > 2/3$.*

6 Conclusion

Stiglitz (1977) introduced the monopoly insurance problem analyzed in this paper, and provided a complete solution to the two-type case using an illuminating graphical analysis that has become standard in economic theory textbooks. But despite the importance of adverse selection in insurance and the difficulties encountered in the literature on its competitive provision, the monopoly case has so far received little attention.

Besides extending Stiglitz's results to an arbitrary number of types with an arbitrary distribution, this paper has provided several new results, including a thorough analysis of wealth effects, complete sorting, optimal exclusion, and quantity discounts. Moreover, our arguments give a unified treatment that eases the transition between a finite and a continuum of types, and between the insurance setting and the standard model with quasilinear preferences, often analyzed with different tools.

We have focused on the standard textbook model of monopoly insurance provision under adverse selection. But there are several extensions and variations of the model that deserve further attention. Two that are worth mentioning are the case with unobservable risk aversion (both with and without unobservable riskiness), and the multi-period case, which raises a host of new issues such as renegotiation and experience rating.

A Appendix

A.1 Proof of Theorem 1 (Properties of an Optimal Mechanism)

In the proof of Theorem 1 we will use the following result:

Lemma 4 (Indirect Utility Function) *Let* $(x(\cdot), t(\cdot))$ *be bounded and satisfy (IC), with* $x(\theta) \leq \ell$ *for all* $\theta \in \Theta$*. Then* $U(x(\theta), t(\theta), \theta)$ *is decreasing and continuous in* θ*.*

Proof. Let $\theta' > \theta$. We have $U(x(\theta), t(\theta), \theta) \geq U(x(\theta'), t(\theta'), \theta) \geq U(x(\theta'), t(\theta'), \theta')$, where the first inequality follows from (IC) and the second from $x(\theta') \leq \ell$. Hence $U(x(\theta), t(\theta), \theta)$ is decreasing in θ.

Monotonicity implies that the left and right limits exist at any $\theta \in \Theta$. Let $\theta' \in \Theta$. We will show that the left and right limits of $U(x(\theta), t(\theta), \theta)$ are equal at $\theta = \theta'$. Consider any sequence θ_n approaching θ' from below and let $t_n = t(\theta_n)$ and $x_n = x(\theta_n)$. We have

$$0 \geq U(x(\theta'), t(\theta'), \theta') - U(x_n, t_n, \theta_n) \geq U(x_n, t_n, \theta') - U(x_n, t_n, \theta_n), \tag{23}$$

where the first inequality follows from the monotonicity in θ, and the second from (IC). But $U(x_n, t_n, \theta') - U(x_n, t_n, \theta_n) = (\theta' - \theta_n)(u(w - \ell - t_n + x_n) - u(w - t_n))$. Since $(x(\cdot), t(\cdot))$ is bounded and $u(\cdot)$ continuous, $U(x_n, t_n, \theta') - U(x_n, t_n, \theta_n)$ tends to 0 as θ_n tends to θ', so by (23), $U(x(\theta), t(\theta), \theta)$ is left-continuous at $\theta = \theta'$.

Now consider any sequence θ_n approaching θ' from above. For every n,

$$0 \geq U(x_n, t_n, \theta_n) - U(x(\theta'), t(\theta'), \theta') \geq U(t', x', \theta_n) - U(x(\theta'), t(\theta'), \theta'),$$

where the first inequality follows from monotonicity in θ and the second from (IC). But again $U(x', t', \theta_n) - U(x(\theta'), t(\theta'), \theta')$ tends to zero as θ_n tends to θ', so $U(x_n, t_n, \theta_n)$ converges to $U(x(\theta'), t(\theta'), \theta')$, proving that $U(x(\theta), t(\theta), \theta)$ is left-continuous at $\theta = \theta'$. Since θ' was arbitrary, it follows that $U(x(\theta), t(\theta), \theta)$ is continuous in θ. $\square$

We now prove Theorem 1 in several steps, illustrating the formal arguments with pictures. Most of the proofs are by contraposition: we show that if the property fails for a menu, then there is another feasible menu with higher profit (relying on Lemma 1).

(*i*) PREMIUM, INDEMNITY, AND NET INDEMNITY ARE CO-MONOTONE. Wlog, let $\theta' > \theta$. From (IC), $U(x(\theta), t(\theta), \theta) \geq U(x(\theta'), t(\theta'), \theta)$ and $U(x(\theta'), t(\theta'), \theta') \geq U(x(\theta), t(\theta), \theta')$, so that either (*a*) $x(\theta') \geq x(\theta)$ and $t(\theta') \geq t(\theta)$, or (*b*) $x(\theta') \leq x(\theta)$ and $t(\theta') \leq t(\theta)$. Unless the contracts are the same, case (*b*) is ruled out by the single crossing property. Thus, $x(\theta') \geq x(\theta)$ and $t(\theta') \geq t(\theta)$. But then $x(\theta') - t(\theta') \geq x(\theta) - t(\theta)$; otherwise, $U(x(\theta'), t(\theta'), \theta') < U(x(\theta), t(\theta), \theta')$ as $u(w - t(\theta)) \geq u(w - t(\theta')$ and $u(w - \ell + x(\theta) - t(\theta)) > u(w - \ell + x(\theta') - t(\theta'))$, violating (IC) for θ'.

(*ii*) NO OVERINSURANCE. Suppose that $x(\theta) > \ell$ on a set of types with positive measure. By monotonicity, every type above $\theta^l = \inf\{\theta | x(\theta) > \ell\}$ must have an indemnity greater than the loss. There are two cases.

(*a*) $\theta^l \notin \{\theta | x(\theta) > \ell\}$: Then $x(\theta^l) \leq \ell$. For $\theta \geq \theta^l$, let $\tau(\theta)$ be the unique solution to

$$U(\ell, \tau(\theta), \theta) \;\; = \;\; U(x(\theta), t(\theta), \theta).$$

Define $\tau^* = \sup \tau(\theta)$, where the supremum is taken over types $\theta \geq \theta^l$. Then pooling all types $\theta > \theta^l$ at (ℓ, τ^*) (and leaving the contract for all other types the same) increases profit by Lemma 1, and the modified menu satisfies (IC) and (P).

(*b*) $\theta^l \in \{\theta | x(\theta) > \ell\}$: In this case $x(\theta^l) > \ell$. For each θ with $x(\theta) \leq \ell$, let $\tilde{t}(\theta)$ be the unique solution to

$$U(\ell, \tilde{t}(\theta), \theta) \;\; = \;\; U(x(\theta), t(\theta), \theta).$$

Define $t^l = \sup \tilde{t}(\theta)$, where the supremum is taken over types θ with $x(\theta) \leq \ell$. Pool all types $\theta \geq \theta^l$ at the contract (ℓ, t^l), and leave the contract for all other types unchanged. This menu increases profit by Lemma 1, and it satisfies (IC) and (P).

Insert Figure Here

(*iii*) PREMIUM, INDEMNITY, AND NET INDEMNITY ARE NONNEGATIVE. It is enough to show that $t(\theta) \geq 0$ for almost all θ, for then (P) implies $x(\theta) \geq 0$ and $x(\theta) - t(\theta) \geq 0$ for almost all θ. Suppose that $t(\theta) < 0$ on a set of types with positive measure. By monotonicity, it contains all types below $\theta^s = \sup\{\theta | t(\theta) < 0\}$. There are two cases.

(a) $\theta^s \notin \{\theta | t(\theta) < 0\}$: In this case $t(\theta^s) \geq 0$. For each θ such that $t(\theta) < 0$, let $(\hat{x}(\theta), \hat{t}(\theta))$ be the unique solution to

$$\begin{aligned} U(\hat{x}(\theta), \hat{t}(\theta), \theta) &= U(x(\theta), t(\theta), \theta) \\ U(\hat{x}(\theta), \hat{t}(\theta), \theta^s) &= U(x(\theta^s), t(\theta^s), \theta^s), \end{aligned}$$

and let $(x^s, t^s) = \sup(\hat{x}(\theta), \hat{t}(\theta))$, where the supremum is taken over types θ with $t(\theta) < 0$. Also, let $\tilde{x}$ be the unique solution to

$$U(\tilde{x}, 0, \theta^s) = U(x(\theta^s), t(\theta^s), \theta^s).$$

If $t^s \geq 0$, then pool all types $\theta < \theta^s$ at (x^s, t^s), and if $t^s < 0$, do so at $(\tilde{x}, 0)$. In either case, profit increases by Lemma 1 and *(ii)*, and the new menu satisfies (IC) and (P).

(b) $\theta^s \in \{\theta | t(\theta) < 0\}$: In this case $t(\theta^s) < 0$. For each θ such that $t(\theta) \geq 0$, now let $\hat{x}(\theta)$ be the unique solution to

$$U(\hat{x}(\theta), 0, \theta) = U(x(\theta), t(\theta), \theta),$$

and let $x^l = \inf x(\theta)$, where the infimum is taken over types θ with $t(\theta) \geq 0$. Also, let $\tilde{x}$ be the unique solution to

$$U(\tilde{x}, 0, \theta^s) = U(x(\theta^s), t(\theta^s), \theta^s).$$

Set $x^m = \min\{x^l, \tilde{x}\}$. Pool all types $\theta \leq \theta^s$ at $(x^m, 0)$. By Lemma 1 and *(ii)*, profit increases, and the modified menu satisfies (IC) and (P).

Insert Figure Here

(iv) PARTICIPATION BINDS AT THE BOTTOM. Suppose that $U(x(\underline{\theta}), t(\underline{\theta}), \underline{\theta}) - U(0, 0, \underline{\theta}) = K > 0$. Since expected utility is continuous in θ, there exists a nonnegative real number ε such that $[\underline{\theta}, \underline{\theta} + \varepsilon]$ has positive measure and for all $\theta \in [\underline{\theta}, \underline{\theta} + \varepsilon] \cap \Theta$

$$U(x(\underline{\theta}), t(\underline{\theta}), \theta) - U(0, 0, \theta) > 0.$$

(If $\underline{\theta}$ has positive measure, then ε can be zero.) It then follows from (IC) that

$$U(x(\theta), t(\theta), \theta) - U(0,0,\theta) \geq U(x(\underline{\theta}), t(\underline{\theta}), \theta) - U(0,0,\theta)$$

for all $\theta \in [\underline{\theta}, \underline{\theta}+\varepsilon] \cap \Theta$. But since the difference on the right side of this last inequality tends to K as θ tends to $\underline{\theta}$, there is a nonnegative real number η such that $[\underline{\theta}, \underline{\theta}+\eta]$ has positive measure and, for all $\theta \in [\underline{\theta}, \underline{\theta}+\eta] \cap \Theta$

$$U(x(\theta), t(\theta), \theta) - U(0,0,\theta) \geq K/2 > 0.$$

By (iii), there is a $\theta' \in [\underline{\theta}, \underline{\theta}+\eta] \cap \Theta$ with $(x(\theta') - t(\theta'), t(\theta')) \geq 0$, so $U(x(\theta'), t(\theta'), \theta) - U(0,0,\theta)$ is increasing in θ. Combine this fact with (IC) to get that

$$U(x(\theta), t(\theta), \theta) - U(0,0,\theta) \geq U(x(\theta'), t(\theta'), \theta) - U(0,0,\theta) \geq K/2 > 0$$

for all $\theta \in (\underline{\theta}+\eta, \bar{\theta}] \cap \Theta$, so $U(x(\theta), t(\theta), \theta) - U(0,0,\theta) \geq K/2 > 0$ for all $\theta \in \Theta$. Now change the menu to reduce the utility of each type by $K/2$ in each state. Both (P) and (IC) continue to hold, but the principal's profit increases.

Insert Figure Here

(v) Full Insurance at the Top. We show that the essential supremum of $x(\cdot)$ is ℓ. Suppose that $x(\bar{\theta}) < \ell$. We will show that profit rises if all 'sufficiently' high types are pooled at full insurance.

For every $\varepsilon \in (0, \bar{\theta} - \underline{\theta})$, the set $[\bar{\theta}, \bar{\theta} - \varepsilon] \cap \Theta$ has positive probability. For such an ε, define θ_ε to be the smallest number in $[\bar{\theta}, \bar{\theta} - \varepsilon] \cap \Theta$. Define τ_ε to be the largest premium that the insurer could charge a type-θ_ε for full insurance that leaves type-θ_ε at least as well off as with $(x(\theta_\varepsilon), t(\theta_\varepsilon))$; i.e., τ_ε is the unique solution to

$$U(\ell, \tau_\varepsilon, \theta_\varepsilon) = U(x(\theta_\varepsilon), t(\theta_\varepsilon), \theta_\varepsilon).$$

It follows by Lemma 4 that $\lim_{\varepsilon \to 0} \tau_\varepsilon = \tau_0$, which solves $U(\ell, \tau_0, \bar{\theta}) = U(x(\bar{\theta}), t(\bar{\theta}), \bar{\theta})$.

We now show that there exists an $\varepsilon \in (0, \bar{\theta} - \underline{\theta})$ such that $\tau_\varepsilon - \theta\ell > t(\theta) - \theta x(\theta)$ for all $\theta \in [\bar{\theta} - \varepsilon, \bar{\theta}] \cap \Theta$. Suppose to the contrary that, for all $\varepsilon \in (0, \bar{\theta} - \underline{\theta})$, there is a type

$\theta(\varepsilon) \in [\bar{\theta} - \varepsilon, \bar{\theta}] \cap \Theta$ such that

$$t(\theta(\varepsilon)) - \theta(\varepsilon)x(\theta(\varepsilon)) \geq \tau_\varepsilon - \theta(\varepsilon)\ell. \tag{24}$$

By Lemma 1, monotonicity, and $x < \ell$, we have that

$$\tau_0 - \theta(\varepsilon)\ell > t(\bar{\theta}) - \theta(\varepsilon)x(\bar{\theta}) \geq t(\theta(\varepsilon)) - \theta(\varepsilon)x(\theta(\varepsilon)), \tag{25}$$

for the principal makes more profit if $\theta(\varepsilon)$ gets the contract offered to type $\bar{\theta}$, and even more profit if $\theta(\varepsilon)$ gets (τ_0, ℓ). Letting ε go to zero in (24) and (25) yields $\tau_0 - \bar{\theta}\ell > \tau_0 - \bar{\theta}\ell$, a contradiction. So for some $\varepsilon \in (0, \bar{\theta} - \underline{\theta})$, $\tau_\varepsilon - \theta\ell > t(\theta) - \theta x(\theta)$ for all $\theta \in [\bar{\theta} - \varepsilon, \bar{\theta}] \cap \Theta$.

Fix such an ε and consider the following alternative menu $(\hat{x}(\theta), \hat{t}(\theta))_{\theta \in \Theta}$: for $\theta \in [\bar{\theta} - \varepsilon, \bar{\theta}] \cap \Theta$, set $\hat{x}(\theta) = \ell$ and $\hat{t}(\theta) = \tau_\varepsilon$; otherwise set $(\hat{x}(\theta), \hat{t}(\theta)) = (x(\theta), t(\theta))$. By construction, expected profit is higher for $(\hat{x}(\theta), \hat{t}(\theta))_{\theta \in \Theta}$ than for $(x(\theta), t(\theta))_{\theta \in \Theta}$, and the alternative menu also satisfies (IC) and (P).

(vi) No Pooling at the Top. Suppose that $\{\theta \in \Theta - \{\bar{\theta}\} | x(\theta) = \ell\}$ is of positive measure. Let θ^l be the infimum of this set. By (IC), any agent with full insurance is charged the same premium, call it τ. By monotonicity and $x \leq \ell$, $x(\theta) < \ell$ if $\theta < \theta^l$ and $x(\theta) = \ell$ if $\theta > \theta^l$. By Lemma 4, $\lim_{\theta \to \theta^l} U(x(\theta), t(\theta), \theta) = U(\ell, \tau, \theta^l)$. Wlog, we can set $x(\theta^l) = \ell$ and $t(\theta^l) = \tau$, for expected profit does not fall and (IC) and (P) still hold.

Fix $\varepsilon \in (0, \bar{\theta} - \theta^l)$ for the rest of the proof. Let $\tau(\varepsilon)$ be the largest premium that satisfies (P) for type $\theta^l + \varepsilon$ when given full insurance; i.e., $U(\ell, \tau(\varepsilon), \theta^l + \varepsilon) = U(0, 0, \theta^l + \varepsilon)$. Define, for each $\delta \in (0, \tau(\varepsilon))$, a contract (x_δ, t_δ) as the unique solution to

$$\begin{aligned} U(\ell, \tau, \theta^l) &= U(x_\delta, t_\delta, \theta^l) \\ U(\ell, \tau + \delta, \theta^l + \varepsilon) &= U(x_\delta, t_\delta, \theta^l + \varepsilon) \end{aligned}$$

Define a menu of contracts $(\hat{x}(\theta), \hat{t}(\theta))_{\theta \in \Theta}$ as follows: $(\hat{x}(\theta), \hat{t}(\theta)) = (\ell, \tau + \delta)$ if $\theta \in \Theta \cap (\theta^l + \varepsilon, \bar{\theta}]$; $(\hat{x}(\theta), \hat{t}(\theta)) = (x_\delta, t_\delta)$ if $\theta \in \Theta \cap [\theta^l, \theta^l + \varepsilon]$; and $(\hat{x}(\theta), \hat{t}(\theta)) = (\tilde{x}_\delta(\theta), \tilde{t}_\delta(\theta))$ if $\theta \in \Theta \cap [\underline{\theta}, \theta^l)$, where $(\tilde{x}_\delta(\theta), \tilde{t}_\delta(\theta))$ is equal to the best of the two contracts $(x(\theta), t(\theta))$ or (x_δ, t_δ). The menu $(\hat{x}(\theta), \hat{t}(\theta))_{\theta \in \Theta}$ satisfies (IC) and (P) for every $\delta \in (0, \tau(\varepsilon))$, and its expected profit is

$$\int_{(\theta^l+\varepsilon,\overline{\theta}]}[\tau+\delta-\theta\ell]dF(\theta)+\int_{[\theta^l,\theta^l+\varepsilon]}[t_\delta-\theta x_\delta]dF(\theta)+\int_{[\underline{\theta},\theta^l)}[\tilde{t}_\delta(\theta)-\theta\tilde{x}_\delta(\theta)]dF(\theta). \quad (26)$$

We now show that each of the three expressions is differentiable in δ at $\delta=0$ and that the derivative of the sum is positive. The first term is differentiable and the derivative at $\delta=0$ is $\int_{(\theta^l+\varepsilon,\overline{\theta}]}dF(\theta)>0$. Next, note that, since u is C^1, x_δ and t_δ are differentiable in δ for every $\delta\in[0,\tau)$, $t'_\delta-\theta x'_\delta$ is bounded on $[0,\tau)\times\Theta$, and $t'_0-\theta x'_0=(-\theta^l+\theta)/\varepsilon\geq 0$ for $\theta\in[\theta^l,\theta^l+\varepsilon]$. Thus the derivative of the second integral in (26) with respect to δ exists at $\delta=0$ and equals $\int_{[\theta^l,\theta^l+\varepsilon]}\frac{\theta-\theta^l}{\varepsilon}dF(\theta)\geq 0$. Finally, consider the third integral in (26). The integrand is differentiable in δ at $\delta=0$ and the value of the derivative is zero for every $\theta\in\Theta\cap[\underline{\theta},\theta^l)$. Moreover, since x_δ and t_δ are decreasing in δ, it follows that, for each $(\delta,\theta)\in(0,\tau(\varepsilon))\times\big(\Theta\cap[\underline{\theta},\theta^l)\big)$, $|\frac{\tilde{x}_\delta(\theta)-x(\theta)}{\delta}|\leq|\frac{x_\delta-x_0}{\delta}|$ and $|\frac{\tilde{t}_\delta(\theta)-t(\theta)}{\delta}|\leq|\frac{t_\delta-t_0}{\delta}|$, which shows that the derivatives of x_δ and t_δ are bounded on $[0,w-\tau]$. Hence, the third integral is differentiable at $\delta=0$ and the derivative equals zero.

Insert Figure Here

(*vii*) The Principal Makes Positive Profit. Let $\varepsilon>0$, and consider a menu in which each type chooses either $(0,0)$ or $(\ell,\overline{\theta}\ell+\varepsilon)$ to maximize expected utility. By construction, $\overline{\theta}\ell>\theta\ell$, for any $\theta<\overline{\theta}$, so expected profit is positive from any type less than $\bar{\theta}$ who chooses $(\ell,\overline{\theta}\ell+\varepsilon)$. For a small enough $\varepsilon>0$, a positive measure of types in $[\underline{\theta},\bar{\theta})$ will choose $(\ell,\overline{\theta}\ell+\varepsilon)$, so expected profit of this menu is positive. □

A.2 Proof of Continuous Differentiability of $U(\theta)$

We will show that if $f(\cdot)$ is continuous on θ and positive in its interior, then $U(\cdot)$ is C^1 on $(\underline{\theta},\overline{\theta})$. We proceed in three steps.

No Indifference at an Optimal Menu. We first show that if $(x(\theta),t(\theta))_{\theta\in\Theta}$ is an optimal menu, then each type strictly prefers his contract to any other one. To prove this, suppose that there are types θ', θ'' in Θ with $(x(\theta''),t(\theta''))\neq(x(\theta'),t(\theta'))$ and $U(\theta'')=U(x(\theta'),t(\theta'),\theta'')$. We will show that the set of types between θ' and θ'' is of zero measure, which implies that $\theta'=\theta''$ as any interval has positive measure.

Towards a contradiction, suppose instead that this set is of positive measure.

There are two cases to consider: (*i*) $\theta'' > \theta'$; and (*ii*) $\theta' > \theta''$.

Consider case (*i*). First note that (IC) and the single crossing property imply that types in (θ', θ'') must be pooled at $(x(\theta'), t(\theta'))$ (each type prefers $(x(\theta'), t(\theta'))$ to $(x(\theta''), t(\theta''))$, and any other distinct contract that a type $\theta \in (\theta', \theta'')$ likes as least as much as $(x(\theta'), t(\theta'))$ will be envied by either θ' or θ''). For any $\varepsilon \in (0, \theta'' - \theta')$, give all types in $[\theta' + \varepsilon, \theta'')$ a contract $(x(\theta') + \delta, \tau(\delta))$ satisfying $\delta \in (0, x(\theta'') - x(\theta'))$ and $U(x(\theta') + \delta, \tau(\delta), \theta'') = U(\theta'')$. Also, give types in $[\theta', \theta' + \varepsilon)$ a contract $(x(\theta') + d(\delta), p(\delta))$ satisfying the following equations:

$$\begin{aligned} U(x(\theta') + d(\delta), p(\delta), \theta' + \varepsilon) &= U(x(\theta') + \delta, \tau(\delta), \theta' + \varepsilon) \\ U(x(\theta') + \delta, \tau(\delta), \theta') &= U(\theta'). \end{aligned}$$

For all other types, the contract is unchanged. It is straightforward to show that the new menu satisfies (IC) and (P) for each $\varepsilon \in (0, \theta'' - \theta')$ and each δ satisfying the given equalities. The change in expected profit is

$$\int_{[\theta' + \varepsilon, \theta'')} [\tau(\delta) - t(\theta') - \theta\delta] dF(\theta) + \int_{(\theta', \theta' + \varepsilon)} [p(\delta) - t(\theta') - \theta d(\delta)] dF(\theta).$$

Each of the functions τ, p, and d are continuously differentiable in δ with derivatives uniformly bounded on $(0, x(\theta'') - x(\theta')) \times [\theta', \theta'']$. Hence the change in expected profit is differentiable (for each $\varepsilon \in (0, \theta'' - \theta')$) and equals

$$\int_{[\theta' + \varepsilon, \theta'')} [\tau'(0) - \theta] dF(\theta) + \int_{(\theta', \theta' + \varepsilon)} [p'(0) - \theta d'(0)] dF(\theta).$$

By Lemma 1, the first term is positive. As $\varepsilon \to 0$, the second term vanishes, while the first tends to $\int_{[\theta', \theta'')} [\tau'(0) - \theta] dF(\theta) > 0$, so $(x(\theta), t(\theta))_{\theta \in \Theta}$ cannot be optimal.

Now consider case (*ii*). As in case (*i*) any types in (θ'', θ') must be pooled at $(x(\theta'), t(\theta'))$. Let $\varepsilon \in (0, \theta' - \theta'')$. For each type in $[\theta'' + \varepsilon, \overline{\theta}]$, consider a new menu that lowers the utility in each state by $\delta > 0$ satisfying $U(\theta'' + \varepsilon) - \delta > U(x(\theta'), t(\theta'), \theta'' + \varepsilon)$. Give types in $(\theta', \theta' + \varepsilon)$ the contract with consumption plan at the intersection of the indifference set of type θ' (at $(x(\theta'), t(\theta'))$) and type $\theta'' + \varepsilon$ (at the new contract for it).

The contract remains the same for all other types. By the single crossing property and feasibility of $(x(\theta), t(\theta))_{\theta\in\Theta}$, the new contract will satisfy (P) and (IC). Moreover, on the set $(\theta'' + \varepsilon, \overline{\theta}]$, expected profit rises and has a positive derivative with respect to δ at $\delta = 0$ for every $\varepsilon \in [0, \theta' - \theta'')$. As in case ($i$), the derivative of expected profit conditional on the complement of types wrt δ at $\delta = 0$ tends to zero as ε tends to zero. Since $(\theta'', \overline{\theta}] = \cup_{\varepsilon>0}[\theta'' + \varepsilon, \overline{\theta}]$, $(\theta'', \theta') \subset (\theta'', \overline{\theta}]$, and by hypothesis (θ'', θ') is of positive measure, we have that for some $\varepsilon \in (0, \theta' - \theta'')$, the new contract gives the principal higher expected profit than $(x(\theta), t(\theta))_{\theta\in\Theta}$, contradiction.

$U(\cdot)$ IS DIFFERENTIABLE. By the previous result, for each $\theta \in \Theta = [\underline{\theta}, \overline{\theta}]$, the contract in $\{x(\theta'), t(\theta') | \theta' \in \Theta\}$ which maximizes $U(\cdot, \cdot, \theta)$ is unique. Since $U(\chi, \tau, \theta)$ is affine in θ, $U'(\theta)$ exists and equals $-\Delta(\theta)$ (e.g. Milgrom and Segal (2002), Theorem 3).

$U(\cdot)$ IS CONTINUOUSLY DIFFERENTIABLE. Since $U(\cdot)$ is convex and differentiable, it follows from Rockafellar (1970) (Corollary 25.5.1) that it is C^1 on $(\underline{\theta}, \overline{\theta})$. $\square$

A.3 Proof of Lemma 2 (Complete Sorting)

Using (13)-(14) to eliminate $\dot{\lambda}(\theta)$, rewrite the numerator of (17) as $\lambda(\theta)B(\theta)$, where

$$B(\theta) = f(\theta)\left[\frac{f'(\theta)}{f(\theta)}\theta(1-\theta) + (1-3\theta) + \frac{1}{1 - \frac{h'(U(\theta)-(1-\theta)\Delta(\theta))}{h'(U(\theta)+\theta\Delta(\theta))}}\right].$$

Since

$$\frac{1}{1 - \frac{h'(U(\theta)-(1-\theta)\Delta(\theta))}{h'(U(\theta)+\theta\Delta(\theta))}} = 1 + \frac{1}{\frac{h'(U(\theta)+\theta\Delta(\theta))}{h'(U(\theta)-(1-\theta)\Delta(\theta))} - 1}$$

and $\lambda(\theta) \leq 0$, it follows that $\dot{\Delta}(\theta) \leq 0$ *if and only if*

$$\frac{f'(\theta)}{f(\theta)}\theta(1-\theta) + (2-3\theta) + \frac{h'(U(\theta)-(1-\theta)\Delta(\theta))}{h'(U(\theta)+\theta\Delta(\theta)) - h'(U(\theta)-(1-\theta)\Delta(\theta))} \geq 0,$$

or, equivalently,

$$\frac{f'(\theta)}{f(\theta)} \geq \frac{3\theta - 2 - b(\theta)}{\theta(1-\theta)},$$

where $b(\theta) = h'(U(\theta) - (1-\theta)\Delta(\theta))/[h'(U(\theta) + \theta\Delta(\theta)) - h'(U(\theta) - (1-\theta)\Delta(\theta))]$. $\square$

A.4 Proof of Theorem 3 (Complete Sorting: Sufficiency)

(i) We show that the following condition is sufficient for complete sorting:

$$\frac{f'(\theta)}{f(\theta)} \quad > \quad \frac{3\theta - 1}{\theta(1-\theta)} - \frac{f(\theta)}{1-F(\theta)}. \tag{27}$$

Fix $\hat{\theta} \in [\underline{\theta}, \bar{\theta})$. We first claim that if $\dot{\Delta}(\tau) < 0$ for all $\tau \in (\hat{\theta}, \bar{\theta})$, and condition (27) holds, then $\dot{\Delta}(\hat{\theta}) < 0$. To establish this claim, we show that

$$b(\hat{\theta}) > -1 + \frac{f(\hat{\theta})}{1-F(\hat{\theta})}\hat{\theta}(1-\hat{\theta}), \tag{28}$$

implying that the necessary and sufficient condition (18) holds at $\hat{\theta}$.

Let $h'_n(\theta) = h'(U(\theta) - \theta\Delta(\theta))$. Since $\dot{\Delta}(\tau) < 0$ and $f(\tau) > 0$ for all $\tau \in (\hat{\theta}, \bar{\theta})$,

$$f(\tau)h'_n(\hat{\theta}) > f(\tau)h'_n(\tau) > f(\tau)a(\tau), \tag{29}$$

for all $\tau \in (\hat{\theta}, \bar{\theta})$, with equalities at $\bar{\theta}$, where $a(\cdot)$ is defined in equation (16).

Integrate both sides of (29) from $\hat{\theta}$ to $\bar{\theta}$ and divide by $1 - F(\hat{\theta})$ to obtain

$$h'_n(\hat{\theta}) > \frac{1}{1-F(\hat{\theta})}\int_{\hat{\theta}}^{\bar{\theta}} a(\tau)f(\tau)d\tau = (\Delta h)'(\hat{\theta})\frac{\hat{\theta}(1-\hat{\theta})f(\hat{\theta}),}{1-F(\hat{\theta})} \tag{30}$$

where $(\Delta h)'(\theta) = h'(U(\theta) + \theta\Delta(\theta)) - h'(U(\theta) - (1-\theta)\Delta(\theta))$ and we have used (16). Adding $-\hat{\theta}(\Delta h)'(\hat{\theta})$ to both sides of (30) and rearranging yields (28), implying that the necessary and sufficient condition (18) holds at $\hat{\theta}$, thus proving the claim.

It now follows that $\dot{\Delta}(\theta) < 0$ for all θ under condition (27). For suppose not; then there would be a largest $\theta \in [\underline{\theta}, \bar{\theta})$ with $\dot{\Delta}(\theta) \geq 0$, since $\dot{\Delta}(\cdot)$ is continuous and $\limsup_{\theta \to \bar{\theta}} \dot{\Delta}(\theta) < 0$. By the claim, condition (27) would fail.

(ii) Let $f'(\cdot)/f(\cdot)$ be decreasing and suppose that $\overline{\theta} \leq 1/2$. Assume $\dot{\Delta}(\theta) \geq 0$ for some θ, say $\hat{\theta}$. We will show that this leads to a contradiction.

We know from Lemma 3 that a necessary and sufficient condition for complete sorting is $f'(\theta)/f(\theta) > (3\theta - 2 - b(\theta))/\theta(1-\theta)$ for all θ. By the MLRP, the left side of this expression is strictly decreasing in θ. Regarding the right side, it is easy to check that if

$\overline{\theta} \leq 1/2$, then it is strictly increasing at θ if $b'(\theta) < 0$. Now,

$$b'(\theta) = -\dot{\Delta}(\theta)\left[\frac{h_n' h_\ell''(1-\theta) + h_n'' h_\ell' \theta}{(h_n' - h_\ell')^2}\right],$$

which is nonpositive at $\hat{\theta}$ since $\dot{\Delta}(\hat{\theta}) \geq 0$. Since $(3\theta - 2 - b(\theta))/\theta(1-\theta)$ is strictly increasing at $\hat{\theta}$, the same holds in a neighborhood of $\hat{\theta}$, and thus $\dot{\Delta}(\cdot) > 0$ in that neighborhood. As this happens for any arbitrary $\hat{\theta}$ for which $\dot{\Delta}(\hat{\theta}) \geq 0$, then if one such $\hat{\theta}$ exists, $\dot{\Delta}(\theta) > 0$ for all $\theta > \hat{\theta}$. But this is a contradiction, since $\dot{\Delta}(\cdot) < 0$ near $\overline{\theta}$.

Let $f'(\theta) \geq 0$ for all θ and $\overline{\theta} \leq 1$. Assume $\dot{\Delta}(\theta) \geq 0$ for some θ, say $\hat{\theta}$. Rewrite the necessary and sufficient condition for complete sorting as $(1-\theta)f'(\theta)/f(\theta) > (3\theta - 2 - b(\theta))/\theta$. It is easy to check that MLRP plus $f'(\cdot) \geq 0$ implies that the left side is strictly decreasing in θ. In turn, the right side is strictly increasing at θ if $b'(\theta) < 0$, which holds when $\dot{\Delta}(\theta) \geq 0$. Following the same steps as before we reach a contradiction. □

A.5 Proof of Proposition 2 (No Exclusion)

Type $\underline{\theta}$ is not excluded from the optimal menu of contracts if $\Delta(\tilde{\theta}) > \Delta_0$. From equation (16) and the concavity of the optimal control problem, this is tantamount to showing that the marginal benefit of providing insurance to $\tilde{\theta}$ starting from no insurance is bigger than the marginal cost of doing so. Formally,

$$f(\underline{\theta})\underline{\theta}(1-\underline{\theta})[h'(u(w)) - h'(u(w-\ell))] > \int_{\underline{\theta}}^{\overline{\theta}} a(s)f(s)ds. \tag{31}$$

Assume first that $-u'''(\cdot)/u''(\cdot) < -3u''(\cdot)/u'(\cdot)$, which holds for a large class of utility functions. It is easy to show that in this case $a'(\cdot) < 0$. Hence,

$$\int_{\underline{\theta}}^{\overline{\theta}} a(s)f(s)ds < a(\underline{\theta}),$$

and (31) holds if $f(\underline{\theta})\underline{\theta}(1-\underline{\theta})[h'(u(w)) - h'(u(w-\ell))] > a(\underline{\theta})$. Since $a(\underline{\theta}) < (1-\underline{\theta})h'(u(w)) + \underline{\theta}h'(u(w-\ell))$, it follows that

$$f(\underline{\theta})\underline{\theta}(1-\underline{\theta})[h'(u(w)) - h'(u(w-\ell))] > (1-\underline{\theta})h'(u(w)) + \underline{\theta}h'(u(w-\ell)),$$

which is equivalent to

$$f(\underline{\theta}) > \frac{(1-\underline{\theta})h'(u(w)) + \underline{\theta}h'(u(w-\ell))}{\underline{\theta}(1-\underline{\theta})[h'(u(w)) - h'(u(w-\ell))]}. \tag{32}$$

Thus, type $\underline{\theta}$ is not excluded from the optimal menu if $f(\underline{\theta})$ is larger than the right side of (32). But Theorem 3 (*i*) or (*ii*) imply that $\theta(1-\theta)f(\theta)/(1-F(\theta))$ is increasing in θ, thereby showing that no $\theta \geq \underline{\theta}$ will be excluded from the optimal menu of contracts.

Without imposing $-u'''(\cdot)/u''(\cdot) < -3u''(\cdot)/u'(\cdot)$, a stronger sufficient condition holds, with the numerator on the right side of (32) replaced by $h'(u(w))$. $\square$

A.6 Proof of Proposition 3 (Exclusion)

Type $\tilde{\theta}$ is excluded from the optimal menu of contracts if $\Delta(\tilde{\theta}) = \Delta_0$. From equation (16), we must show that the marginal benefit of providing insurance to $\tilde{\theta}$ starting from no insurance is less than the marginal cost of doing so. Formally,

$$f(\tilde{\theta})\tilde{\theta}(1-\tilde{\theta})[h'(u(w)) - h'(u(w-\ell))] < \int_{\tilde{\theta}}^{\overline{\theta}} a(s)f(s)ds. \tag{33}$$

Assume first that $-u'''(\cdot)/u''(\cdot) < -3u''(\cdot)/u'(\cdot)$. In this case,

$$\int_{\tilde{\theta}}^{\overline{\theta}} a(s)f(s)ds > a(\overline{\theta})(1-F(\tilde{\theta})),$$

and (33) holds if $f(\tilde{\theta})\tilde{\theta}(1-\tilde{\theta})[h'(u(w)) - h'(u(w-\ell))] < a(\overline{\theta})(1-F(\tilde{\theta}))$. Since $a(\overline{\theta}) = h'(U(\overline{\theta})) \geq h'((1-\overline{\theta})u(w) + \overline{\theta}u(w-\ell))$, type $\tilde{\theta}$ is excluded from the optimal menu if

$$\frac{f(\tilde{\theta})}{(1-F(\tilde{\theta}))} < \frac{h'((1-\overline{\theta})u(w) + \overline{\theta}u(w-\ell))}{\tilde{\theta}(1-\tilde{\theta})[h'(u(w)) - h'(u(w-\ell))]}. \tag{34}$$

But Theorem (*i*) or (*ii*) imply that $\theta(1-\theta)f(\theta)/(1-F(\theta))$ is increasing in θ, which shows that any $\theta \leq \tilde{\theta}$ will be excluded from the optimal menu of contracts as well.

Without imposing $-u'''(\cdot)/u''(\cdot) < -3u''(\cdot)/u'(\cdot)$, a similar, but stronger, sufficient condition for exclusion holds, with the numerator of (34) replaced by $h'(u(w-\ell))$. $\square$

A.7 Proof of Lemma 3 (Curvature)

Let $(U(\theta), \Delta(\theta))$ solve the optimal control problem with $\dot{\Delta}(\cdot) < 0$ everywhere. Since $u(\theta) = U(\theta) + \theta\Delta(\theta)$ for all θ, we can use (5)-(6) to recover the optimal menu $(x(\theta), t(\theta))_{\theta\in\Theta}$.

Recalling that $\dot{U}(\theta) = -\Delta(\theta)$, we obtain $\dot{u}(\theta) = \theta\dot{\Delta}(\theta)$. Differentiating (5)-(6) yield

$$\dot{t}(\theta) = -\frac{\theta\dot{\Delta}(\theta)}{u'(w - t(\theta))} \tag{35}$$

$$\dot{x}(\theta) = -\frac{\dot{\Delta}(\theta)[(1-\theta)u'(w - t(\theta)) + \theta u'(w - \ell + x(\theta) - t(\theta))]}{u'(w - \ell + x(\theta) - t(\theta))u'(w - t(\theta))}. \tag{36}$$

where we have used $u(\theta) = u(w - t(\theta))$ and $h'(\cdot) = 1/u'(h(\cdot))$.

Since we are assuming complete sorting, $\dot{x}(\cdot) > 0$, and thus the inverse of $x(\cdot)$, call it $z(\cdot)$, is well defined (i.e., $\theta = z(x)$). We can now represent the optimal mechanism as a *nonlinear premium schedule* $T(x) = t(z(x))$. Then, (35)-(36) and $\theta = z(x)$ yield

$$\dot{T}(x) = \dot{t}(z(x))\dot{z}(x) = \frac{\dot{t}(z(x))}{\dot{x}(z(x))} = \frac{\theta u'(w - \ell + x(\theta) - t(\theta))}{(1-\theta)u'(w - t(\theta)) + \theta u'(w - \ell + x(\theta) - t(\theta))}.$$

Differentiating $\dot{T}(\cdot)$ and using $\theta = z(x)$ yields, after some algebra,

$$\ddot{T}(x) = \frac{1}{\theta u'_\ell + (1-\theta)u'_n}\left\{\frac{u'_n u'_\ell}{\dot{x}(\theta)} + \theta(1-\theta)\left[u''_n u'_\ell \frac{\dot{t}(\theta)}{\dot{x}(\theta)} + u''_\ell u'_n \left(1 - \frac{\dot{t}(\theta)}{\dot{x}(\theta)}\right)\right]\right\}.$$

Inserting (35)-(36) into (37) and manipulating the resulting expression, reveals that $\ddot{T}(x) \gtrless 0$ if and only if $\dot{\Delta}(\theta) \gtrless \left(\theta(1-\theta)\left[\theta\frac{u''_n}{u'^2_n} + (1-\theta)\frac{u''_\ell}{u'^2_\ell}\right]\right)^{-1}$. Inserting (17) in the left side of this expression yields (19), after tedious algebra. ☐

References

ARNOTT, R., AND J. STIGLITZ (1988): "Randomization with Asymmetric Information," *Rand Journal of Economics*, 19, 344–362.

BARLOW, R., A. MARSHALL, AND F. PROSCHAN (1963): "Properties of Probability Distributions with Monotone Hazard Rate," *The Annals of Mathematical Statistics*, 34, 375–389.

Biais, B., D. Martimort, and J.-C. Rochet (2000): “Competing Mechanisms in a Common Value Environment,” *Econometrica*, 68, 799–37.

Cawley, J., and T. Philipson (1999): “An Empirical Examination of Information Barriers to Trade in Insurance,” *American Economic Review*, 89, 827–846.

Chiappori, P.-A., and A. Bernardo (2003): “Bertrand and Walras Equilibria under Moral Hazard,” *Journal of Political Economy*, 111, 785–817.

Chiappori, P.-A., B. Jullien, B. Salanié, and F. Salanié (2007): “Asymmetric Information in Insurance,” *Rand Journal of Economics*, forthcoming.

Hellwig, M. (2006): “Incentive Provision with Hidden Characteristics: A Unified Approach,” SSRN WP.

Jewitt, I. (1987): “Risk Aversion and the Choice between Risky Prospects: The Preservation of Comparative Statics,” *Review of Economic Studies*, 54, 73–85.

Jullien, B. (2000): “Participation Constraints in Adverse Selection Models,” *Journal of Economic Theory*, 93, 1–47.

Maskin, E., and J. Riley (1984): “Monopoly with Incomplete Information,” *Rand Journal of Economics*, 15, 171–196.

Matthews, S., and J. Moore (1987): “Monopoly Provision of Quality and Warranties: An Exploration in the Theory of Multidimensional Screening,” *Econometrica*, 55, 441–67.

Milgrom, P., and I. Segal (2002): “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70, 583–601.

Mussa, M., and S. Rosen (1978): “Monopoly and Product Quality,” *Journal of Economic Theory*, 18, 301–17.

Noldeke, G., and L. Samuleson (2007): “Optimal bunching without optimal control,” *Journal of Economic Theory*, 134, 405–20.

Prescott, E., and R. Townsend (1984): “Pareto Optima and Competitive Equilibrium with Adverse Selection and Moral Hazard,” *Econometrica*, 52, 21–46.

Rockafellar, T. (1970): *Convex Analysis.* Princeton University Press, Princeton, New Jersey.

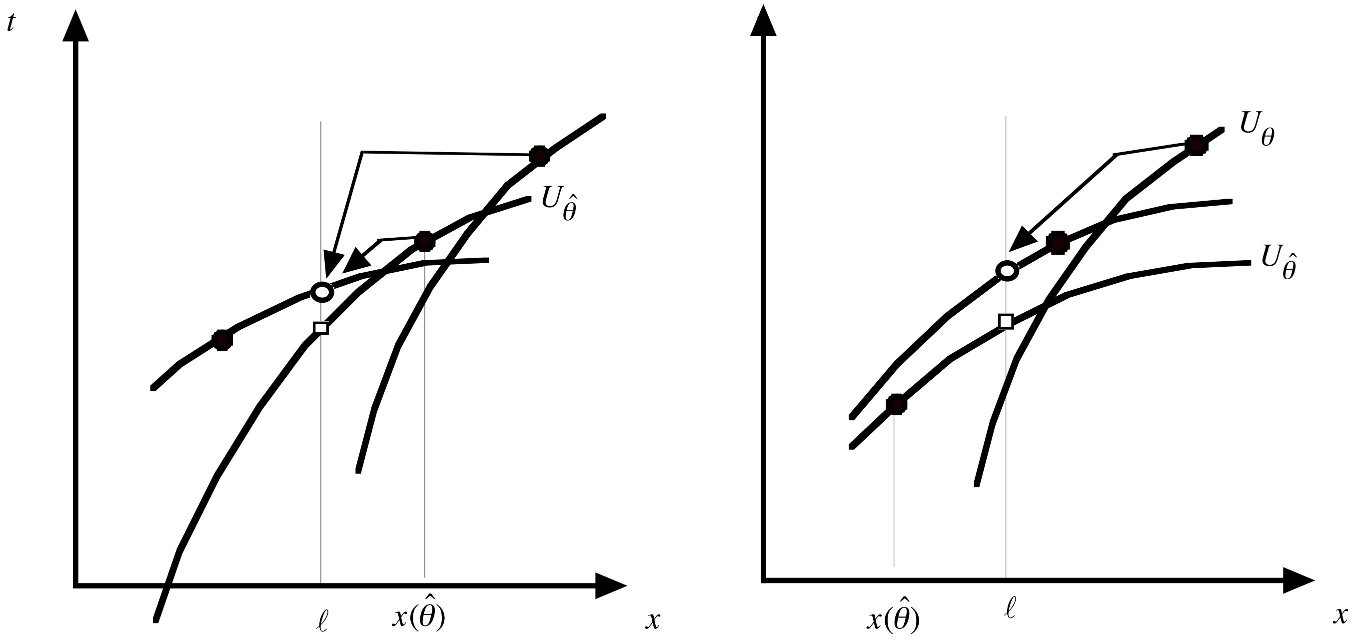
Rothschild, M., and J. Stiglitz (1976): “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information,” *Quarterly Journal of Economics*, 90, 629–650.

Shannon, C. (1995): “Weak and Strong Monotone Comparative Statics,” *Economic Theory*, 5, 209–227.

Stiglitz, J. (1977): “Monopoly, Non-linear Pricing and Imperfect Information: The Insurance Market,” *Review of Economic Studies*, 44, 407–430.

Thiele, H., and A. Wambach (1999): “Wealth Effects in the Principal Agent Model,” *Journal of Economic Theory*, 89, 247–260.

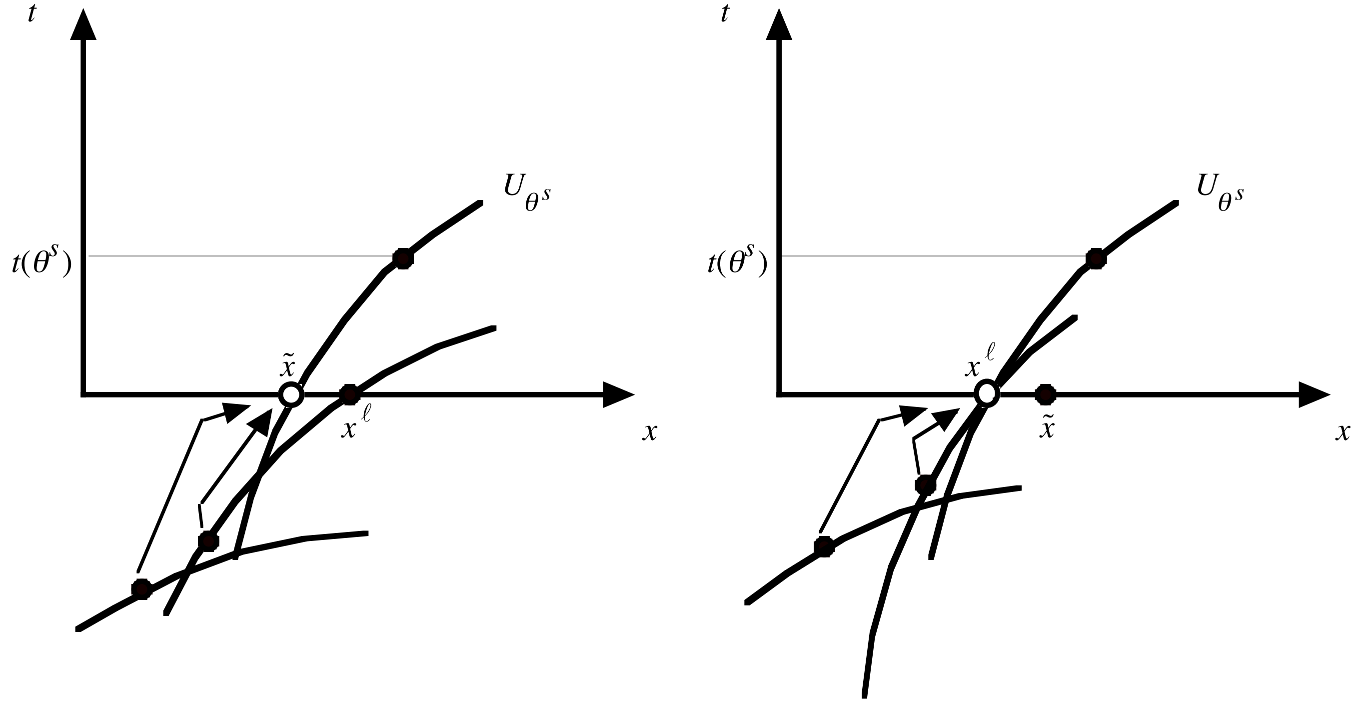
Wilson, R. (1993): *Nonlinear Pricing.* Oxford University Press, New York.



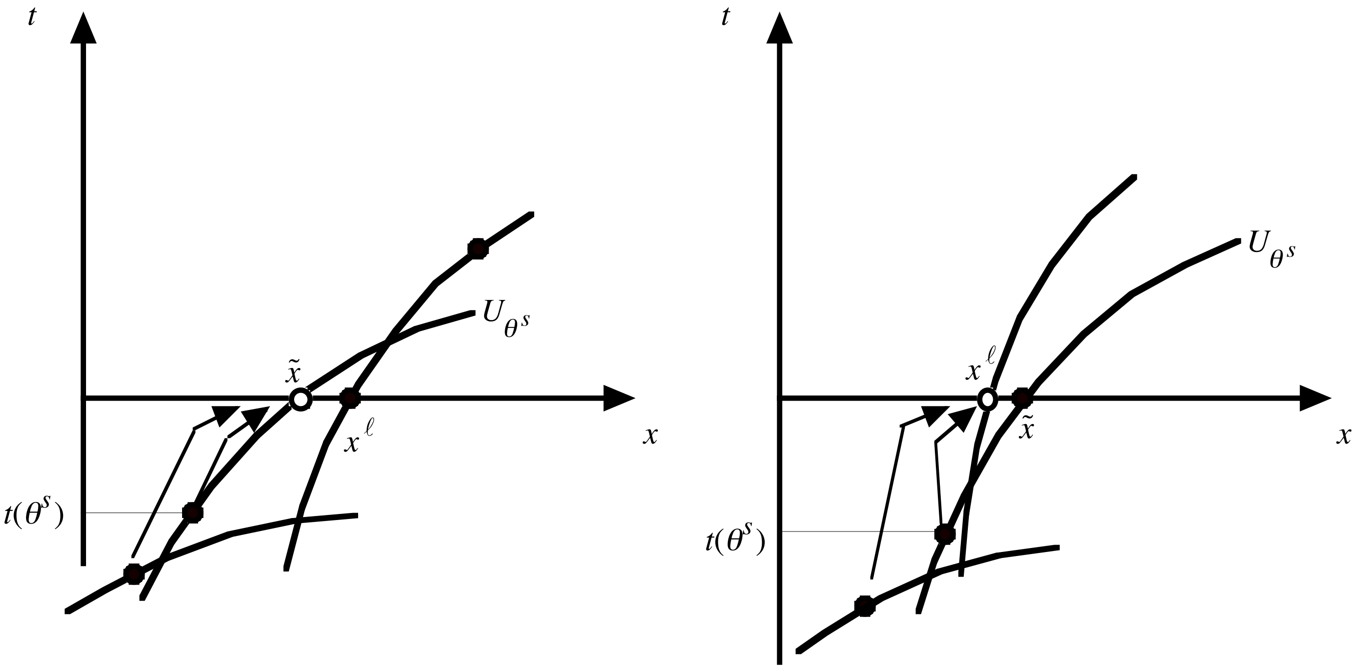
$\hat{\theta} \in \{\theta \mid x(\theta) > \ell\}$

$\hat{\theta} \notin \{\theta \mid x(\theta) > \ell\}$

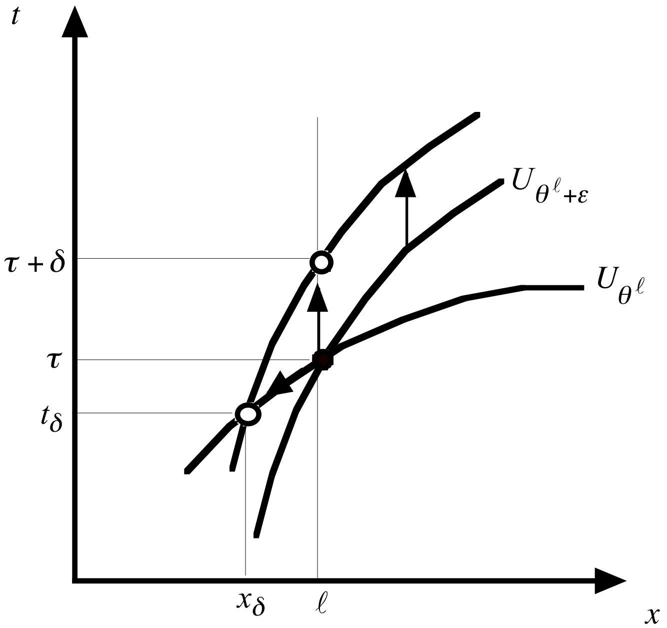
(ii) No Overinsurance



(iiia) Nonnegativity, case (a): $\theta^s \notin \{\theta \mid t(\theta) < 0\}$



(iiib) Nonnegativity, case (b): $\theta^s \in \{\theta \mid t(\theta) < 0\}$



(vi): No pooling at the top