Coverage Denied: Excluding Bad Risks, Inefficiency, and Pooling in Insurance[*]

(Preliminary and Incomplete: Do Not Circulate)

Hector Chade[†]
Arizona State University
Department of Economics

Edward Schlee[‡]
Arizona State University
Department of Economics

September 15, 2013

Abstract

Models of insurance with adverse selection predict that only the best risks—those *least* likely to suffer a loss—are uninsured, a prediction at odds with coverage denials for pre-existing conditions. They also typically assume that insurance provision is costless: an insurer's only cost is payment of claims. We introduce costly insurance provision in an otherwise standard monopoly insurance model with adverse selection. We show that with loading or a fixed cost of claims processing, the insurer denies coverage *only* to the worst risks, those most likely to suffer a loss. We also show that loading overturns three classic textbook properties of monopoly pricing models: no one is pooled with the highest consumer type; the highest type gets an efficient contract; and all other types get contracts distorted downwards from their efficient contracts.

Keywords: Adverse selection, Insurance, Loading, Fixed Cost, Coverage Denials, Non-responsiveness.

[*]We thank Dan Silverman for pointing out Hendren's work to us. We thank Dan, Nathaniel Hendren, François Salanié and participants at the 2013 SAET conference in Paris for comments and discussions. Schlee's research was supported in part by 2010 and 2011 Summer Grants from the W.P. Carey School at Arizona State University, and by a stipend from the Toulouse School of Economics, where he visited in June, 2010. Part of this work was done during Schlee's visit to Johns Hopkins University in the Spring of 2012.

[†]Email: `hector.chade@asu.edu`

[‡]Email: `edward.schlee@asu.edu`

1 Introduction

A striking prediction of monopoly models of insurance under adverse selection is that only the best risks—those *least* likely to suffer a loss—are uninsured (Stiglitz (1977); Chade and Schlee (2012)). The uninsured moreover go without coverage voluntarily: the insurer offers each consumer a menu of contracts, but the best risks simply chose zero coverage. Casual evidence—and non-casual evidence, such as Gruber (2008), Hendren (2013), and McFadden, Noton, and Olivella (2012)—suggests that some consumers are *involuntary uninsured* in the sense that they are not offered any (nonzero) contracts. For example, some insurers refuse to write health care policies for consumers with 'pre-existing' adverse health conditions.[1] And these involuntarily uninsured are those that insurers believe to be bad risks: we know of no evidence, casual or otherwise, that those believed to be good risks—those least likely to file a claim—are denied coverage.

To write the obvious, insurers deny coverage to a consumer because they expect to lose money from any policy that the consumer would accept. Such a belief presumably comes from observing an attribute of a consumer, for example a medical history. We model this attribute as a *signal* that is correlated with the consumer's loss chance. But adding a signal to the standard monopoly insurance model is not enough: Chade and Schlee (2012) find that a monopolist insurer always makes positive expected profit: there are always gains to trade between the insurer and the consumer.

The standard model assumes that the insurer's only cost is payment of claims: *insurance provision itself is costless.* We find that adding insurance provision costs can account for coverage denials for only the worst risks. We consider the three most common costs discussed in the complete information insurance literature: *loading*; a fixed cost of *claims processing*; and a fixed cost of *entry* into a line of insurance.[2]

We first present a general no-trade comparative static result for monopoly pricing models: under an assumption on the monopolist's set of feasible expected profits, we

[1] Appendix F in Hendren (2013) contains an excerpt from Genworth Financial's underwriting guidelines for long-term care insurance. There are two pages of 'uninsurable conditions.' Lists of uninsurable conditions from other insurers can be found at Hendren's research page, http://scholar.harvard.edu/hendren/publications. See also the discussion on 'lemon dropping' (exclusion of bad risks) in the recent survey by Einav, Finkelstein, and Levin (2010).

[2] The cost of insurance provision is often referred to as an *administrative* cost. Both the empirical insurance literature and practitioners extensively discuss these costs. Gruber (2008) for example writes that administrative costs average 12% of the premium paid by consumers in the US health insurance industry. And they are mentioned prominently in standard textbook and survey treatments of insurance under complete information (e.g. Rees and Wambach (2008)).

show that if there are no gains to trade with a consumer at a belief about the consumer's type, then there are no gains to trade at any belief that is worse in the sense of *likelihood ratio* dominance. The assumption is satisfied in our insurance model if there is no fixed cost of entry (that is, the only provision costs are loading and a fixed claims cost). At one level, this result answers our question about excluding only bad risks. But the comparative static result does not pin down who is denied coverage and who isn't. In the special case of pure loading, we also derive a necessary and sufficient condition for coverage denials that includes as a special case a recent result by Hendren (2013). Although the condition potentially requires that many inequalities be checked, we also give a simple sufficient condition for coverage denials consisting of a single inequality.

Turning to a fixed entry cost, we derive a necessary and sufficient condition for the insurer to deny coverage only to the worst risks. But our message here is negative: the condition is that complete-information profit is zero for every type except the lowest (the type least likely to suffer a loss).

These results demarcate when there are gains to trade or not when there are insurance provision costs. Chade and Schlee (2012) confirm that, absent provision costs, three classic contracting properties hold for monopoly insurance under adverse selection: no consumer type pools with the 'highest' type (*no-pooling at the top*); the highest type gets an efficient contract (*efficicency at the top*); and each of the other types gets a contract that is distorted downwards from its efficient contract (*downward distortions elsewhere*). Although fixed costs do not affect these properties, *all* three can fail with loading. They fail since, with loading, the complete information contract can be *decreasing* in the loss chance, though incentive compatibility requires menus to be increasing in the loss chance: this conflict between incentive compatibility and complete-information contracts is called *nonresponsiveness* (Guesnerie and Laffont (1984), Morand and Thomas (2003)).

Our paper is obviously related to the large theoretical and empirical literature on insurance under adverse selection (see Chade and Schlee (2012) and the papers they cite). We are not aware of any systematic analysis of the effects of insurance provision costs on monopoly contracts. The closest paper on coverage denials is Hendren (2013). He provides a sufficient and necessary condition for there to be no gains to trade between an insurer and a consumer. The condition puts restrictions both on the distribution of types and the consumer's risk aversion: roughly, the consumer is not too risk averse; and the distribution is shifted far enough to the right—in particular the support *must* include a type who suffers a loss with probability one. Our results for coverage denials with loading or a fixed claims cost (Propositions 2 and 3) deliver his conclusion without requiring the existence of a type who suffers a loss for sure. A major goal of Hendren (2013) is to test a model of coverage denials: he estimates the distribution of loss types for consumers rejected for coverage and those which are not for three insurance markets, and finds that the type distributions for rejected consumers have fatter right tails. Our general comparative statics result (Proposition 1) provides a new theoretical foundation for his procedure.

The paper is also related to the broader literature on contracts in principal-agent models where the agent has private information. Hellwig (2010) analyzes a general version of such a screening problem that includes, among others, the seminal Mirlees (1971) model as a special case. He shows that the three classic contracting properties mentioned above hold in his model. In our insurance context with costly provision, we show that they fail (Proposition 6).

We consider monopoly insurance for three reasons. First, it is the most challenging market structure for explaining coverage denials. Since a monopolist earns the highest possible profit, it follows that if there are no gains to trade with a monopolist insurer, then there are no gains to trade with other market structures. Second, there is evidence that insurance markets are not competitive (e.g., see McFadden, Noton, and Olivella (2012) p.10 for the case of health insurance in the US and the references cited in Chade and Schlee (2012)). Third, there is no agreement about the "right" model of an imperfectly competitive insurance market (see the discussion on this point in Section 6 of Einav, Finkelstein, and Levin (2010)); monopoly seems to be a useful place to start thinking about imperfectly competitive insurance markets.

After explaining our model of insurance with adverse selection and costly provision, we show how these costs can reconcile adverse selection with coverage denials for *only* bad risks (Section 3). We then show how loading overturns some classic properties of insurance with adverse selection (Section 4). All the proofs are in the Appendix.

2 A Model of Costly Insurance Provision

The model is the standard monopoly insurance model (Stiglitz (1977)) except for insurance provision costs. A consumer has initial wealth of $w > 0$, faces a potential loss of $\ell \in (0, w)$ with chance $p \in \mathcal{P} \subset (0, 1]$, and has preferences represented by a differentiable, strictly concave von Neumann-Morgenstern utility function u on $\mathbb{R}_+$, with $u' > 0$. The loss chance p, from now on the consumer's *type*, is known privately to the consumer.

The monopoly insurer is risk neutral. It has a belief ρ about the consumer's type with support in $\mathcal{P}$. It chooses, for each $p \in \mathcal{P}$, a contract $(x, t) \in \mathbb{R}$ consisting of a premium t and an indemnity payment x in the event of a loss. We denote the resulting *menu* of contracts by $\{(x(p), t(p)\}_{p \in \mathcal{P}}$. The expected utility of a type-p consumer for a contract (x, t) is $U(x, t, p) = pu(w - \ell + x - t) + (1 - p)u(w - t)$.

We use the classic two-type case for examples and intuition. In this special case, $\mathcal{P} = \{p_L, p_H\}$, where $0 < p_L < p_H \leq 1$, and we let $\rho \in [0, 1]$ denote the insurer's belief that the consumer the high risk type, p_H.

Except for allowing $p = 1$ in the type support, so far the model is as in Chade and Schlee (2012). We change the model in two important ways.

First, we assume that insurance provision is costly. The insurance literature has mentioned three kinds of such costs (often called 'administrative' costs in that literature): *loading* (expected marginal cost of coverage exceeds the loss chance); *claims processing* costs (which occur only in the even of a loss); and a cost of *entering an insurance line* (which Shavell (1977) calls the cost of *opening a policy*).[3] We assume that the insurer's expected cost from a contract (x, t) given to a type-p consumer in a menu with some nonzero contracts is

$$c(x, p, \lambda, k, K) = \begin{cases} \lambda p x + kp + K & \text{if } x > 0 \\ K & \text{if } x = 0 \end{cases}$$

where $\lambda \geq 1$ is a *loading factor*, $k \geq 0$ is the (fixed) cost incurred when a claim is made, and $K \geq 0$ is a *fixed cost* of entry. The expected cost is 0 for a no-trade menu.[4]

Chade and Schlee (2012) prove that the insurer offers nonnegative contracts with coverage not greater than the loss and premium not greater than the consumer's wealth. Here we simply impose these properties as a constraint. Let $\mathcal{C}$ be the set of menus satisfying $0 \leq x(p) \leq \ell$ and $0 \leq t(p) \leq w$ for every $p \in \mathcal{P}$.[5]

The second modification is that, before offering a menu, the insurer observes a *signal* correlated with the consumer's loss chance. The signal could be either costly or costless, and its informativeness could be either exogenous or endogenous. An important example

[3] See Boland (1965), Lees and Rice (1965), Shavell (1977), Arrow (1965), and Diamond (1977).

[4] Our results all go through for any differentiable cost function $c(x, p)$ convex in x for every $p \in \mathcal{P}$; $c_x(x, p) > p$ for every $x \geq 0$ and $p \in \mathcal{P}$ and $x < \ell$; and $c_x(x, p)/p$ is increasing in p for every $x \geq 0$.

[5] The constraint $x \leq \ell$ does not bind here. And for what we argue are reasonable specifications of the cost function for $x < 0$, the constraint $(x, t) \geq 0$ does not bind either.

of an endogenous, costly signal in insurance is *underwriting*.[6] After observing the signal realization, the insurer updates beliefs about the consumer's type and decides what nonzero contracts, if any, to offer. We interpret ρ as the insurer's posterior belief after observing the realization of the signal.[7]

Since the fixed entry cost only affects *whether* any nonzero contracts are offered, we first write down the problem without it (by setting $K = 0$). For a given belief ρ, the insurer's problem is to choose a (measurable) menu of contracts to solve

$$V(\rho, \lambda, k) = \max_{\{(x(\cdot),t(\cdot))\}\in\mathcal{C}} E_\rho \left[t(p) - c(x(p), p, \lambda, k, 0)\right] \tag{1}$$

subject to

$$U(x(p), t(p), p) \geq U(x(p'), t(p'), p) \qquad \text{for } p, p' \in \mathcal{P}, \quad \text{(IC)}$$

$$U(x(p), t(p), p) \geq U(0, 0, p) \qquad \text{for } p \in \mathcal{P}, \quad \text{(P)}$$

where $E_\rho[\cdot]$ is the expectation taken over the type set $\mathcal{P}$ using the insurer's belief ρ. Condition (IC) summarizes the incentive compatibility constraints and (P) the participation constraints.[8] If $V(\rho, \lambda, k) \leq K$, then there are no gains to trade and we say that *coverage is denied*. (Recall that $K \geq 0$.)

As in the case of costless insurance, any menu of contracts that satisfies (IC) is *increasing*: if $p' > p$, then $x(p') \geq x(p)$ and $t(p') \geq t(p)$. Monotonicity follows from (IC) since U satisfies the strict single crossing property in (x, t) and p. Namely, if a low type prefers a higher contract to any distinct lower contract, then a higher type strictly prefers the higher contract.[9]

To keep the effects of different costs clear, we first present the main results for the case of loading without any fixed costs, and then for fixed costs without loading. In most cases it will be clear how the results change when all of the costs are present.

[6] We do not model the cost of a signal since we envision it as already sunk when the insurer offers a menu (as is natural in the case of underwriting).

[7] Rather than formally introducing signals, we work directly with posterior beliefs. There is no loss of generality in doing so, since 'observing the posterior' is equivalent to 'observing the signal realization.' See for example the Bayesian account of Blackwell's theorem in Kihlstrom (1984). We assume realistically that the signal is observed before writing a menu and it is not contractible.

[8] Since u is strictly concave, it follows immediately that the solution is unique in the finite type case.

[9] In calculus, the marginal rate of substitution $MRS(x, t, p) = -U_x(x, t, p)/U_t(x, t, p)$ is strictly increasing in p (e.g, see Figure 1).

3 Coverage Denials: Excluding Only Bad Risks

With costless insurance provision, Chade and Schlee (2012) show that there are always gains from trade between the insurer and the consumer.[10] It is trivial that there are no gains to trade if these costs are large enough. What is nontrivial is to find conditions under which the insurer will deny coverage *only* to those who are likely to be bad risks, in the sense that if the insurer denies coverage at a belief about the consumer's type, then it does so for 'worse' beliefs (suitably defined).

3.1 A No-Trade Comparative Statics Result

We begin with a simple but general no-trade comparative statics result that is of independent interest. The result is easier to present and prove after a change of variables: rather than choosing a menu of feasible contracts, we have the insurer choosing a menu of feasible expected profits. For any menu $\{(x(p), t(p)\}_{p \in \mathcal{P}}$ in $\mathcal{C}$ that satisfies (IC) and (P), there is a function $\pi : \mathcal{P} \to \mathbb{R}$ that gives the expected profit for each type $\pi(p) = t(p) - c(x(p), p, \lambda, k, K)$, for every $p \in \mathcal{P}$. Let Φ be the set of such functions. The next assumption says that if a menu of expected profits is feasible, then so is a menu that sets expected profit of all types below some threshold type equal to zero, and leaves the expected profit from each of the other types unchanged.

Assumption 1. *If $\pi(\cdot) \in \Phi$, then for any $p' \in \mathcal{P}$, it follows that $\pi'(\cdot) \in \Phi$, where $\pi'(p) = \pi(p)$ for $p > p'$ and $\pi'(p) = 0$ otherwise.*

We discuss this assumption at the end of this subsection. For now we just note that it holds in the monopoly insurance models of Stiglitz (1977) and Chade and Schlee (2012).

Milgrom (1981) shows that *Likelihood Ratio* (LR) dominance can be interpreted as receiving better or worse news about a unknown parameter.[11] In a standard abuse of Milgrom's definition, we will say that if belief ρ' about the consumer's loss chance LR dominates ρ, then ρ' is worse news about the consumer's loss chance than ρ.[12]

[10] They assume that the type set excludes $p = 1$.

[11] Let F be the cumulative distribution function (cdf) with upper bound for its support equal to $p_H \leq 1$. A cdf G *likelihood ratio* dominates F if there is a nondecreasing function h on $[0, p_H]$ such that $G(p) = \int_0^p h(p)dF(p)$ for every $p \in [0, p_H]$. If F and G are either differentiable or simple functions, then the condition is equivalent to the condition that the likelihood ratio g/f is nondecreasing, where g and f are the density or mass functions for the cdfs G and F.

[12] It is the conclusion of one of his propositions, not the definition.

Proposition 1. *Suppose Assumption 1 holds and that there are no gains to trade at a belief about the consumer's type. Then there are no gains to trade at any belief with the same support that LR dominates it.*

Recall that we interpret the insurer's belief to be a posterior belief after observing a signal correlated with the consumer's loss chance. The result implies that, if the set of possible posteriors is ordered by LR dominance—by better or worse news in Milgrom's (1981) sense—then only consumers with type distributions shifted most to the right are denied coverage. In that sense, coverage is denied *only* to the worst risks.

The argument for Proposition 1 is straightforward in the two-type case. Suppose that the insurer cannot earn positive profit at some belief about the likelihood of the types. By Assumption 1, the insurer must lose money on the high risk type (otherwise the insurer would get positive profit by giving the low-risk type zero coverage). And the complete-information profit cannot be positive for the high risk type, since it is always possible to give the low risk type zero coverage and the high risk its complete information contract. So if the insurer cannot earn positive profit at some belief, it must be that any profit from the low risk type cannot make up for losses from the high risk. If the high risk type now becomes more likely (a LR change in the insurer's belief), then it is all the more the case that low-risk profit cannot make up for high-risk losses.

It is worth pointing out that the proof does not use any details of the insurance problem: it only uses LR dominance and Assumption 1.[13] And Assumption 1 holds if i) the strict single crossing property holds for the consumer; ii) contracts are nonnegative; iii) the no-trade contract $(0, 0)$ is feasible; and iv) expected profit from the no-trade contract is 0 for every type. Under these four conditions, we can give the null contract to every type less than or equal the threshold type p' and leave the contracts to other types the same; by the single-crossing property, the new menu is feasible, and it gives the insurer zero expected profit for all $p \leq p'$. So Proposition 1 holds across a range of monopoly pricing problems with privately informed consumers. In particular, these properties hold in our insurance model if there is no fixed cost of entry ($K = 0$); but with an entry cost Assumption 1 fails, and so does the conclusion of Proposition 1.

An important application of Proposition 1 to insurance with adverse selection is to Hendren (2013), who determines whether the estimated loss distribution of potential consumers who are denied have fatter right tails than consumers who get coverage. Our

[13] In our insurance model we can drop the common support restriction in Proposition 1.

Proposition 1 gives a general foundation for comparing the right tails of the distribution of the loss chances for consumers who are and are not denied coverage.

3.2 Loading

To isolate the effects of loading on coverage denials, we set $K = k = 0$ and let $\lambda \geq 1$. As mentioned, Assumption 1, and so Proposition 1, holds in this case. We go beyond that by pinning down *exactly* when there are no gains to trade.

Let $MRS(p) \equiv MRS(0,0,p) = -U_x(0,0,p)/U_t(0,0,p) = pu'(w-\ell)/(pu'(w-\ell)+(1-p)u'(w))$ be type-p's marginal rate of substitution of x for t at the no-trade contract (graphically, it is the slope of type-p's indifference curve in the (x,t) space at the origin). Also, let $E_\rho[p|p \geq \hat{p}]$ be the conditional expectation of the consumer's type p given the event $\{p \geq \hat{p}\}$ when the insurer's belief is give by ρ.

Proposition 2 (Loading and No Trade)**.** *(i) Let $\lambda \geq 1$, $K = k = 0$, and fix ρ. There are no gains to trade if and only if*

$$MRS(\hat{p}) \leq \lambda E_\rho[p|p \geq \hat{p}] \quad \textit{for all} \quad \hat{p} \in \mathcal{P}. \tag{2}$$

(ii) If $\lambda > 1$, then a sufficient condition for (2) is

$$MRS(E_\rho[p]) \leq \lambda E_\rho[p]. \tag{3}$$

Part (i) says that if the inequalities all hold at a belief ρ, then the consumer is *denied coverage*. In the two-type case, it follows that $E_\rho[p|p \geq \hat{p}]$ is increasing in the probability that the consumer is high risk for any $\hat{p} \in \mathcal{P}$, so we get the comparative statics result we sought: *if the insurer denies coverage at some belief that the consumer is the high type, it does so at any larger belief that the consumer is high risk.* Using Proposition 1, the result extends beyond two types when ρ' LR dominates ρ.

Although LR dominance is well-grounded as a formalization of better or worse news about some parameter, in the case of loading without fixed costs we can use (2) to strengthen the conclusion to *hazard rate* (*HR*) dominance.[14] This result gives us an additional testable implication beyond Proposition 1 for the case of pure loading.

[14] For cumulative distribution functions defined on $[0,1]$, G hazard rate (HR) dominates F if $(1-G(p))/(1-F(p))$ is nondecreasing in p on $[0, p_H]$ where p_H is the upper bound on the support of F. LR dominance implies HR dominance.

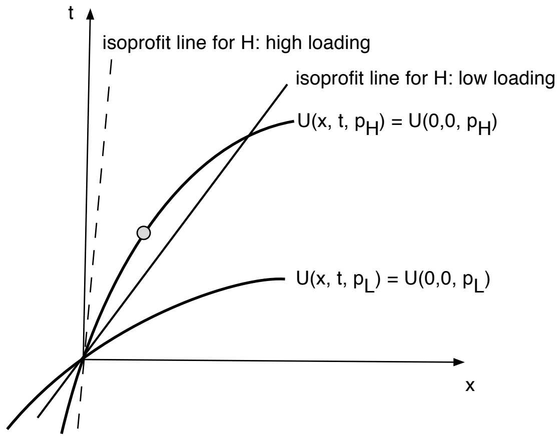

Figure 1: **Loading and Coverage Denials, Two-Type Case: Condition (2) for the High Type.** The high-risk consumer's indifference curve is the steeper of the two. The dashed line is an insurer iso-expected-profit line for sufficiently high loading. For that case condition (2) holds at $\hat{p} = p_H$. For the flatter iso-profit line, it fails, and the insurer can get positive expected profit with the contract depicted on the high-risk consumer's indifference curve. The low-risk consumer voluntarily refuses that contract.

Corollary 1 (HR and Coverage Denials)**.** *Let $\lambda \geq 1$ and $k = K = 0$. If coverage is denied at some belief about types, then it is denied at any belief that HR dominates it.*

In addition, the inequalities in (2) reveal if λ is not too large and the highest type is less than 1, then there are gains to trade for some belief. This result follows from two facts: the insurer's maximum profit V is continuous in λ; and from Theorem 1 (*vi*) in Chade and Schlee (2012), which asserts that, without loading (i.e., with $\lambda = 1$), the insurer's expected profit is positive when the highest loss chance is less than 1. Of course if $\lambda p > 1$ for all types, then there are no gains to trade.

Figures 1 and 2 depict a graphical explanation of Proposition 2 (*i*) for the two-type case, and illustrates a profitable contract for the insurer when (2) fails.

When $\lambda = 1$ then Proposition 2 (*i*) specializes to Theorem 1 in Hendren (2013). Thus, our Proposition 2 (*i*) extends his result to the case of loading.[15] Note that if $\lambda \hat{p} \geq 1$, then condition (2) must hold, so loading can generate coverage denials *without* requiring that the riskiest consumer suffer a loss with probability 1. By Theorem 1 (vi)

[15]Indeed, we could prove Proposition 2 (*i*) largely by changing notation in Hendren (2013) to include loading; for completeness, we include a short and direct proof.

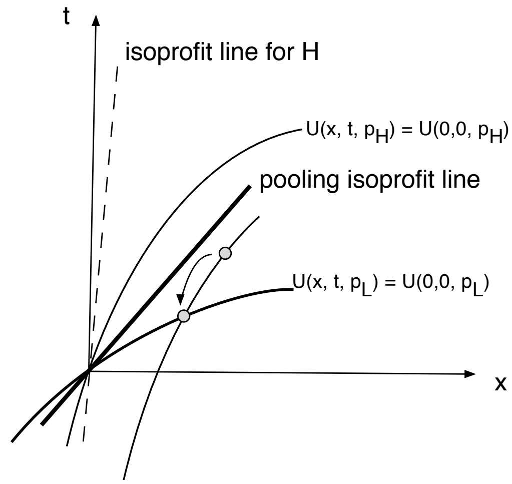

Figure 2: **Loading and Coverage Denials: Condition (2) for the Low Type.** Inequality (2) holds for $\hat{p} = p_L$ and for $\hat{p} = p_H$. Expected profit from the high risk consumer is negative for any contract this consumer is willing to buy. To make positive expected profit, the low-risk type must buy a positive contract. By (IC) and (2), the best hope for positive profit is by offering a pooling contract, which must lie on the low-risk consumer's indifference curve through $(0,0)$. But since (2) holds at $\hat{p} = p_L$, no such pooling contract can be profitable.

in Chade and Schlee (2012), a *necessary* condition for coverage denials with costless insurance is that the highest type suffers a loss with probability 1. Since loading simply multiplies the loss chance in the insurer's cost, one might conjecture that a necessary condition for coverage denials with loading is that $\lambda p \geq 1$ with positive probability. Scrutiny of equation (2) reveals that the conjecture is false. Consider the two type case $\mathcal{P} = \{p_L, p_H\}$ with $0 < p_L < p_H < 1/\lambda$ but $\lambda > 1$. Now let the loss size ℓ become small. It is easy to show that $\lim_{\ell \to 0} MRS(p) = p$, so in the limit $MRS(\hat{p}) = \hat{p} < \lambda E_\rho[p|p \geq \hat{p}]$ for $\hat{p} \in \{p_L, p_H\}$ and equation (2) holds as a strict inequality for all sufficiently small losses. Intuitively, when the loss size is small, the consumer's demand for insurance is low, and loading makes insurance unprofitable;[16] the conclusion holds even though $\lambda p < 1$ with probability 1.[17]

Proposition 2 (*ii*) asserts that *all* the inequalities in (2) hold if the *single* inequality (3) holds. This can be useful when $\mathcal{P}$ contains a large number of types. For an intuition, suppose that uncertainty is *symmetric*, in the sense that neither the insurer nor the

[16] A point emphasized by Lees and Rice (1965) in their comment on Arrow (1963).

[17] Similarly, there are no gains to trade with any consumer whose risk aversion is (uniformly) small enough. At the other extreme, if risk aversion is high enough at $w - \ell$, then there will be gains to trade for any fixed loading factor.

consumer know the consumer's loss chance. Then the insurer offers a contract tailored to the mean loss chance. One can easily verify that there is trade in the symmetric case if and only if $MRS\,(E_\rho[p]) > \lambda E_\rho[p]$. So one can interpret inequality (3) as follows: *If there is no trade under symmetric uncertainty, then there is no trade under adverse selection.* Note that this condition bites *only* with loading, since it fails with $\lambda = 1$.[18]

3.3 Fixed Costs

We now turn to the analysis of fixed costs and coverage denials. Notice that the *sufficiency* part of Proposition 2 (i) still holds with positive fixed costs: if a no-trade menu solves the insurers problem for $K = 0 = k$, then it solves it for $(K, k) \geq 0$. But *necessity* fails: if condition (2) fails, a no-trade menu can still maximize profit. It is worth pointing out that the empirical tests in Hendren (2013) do not involve direct tests of condition (2); rather he tests whether the estimated distribution of loss chances for those who are denied coverage have *fatter right tails* than those not denied. We ask: can his empirical findings be made consistent with fixed costs?

When there is a fixed claims cost ($k > 0$) and no entry cost ($K = 0$), Proposition 1 holds (with $\lambda \geq 1$). It is straightforward to show that the conclusion holds nontrivially, that is, a no-trade menu sometimes does and sometimes does not solve the insurer's problem. To illustrate this point, set $\lambda = 1$ so that complete information expected profit is $\Pi(p, 1) - pk$—the risk premium minus the expected claim cost. Since $\Pi(1, 1) = 0$ and the risk premium is strictly concave in p, there is a $\hat{p} < 1$ such that $\Pi(p, 1) - pk < 0$ if and only if $p > \hat{p}$. So as long as the insurer's beliefs put enough weight on types above $\hat{p}$, then a no-trade menu maximizes profit. Moreover if k and λ are not too large, then there are gains to trade for beliefs that put enough weight on types sufficiently close to 0. But unlike Proposition 2 for loading, we have not found a tight necessary and sufficient condition for no trade.

Proposition 3 (Fixed Claims Cost)**.** *Suppose that $K = 0$ and $0 < k < u'(w)\ell/[u(w) - u(w - \ell)]$. Let p_L be the smallest and p_H the largest element of $\mathcal{P}$. There are values of $(p_L, p_H) \in (0, 1)^2$ and λ close to 1 such that there are gains to trade for some insurer beliefs and there are no gains to trade for other beliefs.*

[18] With costless provision, the insurer's expected profit is higher when no one knows the consumer's type than if both do—since the risk premium is concave in the loss chance—and it is even lower under adverse selection. Another effect of loading is that the first comparison does not hold since the complete-information value function, given in equation (4), cannot be concave in p on $[0, 1]$ when $\lambda > 1$.

Conditions for no trade with a fixed entry cost are more difficult to pin down from primitives. In what follows we set $k = 0$. To understand the problem, consider first the complete-information problem for an insurer selling to a consumer with p when $K = 0$:

$$\Pi(p, \lambda) = \max_{x \geq 0} T(x, p) - \lambda x p, \tag{4}$$

where T is the willingness-to-pay of a type-p consumer for coverage x, defined by $U(x, T, p) = U(0, 0, p)$. If $\lambda = 1$, then Π is the risk premium, since the solution to (4) is full coverage. Notice that $\Pi(0, 1) = 0 = \Pi(1, 1)$. So if we add a fixed entry cost $K > 0$, then both very good *and* very bad risks can be denied coverage—though 'moderate' risks might not.

Turning to asymmetric information, there are gains to trade if and only if $0 \leq K < V(\rho, \lambda, 0)$. Consider the two-type case. If a monopolist is to deny coverage *only* to the worst risks for *any* $K > 0$, then V should be decreasing in the probability ρ that the consumer is the high-risk. We show in the Appendix (Section A.4) that the derivative $V_\rho(1, \lambda, 0)$ is nonnegative, and it is positive if and only if $\Pi(p_H, \lambda) > 0$. So if $\Pi(p_H, \lambda) > 0$, the insurer's value function $V(\cdot, \lambda, 0)$ *cannot* be decreasing in ρ. We show that in this case $V(\cdot, \lambda, 0)$ is decreasing in ρ *if and only if* $\Pi(p_H, \lambda) = 0$, and strictly so whenever V is positive. Absent loading, of course, complete information expected profit is positive. The two-type case suggests that perhpas the interaction between loading and a fixed entry cost can explaining coverage denials only to the worst risks. The extension to many types is far more negative. We will say that *coverage is denied only to bad risks* if the conclusion of Proposition 1 holds: If coverage is denied at some insurer belief, then it is denied at any insurer belief that LR dominates it (with the same support).

Proposition 4 (Fixed Entry Cost)**.** *Let $\mathcal{P} = \{p_1, ..., p_n\}$ with $n \geq 2$ and $p_1 < ... < p_n$ and fix $k = 0$ and $\lambda \geq 1$. Assume beliefs have full support. Then only bad risks are denied coverage for every $K \geq 0$ if and only if $V(\cdot, \lambda, 0)$ is decreasing in the LR; and this last condition holds if and only if $\Pi(p, \lambda) = 0$ for every type p except the lowest type.*

This result suggests that we cannot reconcile coverage denials for only bad risks with a fixed entry cost—or at least not for plausible assumptions on the set of types.

More positively, we show (Section A.6) that, as risk aversion increases uniformly without bound, the insurer value function V converges pointwise to a function that *is* decreasing in the LR order: indeed the limiting value function is just $\max\{0, \ell(1 -$

$\lambda E_\rho[p])\}$, where $E_\rho[p]$ is the mean loss chance of the consumer. This result, however, is merely a limiting one: short of the limit, the conclusion of Proposition 4 applies.

4 Pooling and Efficiency

As mentioned, three classic properties of contracting menus with private information, but costless insurance provision, are that no type pools with the highest type (*no pooling at the top*); the highest type gets an efficient contract (*efficiency at the top*); and all other types get coverage smaller than the efficient level (*downward distortions elsewhere*).[19] We now consider whether these three properties hold with costly insurance provision.

4.1 Loading

As with coverage denials, we first assume $K = k = 0$, and $\lambda \geq 1$ to isolate the effects of loading on menus when there are gains to trade. Loading can have dramatic effects on insurance. We begin with complete-information insurance.

COMPLETE-INFORMATION INSURANCE. With complete information and costless provision, the profit-maximizing menu is simple: each type gets full insurance and is charged a premium that extracts all the surplus from that type, so the insurer's expected profit is just that type's risk premium. With loading *no* type gets full insurance, and the exact amount depends on the loss chance. More dramatically, the monopoly complete-information menu can be strictly *decreasing* in the loss chance.

Example 1. *Assume the consumer has a CARA utility function $u(z) = -e^{-rz}$, where $r > 0$ is the consumer's coefficient of absolute risk aversion. In the Appendix (Section A.6) we show that the profit-maximizing coverage is given by*

$$x^*(p) = \max\left\{0\,,\ \ell + \frac{1}{r}\log\frac{(1-p\lambda)}{(1-p)\lambda}\right\} < \ell$$

for every $p \in (0, 1/\lambda)$, and equals 0 for $p \geq 1/\lambda$. The coverage is nonincreasing in p on $[0, 1]$, and strictly decreasing on the set of loss chances for which coverage is positive.

More generally, non-decreasing risk aversion suffices for the conclusion.

[19]Theorem 1 in Chade and Schlee (2012). Hellwig (2010) shows these properties hold in a general private-values Principal-Agent model.

Proposition 5. *If the consumer's preferences satisfy non-decreasing absolute risk aversion then the complete information menu is nonincreasing, and strictly decreasing on the set of types with positive coverage.*

A simple intuition for this result comes from considering analogues of substitution and wealth effects of an increase in p. The first order conditions for the monopolist are

$$\frac{pu'(w-\ell-t+x)}{pu'(w-\ell-t+x)+(1-p)u'(w-t)} = \lambda p \tag{5}$$

and $U(x,t,p) = U(0,0,p)$, the participation constraint, where the left side of (5) is $-\frac{U_x}{U_t}(x,t,p)$. Let $p_1 > p_0$ and let (x_0, t_0) be the complete-information contract for type p_0. *Fixing the contract* (x_0, t_0), an increase in p from p_0 to p_1 raises the right side of (5) *more* than the left: if the insurer chooses (x,t) to maximize expected profit subject to $U(x,t,p_1) = U(x_0,t_0,p_1)$, then coverage would *decrease.* We identify this change as a *a substitution effect* (with a Slutsky compensation for the consumer). Now however the participation constraint is slack (see Figure 3). Increase t until $U(x_0,t,p_1) = U(0,0,p_1)$. If risk aversion decreases with t (increasing absolute risk aversion, IARA), then the left side *decreases*, a wealth effect. Under IARA, the substitution and wealth effects work in same direction, but under decreasing absolute risk aversion, they work in opposite directions. Figure 3 illustrates these two effects.

The possibility of strictly decreasing menus under complete information contrasts with increasing menus in the incomplete information case (an implication of (IC) and the strict single-crossing property of consumer preferences in the contract and loss chance).

Insurance with Adverse Selection. As mentioned, three classic properties of profit-maximizing menus with adverse selection costless provision are *no pooling at the top*, *efficiency at the top*, *downward distortions elsewhere.* In the two-type case, the three properties mean that the high type gets full insurance and the profit-maximizing menu exhibits complete sorting of the two types, with the low type getting coverage that is less than its first-best coverage (which is full coverage). We show that each of these three properties can fail with loading.

With costless provision, a contract is efficient if and only if it gives full coverage. With loading, efficiency does not imply full coverage. Here we say that a *contract* $(x', t') \in \mathcal{C}$ given to a type-p consumer is *efficient* if it maximizes the insurer's expected profit $t - \lambda px$ on $\{(x,t) \in \mathcal{C} \mid U(x,t,p) = U(x',t',p)\}$, the set of contracts that are indifferent to (x',t') for a type-p consumer. A contract (x'', t'') is *distorted downwards* from an

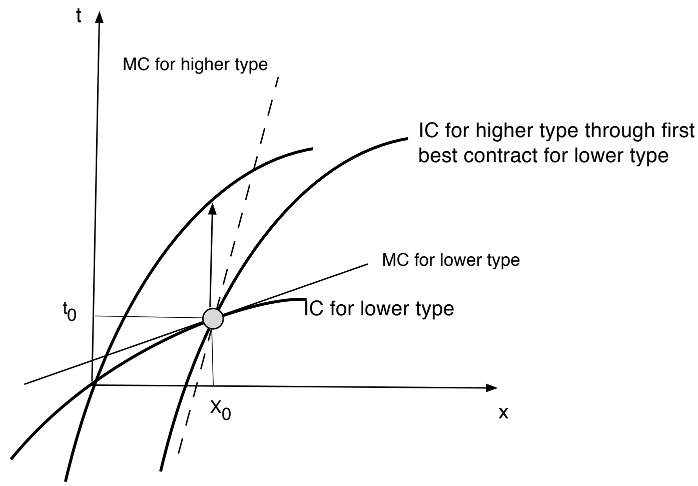

Figure 3: **Decreasing Complete-Information Menus: Substitution and Wealth Effects** The point (x_0, t_0) is the complete information contract for type p_0. If $p_1 > p_0$, then the slope the marginal cost rises faster than the slope of the indifference curve through (x_0, t_0), a substitution effect. If t increases so that P binds, the indifference curve becomes flatter under non-decreasing risk aversion—a wealth effect—reinforcing the substitution effect.

efficient contract (x', t') for type p if $(x'', t'') > (x', t')$ and $U(x'', t'', p) = U(x', t', p)$. If the inequality is reversed, then (x'', t'') is *distorted upwards* from an efficient contract.

The proof of Proposition 2 (*i*) for the two-type case reveals that, if $\lambda p_H > 1$—so that any efficient contract for the high type gives 0 coverage—and the no-trade condition (2) fails for p_L, then the profit-maximizing menu is *pooling* at a positive contract, and the high-type contract is distorted *upwards*. Given $\lambda p_H > 1$, this case occurs if p_L and the probability that the type is low are both low enough. The next result is more general and does not rely on the insurer believing that $\lambda p > 1$ with positive probability.

Proposition 6 (Failure of Classic Contracting Properties)**.** *Suppose that the consumer's preferences satisfy non-increasing absolute risk aversion, set $k = K = 0$, and fix $\lambda \in (1, u'(w - \ell)/u'(w))$. Then there is a type set $\mathcal{P} \subset (0, 1/\lambda)$ such that (a) there are gains to trade for some full-support insurer belief; and (b) at any such belief the three classic properties fail. In particular, some types are pooled with the highest type; the highest type gets an inefficient contract; and some other types get a contract that is distorted upwards from efficiency.*[20]

[20] As the proof reveals, the only properties we require of the type set are that its smallest element be close enough to 0; and that it contains two other distinct elements that are close enough to $1/\lambda$ (how close in each case depends on the consumer's preferences).

To prove Proposition 6 we construct an insurance problem in which the efficient contract for all high-enough types is zero. The next example solves for a profit-maximizing menu in the two-type case that involves pooling for the case of CARA preferences. Besides illustrating Proposition 6 it shows that pooling and inefficiency at the top can occur even when efficient contracts for the high-risk consumer are positive.

Example 2. *Consider the CARA case with two types ($\mathcal{P} = \{p_L, p_H\}$ and risk aversion equal to r, and suppose that $\lambda < p_H$ and r is sufficiently high so that every efficient contract for the high type is positive. We show in the Appendix (Section A.9) that the following pooling contract (x, t) maximizes expected profit if the high type is sufficiently likely, or if λ and either r or ℓ are sufficiently high:*

$$
\begin{aligned}
x &= \ell + \frac{1}{r} \log \frac{p_L(1 - E_\rho[p]\lambda)}{(1 - p_L)E_\rho[p]\lambda} \\
t &= \frac{1}{r} \log \frac{(1 - E_\rho[p]\lambda)}{(1 - p_L)} (p_L e^{r\ell} + (1 - p_L)).
\end{aligned}
$$

Note that the low-risk contract is distorted upwards and the the high risk contract distorted downward compared with the efficient contracts since $p_L < E_\rho[p] < p_H$.

What drives the failure of the three standard properties when there is loading? Intuitively, one reason for the pooling in Proposition 6 and Example 2 is because complete-information efficient contracts are decreasing in the loss chance while incentive compatibility requires contracts to be increasing in the loss chance under incomplete information. This conflict is called *nonresponsiveness* in the contracting literature (see Guesnerie and Laffont (1984), and Morand and Thomas (2003)).

5 Concluding Remarks

Most adverse selection models of monopoly insurance with costless insurance provision exhibit gains from trade, so the insurer is always willing to sell insurance to a consumer. Moreover, only consumers with the smallest loss chances go uninsured and this lack of coverage is voluntary, which does not fit the evidence on the large number of people who are denied coverage for being bad risks. We show that costly insurance provision in the form of loading or a fixed claims cost can account for coverage denials only to bad risks, but a fixed entry cost cannot. We also analyze how these costs affect contracts when there are gains from trade, and in particular show that loading can lead to dramatically different predictions. We know little about the size and kind of these provision costs. (Einav, Finkelstein, and Levin (2010), p. 322 point out the difficulty of measuring them.) Since they matter so much, it would be useful to measure them better.

One open theoretical question is how loading affects quantity discounts/premia. Chade and Schlee (2012) pin down the shape of the optimal menu for the the case of a continuum of loss chances: decreasing absolute risk aversion and log-concave density of types imply that the optimal premium is a 'backwards-S' shaped, first concave, then convex. This shape is consistent with global quantity discounts in insurance (i.e., $t(p)/x(p)$ is decreasing in p). This result is important since several scholars find evidence of quantity discounts in insurance, and often interpret them as evidence against adverse selection. We conjecture that the optimal menu with loading could still be consistent with global quantity discounts, but leave this conjecture for future work.

A Appendix

A.1 Proof of Proposition 1

Let G and F be two possible insurer beliefs—cumulative distribution functions—and suppose that G LR dominates F. The statement of the proposition requires that they have the same support. We prove a slightly more general result and allow the support of G to be a subset of the support of F. Let p_L be the infimum and p_H the supremum of the support of G. Since G LR dominates F, the supremum of the support of F must equal p_H. From the definition of LR dominance, there is a nonnegative, nonincreasing function $h(\cdot)$ on on $[p_L, p_H]$ such that

$$F(p) = \int_{p_L}^{p} h(q) dG(q)$$

for every $p \in [p_L, p_H]$. Let π^G maximize expected profit on Φ at belief G, and let π^F maximize expected profit on Φ at belief F subject to the *additional* constraint that $\pi^F(p) = 0$ for $p < p_L$. We prove the contrapositive. Suppose that $\int \pi^G dG > 0$. Since $\max_{\pi \in \Phi} \int \pi dF \geq \int \pi^F dF \geq \int \pi^G dF$, it suffices to show that $\int \pi^G dF > 0$. Using the continuity of the integral $\int_{[p_L,p]} \pi^G(q)dG(q)$, integrate by parts to find[21]

$$\int_{[p_L,p_H]} \pi^G(p)dF(p) = \int_{[p_L,p_H]} \pi^G(p)h(p)dG(p) = \tag{6}$$

$$h(p_H)\int_{[p_L,p_H]} \pi^G dG + \int_{[p_L,p_H]} \left[\int_{[p_L,p]} \pi^G(q)dG(q)\right] d(-h(p)). \tag{7}$$

Consider the two terms in (7). The first term is nonnegative by hypothesis. Assumption 1 implies that $\int_{[p_L,p]} \pi^G(q)dG(q) \geq 0$ for every $p \in [p_L, p_H]$: if for some $\pi \in \Phi$, there is a $p \in [p_L, p_H]$ with $\int_{[p_L,p]} \pi(q)dG(q) < 0$, then by setting $\pi'(p) = \mathbb{1}_{(p,p_H]}\pi(p)$ (where $\mathbb{1}_{(\cdot)}$ is the indicator function), $\int \pi' dG > \int \pi dG$, and π cannot solve the insurer's problem at belief G. Since h is nonincreasing, it follows that the second term in (7) is nonnegative. We are done if at least one of the terms in (7) is positive. There are two possibilities. First, $h(p_H) > 0$, in which case the first term in (7) is positive. Second, $h(p_H) = 0$. Since p_H is the supremum of the support of G and F, $1 - F(p) > 0$ for every $p \in [p_L, p_H)$. Integrate by parts to find that, for every $p \in [p_L, p_H)$

$$0 < 1 - F(p) = \int_{(p,p_H]} h(q)dG(q) = h(p_H) - G(p_L)h(p_L) - \int_{(p,p_H]} G(q)dh(q)$$

$$= -G(p_L)h(p_L) + \int_{(p,p_H]} G(q)d(-h(q)), \tag{8}$$

so that $\int_{(p,p_H]} G(q)d(-h(q)) > 0$ for every $p \in [p_L, p_H)$. It follows that the cumulative distribution function $-h$ puts positive measure on all sets of the form $(p, p_H]$ for $p < p_H$. Since $\int_{p_L}^{p} \pi^G(q)dG(q) \geq 0$ with a strict inequality at $p = p_H$, and the integral is continuous in p, the second term in (7) is positive. □

A.2 Proof of Proposition 2

(i) It will be useful to rewrite the condition $MRS(0, 0, \hat{p}) \leq \lambda E_\rho[p|p \geq \hat{p}]$ as follows:

$$\frac{\hat{p}u'(w - \ell)}{(1 - \hat{p})u'(w)} \leq \frac{\lambda E_\rho[p|p \geq \hat{p}]}{1 - \lambda E_\rho[p|p \geq \hat{p}]}. \tag{9}$$

To prove that (2) is sufficient for a null menu maximizing the insurer's expected

[21] The function $-h$ is nondecreasing but need not be right-continuous. But there is a unique function h^* which is increasing, right-continuous and agrees with $-h$ whenever it is right continuous. The integral $\int f d(-h)$ is defined to be equal to $\int f dh^*$ (Royden (1968), p. 263.).

profit, assume that (9) holds for all types. Suppose first that F has finite support $\{p_1, ..., p_n\}$ with $p_n > ... > p_1$. Recall that with finite types, we can reduce the insurer's problem to one of maximizing expected subject to $x_1 \le x_2 \le ... \le x_n$, the binding participation constraint of the lowest type, and the binding local downward constraints. We split the problem into a series of programs that can be solved recursively starting from the highest type.

We first show by induction that any solution to the monopolist problem involves pooling, namely, $x_1 = x_2 = ... = x_n$ and $t_1 = t_2 = ... = t_n$. Fix $\{((x_1, t_1), ..., (x_{n-1}, t_{n-1})\}$ with each contract nonnegative and $x_i \ge t_i$ for $i = 1, ..., n-1$. Consider the problem of choosing (x_n, t_n) to maximize $t_n - \lambda p_n x_n$ subject to the constraints that $x_n \ge x_{n-1}$ and $U(x_n, t_n, p_n) = U(x_{n-1}, t_{n-1}, p_n)$. By (2), the strict concavity of u, and $x_{n-1} \ge t_{n-1} \ge 0$ it follows that

$$\lambda p_n \ge \frac{p_n u'(w-\ell)}{(1-p_n)u'(w)} \ge \frac{p_n u'(w-\ell+x_{n-1}-t_{n-1})}{(1-p_n)u'(w-t_{n-1})} = -\frac{U_x}{U_t}\Big(x_{n-1}, t_{n-1}, p_n\Big).$$

Let m equal the last expression, the marginal rate of substitution of the type-n consumer at the contract for a type $n-1$ consumer. Now consider any (x, t) satisfying the constraint for the insurer's problem for type n. Since $U(\cdot, p)$ is strictly concave for every p, and $(x, t) \ge (x_{n-1}, t_{n-1})$ it follows that $t - t_{n-1} \le m(x - x_{n-1})$. Use the inequality $\lambda p_n \ge m$ and rearrange to find that $t_{n-1} - \lambda p_n x_{n-1} \ge t - \lambda p_n x$ so $x_n = x_{n-1}$ and $t_n = t_{n-1}$ solves the problem.

Now fix $\{((x_1, t_1), ..., (x_{n-k}, t_{n-k})\}$ nonnegative with $x_i \ge t_i$ for $i = 1, ..., n-k$, and set $x_n = x_{n-1} = ... = x_{n-k+1}$ and $t_n = t_{n-1} = ... = t_{n-k+1}$. Consider the problem

$$\max_{(x_{n-k+1}, t_{n-k+1}) \ge 0} t_{n-k+1} - \lambda x_{n-k+1} E[p | p \ge p_{n-k+1}]$$

subject to that $x_{n-k+1} \ge x_{n-k}$ and $U(x_{n-k+1}, t_{n-k+1}, p_{n-k+1}) = U(x_{n-k}, t_{n-k}, p_{n-k+1})$. By an analogous argument it follows that $x_{n-k} = x_{n-k+1}$ and $t_{n-k} = t_{n-k+1}$. So the only solution to the insurers's problem is a pooling menu. By (9) applied to $\hat{p} = p_1$, that pooling menu must be a null menu.

Now consider an arbitrary type distribution F. Suppose that (9) holds. Consider a sequence of finite support distribution functions F_n which converge weakly to F and such that (9) holds for all n (Hendren (2013) confirms that such a sequence exists). By the preceding argument the profit at each F_n is 0 and the unique optimal menu is null. Since the monopolist's objective is continuous in the weak convergence topology, the constraint set does not depend on the type distribution, and wlog, the constraint set is compact (in either the relaxed or unrelaxed problem), by Berge's Theorem (e.g. Aliprantis and Border, 1999, Theorem 16.31) the maximum profit at F for the relaxed problem is 0 and the unique optimal menu is the null contract $(0,0)$ given to all types.

To prove that (2) is necessary for a null menu to maximize expected profit, follow Hendren (2013) Lemma A.2) and suppose that (9) does not hold for some $p' \in \mathcal{P}$. Construct a two-contract menu that gives $(0,0)$ to every type below p' and a contract $(x,t) \gg 0$ to every type $p \geq p'$ which leaves type p' indifferent between (x,t) and $(0,0)$. If (x,t) is close enough to $(0,0)$, then this menu yields positive profit to the insurer.

(ii) We will show that $MRS\,(E_\rho[p])$ is an upper bound for $\frac{MRS(p)}{E_\rho[p|p\geq p]}$ for all $p \in \mathcal{P}$.

Consider any $\hat{p} \geq E_\rho[p]$ and assume that $MRS\,(E_\rho[p]) \leq \lambda E_\rho[p]$. Then

$$\lambda \geq \frac{MRS\,(E_\rho[p])}{E_\rho[p]} \geq \frac{MRS\,(\hat{p})}{\hat{p}} \geq \frac{MRS\,(p)}{E_\rho[p|p \geq p]},$$

where the second inequality follows from $MRS(z)/z$ decreasing in z and $\hat{p} \geq E_\rho[p]$, and the third one from $\hat{p} \leq E_\rho[p|p \geq \hat{p}]$. Thus,

$$MRS\,(E_\rho[p]) \leq \lambda E_\rho[p] \Rightarrow MRS\,(\hat{p}) \leq \lambda E_\rho[p|p \geq \hat{p}], \quad \forall \hat{p} \geq E_\rho[p].$$

Consider any $\hat{p} < E_\rho[p]$ and assume that $MRS\,(E_\rho[p]) \leq \lambda E_\rho[p]$. Then

$$\lambda \geq \frac{MRS\,(E_\rho[p])}{E_\rho[p]} > \frac{MRS\,(\hat{p})}{E_\rho[p]} \geq \frac{MRS\,(p)}{E_\rho[p|p \geq \hat{p}]},$$

where the second inequality follows from $MRS(z)$ increasing in z and $\hat{p} < E_\rho[p]$, and the third one from $E_\rho[p] \leq E_\rho[p|p \geq \hat{p}]$. Thus,

$$MRS\,(E_\rho[p]) \leq \lambda E_\rho[p] \Rightarrow MRS\,(\hat{p}) \leq \lambda E_\rho[p|p \geq \hat{p}], \quad \forall \hat{p} < E_\rho[p].$$

Combine the two cases considered to complete the proof. □

A.3 Proof of Proposition 3

Set $\lambda = 1$. The right side of the inequality $k < u'(w)\ell/[u(w) - u(w-\ell)]$ is the marginal willingness to pay for coverage at $(x,p) = (0,0)$. If the inequality holds then complete information profit is positive for all p in a neighborhood of $p = 0$. An insurer whose belief is sufficiently concentrated on that neighborhood earns positive profit. Since $k > 0$ and the risk premium $\Pi(p)$ equals 0 at $p = 1$, $\Pi(p) - kp < 0$ for all sufficiently high p. Let $p_H < 1$ satisfy $\Pi(p_H) - kp_H < 0$. Any insurer belief that is sufficiently concentrated on types with $\Pi(p) - kp < 0$ cannot earn positive profit. Since profit is continuous in λ, the statements still hold for $\lambda \geq 1$.

A.4 Two-Type Case: Properties of $V(\cdot, \lambda, k = 0)$.

In Section 3.3 we asserted that the derivative $V_\rho(1, \lambda, 0)$ of V with respect to the chance ρ that the type is high is always nonnegative at $\rho = 1$; and it is positive if and only if the complete information profit for the high type is positive.

To prove these assertions, let $(x_L(\rho), t_L(\rho), x_H(\rho), t_H(\rho))$ solve (), where we have emphasized its dependence on ρ. By the Envelope Theorem

$$V_\rho(\rho, \lambda) = t_H(\rho) - \lambda p_H x_H(\rho) - (t_L(\rho) - \lambda p_L x_L(\rho)).$$

We must show that $t_H(1) - \lambda p_H x_H(1) \geq t_L(1) - \lambda p_L x_L(1)$, and that the inequality is strict if and only if $\Pi(p_H, \lambda) > 0$. Notice that $(x_H(1), t_H(1))$ is the complete information contract for the high type: thus, the participation constraint for p_H, i.e., $U(x_H(1), t_H(1), p_H) = U(0, 0, p_H)$; and expected profit from $(x_H(1), t_H(1))$ is equal to $\Pi(p_h, \lambda)$, which is nonnegative. The result then follows if we show that $(x_L(1), t_L(1)) = (0, 0)$. Since the constraint set is compact, the objective function is continuous in ρ, and the solution is unique, the menu is continuous in ρ, so $(x_L(1), t_L(1), x_H(1), t_H(1)) = \lim_{\rho \to 1}(x_L(\rho), t_L(\rho), x_H(\rho), t_H(\rho))$. Moreover for every $0 < \rho < 1$, $(x_L(\rho), t_L(\rho), x_H(\rho), t_H(\rho))$ is nonnegative, $U(x_L(\rho), t_L(\rho), p_L) = U(0, 0, p_L)$, $U(x_L(\rho), t_L(\rho), p_H) = U(x_H(\rho), t_H(\rho), p_H)$, and these properties are preserved in the limit: $(x_L(1), t_L(1), x_H(1), t_H(1))$ is nonnegative, $U(x_L(1), t_L(1), p_L) = U(0, 0, p_L)$, and $U(x_L(1), t_L(1), p_H) = U(x_H(1), t_H(1), p_H)$. The strict single crossing property and $U(x_H(1), t_H(1), p_H) = U(0, 0, p_H)$ then implies that $(x_L(1), t_L(1)) = (0, 0)$, so $V_\rho(1, \lambda) = t_H(1) - \lambda p_H x_H(1) = \Pi(p_H, \lambda) \geq 0$ and the inequality is strict if and only if $\Pi(p_H, \lambda) > 0$. □

A.5 Proof of Proposition 4

[To be completed...]

We prove a more general result. Let $\mathcal{P}$ be the type support and let p_1 be its smallest element.

Suppose that the complete information profit $\Pi(p, \lambda) = 0$ for all $p \in \mathcal{P} - p_1$. Let F dominate G in the first-order stochastic dominance sense (FOSD), which is implied by LR. Since complete information expected profit is 0 for every type except possibly p_1, the menu is pooling, and if any type higher than p_1 gets a positive contract, expected profit from that type is negative. Let $(x_1(F), t_1(F))$ be the optimal pooling menu at belief F. Then

$$
\begin{aligned}
V(F) &= F(p_1)[t_1(F) - \lambda p_1 x_1(F)] + (1 - F(p_1))t_1(F) - \lambda x_1(F) \int_{\mathcal{P}-\{p_1\}} p dF(p) \\
&= [t_1(F) - \lambda p_1 x_L(F)] - \lambda x_1(F) \int_{\mathcal{P}-\{p_1\}} (p - p_1) dF(p) \\
&\leq [t_1(F) - \lambda p_1 x_1(F)] - \lambda x_1(F) \int_{\mathcal{P}-\{p_1\}} (p - p_1) dG(p) = V(G),
\end{aligned}
$$

where the first equality follows from adding and subtracting $(1 - F(p_1))\lambda x_1(F)$ and rearranging, and the inequality holds since $p - p_1 \geq 0$ on $\mathcal{P} - \{p_1\}$ and F FOSD G implies that $F/(1 - F(p_1))$ FOSD $G/(1 - G(p_1))$. Hence, $(1 - F(p_1)) \int_{\mathcal{P}-\{p_1\}} (p - p_1) dF(p)/(1 - F(p_1)) \geq (1 - G(p_1)) \int_{\mathcal{P}-\{p_1\}} (p - p_1) dG(p)/(1 - G(p_1))$. Thus, $\Pi(p, \lambda) = 0$ for every $p \in \mathcal{P} - \{p_1\}$ implies $V(F) \leq V(G)$ when F FOSD dominates G.

We prove the other direction by contraposition. Suppose that $\Pi(p_j, \lambda) > 0$ for some $1 < j \leq n$, and let m be the *largest* such type. Then there is a distribution μ over the set of types $\{p_m, ..., p_n\}$ that puts positive probability on each type and would give the insurer positive expected profit (before subtracting the fixed entry cost). The profit maximium menu at μ moreover is pooling since the complete information profit for types $m + 1, ..., n$ is 0. Let ν be any distribution that puts positive probability on each type in $\{1, ..., m - 1\}$. Let ρ_α be given by $\rho_\alpha = (1 - \alpha)\nu + \alpha\mu$. [Informal: there is an $\hat{\alpha}$ such that $\alpha > \hat{\alpha}$ implies that types 1,...m-1 get (0,0). on that interval expected profit increases in α.] □

A.6 Limiting V as Risk Aversion Increases

We asserted at the end of Section 3.3 that V converges to a decreasing function as risk aversion increases uniformly without a bound on $[w - \ell, w]$. We now prove this assertion.

Complete Information: Decreasing Π. We first show the result for the complete information case. It suffices to show that the conclusion holds for CARA preferences (since the revenue $T(x,p)$ for any vN-M utility that is uniformly more risk averse than some given CARA utility must lie in between the CARA risk premium and the upper bound $\max\{\ell(1-\lambda p),0\}$). Fix $p \in (0,1]$. If $\lambda p \geq 1$, then clearly $\Pi(x,p)=0$ for any $0 \leq x \leq \ell$, so suppose that $\lambda p \leq 1$. The willingness-to-pay $T(x,p,r)$ for coverage $x \in [0,\ell]$ for a CARA vN-M utility with risk aversion equal to r is

$$T(x,p,r) = \frac{1}{r}\ln\left[\frac{1-p+pe^{r\ell}}{1-p+pe^{r(\ell-x)}}\right] < x$$

for every $p < \lambda^{-1}$ and $r > 0$. Routine calculations confirm that the value of x which maximizes $T(x,p,r) - x\lambda p$ is

$$x^* = \ell - \frac{1}{r}\ln\left[\frac{\lambda(1-p)}{1-\lambda p}\right]$$

so the insurer's complete information value function is

$$\Pi(p,r) = \frac{1}{r}\ln\left[\frac{1-p+pe^{r\ell}}{p\frac{\lambda(1-p)}{1-\lambda p}+1-p}\right] + \frac{\lambda p}{r}\ln\left[\frac{\lambda(1-p)}{1-\lambda p}\right] - \lambda p\ell$$

which converges to $\ell(1-\lambda p)$ as $r \to \infty$. □

Adverse Selection: Decreasing V. We will show that as risk aversion increases without bound, the optimal menu converges to a pooling menu at full insurance, which yields expected profit equal to $\ell(1-\lambda E_\rho[p])$. Consider a sequence u_k of vN-M utilities that have absolute risk aversion of at least k at every point in $[w-\ell,w]$. Let $T_k(x,p)$ be the willingness to pay of a type-$p \in \mathcal{P}$ consumer with vN-M utility u_k for coverage of $0 \leq x \leq \ell$. We have the following inequalities for every k

$$\ell(1-\lambda E_\rho[p]) \geq E_\rho\left[\Pi_k(p)\right] \geq V_k(\rho) \geq T_k(\ell,p_L) - \lambda\ell E_\rho[p], \tag{10}$$

the first since $\Pi_k(p) \leq \ell(1-\lambda p)$ for every p, the second since $\Pi_k(p)$ is the complete-information expected profit that the insurer can extract from p, and the third since it is feasible to pool both types at full insurance with premium equal to the willingness-to-pay of the low-risk consumer, $T_k(\ell,p_L)$. As k goes to infinity the consumer becomes infinitely risk averse in the limit, and it follows from the participation constraint that $\lim_{k\to\infty} T_k(\ell,p_L) = \ell$. Hence, taking limits in (10) yields $\lim_{k\to\infty} V_k(\rho) = \ell(1-\lambda E_\rho[p])$, which is a strictly decreasing function of ρ. $\square$

A.7 Optimal Complete-Information Menus: Proposition 5 and Example 1

To simplify the presentation, let us assume in this result that u is C^2 on R_{++}. For each p, the complete information contract with loading solves $\max_{x,t} t - \lambda p x$ subject to $U(x,t,p) = U(0,0,p)$. Since $U(x,t,p)$ is strictly decreasing in t, we can invert the constraint and write it as $t = T(x,p)$; it is easy to check that T is increasing in p. Then the problem becomes $\max_x T(x,p) - \lambda p x$, and to show that its solution strictly decreases in p it suffices to prove that $T(x,p)$ satisfies strictly decreasing differences in (x,p). Now, $T_x(p,x) = (U_x/U_t)(T(x,p),x,p)$, and it is easy to show that $T_{xp} < 0$ if and only if

$$- T_p p(1-p) u'_\ell u'_n (R_n - R_\ell) \leq 0, \tag{11}$$

where $u_\ell = u(w - \ell + x - t)$, $u_n = u(w-t)$, and R_ℓ and R_n is the coefficient of absolute risk aversion evaluated at the loss and no loss states, respectively.

It follows from (11)that x strictly *decreases* in p if u exhibits increasing absolute risk aversion (e.g., $R_n \geq R_\ell$), which includes CARA as special case.

Using the first-order condition in the CARA case (see Section A.6), it is easy to verify that the optimal reimbursement in the complete information case is the one given in Example 1. $\square$

A.8 Proof of Proposition 6

Since $\lambda < u'(w-\ell)/u'(w)$, there is a point $p_L \in (0,1)$ satisfying $MRS(p_L)/p_L > \lambda$. Moreover since $MRS(\lambda^{-1}) > 1$, there is a $p' \in (p_L, 1/\lambda)$ such that $MRS(p')/p' < \lambda$. Let p_H be any point in $(p', 1/\lambda)$ and let $\mathcal{P}$ be *any* type set containing the three points p_L, p', and p_H (for example $\mathcal{P} = [p_L, p_H]$). Let ρ_n be any sequence of cumulative distribution functions that are each strictly increasing on $\mathcal{P}$ and that converge weakly to the distribution that puts probability 1 on p_L. Since expected profit is continuous in the weak convergence topology and is positive at the limiting distribution, part (a) follows. Now consider any full-support belief ρ with positive expected profit, and let $\{x(\cdot), t(\cdot)\}_{p\in\mathcal{P}}$ maximize expected profit at ρ. Since $MRS(p')/p' < \lambda$, the maximum complete-information profit for any type p' or higher is 0; so in any feasible menu (in particular one in which (P) holds), $t(p)-\lambda px(p)\leq 0$ for any $p\in[p',p_H]$. Since expected profit from the menu is positive, $t(p)-\lambda px(p)>0$ for some $p\in[p_L,p')$, so $x(p)>0$. Since the menu is nondecreasing, $x(p')>0$. And since the complete-information contract for type p' is $(0,0)$ and the consumer satifies non-increasing absolute risk aversion, the efficient contract for type p' that is indifferent to $(x(p'),t(p'))$, must specify 0 coverage; so the contract for p' is distorted upwards from its efficient contract. Finally, by the logic of Lemma 1 in Chade-Schlee (2012), every $p\in[p',p_H]$ is pooled at $(x(p'),t(p'))>0$ (since an efficient contract for any higher type also gives 0 coverage). So there is pooling and inefficiency at the highest type. □

A.9 Optimal Pooling and Distortions: Example 2

Let $v(p,x)=-r^{-1}\log[pe^{r(\ell-x)}+(1-p)]$. The optimal menu in the CARA case solves

$$\max_{x_L,x_H,t_L,t_H}\rho[t_H-\lambda p_H x_H]+(1-\rho)[t_L-\lambda p_L x_L]$$

subject to $v(p_L,x_L)-t_L=v(p_L,0)$, $v(p_H,x_H)-t_H=v(p_H,x_L)-t_L$, and $x_H\geq x_L$.

Use the first two constraints to solve for t_H and t_L and rewrite the problem as

$$\max_{x_L,x_H}\rho[v(p_H,x_H)-v(p_H,x_L)+v(p_L,x_L)-v(p_L,0)-\lambda p_H x_H]+(1-\rho)[v(p_L,x_L)-v(p_L,0)-\lambda p_L x_L]$$

subject to $x_H\geq x_L$. Let us ignore the constraint and solve for x_H and x_L. The first-order conditions of this relaxed problem are

$$v_x(p_H,x_H) = \lambda p_H \tag{12}$$

$$v_x(p_L,x_L) = \rho v_x(p_H,x_L)+(1-\rho)\lambda p_L. \tag{13}$$

If the solution to these equations satisfy the omitted constraint with slack, then the optimal menu entails complete sorting. If violated or satisfied with equality, then the optimal menu pools both types. The goal is to find conditions on the problem's parameters so that the optimal solution is pooling. We will use the following change of variables: $z_i=e^{-r(\ell-x_i)}$, $i=l,h$. Then $x_H\geq x_L\Leftrightarrow z_H\geq z_L$.

Equation (12) reveals that the optimal value for z_H is $z_H^*=(1-p_H\lambda)/(\lambda(1-p_H))$. Since we want an interior solution $x_H\in(0,\ell)$, we need $z_H^*\in(e^{-r\ell},1)$. This holds if $\lambda\in(1,(p_h+(1-p_he^{-r\ell})^{-1})$, a parametric restriction that we henceforth impose.

Consider equation (13). It can be written as

$$\frac{1+\left(\frac{1-p_H}{p_H}\right)z_L}{1+\left(\frac{1-p_L}{p_L}\right)z_L} = (1-\rho)\lambda p_L\left(1+\left(\frac{1-p_H}{p_H}\right)z_L\right)+\rho. \tag{14}$$

It is easy to verify that the left side starts above the right side for low values of z_L and it lower than it for large values. Also, the left side is strictly decreasing in z_L while the right side is strictly increasing. Thus, there is a unique solution z_L^* that solves it.

If we set $\rho = 0$ in equation (14) we obtain the complete information solution for the low type $z_L^* = (1-p_L\lambda)/(\lambda(1-p_L))$, and we know from Example 1 that this is greater than z_H^*. By continuity, this is true for ρ sufficiently small. Therefore, pooling is optimal for values of ρ in a right-neighborhood of $\rho = 0$.

Regarding other parameters of the model, notice that a sufficient condition for pooling to be optimal is that the left side of (14) evaluated at z_H^* be larger than the right side evaluated at that point. This holds if and only if

$$\frac{\lambda(1-p_H)+\left(\frac{1-p_H}{p_H}\right)(1-p_H\lambda)}{\lambda(1-p_H)+\left(\frac{1-p_L}{p_L}\right)(1-p_H\lambda)} \geq (1-\rho)\frac{p_L}{(1-p_H)}\left((1-p_H)\lambda+\left(\frac{1-p_H}{p_H}\right)(1-p_H\lambda)\right)+\rho.$$

Notice that if $\lambda = 1/p_H$ then this inequality strictly holds, but this violates the condition $\lambda < (p_h + (1-p_h e^{-r\ell})^{-1} < 1/p_H$ for an interior x_H. Let $\varepsilon > 0$ be sufficiently small so that the inequality holds if $\lambda \in ((1/p_H)-\varepsilon, 1/p_H)$. Suppose r is large enough so that $(p_h + (1-p_h e^{-r\ell})^{-1} < (1/p_H)-\epsilon$. Then for any λ close to $1/p_H$, pooling is optimal for sufficiently large values of r. Notice that the same can be done for large values of ℓ (this requires adjustments in w to keep $w > \ell$, which can be done in the CARA case).

Set $x = x_H = x_L$ and $t = t_H = t_L$ into the insurer's problem and solve for the optimal (x,t) to get the equations given in Example 2. □

References

Arrow, K. (1963): "Uncertainty and the Economics of Medical Care," *American Economic Review*, 53, 941–973.

——— (1965): "Uncertainty and the Welfare Economics of Medical Care: Reply

(The Implications of Transaction Costs and Adjustment Lags)," *American Economic Review*, 55, 154–158.

Boland, V. (1965): "Uncertainty and the Welfare Economics of Medical Care: Comment," *American Economic Review*, 55, .

Chade, H., and E. Schlee (2012): "Optimal Insurance in Adverse Selection," *Theoretical Economics*, 7, 571–607.

Diamond, P. (1977): "Insurance theoretic aspects of workers' compensation," in *Natural resources, uncertainty, and general equilibrium systems : essays in memory of Rafael Lusky*, ed. by A. S. B. Rafael Lusky, and P. Friedman. Academic Press.

Einav, L., A. Finkelstein, and J. Levin (2010): "Beyond Testing: Empirical Models of Insurance Markets," *Annual Reviews of Economics*, 2, 311–336.

Gruber, J. (2008): "Covering the Uninsured in the United States," *Journal of Economic Literature*, 46, 571–606.

Guesnerie, R., and J.-J. Laffont (1984): "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-Managed Firm," *The Journal of Public Economics*, 25, 329–369.

Hellwig, M. (2010): "Incentive problems with unidimensional hidden characteristics: a unified approach," *Econometrica*, 78, 1201–1237.

Hendren, N. (2013): "Private Information and Insurance Rejections," *Econometrica*, forthcoming, 407–430.

Kihlstrom, R. E. (1984): "A 'Bayesian' Exposition of Blackwell's Theorem on the Comparison of Experiments," in *Bayesian models of economic theory*, ed. by M. Boyer, and R. E. Kihlstrom, pp. 13–31. Elsevier.

Lees, D., and R. Rice (1965): "Uncertainty and the Economics of Medical Care: Comment," *American Economic Review*, 55, 140–154.

McFadden, D., C. Noton, and P. Olivella (2012): "Remedies for Sick Insurance," NBER Working Paper.

MILGROM, P. (1981): "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics*, 12, 303–350.

MIRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

MORAND, P.-H., AND L. THOMAS (2003): "On Non-responsiveness in Adverse Selection Models with Common Value," *The B.E. Journals in Theoretical Economics: Topics in Theoretical Economics*, 3, .

REES, R., AND A. WAMBACH (2008): "The Micoreconomics of Insurance," *Foundations and Trends in Microeconomics*, 4, 1–163.

ROYDEN, H. (1968): *Real Analysis*. Macmillan, Cambridge, MA.

SHAVELL, S. (1977): "Theoretical Issues in Medical Malpractice," in *The Economics of Medical Malpractice*, ed. by R. Helms. American Enterprise Institute.

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