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

Hector Chade[†]
Arizona State University
Department of Economics

Edward Schlee[‡]
Arizona State University
Department of Economics

September 22, 2014

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: the only cost is claims payment. We introduce costly insurance provision into a 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 those likely to be the worst risks. Unexpectedly, it turns out that loading also 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. Indeed all types can be pooled on a single contract. Finally we show that both loading and a fixed claims cost do *not* affect qualitative properties of (Rothschild-Stiglitz) competitive equilibrium (when it exists), so these costs generate potentially testable implications of competitive vs monopoly insurance: for example, no competitive equilibrium can be pooling.

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

[*]We thank Dan Silverman for pointing out Hendren's work to us. We also thank Dan, Rohan Dutta, Liran Einav, Nathaniel Hendren, Ilse Lindenlaub, Asaf Plan, François Salanié and participants at the 2013 SAET conference in Paris, the Economic Theory Workshop at the University of Montreal, the University of Arizona, and Universitat Pompeu Fabra 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, and Chade's visit to Pompeu Fabra in the Summer 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 choose 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* correlated with the consumer's loss chance. But adding a signal to the standard monopoly insurance model is not enough to account for coverage denials: Chade and Schlee (2012) find that the monopolist *always* makes positive expected profit: there are *ex ante* gains to trade between the insurer and the consumer, no matter how pessimistic the insurer's beliefs about the consumer's loss chance.

In the standard model, the insurer's only cost is payment of claims: *insurance provision itself is costless*. Yet both the empirical insurance and practitioner literatures extensively discuss provision costs.[2] So besides the signal, we introduce realistic insurance provision costs into an otherwise standard model. We find that adding these costs can account for coverage denials for *only* those believed to be 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

[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. 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)).

of insurance.

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 show that if there are no gains to trade at a belief about a consumer's loss chance, 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 (the only provision costs are loading and a fixed claims cost).

At one level, this comparative static result answers our question about excluding only bad risks; it does not, however, pin down who is denied coverage and who is not. But in the special case of pure loading, we give 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.

We also consider a fixed entry cost, but the result here is more negative: the necessary condition for excluding only the worst risks is quite strong (Proposition 4).

These results demarcate when there are gains to trade or not with insurance provision costs. A natural question to ask is how these costs affect properties of optimal menus when there *are* gains to trade. Chade and Schlee (2012) confirm that, without provision costs, three classic contracting properties hold for monopoly insurance under adverse selection: no consumer risk type pools with the 'highest' type, the one with the highest loss chance (*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 fail with loading. Intuitively, 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)). Finally, we show that the classic properties hold if instead of monopoly we consider the competitive model of Rothschild and Stiglitz (1976).

Our results show that a modest and realistic departure from the standard model dramatically affects its predictions. They might prove useful in applied work that tests the implications of insurance models with adverse selection. For instance, the exclusion of only bad risks that we can account for is of obvious empirical relevance (see Hendren

(2013) for a thorough discussion and evidence on this point). Also, that pooling emerges in monopoly and not under competition provides an additional test for models with adverse selection. The emergence of pooling under provision costs also might help explain why menus often have one or only a few contracts in reality, without resorting to a behavioral explanation. Finally, these results could be useful for policy discussions on health care reforms, where market structure, exclusion, and administrative costs are part of the debate.

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 under asymmetric information. 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). At the end of Section 3, we consider a potential rival explanation for coverage denials, regulation. 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 (a cumulative distribution function) 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 abuse notation by letting $\rho \in [0, 1]$ denote the insurer's belief that the consumer is high risk, 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 event 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

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

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 the insurer observes a *signal* correlated with the consumer's loss chance before offering a menu. The signal is *public* (i.e., both parties observe its realization), and could be either costly or costless, and its informativeness could be either exogenous or endogenous. An important example 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 signal realization.[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

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

[4] That is, the menu that offers the contract $(x, t) = (0, 0)$ to all types. Our results all go through for any differentiable cost function $c(x, p)$ that is convex in x for every $p \in \mathcal{P}$; $c_x(x, p) > p$ for every $x \geq 0$, $p \in \mathcal{P}$, 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.

[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). Also, we assume realistically that the signal is observed before writing a menu and that it is public. If it were privately observed by the principal, then the model becomes a complex informed principal problem with two sided private information and nonlinear utility on the agent's side. And if it were revealed after the mechanism is played, then the principal could benefit from the correlation between the signal and the agent's type by conditioning the contracts on both. Although these variations are interesting theoretical extensions, they do not seem as relevant for the problem at hand as our formulation.

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 present the main results for the case of loading without any fixed costs, and for fixed costs without loading. In most cases it will be clear how the results change when all of the costs are present.

3 Coverage Denials: Excluding Only Bad Risks

As mentioned, Chade and Schlee (2012) show that there are always *ex ante* gains to trade between the insurer and the consumer when insurance provision is costless.[10] Obviously, there are no gains to trade if these costs are large enough. What is not obvious is what restrictions on these costs lead the insurer to 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).

[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).

[10] They assume that the type set does not contain $p = 1$.

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 does not lower aggregate expected profit from the other types.

Assumption 1. *If $\pi(\cdot) \in \Phi$, then for any $p' \in \mathcal{P}$, there is a $\pi'(\cdot) \in \Phi$ with $\pi'(p) = 0$ for $p < p'$ and $\int_{[p',1]\cap\mathcal{P}} \pi'(r)d\rho(r) \geq \int_{[p',1]\cap\mathcal{P}} \pi(r)d\rho(r)$.*

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); more generally it holds in an adverse selection insurance model in which the cost of any no-trade contract is zero.

Milgrom (1981) proves that one signal is better or worse news than another—in the sense that the resulting posterior beliefs are ordered by first-order stochastic dominance for *every* prior—if and only if the signal distributions are ordered by *Likelihood Ratio* (LR) dominance. For the special case of simple distributions or density (wrt Lesbegue measure) functions with the same support, g *LR* dominates f if g/f is increasing. We give a more general formal definition in Appendix A.1, right before the proof of Proposition 1.

Proposition 1 (Bad Risks and Coverage Denials)**.** *Suppose Assumption 1 holds and that there are no gains to trade at a beliefabout the consumer's type. Then there are no gains to trade at any belief that LR dominates it with the same support.*

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 those that are believed to be the worst risks after observing the realization of the signal. We will show in what follows that the premise of Proposition 1 holds non trivially with loading or a fixed claims cost. Thus, a parsimonious and realistic departure from the standard model such as the introduction of these costs can rationalize the exclusion of bad risks from insurance.

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 consumer's type. By Assumption 1, the insurer loses money on the high type (otherwise the insurer would get positive profit by giving the low-risk type zero coverage). So if the insurer cannot earn positive profit at some belief, it must be that any profit from the low risk consumer is not enough to make up for losses from the high risk. If the high risk consumer 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 uses only LR dominance and Assumption 1. Assumption 1 holds if i) the consumer's preferences satisfy the strict single crossing property in p and (x, t); 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 in any menu. Under these four conditions, we can give the null contract to every type less than the threshold type p'. Then either the menu which leaves contracts to other types the same is feasible (in which case we are done); or it is not. If not, by the single-crossing property and ii), the new menu is downward incentive compatible: so the only violation of IC is that some types with null contracts envy some contracts given to higher types and (P) is slack for type p' (and higher types by ii). But then problem of picking a menu for types greater than or equal to p' subject to (P) imposed only on those types gives higher profit. So Proposition 1 holds across a range of monopoly pricing problems with privately informed consumers. Since our insurance model satisfies i)-iii) Assumption 1 holds if the expected cost for $x = 0$ is 0 for every p and menu $x(\cdot)$; but with a fixed entry cost Assumption 1 fails.

Proposition 1 says that if expected profit is 0 at some belief, it remains 0 for any rightward LR shift in beliefs. One might conjecture that the monopolist's profit is globally decreasing in the LR order, a great simplification. Unfortunately this monotonicity holds only under an extremely restrictive assumption, as we confirm in Proposition 4.

An important application of Proposition 1 is to Hendren (2013), who determines whether the estimated loss distribution of potential consumers who are denied coverage have fatter right tails than consumers who get coverage. Our 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 given 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) implies that there are no gains to trade—coverage is denied—if and only if there are no gains to trade with a *two*-contract menu, with one the zero contract and the other a small, positive contract, since the right side of (2) is the slope of the insurer's isoprofit line from a contract that pools all types above $\hat{p}$. In the two-type case, $E_\rho[p|p \geq \hat{p}]$ 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 of Proposition 1 to *hazard rate* (*HR*) dominance.[11] This

[11] For cumulative distribution functions defined on $[0, 1]$, G hazard rate (HR) dominates F if $(1 -$

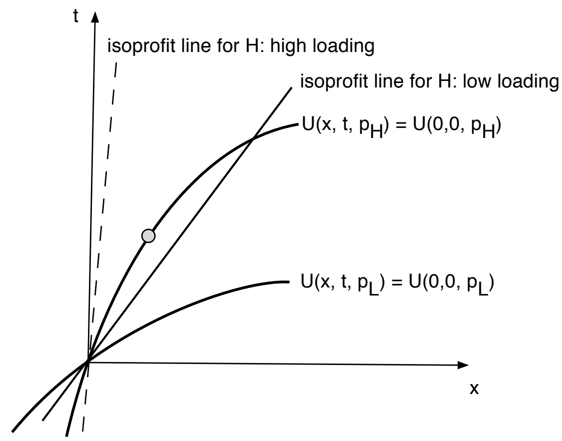

Figure 1: **Loading and Coverage Denials, Two-Type Case: Condition (2) for the High Type.** The high-risk consumer's indifference curve is the steeper of the two. The dashed line is an insurer iso-expected-profit line for sufficiently high loading. For that case condition (2) holds at $\hat{p} = p_H$.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

result gives us an additional testable implication beyond Proposition 1 for the case of pure loading. It follows from Fact 1.B.12 in Shaked and Shanthikumar (2007).

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 summarizes the content of Proposition 2 (*i*) for the two-type case. The figures also reveal the economic principle underlying condition (2): given the consumer's preferences and the loading factor, it is unprofitable for the insurer to provide

$G(p))/(1 - F(p))$ 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.

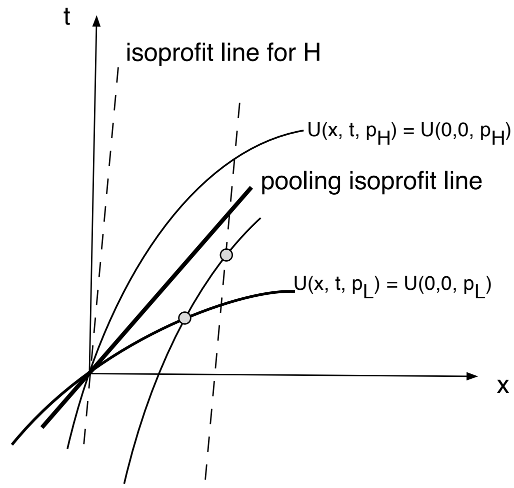

Figure 2: **Loading and Coverage Denials: Condition (2) for the Low Type.** Inequality (2) holds for $\hat{p} = p_L$ and for $\hat{p} = p_H$. Expected profit from the high risk consumer is negative for any contract this consumer is willing to buy. To make positive expected profit, the low-risk type must buy a positive contract. Since (2) holds at $\hat{p} = p_H$ and u is strictly concave, the best hope for positive profit is a pooling contract: concavity of u implies that the marginal rate of substitution for the high type at any contract (x, t) with $x > t > 0$ is less than $MRS(p_H)$. But since (2) holds at $\hat{p} = p_L$, no such pooling contract can be profitable.

even a small amount of coverage to a given type when it pools all types above it. Figure 2 also suggests why pooling is without loss of generality.

When $\lambda = 1$, Proposition 2 (i) specializes to Theorem 1 in Hendren (2013); our Proposition 2 (i) extends his result to the case of loading.[12] 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) 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 just multiplies the loss chance in the insurer's cost, one might conjecture that a necessary condition for coverage denials with loading is that $\lambda p \geq 1$ for the highest type. Equation (2) shows 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 (2) holds as a strict inequality for all sufficiently small losses. Intuitively, when the loss size is small, the

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

consumer's demand for insurance is low, and loading makes insurance unprofitable;[13] the conclusion holds even though $\lambda p < 1$ for every type.[14]

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 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 $\lambda = 1$.[15]

This proposition is also useful for distinguishing our explanation for coverage denials with an alternative explanation: regulation of premiums. A conjecture is that an insurer might deny coverage to the worst risks if premiums are capped by regulation (Tennyson (2007)). But Proposition 2 shows that there are gains to trade if and only if there are gains to trade with a single, "small" contract. Even without explicit premium regulation, insurers might voluntarily cap premiums out of mere fear of a regulatory response, or for reputational reasons. But again, Proposition 2 implies that a high-premium contract is profitable only if a "small" contract is profitable. Hence, explicit regulation of premiums or voluntary premium caps alone cannot explain coverage denials. If a minimum coverage requirement is added to a premium cap, then coverage denials only to the worst risks can be explained without provision costs. But the recent cancellation of private health insurance policies because they did not meet the minimum coverage requirements of the Affordable Care Act highlights the existence of low-premium, low-coverage policies in this market.[16] Whatever the fraction of coverage denials that regulation might account for, our modest and realistic departure from the standard insurance model explains why coverage denials would occur in the complete absence of price or quantity regulation.

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

[14] 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.

[15] 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 $\lambda > 1$.

[16] The National Association of Insurance Commissioners (2011) report that 31 states in the U.S. merely required that premiums be actuarially justified in the private health insurance market: high premia were allowed for bad risks provided that firms can show that they are indeed bad risks.

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? More precisely, we ask whether our foundational result, Proposition 1, holds 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 $0 < k < ([u(w) - u(w - \ell)]/u'(w)) - \ell$. 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.*

Conditions for no trade with a fixed entry cost are more difficult to pin down. In what follows we set $k = 0$. To understand the difficulty, 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 under complete information 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.

Proposition 4 (Fixed Entry Cost)**.** *Fix $k = 0$ and $\lambda \geq 1$, and assume beliefs have full support on $\mathcal{P}$. Then (i) coverage is denied only to those likely to be the worst risks for every $K \geq 0$ if and only if monopoly profit $V(\cdot, \lambda, 0)$ is decreasing in the LR order on $\mathcal{P}$; and (ii) the last condition holds* ***only if*** *complete-information monopoly profit is 0 for every type p except the lowest type in $\mathcal{P}$.*

The result implies that we cannot reconcile coverage denials for only bad risks with a fixed entry cost—at least not for plausible assumptions on the set of types. Indeed if the support of types is an interval, Proposition 4 implies that V cannot be decreasing in the LR order unless there are never gains to trade: since complete information profit is continuous in p, it cannot be 0 for all but one type.

More positively, as risk aversion increases uniformly without bound, the insurer value function V converges pointwise to a function that *is* decreasing in the *LR* order (Section A.6): 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 general contracting menus with private information 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*). Chade and Schlee (2012) confirm that these hold with costless insurance provision.[17] We now consider whether they hold with costly insurance provision.

Since fixed costs do not affect those properties, we consider pure loading and no fixed costs. If a nonzero menu maximizes expected profit when $K > 0$, then clearly the same menu maximizes expected profit when $K = 0$. Thus, nonzero menus are unaffected by an entry cost. Although a fixed claims cost $k > 0$ can affect the profit maximizing menu, the three classic properties hold in the absence of loading: if $k > 0$, and $\lambda = 1$, then any profit-maximizing menu satisfies efficiency at the top, no pooling at the top, and downward distortions elsewhere. The proof is almost identical to the proof of Theorem 1 in Chade and Schlee (2012).

Assume from now on that $K = k = 0$ and $\lambda \geq 1$. We begin with the analysis of the complete-information case.

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 strikingly, 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 or large enough loading factor suffice for the conclusion.

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

Proposition 5 (Complete Information)**.** *The complete information menu is nonincreasing, and strictly decreasing on the set of types with positive coverage, under either one of these conditions.*

(i) The consumer's preferences satisfy non-decreasing absolute risk aversion.

(ii) The loading factor is large enough (how large depends on u, w and ℓ).

A simple intuition for part (i) comes from considering analogues of substitution and wealth effects for 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, and the result holds for λ sufficiently large. Figure 3 illustrates these effects.

To understand part (ii), note first that an increase in loading λ decreases coverage. With small coverage the extra premium that the insurer can extract from an increase in p is small, so the wealth effect is small; in the limit as coverage tends to zero as λ becomes large, the wealth effect vanishes. The substitution effect however does not vanish as coverage shrinks to zero.

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. In the textbook analysis of the two-type case, the profit-maximizing menu gives the high type full insurance and sorts the two types, with the low type getting coverage that is less than its first-best (full) coverage. These three classic properties—*no pooling at the top*, *efficiency at the top*, *downward*

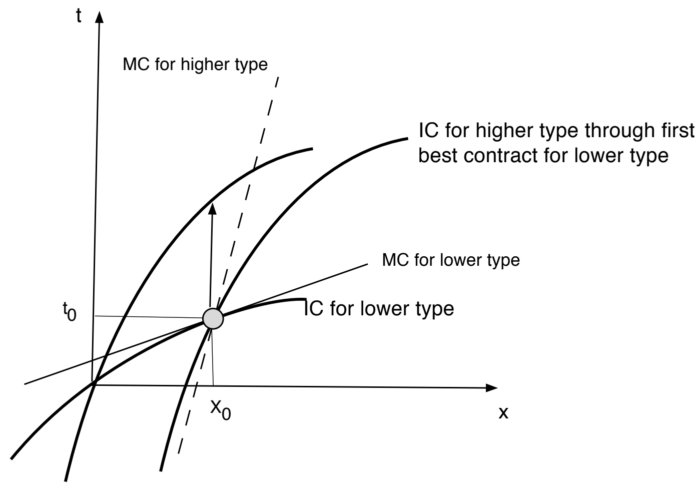

Figure 3: **Decreasing Complete-Information Menus: Substitution and Wealth Effects** The point (x_0, t_0) is the complete information contract for type p_0. If $p_1 > p_0$, then the slope the marginal cost rises faster than the slope of the indifference curve through (x_0, t_0), a substitution effect.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

distortions elsewhere—hold for an arbitrary type distribution when insurance provision is costless. Surprisingly, each of these three properties *fail* with loading, as we now show.

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 $t - \lambda p x$ on $\{(x, t) \in \mathcal{C} \,|\, U(x, t, p) = U(x', t', p)\}$, the set of contracts that are indifferent to (x', t') for a type-p consumer. A contract (x'', t'') is *distorted downwards* from an 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)**.** *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.*[18]

To prove Proposition 6 we construct an insurance problem in which the complete-information 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 complete-information 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 $\lambda p_H < 1$ 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)).

[18] 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).

5 Competition

Since provision costs dramatically affect monopoly insurance, it is natural to ask how they affect competitive insurance contracts. Here we consider the competitive model of Rothschild and Stiglitz (1976). They define a competitive equilibrium to be a set of contracts such that, when consumers choose contracts to maximize expected utility, no contract in that set makes negative profit; and, given this set, no contracts outside the set would earn positive profit if offered.[19]

With a fixed entry cost ($K > 0$) that is sunk before a menu is offered, at most one firm enters and the equilibrium outcome is the same as our monopoly model: since post-entry competition gives zero expected profit, the fixed entry cost cannot be recovered if more than one firm enters. So from now on we set the entry cost to zero.

When an equilibrium exists in the Rothschild-Stiglitz model (with two types), (i) there is no pooling; (ii) the high type gets an efficient contract (full insurance in their case); (iii) the low risk consumer's contract is distorted downwards from efficiency. These properties straightforwardly extend to more than two types—if in condition (iii), the word"low" is replaced by "each lower" (see Appendix A.10).

Properties (i) - (ii) follow from two facts: consumer preferences satisfy the strict single-crossing property; and at any given nonegative contract expected profit is lower for higher types. These properties hold under loading or a fixed claims cost. *So properties* (i) - (ii) *continue to hold for a competitive equilibrium with these provision costs* (as we confirm in Appendix A.10). Property (iii) continues to hold if complete information coverage is nonincreasing in type (including the nonresponsive case), but can fail if it is strictly increasing in type: a competitive equilibrium can be *complete information* Pareto optimum.

The stark difference between how provision costs affect on monopoly and competitive insurance is potentially important for *empirical* work on the subject. For example, one interesting implication of these facts is a way to distinguish empirically between (Rothschild-Stiglitz) competition and monopoly under adverse selection when insurance provision is costly: observing a menu with just one nonzero contract (pooling) is consistent with monopoly, but *not* with competition under adverse selection.

Another empirical test emerges from the exclusion of only bad risks. Consider prop-

[19]This equilibrium notion can be reformulated as 'Bertrand competition' with two or more firms simultaneously offering a set of contracts. See Kreps (1990), p. 649.

erty (*ii*) just alluded to, that the highest risk type gets an efficient contract. It implies that, if a competitive equilibrium exists, there are nonzero contracts if and only if the complete information (zero-profit) contract for the highest risk type is nonzero: *provided that the lowest type remains the same, the distribution of risk types is irrelevant for whether or not there are gains to trade.* This fact makes it unlikely that we can reconcile coverage denials for only the worst risks with (Rothschild-Stiglitz) competition.

6 A Summing Up

There are always gains to trade in the standard monopoly insurance model with costless insurance provision, in the sense that the insurer is willing to trade with some consumer risk types. Some risk types go uninsured, but only those with the smallest loss chances, and they do so voluntarily. This voluntary uninsurance does not fit the evidence on the large number of people who are denied coverage because they are perceived to be 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 the kind of cost matters so much, it would be useful to measure each better.

We stress that our results have important implications for empirical work on insurance, which we have pointed out throughout the text. One open theoretical question also of empirical relevance is how provision costs affect 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 decreases whenever $x(p)$ increases). This result is important since several scholars find evidence of quantity discounts in insurance, and often interpret them as evidence against adverse selection. A fixed entry cost clearly has no affect on quantity discounts. We conjecture that the optimal menu with loading or a claims cost could still be consistent with global quantity discounts, but leave this conjecture for future work.

A Appendix

A.1 Proof of Proposition 1

When two beliefs have the same support and each has either a probability mass function or density (wrt Legesgue measure), then the two beliefs are mutually absolutely continuous. Our definition of LR dominance imposes this condition explicitly.[20]

Definition 1. *Let ρ_1, ρ_2 be two beliefs (cdfs). The belief ρ_2* likelihood ratio (LR) dominates ρ_1 *if there are points $0 \leq p_L < p_H \leq 1$ such that*

(i) $\rho_1(p_H) = 1$ and $\rho_2(p) = 0$ for $p < p_L$;

(ii) the beliefs are mutually absolutely continuous on $[p_L, p_H]$;

*(iii) and in particular there is a **decreasing** nonnegative function h on $[p_L, p_H]$ such that for every $p \in [p_L, p_H]$*

$$\rho_1(p) = \int_{[p_L,p]} h(p) d\rho_2(p) + \rho_1(p_L^-).$$

where $\rho_1(p_L^-) = \lim_{p \to p_L^-} \rho_1(p)$.

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. From the definition of LR dominance, there is a nonnegative, nonincreasing function $h(\cdot)$ and nondegenerate interval $[p_L, p_H]$ such that

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

for every $p \in [p_L, p_H]$ where the support of F lies in $[0, p_H]$ and the support of G lies in $[p_L, 1]$. We take p_H to be the supremum of the support of F and p_L to be the infimum of the support of G. Since the support of G lies in the support of F, p_H is also the supremum of the support of F. Let π^G maximize expected profit on Φ at belief G, let

[20] Some definitions (e.g. Athey (1996), p. 213) do not imply mutual absolute continuity when the supports have a nonempty intersection, but we use this property in our proof.

π^F maximize expected profit on Φ at F, and let $\hat{\pi}^F$ maximize expected profit on Φ at F subject to the *additional* constraint that $\hat{\pi}^F(p) = 0$ for $p < p_L$. We must prove that if $\int \pi^F dF = 0$ then $\int \pi^G dF = 0$. We prove the contrapositive. Suppose that $\int \pi^G dG > 0$. Since $\int \pi^F dF \geq \int \hat{\pi}^F dF \geq \int \pi^G dF$, it suffices to show that $\int \pi^G dF > 0$. By Theorem 7 in Border (1996) the following integration-by-parts formula holds[21]

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

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

where the p^- in the integrand means the limit is taken from the left. 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 any $\pi \in \Phi$ there is a $p' \in [p_L, p_H]$ with $\int_{[p_L,p'^-]} \pi(q)dG(q) < 0$, then by Assumption 1 there is a $\pi'(p) \in \Phi$ with $\pi'(p) = 0$ for $p < p'$ and $\int_{[p',p_H]} \pi' dG \geq \int_{[p',p_H]} \pi dG$, and π does 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 the beliefs are mutually absolutely continuous, p_H is not point of positive measure for *either* belief. 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 (again using the version in Border (1996), Theorem 7) to find that, for every $p \in [p_L, p_H)$

$$\begin{aligned} 0 < 1 - F(p) &= \int_{(p,p_H]} h(q)dG(q) = h(p_H) - G(p_L)h(p_L) - \int_{(p,p_H]} G(q)dh(q) \\ &= -G(p_L)h(p_L) + \int_{(p,p_H]} G(q^-)d(-h(q)), \end{aligned} \tag{8}$$

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 at p_H,[22] the second term in (7) is positive. $\square$

[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.).

[22] We have $|\int_{[p_L,p_H]} \pi^G dG - \int_{[p_L,p-]} \pi^G dG| \leq |\int_{[p,p_H]} \pi^G dG|$. Since π^G is bounded and $1 - G(p-)$

A.2 Proof of Proposition 2

(i) To prove that (2) is sufficient for a null menu maximizing the insurer's expected profit, assume that it 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 profit subject to $x_1 \leq x_2 \leq ... \leq x_n$, 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 $x_i \geq t_i$ 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 $x_n \geq x_{n-1}$ 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 $x_{n-1} \geq t_{n-1} \geq 0$ it follows that (recall that $MRS(x, t, p) = -U_x/U_t$ evaluated at (x, t, p).)

$$\lambda p_n \geq MRS(0, 0, p_n) \geq MRS(x_{n-1}, t_{n-1}, p_n).$$

Let $m = MRS(x_{n-1}, t_{n-1}, p_n)$. 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 $(x, t) \geq (x_{n-1}, t_{n-1})$ it follows that $t - t_{n-1} \leq m(x - x_{n-1})$. Use the inequality $\lambda p_n \geq m$ and rearrange to find that $t_{n-1} - \lambda p_n x_{n-1} \geq t - \lambda p_n x$ 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 $x_i \geq t_i$ 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

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

subject to that $x_{n-k+1} \geq x_{n-k}$ 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 (2) applied to $\hat{p} = p_1$, that pooling menu must be a null menu.

Now consider an arbitrary type distribution F. Suppose that (2) holds. Consider a sequence of finite support distribution functions F_n which converge weakly to F and

tends to zero as p converges to p_H, the integral is continuous at p_H

such that (2) 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 (2) 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

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

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

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

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 firm's complete-information profit from type p is $\max 0, \Pi(p) - kp$, where Π is the risk premium. We begin by showing that the inequality $k < ([u(w) - u(w-\ell)]/u'(w)) - \ell$ implies that complete information profit is positive for p close to 0. For this it is enough to show that the derivative of $\Pi(p) - kp$ with respect to p is positive at $p = 0$. The right side of the inequality, $([u(w) - u(w-\ell)]/u'(w)) - \ell > 0$ is the derivative of the risk premium at $p = 0$. If k is less than this value, then complete information profit is positive for all p in a neighborhood of $p = 0$. An insurer whose belief is sufficiently concentrated on a type in 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 binds, 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 $\Pi(p_H, \lambda) > 0$. $\square$

A.5 Proof of Proposition 4

Let p_L be the smallest and p_H the largest element of the support $\mathcal{P}$ of beliefs. Fix $\lambda \geq 1$, $k = 0$ and suppress these in what follows. Part (i) is immediate. For part (ii), suppose there is a $\hat{p}$ in $\mathcal{P}$ such that complete-information profit at $\hat{p}$ is positive but $\hat{p} > p_L$. We will show that V cannot be decreasing in the LR order on $\mathcal{P}$. There is a distribution μ with support equal to $[\hat{p}, p_H] \cap \mathcal{P}$ that puts enough probability on types near $\hat{p}$ so that expected profit, $V(\mu)$, is positive. Let ν be any distribution on $[p_L, \hat{p}] \cap \mathcal{P}$ with $\nu([p_L, \hat{p})) = 1$. For $\alpha \in (0, 1]$, define ρ_α by $\rho_\alpha = (1 - \alpha)\nu + \alpha\mu$. The family of distributions $\{\rho_\alpha\}_{\alpha \in [0,1]}$ is ordered by likelihood ratio dominance. Let $g(\alpha) = V(\rho_\alpha)$, and let $\pi^\alpha(\cdot)$ be any profit-maximizing menu of expected profit at belief ρ_α (using the change of variables introduced at the beginning of Section 3.1), so $g(\alpha) = \int \pi^\alpha d\rho_\alpha$. If g were differentiable at $\alpha = 1$, then it would follow from the first Envelope Theorem (Theorem 1) in Milgrom and Segal (2002) that

$$g'(1^-) = \int \pi^1 d\mu = V(\mu) > 0, \tag{9}$$

implying that V is not decreasing in the LR order. We prove *directly* that g cannot be decreasing, stopping short of confirming differentiability of g; a crucial step in our argument establishes that the solution correspondence is continuous at $\alpha = 1$. One can use this fact to confirm that g is indeed differentiable at $\alpha = 1$ from which (9) follows. (See the Remark after our proof.)

To simplify the notation further, note that, for every feasible menu $\pi(\cdot)$ of expected profits, there is a pair of conditional expected profits, (y, z), with $y = \int_{[p_L,\hat{p})} \pi d\nu$ and $z = \int_{[\hat{p},p_H]} \pi d\mu$; let Z be the set of such feasible conditional expected profits. It follows that

$$g(\alpha) = \max_{(y,z)\in Z} \alpha z + (1 - \alpha)y. \tag{10}$$

Since the constraint set in the original contract space is compact, and expected profit is continuous in the contract, Z is compact, so g is continuous. And since we are considering the firm's problem before subtracting the fixed cost, Assumption 1 holds: $(0,0) \in Z$ and if $(y,z) \in Z$, then $(0,z) \in Z$. Let $(y^*(\cdot), z^*(\cdot))$ be *any* selection from the solution correspondence to (10). It follows that

$$\frac{g(1) - g(\alpha)}{1-\alpha} = \frac{z^*(1) - \alpha z^*(\alpha) - (1-\alpha) y^*(\alpha)}{1-\alpha}. \tag{11}$$

Add and subtract $z^*(1)$ to the right side and rearrange to find

$$\frac{g(1) - g(\alpha)}{1-\alpha} = z^*(1) - y^*(\alpha) + \alpha \frac{z^*(1) - z^*(\alpha)}{1-\alpha}. \tag{12}$$

A simple comparative statics argument shows that z^* is nondecreasing and y^* is nonincreasing in α.[23] Since z^* is nondecreasing, $z^*(1) - z^*(\alpha) \geq 0$ for every $\alpha \in [0,1]$ and the third term on the right side of (12) is nonnegative for every $\alpha \in [0,1)$.

Next we show that $\lim_{\alpha \to 1} y^*(\alpha) = 0$. Since Assumption 1 holds, y^* is bounded below by 0. Since it decreases in α, $\lim_{\alpha \to 1} y^*(\alpha)$ exists and by Assumption 1 it is nonnegative; denote it by $\underline{y}$. Define $\tilde{z}(y) = \max\{z | (y,z) \in Z\}$ on the set of numbers y such that $(y,z) \in Z$. Since Z is compact, $\tilde{z}$ is well-defined on its domain and it is continuous. Note that $z^*(\alpha) = \tilde{z}(y^*(\alpha))$. Since g is continuous at $\alpha = 1$, $\lim_{\alpha \to 1} z^*(\alpha) = z^*(1)$.[24] But then $\lim_{\alpha \to 1} \tilde{z}(y^*(\alpha)) = \tilde{z}(\underline{y}) = z^*(1)$. Since the participation constraint binds for type $\hat{p}$ when $\alpha = 1$, it follows that $\underline{y} = 0$ (if $\underline{y} > 0$, then some contracts in $[p_L, \hat{p})$ would be positive at $\alpha = 1$ and the participation constraint would be slack for type $\hat{p}$ by the strict single crossing property). Since $z^*(1) > 0$ it now follows that the first term on the right side of (12) converges to a positive number. Recalling that the second term is nonnegative it follows in turn that

$$\liminf_{\alpha \to 1^-} \frac{g(1) - g(\alpha)}{1-\alpha} > 0$$

and g cannot be decreasing in a neighborhood of $\alpha = 1$ in $(0,1]$. $\square$

[23] It follows from a 'revealed preference' argument. Suppose that (y,z) maximizes profit at α and $(y+\Delta y, z+\Delta z)$ maximizes profit at $\alpha+\Delta\alpha$: in particular $\alpha z + (1-\alpha)y \geq \alpha(z+\Delta z) + (1-\alpha)(y+\Delta y)$ and $(\alpha+\Delta\alpha)(z+\Delta z) + (1-\alpha-\Delta\alpha)(y+\Delta y) \geq (\alpha+\Delta\alpha)z + (1-\alpha-\Delta\alpha)y$. Rearrange the inequalities to find that $\Delta\alpha\Delta z \geq \Delta\alpha\Delta y$. Since profit maximization implies that $\Delta y > 0$ if and only if $\Delta z < 0$—if y could strictly increase while z not decrease, then the starting pair could not maximize profit—the conclusion follows: $\Delta\alpha > 0$ implies $\Delta z \geq 0$ and $\Delta y \leq 0$.

[24] We have $z^*(1) = \lim_{\alpha \to 1} g(\alpha) = \lim_{\alpha \to 1}[\alpha z^*(\alpha) + (1-\alpha) y^*(\alpha)] = \lim_{\alpha \to 1} \alpha z^*(\alpha)$, where the last equality uses the boundedness of y^*.

Remark 1. *Note that the our proof establishes that the solution correspondences are continuous (and single-valued) at $\alpha = 1$. It is easy to verify from this that the sufficient conditions of Theorem 3 of Milgrom and Segal are met for problem (10) at $\alpha = 1$. It follows that g is differentiable and equation (9) holds.*

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

$$\ell(1 - \lambda E_\rho[p]) \geq E_\rho[\Pi_k(p)] \geq V_k(\rho) \geq T_k(\ell, p_L) - \lambda \ell E_\rho[p], \tag{13}$$

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 (13) yields $\lim_{k\to\infty} V_k(\rho) = \ell(1 - \lambda E_\rho[p])$, which is a strictly decreasing function of ρ. □

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

(i) To simplify the presentation, let us assume in this result that u is C^2 on R_{++} with $u'' < 0$ and that $\mathcal{P} \subset [0, 1)$. Let $x^*(p, \lambda)$ be the complete information coverage, the solution to $\max_{x,t} t - \lambda p x$ 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 whenever it is positive, we show that $T(x, p) - \lambda p x$ satisfies the strict single crossing property in $(x, -p)$. Now, $T_x(p, x) = -(U_x/U_t)(T(x, p), x, p)$, and the first-order necessary condition for an interior maximum is $T_x - \lambda p = 0$. Let

$$m(x, t, p) = -\frac{U_x}{U_t}(x, t, p)\frac{1}{p}$$

and rewrite the first-order condition as $m(x, T(x, p), p) - \lambda = 0$. Some calculation reveals that $m = \lambda > 1$ implies that $m_x + pmm_t < 0$, as the implicit function theorem requires. At any (p, λ) with $x^*(p, \lambda) > 0$, apply the implicit function theorem to find that

$$\frac{\partial x^*(p, \lambda)}{\partial p} = \frac{m_t}{-(m_x + pm_t m)} T_p + \frac{m_p}{-(m_x + pm_t m)}. \tag{14}$$

where the right side is evaluated at $(x^*(p, \lambda), T(x^*(p, \lambda), p)$. Specifically

$$m_t = -(1 - p)\frac{u'_\ell u'_n (R_n - R_\ell)}{E[u']^2}, \tag{15}$$

where $u_\ell = u(w - \ell + x - t)$, $u_n = u(w - t)$, $E[u'] = pu'_\ell + (1-p)u'_n$, and R_ℓ and R_n are the coefficients of absolute risk aversion evaluated at the loss and no loss state wealths; and

$$T_p = \frac{U_p(x^*, T(x^*, p), p) - U_p(0, 0, p)}{E[u']} > 0, \tag{16}$$

$$m_x = -\frac{(1-p)u'_\ell u'_n R_\ell}{E[u']^2} < 0, \quad \text{and} \tag{17}$$

$$m_p = \frac{-u'_\ell(u'_\ell - u'_n)}{E[u']^2} < 0. \tag{18}$$

It follows from (14)-(17) that x strictly *decreases* in p (whenever it is positive) if u exhibits non-decreasing absolute risk aversion (that is, $R_n \geq R_\ell$), which includes CARA.

(ii) Note that the functions m, m_t, T_p, m_x, and m_p are uniformly continuous (since they are continuous on the compact set $\mathcal{P} \times x^*(\mathcal{P}) \times T(x^*(\mathcal{P}), \mathcal{P})$).

Let λ^* be the smallest value of λ such that $x(p, \lambda) = 0$ for every $p \in \mathcal{P}$. It is immediate from (5) that $x(p, \cdot)$ is decreasing for every $p \in \mathcal{P}$. It follows that $x^*(\cdot, \lambda)$ converges uniformly to the zero function as $\lambda \to \lambda^{*+}$. From (16) it follows that T_p converges uniformly to the zero function and, since m_x and m_t are uniformly continuous, $|m_t T_p/(m_x + pm_t m)|$ converges uniformly to the zero function as $\lambda \to \lambda^{*+}$. Finally it is easy to verify that m_p converges uniformly to a function which is negative for every $p \in \mathcal{P}$. These facts and (14) give us the conclusion. $\square$

Using the first-order condition in the CARA case (see Section A.6), simple algebra reveals that the optimal reimbursement in this case is the one given in Example 1.

A.8 Proof of Proposition 6

Since $\lambda < u'(w - \ell)/u'(w)$, there is a point $p_L \in (0, 1/\lambda)$ satisfying $MRS(p_L)/p_L > \lambda$. Moreover, since $MRS(\lambda^{-1}) < 1$, there is a $p' \in (p_L, 1/\lambda)$ such that, for all $p \geq p'$, $MRS(p)/p < \lambda$. 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. For part (b), 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), $t(p)-\lambda p x(p) \leq 0$ for any $p \in [p', p_H]$. Since expected profit from the menu is positive, $t(p)-\lambda p x(p) > 0$ for some $p \in [p_L, p')$, so for that p, $t(p) > 0$. Since the menu is nondecreasing, $t(p') > 0$ and by the participation constraint, $x(p') > t(p')$. And since the complete-information contract for type p' is $(0,0)$ and u is strictly concave, the the marginal rate of substitution for the type p' at any contract $(x(p'), t(p')) >> (t(p'), 0)$ is *less* than $MRS(p')$, so efficient contract (x^*, t^*) for type p' that is indifferent to $(x(p'), t(p'))$, is lower: $(x^*, t^*) << (x(p'), t(p'))$; so the contract for p' is distorted upwards from its efficient contract. By similar reasoning, every type $p \in [p', p_H]$ is pooled at $(x(p'), t(p')) > 0$ and their contracts are distorted upwards. So there is pooling and inefficiency at the highest type. $\square$

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

$$\begin{aligned} v_x(p_H, x_H) &= \lambda p_H && (19) \\ v_x(p_L, x_L) &= \rho v_x(p_H, x_L) + (1-\rho)\lambda p_L. && (20) \end{aligned}$$

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 (19) 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 $\lambda \in (1, (p_H + (1 - p_H e^{-r\ell})^{-1})$, a parametric restriction that we henceforth impose.

Consider equation (20). It can be written as

$$\frac{1 + \left(\frac{1-p_H}{p_H}\right) z_L}{1 + \left(\frac{1-p_L}{p_L}\right) z_L} = (1-\rho)\lambda p_L \left(1 + \left(\frac{1-p_H}{p_H}\right) z_L\right) + \rho. \tag{21}$$

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 (21) 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 (21) 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. □

A.10 Proof of Competition Equilibrium Properties

Set the entry cost K equal to 0, and fix $\lambda \geq 1$, $k \geq 0$. In a competitive equilibrium with a finite number of risk types, we first argue that (i) there is no pooling; (ii) the highest risk type gets an efficient contract; (*iii*) if complete information zero-profit contracts are nonincreasing in types, then all types other than the highest buy contracts distorted downwards from these efficient contracts. The arguments follow essentially from those in Rothschild and Stiglitz (1976), so we present them in outline.

First, if there is an equilibrium, each contract makes zero expected profit: if any contract (x, t) earns positive expected profit, then a new contract (x', t') could be introduced that all types buying (x, t) prefer, earns positive profit from them, but that higher types do not prefer to the contracts they buy in the original set. (If lower types buy it, then these contracts continue to earn positive profit.) Second, there is no pooling: if more than one risk type buys a contract (x, t) that makes zero expected profit, then, by the single-crossing property, another contract could be offered that a) only the lowest risk that bought (x, t) would prefer to it; b) is sufficiently close to (x, t), so expected profit from the new contract is positive (see Figure II, p. 635 in Rothschild and Stiglitz (1976)). So property (*i*) holds. Third, the highest risk type gets an efficient contract: if the only contracts bought by the highest risk are inefficient and earn zero profit, then another contract could be offered that makes positive expected profit if bought by the highest risk types (and so if bought by anyone; see Figure III on p. 636 in Rothschild and Stiglitz (1976) and the discussion surrounding it). So (*ii*) holds.

If complete-information competitive contracts are nonincreasing in risk type (including the nonresponsive case), then (*iii*) all types below the top get contracts that are distorted downward. This follows from complete sorting, the strict single crossing property, and expected utility maximization. If however complete-information competitive contracts are strictly increasing then each type might strictly prefer its complete information contract to the complete information contract of any other type.

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*, 81, 1713–1762.

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.

Kreps, D. (1990): *A Course in Microeconomic Theory*. Princeton University Press, Princeton, NJ.

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.

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

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, .

National Association of Insurance Commissioners (2011): "Issue Brief–Health Insurance Rate Regulation," `www.naic.org/documents/topics_health_insurance_rate_regulation_brief.pdf`, Accessed: 2014-03-03.

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

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.

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

Shaked, M., and J. G. Shanthikumar (2007): *Stochastic Orders*. Springer.

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.

Tennyson, S. (2007): "Efficiency Consequences of Rate Regulation in Insurance Markets," Networks Financial Institute Policy Brief.