Adverse Selection and Insurance Contracting: A Non-Expected Utility Analysis*

Matthew J. Ryan† and Rhema Vaithianathan‡

December, 2000

Abstract

We analyse the implications of the Rank Dependent Utility (RDU) model for insurance markets with asymmetric information. Experimental evidence consistently rejects expected utility in favour of a version of RDU in which probabilities have a non-linear impact on risky choice. Using simple Hirshleifer-Yaari diagrams, we can show that this behaviour has novel implications for insurance contracts subject to adverse selection. Contra the well-known results of Stiglitz (1977), a monopoly

---

*Earlier versions of the present paper have circulated under the title "Health Insurance with Rank-Dependent Expected Utility". The authors would particularly like to thank Simon Grant, Rohan Pitchford, John Quiggin and Peter Wakker for helpful comments and suggestions on earlier drafts. The paper has also benefitted from the input of seminar audiences at The Australian National University, University of Auckland, University of Melbourne and University of Sydney. The usual disclaimer applies.

†Department of Economics, Faculty of Economics and Commerce, The Australian National University, Canberra ACT 0200, Australia. Email: matthew.ryan@anu.edu.au

‡Economics Programme, Research School of Social Sciences, The Australian National University, Canberra ACT 0200, Australia. Email: rhema.vaithianathan@anu.edu.au.

insurer may choose to pool high and low risks; or there may be complete market failure, with neither type being served. Pooling may involve low risks cross-subsidising high risks, or both types may generate positive profits. In the latter case, a common *full insurance* contract may be offered, so a Pareto optimal outcome results despite the informational asymmetry.

1. Introduction

A large body of experimental work on risk-taking behaviour has rejected the empirical validity of classical expected utility (EU) theory. In an attempt to mathematically represent the sort of behaviour observed by these experimenters, extensions to the EU framework have been developed. One such extension is rank-dependent utility (RDU), originally formulated by John Quiggin (1982)[1]. In this paper, we explore the implications of RDU behaviour for insurance contracting under asymmetric information.

Several authors have attested to the robustness of insurance theory to deviations from EU. Machina (1995) provides a useful survey. However, Machina (1995, p.36) notes that the area of insurance under asymmetric information – moral hazard and adverse selection – has yet to be thoroughly explored from a non-EU perspective. The present paper shows that some standard results on insurance contracting under adverse selection are over-turned when EU is replaced by the version of RDU having the strongest experimental support. A companion paper (Ryan and Vaithianathan (2000)) reaches similar conclusions for the case of moral hazard.

In particular we show that a monopoly insurer may choose to pool different risk classes. This contrasts with the well-known result of Stiglitz (1977), obtained under an EU assumption on consumer behaviour, that pooling contracts are never observed.

[1] Quiggin called his theory "anticipated utility".

While we focus attention on a class of preferences that are not globally risk averse, we show that pooling is also possible for globally risk averse agents as well (section 4.2).

We do not impose global risk aversion because its imposition is not supported by the experimental evidence. This evidence is reviewed in section 2.2, but let us here recall Tversky and Kahneman's (1992, p.298) observation that "risk-seeking choices are consistently observed ... when people must choose between a sure loss and a substantial probability of a larger loss". This immediately suggests that a person with a high risk of suffering some loss may be reluctant to fully insure against it (i.e. suffer instead the certain loss of the associated insurance premium), even on actuarially fair terms.

In the RDU model, risk attitude depends jointly on the curvature of the utility function and the shape of a probability distortion function. Even if utility is linear in money, agents may be risk averse (respectively, risk loving) if they distort probabilities in a "pessimistic" (respectively, "optimistic") fashion[2].

We assume concave utility functions throughout the paper. Hence, deviations from risk aversion are driven entirely by distortions of the underlying probabilities in decision-making. The class of RDU preferences we consider preserve the usual single-crossing property, so the separation of high risks and low risks in insurance markets subject to adverse selection requires that the former be offered "more" insurance than the latter. This is costly to the insurer in the following two situations. First, if high risks are risk-lovers who are unwilling to purchase actuarially fair insurance, then they need to be "subsidised" in order to buy insurance. Separation may therefore be less profitable than pooling. In fact, if high risks are a large fraction of the population, the entire market may break down, with neither type being served (the case of a pooled "null" contract). Second, if both high and low risks are risk averse, but low risks distort

[2] See Chew, Karni and Safra (1987). The optimism/pessimism terminology is borrowed from Wakker (forthcoming).

probabilities in a more pessimistic fashion and are sufficiently more numerous than high risks, then it may be optimal to fully insure both types, since the sale of insurance to low risks is particularly lucrative. In this case, all agents may be receiving Pareto optimal contracts.[3] Thus, the class of RDU preferences we study is compatible with pooling in a monopolised insurance market subject to adverse selection, including the extreme cases of complete market failure (neither type insured) and Pareto optimal contracting (both types fully insured).

The outline of the rest of the paper is as follows. In the next section we review the theory of RDU maximisation, and the experimental evidence in its favour. In section 3 we examine the benchmark case of insurance contracting under symmetric information. Sections 4.1 and 4.2 contain the main results on insurance under adverse selection. Section 5 discusses related literature, and section 6 concludes. Proofs are contained in an appendix.

2. RDU preferences

2.1. Theory

Consider a lottery

$$L = (x_1, p_1;\, x_2, p_2;\, ...;\, x_n, p_n)\,.$$

Each $x_i \in \mathbb{R}_+$ is a monetary outcome, and $p_i \in (0,1]$ is the probability of this outcome. As usual, we assume that $\sum_i p_i = 1$. For the RDU model, an outcome's *rank* in L is also important, so we adopt the convention that lottery prizes are listed in descending order. That is:

$$x_1 \geq x_2 \geq \cdots \geq x_n.$$

[3] This result is similar to Ryan and Vaithianathan (2000), where it is shown that RDU preferences may allow Pareto optimal contracting in the presence of moral hazard.

The RDU preference representation involves a utility function $u : \mathbb{R}_+ \to \mathbb{R}$, and a strictly increasing function $w : [0,1] \to [0,1]$, satisfying the normalisation conditions $w(0) = 0$ and $w(1) = 1$. The *probability transformation* function w is uniquely determined from preferences, but the utility function is unique only up to positive affine transformations[4]. The lottery L is evaluated as follows[5]:

$$\sum_{i=1}^{n} u(x_i) \left[w\left(\sum_{j=1}^{i} p_j\right) - w\left(\sum_{j=1}^{i-1} p_j\right)\right] \tag{2.1}$$

Observe that

$$\sum_{i=1}^{n} \left[w\left(\sum_{j=1}^{i} p_j\right) - w\left(\sum_{j=1}^{i-1} p_j\right)\right] = 1,$$

so RDU maximisers do perform expected utility calculations, but using a (rank-dependent) distortion of the true probabilities. Moreover, if w is linear (i.e. the identity function), then (2.1) is a standard expected utility calculation. However, experiments consistently reject the linearity of w.

2.2. Evidence on the shape of w

Wakker (forthcoming, Appendix A) reviews the experimental literature on the shape of w, and concludes that there is "overwhelming evidence" that w has the inverted-S shape of Figure 2.1.[6] Such decision-makers tend to overweight (small) probabilities attached to very highly or very lowly ranked outcomes, and hence to underweight the probabilities of non-extreme outcomes. In particular, the transformation function depicted in Figure 2.1 implies the violation of global risk aversion. The latter requires w to be *convex*

[4] See Chateauneuf and Wakker (1999) for an axiomatisation.

[5] We adopt the convention that $\sum_{i=1}^{0} p_i = 0$.

[6] This is the version of RDU favoured in Quiggin's original proposal: see Quiggin (1982, pp.326 and 334).

(Chew, Karni and Safra (1987)). For example, the over-weighting of small probabilities attached to very high-ranking outcomes is compatible with gambling behaviour.

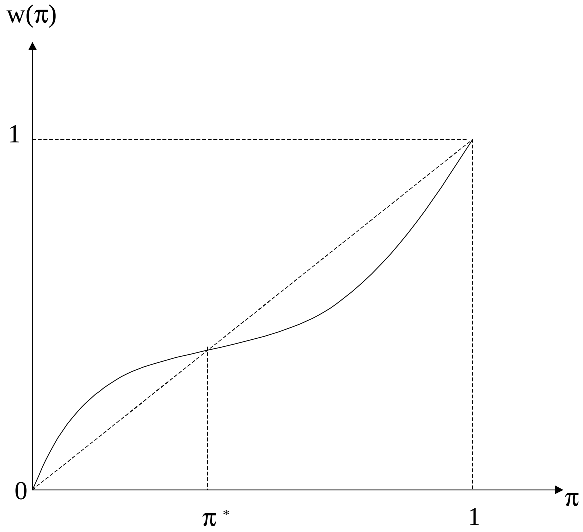

Figure 2.1: Typical probability transformation function

Inverse-S transformation functions are also compatible with the common consequence effect noted by Allais (1953). In fact, Allais' "paradox" is a special case of the phenomenon of *bounded subadditivity* (or SA), in which small probability changes have larger impacts at certainty than away from certainty.[7] In particular, a small probability of a loss (relative to certainty) has a bigger impact than an equal-sized increase in the chance of a loss from a risky starting point. Similarly, a small chance of a gain (relative to certainty) has a bigger impact than a similar increase in the chances of the gain beginning from a risky position. More precise statements may be found in Tversky and Wakker (1995), who have tested for and confirmed the prevalence of SA in risky choice.

[7] Tversky and Kahneman (1992) refer to this phenomenon as the "principle of diminishing sensitivity".

Wu and Gonzalez (1996) observe that SA is consistent with an inverse-S transformation function, but is not equivalent to it, even under a maintained hypothesis of RDU preferences. In particular, SA examines behaviour "at certainty", so it does not impose a conclusive test of the curvature of w on the interior of the probability interval. Wu and Gonzalez therefore develop direct preference-based tests for the concavity and convexity of w. These tests assume RDU, but are independent of the form of the utility function u. Wu and Gonzalez find significant evidence of concavity of w up to a critical probability level of between 0.3 and 0.4, with convexity thereafter.

Wu and Gonzalez (1996) also attempt to econometrically fit the following one-parameter form for w:

$$w(\pi) = \frac{\pi^\gamma}{\left(\pi^\gamma + (1-\pi)^\gamma\right)^{1/\gamma}} \tag{2.2}$$

Their estimation procedure maintains the assumption that $u(z) = z^\alpha$, though α is estimated along with γ. They are able to reject the EU hypothesis ($\gamma = 1$), and obtain estimates of $\alpha = 0.50$ and $\gamma = 0.71$. This implies an inverse-S form with fixed point at $\pi^* = 0.39$. Camerer and Ho (1994) obtain similar results using a different dataset.

3. Insurance with symmetric information

Consider a population of RDU maximising agents with the following common features: wealth level y, a strictly increasing, strictly concave and continuously differentiable utility function u, and transformation function w. We distinguish two types of agent: a "high risk" type, with probability $p^H \in (0,1)$ of suffering a financial loss of $c < y$; and a "low risk" type, with probability $p^L \in (0, p^H)$ of suffering the same loss, c. Let $\theta \in (0,1)$ be the proportion of high risks in the population.

An uninsured high risk agent therefore faces the lottery

$$\left(y, 1-p^H; y-c, p^H\right) \tag{3.1}$$

while low risks face the lottery

$$\left(y, 1-p^L; \, y-c, p^L\right) \tag{3.2}$$

Using (2.1), lottery (3.1) is evaluated as

$$u(y-c) \, + \, w\left(1-p^H\right)[u(y)-u(y-c)] \tag{3.3}$$

Similarly, lottery (3.2) is evaluated as

$$u(y-c) \, + \, w\left(1-p^L\right)[u(y)-u(y-c)] \tag{3.4}$$

Let w be as in Figure 2.1, and suppose that $1-p^H < \pi^* < 1-p^L$. Then, relative to EU maximisers, *high risk* individuals act as if they are *over-optimistic* about their chances of avoiding the loss, while *low risk* individuals act in an *overly pessimistic* fashion. This distortion of probability information will augment the potential gains from trade (relative to the EU case) between the insurance company and a low risk client; but will squeeze potential gains in the case of high risk clients. Indeed, it is possible that gains from trade may vanish altogether in the latter case.

To take a concrete example, let us suppose that w takes the form (2.2) estimated by Wu and Gonzalez (1996). Consider the indifference curves of such a consumer in a Hirshleifer-Yaari diagram. Fix two states, having probabilities p and $1-p$ respectively. The vertical axis in the Hirshleifer-Yaari diagram will measure wealth (z_2) in the state that occurs with probability p; while the horizontal axis measures wealth (z_1) in the other state. The following lemma describes the qualitative features of indifference curves in this diagram.

Lemma 3.1. *If w is given by (2.2) with $\gamma \in (0,1)$, then for any $p \in (0,1)$:*

1. *indifference curves in the Hirshleifer-Yaari diagram are non-differentiable at certainty; and*

2. *upper contour sets (in the same diagram) are convex.*

Indifference curves for $p = 0.25$, $p = 0.5$, and $p = 0.75$ are depicted in Figures 3.1–3.3. In each case, we take $u(z) = \sqrt{z}$. The diagrams include a (dashed) fair bet line, so that one may judge the degree and direction of the probability distortion in each case. Note that only the $p = 0.5$ indifference curves (Figure 3.2) exhibit risk aversion (i.e. lie entirely above the fair bet line). Observe also that each indifference curve has a single point of non-differentiability where it intersects the certainty line (where $z_2 = z_1$). This is due to the change in the outcome ranking of the states as we cross this line: below the line, state 1 delivers a higher payoff than state 2, and conversely above the line.

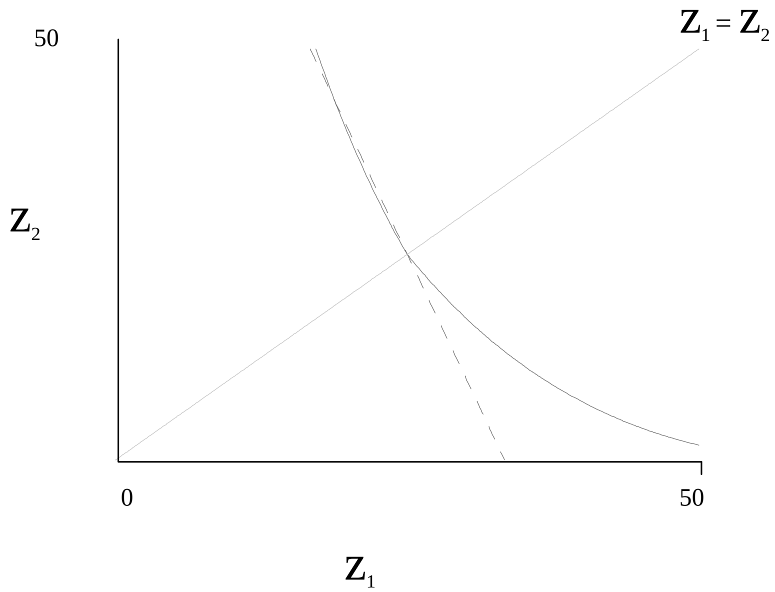

Figure 3.1: Indifference curve with $p = 0.25$ (low risk)

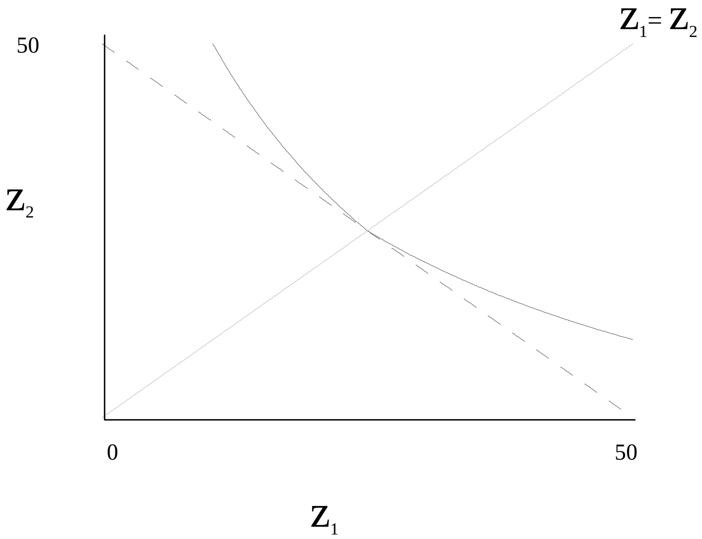

Figure 3.2: Indifference curve with $p = 0.5$

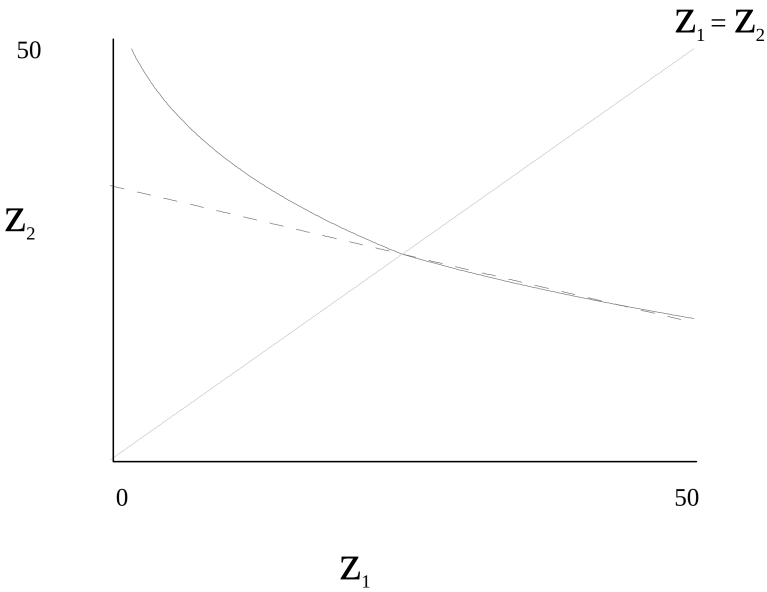

Figure 3.3: Indifference curve with $p = 0.75$ (high risk)

Thus, unlike the risk-averse EU case, in which Pareto optimal risk sharing always requires the consumer to be completely insured, here a Pareto optimal contract may take one of four possible forms: (i) over-insurance; (ii) full insurance; (iii) partial insurance; or (iv) no insurance.

By considering Figures 3.1–3.3 and equation (A.1) in the appendix, we see that over-insurance is optimal if and only if

$$\lim_{z_2 \downarrow z_1} \frac{1-w(p)}{w(p)} \frac{u'(z_1)}{u'(z_2)} = \frac{1-w(p)}{w(p)} < \frac{1-p}{p}.$$

This will be the case if and only if $p < w(p)$. Thus, over-insurance occurs for those types whose probability of loss lies *strictly below* the fixed point π^* of the transformation function (see Figure 2.1).

By symmetric reasoning, less-than-full insurance is Pareto optimal if and only if $1 - p < \pi^*$. In other words, full insurance contracts are offered to those consumers whose probability of loss lies in the range $[\pi^*, 1 - \pi^*]$. For these types, and these types only, the qualitative implications of risk-averse EU and Wu and Gonzalez' version of RDU coincide.[8]

To distinguish between the partial and no insurance cases, we need additional information. No insurance is optimal if

$$\frac{w(1-p)}{[1-w(1-p)]} \frac{u'(y)}{u'(y-c)} \geq \frac{1-p}{p} \tag{3.5}$$

so that the slope of the indifference curve through the "no insurance" point is no smaller than the slope of the fair bet line.[9] Suppose, for example, that $u(z) = \sqrt{z}$, consistent with the estimation results from Wu and Gonzalez (1996). Then (3.5) is equivalent to

[8] Observe that the existence of such types requires $\pi^* \leq 0.5$.

[9] Note that the condition (3.5) implies $p > 1 - \pi^*$.

the condition

$$\frac{c}{y} \leq 1 - k(p)^2$$

where

$$k(p) = \frac{(1-p)\,[1 - w(1-p)]}{pw(1-p)} \tag{3.6}$$

If we define the function $r(p) = 1 - k(p)^2$, then for any $p > 1 - \pi^*$, it is Pareto optimal to be uninsured if and only if $r(p) \geq c/y$. Therefore, a consumer with risk $p > 1 - \pi^*$ will be uninsured if the potential loss is *sufficiently small* as a proportion of total income. Otherwise, partial insurance is Pareto optimal.

Figure 3.4 graphs the function $r(p)$ over $[1 - \pi^*, 1]$ for the case in which $w(p)$ is given by (2.2) with $\gamma = 0.71$ (again, as per the estimation results of Wu and Gonzalez (1996)). The area under the graph is the parameter region for which it is Pareto optimal not to purchase insurance. Observe that the function $r(p)$ is strictly increasing on $[1 - \pi^*, 1]$. Thus, a potential loss of 50% of one's income will go uninsured provided the risk of loss exceeds approximately 0.8; while a 0.7 probability is sufficient for the risk of losing 25% of one's income to go uninsured.

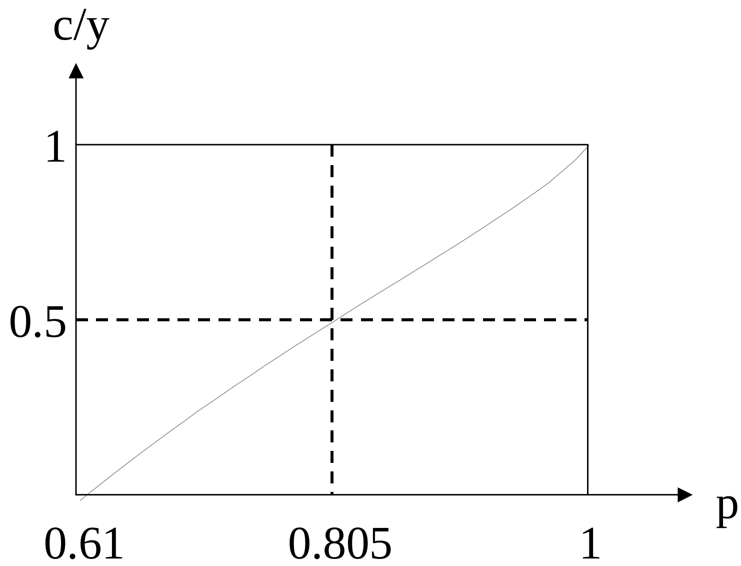

Figure 3.4: $r(p)$ with $\alpha = 0.5$ and $\gamma = 0.71$

To summarise the discussion so far:

(a) If w takes the form (2.2), with $\gamma \in (0, 1)$, then

- **(i)** Over-insurance is Pareto optimal for $p < \pi^*$, where π^* is the fixed point of w.
- **(ii)** Full insurance is Pareto optimal for $p \in [\pi^*, 1 - \pi^*]$.
- **(iii)** Less-than-full insurance is Pareto optimal for $p > 1 - \pi^*$.

(b) Furthermore, if $u(z) = \sqrt{z}$, then

- **(i)** Partial insurance is Pareto optimal if $p > 1 - \pi^*$ and $c/y > r(p)$, where $r(p) = 1 - k(p)^2$, and $k(p)$ is given by (3.6).
- **(ii)** It is Pareto optimal to be uninsured if $c/y \leq r(p)$.

4. Adverse selection

What if the consumer's risk type is private information? With risk-averse EU consumers and a monopoly insurer, it is well-known[10] that the market will either (a) serve high risks but not low risks; or (b) serve both types with a screening menu of contracts in which high risks receive full insurance, but low risks receive only partial cover. In particular, pooling contracts are never observed.

If consumers exhibit the sort of RDU preferences described in the preceding section, then other possibilities arise. First, it is possible to observe pooling contracts under which low risks cross-subsidise high risks.[11] Because of the need for this cross-subsidisation, such contracts are viable only if θ (the proportion of high risks in the population) is sufficiently low. If not, then neither type is served. Second, one may also observe pooling contracts in which both types generate positive profits. In this case, it is even possible to observe a common *full insurance* (Pareto optimal) contract being offered.

4.1. Pooling with cross-subsidisation

Suppose, for example, that both types of consumer have preferences of the Wu and Gonzalez variety, with $u(z) = \sqrt{z}$. Suppose further that $c/y = r(p^H)$ so that high risk consumers could not be profitably served under symmetric information. Figure 4.1 depicts the situation.[12] Point E is the "no insurance" (or endowment) point; L is a

[10] See Stiglitz (1977).

[11] This scenario is in fact quite common in health insurance markets - see Browne (1992) and Browne and Doerpinghaus (1993).

[12] For simplicity, we have not drawn the kink in the indifference curves, since our focus is on points below the certainty line. We similarly omit the kinks in subsequent figures where they play no role in

low-risk indifference curve; and H_1 and H_2 are high-risk indifference curves. Note from (A.1) that the usual single crossing property holds under Wu and Gonzalez' version of the RDU model. Also, it is clear from the fact that w is non-decreasing that any contract meeting the high-risk participation constraint must also meet the low-risk participation constraint.

Included in Figure 4.1 is the line

$$z_2 = \left[1 - r(p^H)\right] z_1 \tag{4.1}$$

Points *above* (respectively, *below*) this line are such that the high type's indifference curve through the point is *steeper* (respectively, *flatter*) than the fair bet line (for p^H probability of loss). Since $c/y = r(p^H)$, the endowment point $E = (y, y - c)$ in Figure 4.1 satisfies (4.1). Thus, the locus of contracts generating zero expected profit (when sold to high risks), is *tangential* to the high risk indifference curve at E. Contracts such as A, which lie above the $z_2 = \left[1 - r(p^H)\right] z_1$ line, are such that the iso-expected profit line through A (for sales to high types only), is flatter that the high type indifference curve H_2 through A.

---

the analysis.

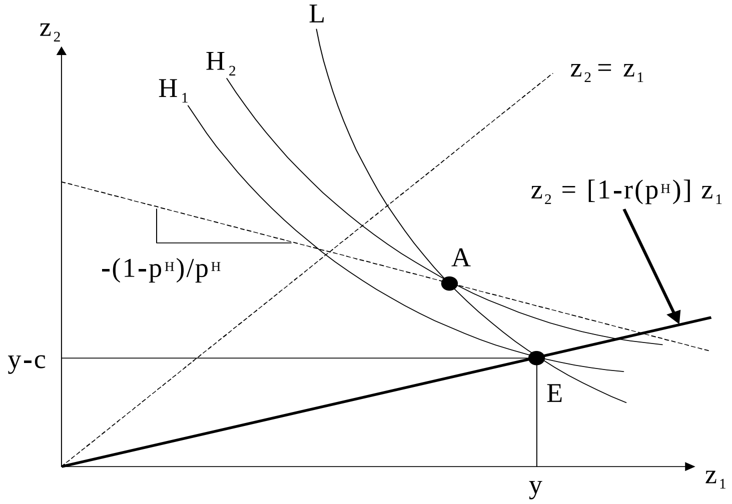

Figure 4.1: Pooling preferable to separation

Suppose that A is the low risk component of a screening menu of contracts. Incentive compatibility requires that the high risk contract must lie to the northwest of A, between the indifference curves labelled L and H_2. Clearly, such a menu will generate lower profit than just offering A alone, since the profit on the high risk contract is lower than it would be if high risks were sold contract A instead. Hence, the usual argument that screening is more profitable than pooling breaks down.

We may now easily prove[13]:

Proposition 4.1. *Let* $u(z) = \sqrt{z}$ *and* w *be given by (2.2), with* $\gamma \in (0, 1)$. *Suppose that*

(i) $c/y \leq r(p^H)$;

[13]Recall that θ is the proportion of high risks in the population.

(ii) *$c/y > r(p^L)$; and*

(iii) $w(1-p^L) \geq 1-p^L$.

Then there exists some $\overline{\theta} \in (0,1)$ such that a monopoly insurer will offer a pooling contract whenever $\theta < \overline{\theta}$, and serve neither type otherwise.

The intuition for the result is as follows. Condition (i) guarantees that high risks cannot profitably be served, while (ii) ensures that there exist gains from trade with low risks. Condition (iii) implies that $p^L \geq 1-\pi^*$. Note that it is possible to find values for c, y, p^H and p^L that satisfy all three conditions if r is strictly increasing, as in Figure 3.4. By the logic of Figure 4.1, if both types are to be served, then offering a single pooling contract is preferable to a screening menu (using condition (i)). Since it is clearly impossible to serve only low risks, a pooling contract will be the most profitable option provided θ is sufficiently close to zero. If θ is too high, the profit maximising strategy is to serve neither type.

4.2. Pooling without cross-subsidisation

In section 4.1 we considered consumers who were sufficiently optimistic about their chances of avoiding the loss c that neither type would be sold full insurance under symmetric information, and high types would not be sold any insurance at all. When type is unobservable, high risks must be served in order to get access to the low risk market. Since the high risk contracts are necessarily loss-making, the insurer's incentive to pool is the minimisation of these losses. However, there are other reasons for pooling RDU consumers. Figure 4.2 illustrates the case in which w is given by (2.2) with $\gamma \in (0,1)$, $1-p^H = \pi^*$ and $1-p^L \in (\pi^*, 1-\pi^*)$. High risks act like risk-averse EU consumers, and full insurance contracts are Pareto optimal for both types. In particular, both segments of the market are profitable under symmetric information. In this case, the incentive to pool comes from the cost of selling "less" insurance to low risks than is being offered to high risks, as the potential gains from trade with former are particularly large.

Consider the full insurance contract P in Figure 4.2. If there were symmetric information, then P would be offered to low risks. Importantly, it yields a level of (first-best) expected profit that is *strictly greater* than if low risks did not distort p^L in decision-making, since $w(1 - p^L) < 1 - p^L$. Thus, under symmetric information, these RDU consumers are more profitable to serve than their EU counterparts (i.e. low risk consumers with the same utility function, u, but a *linear* transformation function, w). Hence, if information is *not* symmetric, it is reasonable to expect that low risks will purchase "more" insurance than in the EU case. In fact, if θ is low enough, both types purchase full insurance contracts, and a Pareto optimal outcome is obtained.

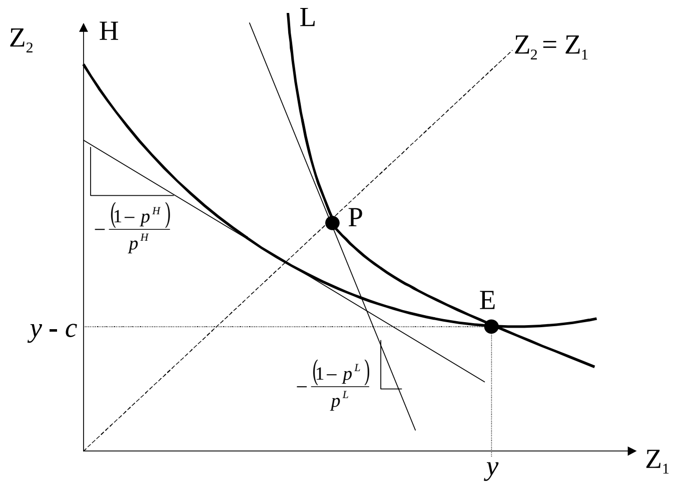

Figure 4.2: Pooling without cross-subsidisation

The basic idea is easily conveyed. In the EU case, low risks receive full insurance *if and only if* $\theta = 0$. For $\theta > 0$ it is strictly more profitable to screen, selling partial insurance to low risks and full insurance to high risks. As $\theta \downarrow 0$, the optimal menu converges smoothly to the first-best full insurance contract for low risks. In the RDU case, the added profitability of insurance sales to low risks implies that convergence is achieved *before* $\theta = 0$. To see why, consider point P in Figure 4.2. Suppose both types are offered this contract. Can expected profit be increased by screening? If the low risk contract is moved down L to point A in Figure 4.3, the insurance company will lose some profits on low risks. On the other hand, the premium charged to high risks for their full insurance contracts can now be increased, thereby increasing profits on sales to this type (see Figure 4.3). In the EU case, the losses on low risks are zero to a first order, so screening is strictly preferable to pooling. However, in the RDU case depicted in Figure 4.3, these losses on low risks are non-zero because of the kink in L. Provided θ is not too large, it is optimal to pool. The details of this sketch argument are provided in the appendix, thereby establishing the following result:

Proposition 4.2. *Let w be given by (2.2) with $\gamma \in (0,1)$, $1 - p^H = \pi^*$ and $1 - p^L \in (\pi^*, 1 - \pi^*)$. For θ sufficiently close to zero, the second-best outcome under asymmetric information is a pooled full-insurance contract.*

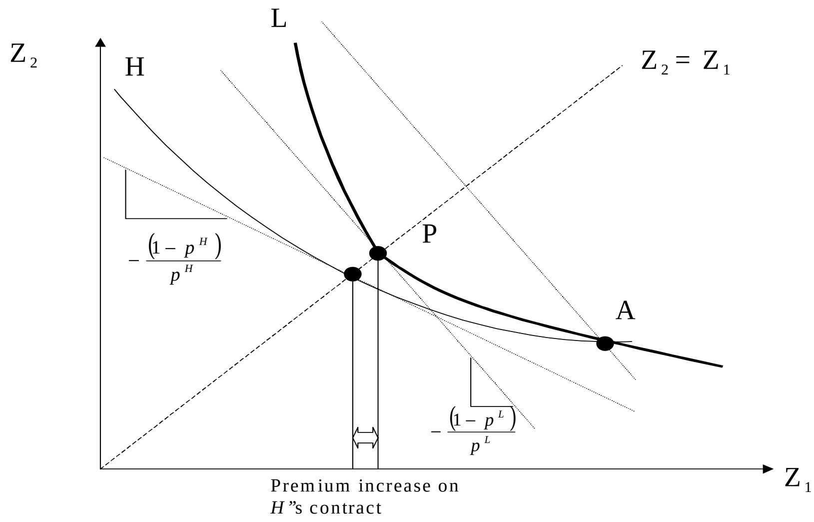

Figure 4.3: Pooling more profitable than screening

Before leaving this section, let us observe that Proposition 4.2 also implies the possibility of pooling even when *all consumers are globally risk-averse*. For example, the proof goes through unchanged if high risk consumers have linear transformation functions (EU), and low risks have convex transformations (the limiting case of an inverse-S in which $\pi^* = 0$). Wakker (forthcoming) refers to the latter as possessing *pessimistic* preferences. Chew, Karni and Safra (1987) verify that such preferences are globally risk averse.

5. Related literature

Machina (1995) studies the robustness of several standard results on insurance demand to deviations from EU. He considers a very wide class of preferences, restricted only by a smoothness condition – Fréchet differentiability (Machina (1982)) – and risk aversion.

Machina shows that most of the conclusions obtained under the stronger EU assumptions continue to hold. Karni (1992) suggests that even weaker smoothness assumptions than Fréchet differentiability (such as Gateaux differentiability) may suffice.

Unlike Machina (1995), we allow for asymmetric information, but only consider robustness of results to a very specific deviation from EU: one that has a high degree of experimental support. It should be noted that the class of RDU preferences we consider are neither Fréchet differentiable, nor globally risk averse (cf. Chew, Karni and Safra (1987)). Karni (1995), commenting on Machina (1995), notes that smoothness assumptions can be restrictive. He highlights in particular the possibility of a non-differentiability at "certainty" under RDU, and suggests that this may have "important implications for insurance" (*ibid.*, p.54). However, as the reader will observe, this non-differentiability plays a relatively minor role in our results. Proposition 4.1 makes no use of it whatsoever, so the possibility of pooling is compatible with smooth indifference curves.[14]

Most studies that apply RDU to insurance fail to impose an inverse-S transformation function, and tend to find only minor deviations from the implications of EU. Doherty and Eeckhoudt (1995) is one example. They assume that the transformation function is convex. They also assume that utility is linear in wealth,[15] and it is only this feature of their model which manages to drive a small wedge between their conclusions and those obtained under EU. Schlee (1995) considers competitive equilibrium in a market for insurance contracts with deductibles. He shows that properties of the equilibrium

[14]It is easy to produce an example along the lines of that in section 4.1 using the alternative transformation function

$$w(p) = \frac{1}{2}(2p-1)^3 + \frac{1}{2},$$

for which $w(p) = 1 - w(1-p)$, thereby eliminating the kink from the indifference curve (A.1).

[15]As in Yaari's (1987) model.

contracts established under EU assumptions remain valid for risk-averse RDU maximisers.

A notable exception in the literature is Wakker, Thaler and Tversky (1997). The authors observe that consumers are strongly averse to “probabilistic insurance” – insurance for which there is some small probability that the insurance company will default on its obligations – requiring substantial premium discounts to accept even very small risks of default. They demonstrate that this evidence is incompatible with risk-averse EU behaviour, but explicable if consumers conform to RDU and exhibit bounded subadditivity.

6. Conclusion

In this paper we adopt the view that consumer behaviour conforms to the RDU model with an inverted-S transformation function. This assumptions has greater experimental support than EU. In fact, since all our results are derived in a two-state model, they are equally compatible with the more general class of “biseparable” preferences (Ghirardato and Marinacci (2000)) with an inverse-S distortion (ibid., p.10).[16]

Unlike many other investigations of the robustness of classical insurance theory, our alteration to the standard model sharply alters its predictions. We have shown that the revised model is compatible with partial or non-insurance of high risks under symmetric information. The non-insurance of high risk types is in fact quite common in the health insurance market. As Newhouse (1996, p.1242) points out:

> “In the Rothschild-Stiglitz model high-risk persons obtain the [health] insurance they wish, and low-risks individuals do not. In reality however, it

[16] We thank Simon Grant for this observation.

is high-risks, not low-risks who tend to have trouble obtaining the desired amount of insurance."

The main result in this paper is that inverted-S RDU preferences are compatible with pooling or complete market failure under adverse selection. Note that this result does not depend on risk-loving behaviour. Even if consumers are globally risk-averse, we have shown that pooling may occur (see the last paragraph of section 4.2). Indeed, in this case, Pareto optimal full insurance contracts may be purchased by both types, eliminating all traces of market failure.

The phenomenon of risk pooling is in fact commonly observed in the health insurance market. For instance, Pauly and Nicholson (1999) argue that managed care organisations are in a pooling equilibrium, with heterogeneous risk types insured by large Health Maintenance Organisations under the same plan. Browne (1992) finds that low-risks subsidise high-risks in the health insurance market, supporting the existence of pooling of the sort discussed in section 4.1.

Our companion paper, Ryan and Vaithianathan (2000), studies insurance under moral hazard, and shows that zero co-insurance is Pareto optimal for low-risk RDU consumers with inverted-S transformation functions. This result, too, is incompatible with risk-averse EU behaviour, but appears consistent with insurance market data. Similarly, Wakker, Thaler and Tversky (1997) have shown that subadditive RDU behaviour may explain the observed aversion of consumers to probabilistic insurance, where EU cannot. It is useful to observe that SA is also the property that drives the results in the present paper.[17]

In summary, RDU does fundamentally alter insurance market predictions when

[17] The result on optimal co-insurance in Ryan and Vaithianathan (2000) uses, in addition, the convexity of w above π^*.

asymmetric information is present. This is so under the assumption of global risk aversion – Proposition 4.2 (cf. Machina (1995, p.36)) – or if the more experimentally supported inverse-S transformation is assumed – Proposition 4.1, and Ryan and Vaithianathan (2000).

Appendix

Proof of Lemma 3.1. (1) Let $h(z_1, z_2, \overline{u})$ denote the slope of the indifference curve through $(\overline{z}, \overline{z})$, where $u(\overline{z}) = \overline{u}$, evaluated at the point (z_1, z_2). When $z_1 \neq z_2$, direct calculation gives

$$h(z_1, z_2, \overline{u}) = \begin{cases} \frac{-(1-w(p))}{w(p)} \frac{u'(z_1)}{u'(z_2)} & \text{if } z_2 > z_1 \text{ and } u(z_1) + w(p)\,[u(z_2) - u(z_1)] = \overline{u} \\ \\ \frac{-w(1-p)}{1-w(1-p)} \frac{u'(z_1)}{u'(z_2)} & \text{if } z_1 > z_2 \text{ and } u(z_2) + w(p)\,[u(z_1) - u(z_2)] = \overline{u} \end{cases} \tag{A.1}$$

Since

$$\frac{1 - w(p)}{w(p)} = \frac{(p^\gamma + (1-p)^\gamma)^{1/\gamma} - p^\gamma}{p^\gamma}$$

and

$$\frac{w(1-p)}{1 - w(1-p)} = \frac{(1-p)^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma} - (1-p)^\gamma}\ ,$$

it clearly suffices to show that

$$(p^\gamma + (1-p)^\gamma)^{1/\gamma} - p^\gamma > (1-p)^\gamma \tag{A.2}$$

for any $p \in (0, 1)$. But $\gamma \in (0, 1)$ implies

$$(p^\gamma + (1-p)^\gamma)^{1/\gamma} > (p^\gamma + (1-p)^\gamma)$$

and hence

$$\begin{aligned}(p^\gamma + (1-p)^\gamma)^{1/\gamma} - p^\gamma &> (p^\gamma + (1-p)^\gamma) - p^\gamma \\ &= (1-p)^\gamma\end{aligned}$$

This confirms (A.2), so (1) is proved.

(2) Since u is strictly concave, $u'(z_1)/u'(z_2)$ is strictly increasing in z_2 and strictly decreasing in z_1. Therefore, indifference curves have negative and diminishing (from left to right) slope, both above and below the certainty line. Since the calculations in the proof of (1) revealed that

$$\frac{1-w(p)}{w(p)} > \frac{w(1-p)}{1-w(1-p)},$$

upper contour sets are therefore convex (see (A.1)). □

Proof of Proposition 4.1: The logic of the argument is illustrated in Figure A.1, which assumes $c/y = r(p^H)$ and $w(1-p^L) > 1-p^L$. Because of (i), if high types are served, the insurance company makes a loss on these contracts. Hence, if high types are served, low types must also be served, with low-type contracts being strictly profitable. Hence, either both types are served, with low-types cross-subsidising high, or neither type is served, since it is not possible to serve low types only.

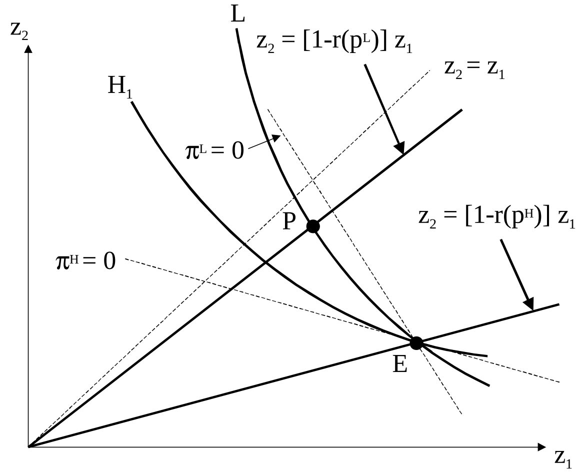

Figure A.1: Pooling with cross-subsidisation

Suppose then that the insurance company serves both types. Since the low-type contract must be strictly profitable and meet the low-risk's participation constraint, it must lie between the line $\pi^L = 0$, which is the set of contracts yielding zero expected profit when sold to low-risks, and the indifference curve labelled L. By (i), any such contract is (on or) above the line $z_2 = [1 - r(p^H)]z_1$, so it is optimal to offer the same contract to high risks.

Thus, if both types are served, a pooling contract must be used, and the profit maximising location for this contract will be somewhere on L between P and E. Note that such a contract leaves no incentive to screen high types (because of (i)), or low types (because of (ii)).

If $\theta = 0$, it is obvious that contract P is optimal. As θ increases, the optimal pooling contract moves down the indifference curve L. If θ becomes high enough, E will be optimal – neither type will be served – since the losses on high risks cannot be recovered on the small population of low risks.

The critical level $\overline{\theta}$ is straightforward to calculate. By considering (A.1), one may easily verify that conditions (i) and (ii) imply

$$p^H \;\geq\; \frac{1 - w(1-p^H)}{1-(1-k)w(1-p^H)} \;>\; \frac{1-w(1-p^L)}{1-(1-k)w(1-p^L)} \;>\; p^L \tag{A.3}$$

where $k = \sqrt{(y-c)/y}$. The first inequality is (i); the second uses $p^H > p^L$; and the third is (ii). Let $p(\theta) = \theta p^H + (1-\theta)p^L$. A strictly profitable pooling contract exists provided

$$p(\theta) \;>\; \frac{1-w(1-p^L)}{1-(1-k)w(1-p^L)} \tag{A.4}$$

This ensures that the iso-expected profit locus for pooling contracts is strictly steeper than the low risk indifference curve at the "no insurance" point (E in Figure A.1). The inequalities in (A.3) ensure that there exists some $\theta \in (0,1)$ which satisfies (A.4). In particular, $\overline{\theta}$ is defined by the condition

$$p(\overline{\theta}) \;=\; \frac{1-w(1-p^L)}{1-(1-k)w(1-p^L)}$$

$$\Leftrightarrow \quad \overline{\theta} \;=\; \frac{\left[1-w(1-p^L)\right] \,-\, p^L\left[1-(1-k)w(1-p^L)\right]}{(p^H-p^L)\left[1-(1-k)w(1-p^L)\right]}$$

For $\theta \geq \overline{\theta}$, the "null" contract E will be offered, and neither type will obtain any insurance. $\square$

Proof of Proposition 4.2: Standard arguments imply that the second-best outcome will involve low risks purchasing a contract on L between the points P and E in Figure 4.2. Let $\hat{p} = \theta p^H + (1-\theta)p^L$. Provided θ is small enough, $\hat{p}$ will be sufficiently close to p^L that P is the most profitable of the pooling contracts – see Figure A.2. Does there exist a contractual outcome that is more profitable than offering the pooling contract P?

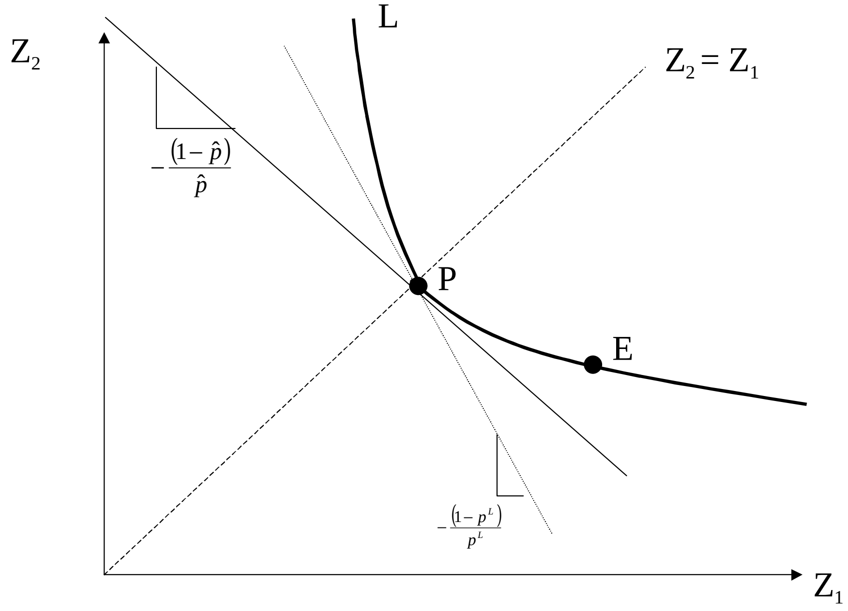

Figure A.2: Optimal pooling contract

The insurance company's options may be usefully divided into the following three categories: (i) offer a single contract attractive only to high risks; (ii) offer a single contract attractive to both types (pooling); and (iii) offer a screening menu of contracts. Suppose we could demonstrate that for θ near zero, there exists a pooling contract that is more profitable than any screening menu. Then, since there clearly exists a $\hat{\theta} > 0$ such that pooling is better than serving only high risks when $\theta < \hat{\theta}$ (recall that low types are profitable to serve under symmetric information), we may conclude that pooling is optimal for θ sufficiently close to zero. Finally, we have already observed that the best pooling contract provides full insurance for θ near zero, so the result would follow.

It therefore suffices to prove that the contract P generates higher expected profit than screening for low values of θ. Consider Figure A.3. By standard Stiglitz (1977) logic, an optimal screening menu will serve low risks with a contract on L between P and E, and high risks with a full insurance contract (as in Figure 4.3). The pooling contract P is a limiting case of such a menu, and is clearly optimal when $\theta = 0$. Let us index all such menus by $z_1 \in [a, b]$, the net wealth of low risks in state 1 (the "good" state) – see Figure A.3. Define the function $f^L : [a, b] \to \mathbb{R}$ by the condition

$$u(f^L(z_1)) + w(1-p^L)\left[u(z_1) - u(f^L(z_1))\right] = u(y-c) + w(1-p^L)\left[u(y) - u(y-c)\right].$$

That is, for each $z_1 \in [a, b]$, $(z_1, f^L(z_1))$ is the corresponding point on the indifference curve L. Since u is strictly increasing and continuously differentiable, f^L is continuously differentiable. Next, define the function $f^H : [a, b] \to \mathbb{R}$ as follows:

$$u(f^H(z_1)) = u(f^L(z_1)) + (1-p^H)\left[u(z_1) - u(f^L(z_1))\right].$$

Thus, $(f^H(z_1), f^H(z_1))$ is the most profitable contract for high risks when $(z_1, f^L(z_1))$ is available for low risks. By the continuous differentiability of f^L and the aforementioned properties of u, f^H is also continuously differentiable.

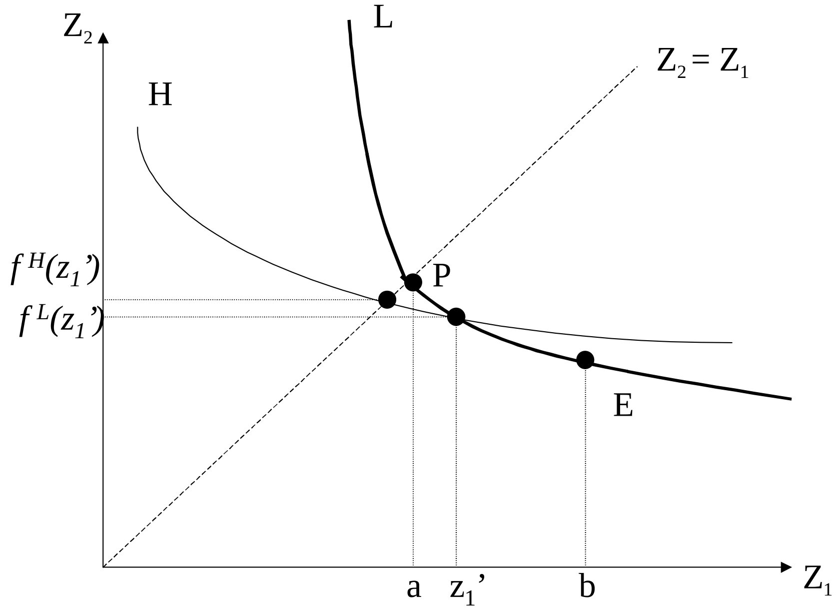

Figure A.3: A screening menu

Finally, let us define the following expected profit function:

$$\Pi(z_1, \theta) \;=\; \theta\left[y - f^H(z_1) - p^H c\right] \;+\; (1-\theta)\left[(y - z_1) + p^L\left(z_1 - c - f^L(z_1)\right)\right] \quad \text{(A.5)}$$

The RHS of (A.5) is the expected profit (per customer) from offering the contract menu $\big((z_1,\, f^L(z_1)),\, (f^H(z_1),\, f^H(z_1))\big)$.

Consider the difference

$$\Pi(z_1, \theta) \;-\; \Pi(a, \theta) \;=$$

$$\theta\left[a - f^H(z_1)\right] \;+\; (1-\theta)\left[\left(1 - p^L\right)(a - z_1) + p^L\left(a - f^L(z_1)\right)\right] \quad \text{(A.6)}$$

This is the net gain from using the menu $\big((z_1,\, f^L(z_1)),\, (f^H(z_1),\, f^H(z_1))\big)$ instead of the pooling contract P. The first term on the RHS represents the average gain on the sales to high risks, since $f^H(z_1) \leq a$ (with equality if and only if $z_1 = a$). The second term represents the average loss on sales to low risks, since $f^L(z_1) \leq a$ (with equality if and only if $z_1 = a$). We now want to show that for θ sufficiently small, $\pi(z,\theta) - \pi(a,\theta) < 0$, for all $z_1 \in (a, b]$. Let

$$\delta \;=\; \max_{z_1 \in [a,b]} \left| \frac{df^H(z_1)}{dz_1} \right| \;>\; 0$$

(taking one-sided derivatives when $z_1 = a$ or $z_1 = b$).[18] Then the profit gain in (A.6)

$$\theta \left[a - f^H(z_1) \right]$$

is bounded above by $\theta\delta(z_1 - a)$. That is :

$$\theta \left[a - f^H(z_1) \right] \;\leq\; \theta\delta(z_1 - a) \tag{A.7}$$

Similarly, we may bound the loss by noting that

$$\max_{z_1 \in [a,b]} \left| \frac{df^L(z_1)}{dz_1} \right| \;=\; \left(f^L\right)'_+(a) \;=\; \frac{w(1-p^L)}{1 - w(1-p^L)}.$$

Hence:

$$\frac{(1-\theta)\left[(1-p^L)(a - z_1) + p^L\left(a - f^L(z_1)\right)\right]}{(z_1 - a)} \;\leq\; (1-\theta)\left[\frac{p^L w(1-p^L)}{1 - w(1-p^L)} \;-\; (1-p^L) \right] \tag{A.8}$$

Let

$$\beta \;=\; \left[(1-p^L) \;-\; \frac{p^L w(1-p^L)}{1 - w(1-p^L)} \right] \;>\; 0.$$

Combining (A.7) and (A.8), a sufficient condition for $\Pi(z_1,\theta) - \Pi(a,\theta) < 0$ for all $z_1 \in (a, b]$ is that

$$\theta \;<\; \frac{\beta}{\delta + \beta}.$$

Hence, for such θ, $\Pi(z_1,\theta) < \Pi(a,\theta)$ for all $z_1 \in (a, b]$, so P is strictly more profitable than any screening menu of contracts. □

[18] The maximum is well defined. For example, it is equal to

$$\max\left\{ \max_{z_1 \in [a,\, a+\frac{1}{2}(b-a)]} \left| \left(f^H\right)'_+(z_1) \right|,\; \max_{z_1 \in [a+\frac{1}{2}(b-a),\, b]} \left| \left(f^H\right)'_-(z_1) \right| \right\},$$

which exists because each one-sided derivative is continuous over the specified compact domain.

References

[1] Allais, M. (1953): "Le Comportement de l'Homme Rationel Devant le Risque, Critique des Postulates et Axiomes de l'Ecole Américaine," *Econometrica* **21**, 503–546.

[2] Browne, M.J. (1992): "Evidence of Adverse Selection in the Individual Health Insurance Market," *Journal of Risk and Insurance* **59(1)**, 13-33.

[3] Browne, M.J. and H.I. Doerpinghaus (1993): "Information Asymmetries and Adverse Selection in the Market for Individual Medical Expense Insurance," *Journal of Risk and Insurance* **60(2)**, 300-312.

[4] Camerer, C. and T.-H. Ho (1994): "Violations of the Betweenness Axiom an Nonlinearity in Probability," *Journal of Risk and Uncertainty* **8**, 167–196.

[5] Chateauneuf, A and P. Wakker (1999): "An Axiomatization of Cumulative Prospect Theory for Decision Under Risk" *Journal of Risk and Uncertainty* **18**, 137-145

[6] Chew, S.H., E. Karni and Z. Safra (1987): "Risk Aversion in the Theory of Expected Utility with Rank Dependent Probabilities," *Journal of Economic Theory* **42**, 370–381.

[7] Cohen, M. (1995): "Risk Aversion Concepts in Expected- and Non-Expected-Utility Models," *Geneva Papers on Risk and Insurance Theory* **20**, 73–91.

[8] Doherty, N.A. and L. Eeckhoudt (1995): "Optimal Insurance Without Expected Utility: The Dual Theory and the Linearity of Insurance Contracts," *Journal of Risk and Uncertainty* **10**, 157–179.

[9] Ghirardato, P and M. Marinacci (2000): "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Social Science Working Paper* 1085, California Institute of Technology.

[10] Karni, E. (1992): "Optimal Insurance: A Nonexpected Utility Analysis," in G. Dionne (ed.) *Contributions to Insurance Economics*, Kluwer Academic.

[11] Karni, E. (1995): "Non-Expected Utility and the Robustness of the Classical Insurance Paradigm: Discussion," *Geneva Papers on Risk and Insurance Theory* **20**, 51–56.

[12] Machina, M. (1982): "'Expected Utility' Analysis without the Independence Axiom," *Econometrica* **50**, 277–323.

[13] Machina, M. (1995): "Non-Expected Utility and the Robustness of the Classical Insurance Paradigm," *Geneva Papers on Risk and Insurance Theory* **20**, 9–50.

[14] Newhouse (1996) "Reimbursing Health Plans and Health Providers: Efficiency in Production versus Selection" *Journal of Economic Literature* 1236-1263.

[15] Pauly, M and Nicholson, S (1999) "Adverse Consequences of Adverse Selection" *Journal of Health Politics, Policy and Law* **24(5)**, 921-930.

[16] Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization* **3**, 323–343.

[17] Ryan, M.J. and R. Vaithianathan (2000): "Medical Insurance With Rank-Dependent Expected Utility," *mimeo*, The Australian National University.

[18] Schlee, E. (1995): "The Comparative Statics of Deductible Insurance in Expected- and Non-Expected-Utility Theories," *Geneva Papers on Risk and Insurance Theory* **20**, 52–72.

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

[20] Tversky, A. and D. Kahneman (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty* **5**, 297–323.

[21] Tversky, A. and P. Wakker (1995): "Risk Attitudes and Decision Weights," *Econometrica* **63**, 1255–1280.

[22] Wakker, P.P. (2000): "Testing and Characterizing Properties of Nonadditive Measures Through Violation of the Sure-Thing Principle," *Econometrica (forthcoming).*

[23] Wakker, P.P., R.H. Thaler and A. Tversky (1997): "Probabilistic Insurance," *Journal of Risk and Uncertainty* **15**, 7–28.

[24] Wu, G. and R. Gonzalez (1996): "Curvature of the Probability Weighting Function," *Management Science* **42**, 1676–1690.

[25] Yaari, M.E. (1987): "The Dual Theory of Choice Under Risk," *Econometrica* **55**, 95–115.