University of
St Andrews

School of Economics & Finance Discussion Papers

Efficient Nash Equilibrium under Adverse Selection

Theodoros M. Diasakos and Kostas Koufopoulos

School of Economics & Finance
Discussion Paper Series
issn 0962-4031
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

School of Economics and Finance Discussion Paper No. 1313

JEL Classification: D86
Keywords: Insurance Market, Adverse Selection, Incentive Efficiency

Efficient Nash Equilibrium under Adverse Selection

Theodoros M. Diasakos[*] Kostas Koufopoulos[†]

August 26, 2013

Abstract

This paper revisits the problem of adverse selection in the insurance market of Rothschild and Stiglitz [28]. We propose a simple extension of the game-theoretic structure in Hellwig [14] under which Nash-type strategic interaction between the informed customers and the uninformed firms results always in a particular separating equilibrium. The equilibrium allocation is unique and Pareto-efficient in the interim sense subject to incentive-compatibility and individual rationality. In fact, it is the unique neutral optimum in the sense of Myerson [22].

Keywords: Insurance Market, Adverse Selection, Incentive Efficiency.
JEL Classification Numbers: D86

---

[*]Corresponding Author: School of Economics & Finance, University of St Andrews, Scotland (E-mail: td31@st-andrews.ac.uk)

[†]Warwick Business School, University of Warwick, U.K. Thus far, we have benefited greatly from the comments of Tasos Dosis, Dino Gerardi, Dominik Grafenhofer, Martin Hellwig, Eric Maskin, Heracles Polemarchakis, and Paolo Siconolfi. All errors are ours.

1 Introduction

This paper readdresses an old but still open question in applied micro-economic theory: how a competitive market will allocate insurance policies when firms cannot distinguish amongst the different risk-classes of customers. To this end, it restricts attention to the simplest depiction that captures the essential features of this question, an economy in which each of a continuum of agents observes a binary parameter indicating the probability of suffering an income loss. For this version of the problem, we propose a game-theoretic structure under which Nash-type strategic interaction between the informed insurees and uninformed insurers delivers the strongest of results. The equilibrium is unique and sorts the two types by maximizing the welfare of the low-risk agents. More importantly, it does so in a way that renders the equilibrium outcome interim incentive efficient (i.e, Pareto-efficient in the interim sense and subject to incentive-compatibility and individual rationality). In fact, and in a sense to be made precise below, it is the most desirable allocation on the interim incentive efficient (IIE) frontier.

Needless to say, we examine the interactions between the market participants under the lenses of non-cooperative game-theory. This has become the standard tool for analyzing markets with adverse selection because of its main virtue, every detail of the economic environment is made explicit. Indeed, a well-defined extensive-form game with incomplete information describes all the institutional details of the market, the information that is available to each of the players and the actions they may take. With respect to studying markets for insurance provision, this approach was first adopted by Rothschild and Stiglitz [28]. Yet, the main implication of their study (as well as of the subsequent generalization by Riley [27]) was that, abstracting away from specific market structures but viewing market participants as engaging in Nash-type strategic behavior, adverse selection can be too cumbersome for competitive markets to function, even under the simplest of settings. It limits the form of contractual arrangements that are consistent with equilibrium, to the extent that it can even preclude its existence altogether.

Specifically, under Nash-type strategic behavior (it is common knowledge amongst all players that no player can influence the actions of any other player), there might be no (pooling) equilibrium arrangement offering a single price per unit of coverage. This occurs when insurance firms have an incentive to charge higher prices for greater coverage because, by doing so, they are able to sort their lower-risk customers from the higher-risk ones (for whom additional coverage yields greater marginal benefits). In this case, the only possible contractual arrangement is separating with each risk-class paying its own premium, equal to its true accident probability. Unfortunately, the Rothchild-Stiglitz (RS) allocation is a viable equilibrium only under limiting conditions.

This alarming observation led to the emergence of a significant body of literature whose principal aim has been to propose allocation mechanisms, along with implementing market structures, which ensure that always some allocation will be supported as competitive equilibrium, under some associated notion of equilibrium. The respective models can be broadly classified into three sets, based upon the extent to which the mechanism allows the players' behavior to be strategic. All but two, however, share an unsatisfactory feature: whenever the RS allocation is not IIE, the same is almost always true also for the suggested equilibria. The two exemptions are Miyazaki [21] and Bisin and Gottardi [2] but either suffers from deficiencies regarding the implementation of the proposed equilibrium.

One class of models has focused on Walrasian mechanisms. In its purest form, this approach was initiated by Prescott and Townsend [25]-[26] and revisited recently by Rustichini and Siconolfi [29]. The central message of these papers is that general economies with adverse selection do not always admit pure Walrasian equilibrium pricing systems and, when they do, the resulting allocations are not necessarily IIE. To guarantee existence, some studies have introduced rationing (Gale [10]-[11], Guerrieri et al. [13]) or suppressed the requirement that firms are profit-maximizing, imposing at the same time quantity constraints on trade (Dubey and Geanakoplos [5], Dubey et al. [6]).

Either of these approaches arrives at some equilibrium that is essentially unique and involves a separating allocation. Typically, however, this is not IIE while uniqueness obtains by restricting the out-of-equilibrium actions and beliefs often in strong ways. Under rationing, the refinement criteria range from subgame perfection (Guerrieri et al. [13]) to the Universal Divinity of Banks and Sobel [1] (Gale [10]) or the Stability of Kohlberg and Mertens [16] (Gale [11]). The latter notion has been deployed also under quantity constraints (Dubey and Geanakoplos [5], Dubey et al. [6]) but seems to be more binding in that environment. As shown in Martin [18], its weakening to something akin to trembling-hand perfection allows for many pooling equilibria which typically Pareto-dominate the separating allocation but are not IIE either.

Instead of constraining the Walrasian mechanism, Bisin and Gottardi [2] enhance it with the implicit presence of institutions that monitor trade appropriately. Restricting attention to the same insurance economy as the one in he present paper, they show that the RS allocation obtains always as the unique Walrasian equilibrium if there are markets for contingent claims in which agents trade only incentive-compatible contracts. To ensure that incentive efficiency is attained whenever the RS allocation is not IIE, they introduce also markets for consumption rights. The ensuing Arrow-Lindahl equilibria internalize the consumption externality due to adverse selection. In fact, by varying the endowment of consumption rights, the authors are able to trace the entire IIE frontier but for one point. The latter, which is no other than the unique equilibrium allocation in the present paper, can be obtained only as the limit of a sequence of equilibria.

Another perspective has been to look at mechanisms in which competitive equilibrium is supported by strategic behavior. This has produced two separate lines of approach. In earlier models, some of the players exhibit strategic behavior which is not of the Nash-type. Specifically, the sellers in Wilson [31], Riley [27], Engers and Fernandez [8], and Miyazaki [21] but also the buyers in Grossman [12] are able to foresee the unraveling of equilibrium Rothschild and Stiglitz warned about and modify their plans so as to prevent it. By contrast, most later studies have been built upon the game-theoretical foundation in Hellwig [14] or its generalization in Maskin and Tirole [19]. Under these structures, whenever the RS allocation is not IIE, a multiplicity of contractual arrangements can be supported as sequential equilibria. Within the richness of the resulting equilibrium set, however, the IIE subset is of negligible size.

Evidently, even though there are by now many views about how a competitive market might allocate insurance policies under adverse selection, a crucial question has been left open: whether, if at all, and under which game-theoretic structure a competitive market whose participants engage in Nash-type strategic behavior may ensure that resources are allocated efficiently. This motivates the present paper which answers this question in the affirmative, and in the strongest sense, using structural elements that are well-known in the literature. Our main message is that a simple extension of Hellwig's game-structure delivers always and uniquely a particular efficient allocation as Nash equilibrium.

As we argue in the sequel, the analysis of Rothschild and Stiglitz can be interpreted by means of a two-stage game in which, at stage 1, the firms make binding offers of insurance contracts while, at stage 2, the customers choose amongst them. By contrast, Hellwig turns the offers firms make at stage 1 non-binding by adding a third stage in which, after observing the other firms' contractual offers at stage 1 and the customers' choices at stage 2, a firm may withdraw any of its own contracts. Regarding this game, we envision expanding the strategy space of the insurance providers along two dimensions. Our firms may subsidize their net income across contracts by offering menus of them at stage 1. They can also publicly pre-commit, if they so wish, to an offer on either of two levels: not withdrawing a contract at stage 3, irrespectively of the history of play at that point (commitment on the contractual level), or not withdrawing an element of a menu unless they withdraw the menu itself (commitment on the policy level).

Under this structural enhancement, the IIE allocation that maximizes the welfare of low-risk customers can always be supported as the unique equilibrium, for a given distribution of the two risk-types in the population, even when one does not exist in the Rothschild-Stiglitz setting. The equilibrium outcome coincides with the RS allocation when the latter is IIE. Otherwise, it involves cross-subsidization across risk-classes but also contracts. Each class pays a different risk premium, the one paid by the high-risk (low-risk) agents being less (more) than their true accident probability. As a result, insurers expect losses on their high-risk customers to be offset by profits from the low-risk ones.

In fact, our equilibrium allocation is the one suggested by Miyasaki who was the first to allow suppliers to offer menus rather than single contracts. His focus, however, was on adverse selection in the labor market and he chose to identify a firm with its wage-structure, its menu of wage-effort contracts. As a result, he viewed free entry and exit in the labor market as dictating that a firm may withdraw its menu but not only a single contract from that menu. This restriction, which is fundamental for Miyasaki's analysis, was heavily criticized in the realm of insurance markets by Grossman. This author pointed out that insurance suppliers more often than not require buyers to submit applications. That is, they may indeed offer menus of contracts but are also able to withdraw specific contracts from these menus by simply rejecting the corresponding applications. Allowing firms to do so, Grossman concluded that the equilibrium contractual arrangement ought to entail pooling, unless it is the RS allocation.

We do take into account this insight but also another equally realistic element of insurance suppliers' behavior: they often choose to send certain customers "pre-approved" applications. As long as it is public belief that the latter term is binding in an enforceable manner, the two elements together permit an insurance supplier not only to withdraw a particular contract from a menu at stage 3, but also to publicly-commit, at stage 1, to not withdraw it. As it turns out, endogenizing in this way the commitment of firms upon their insurance promises restricts dramatically the equilibrium set. In conjunction with endogenous commitment on the policy level, it delivers always the same singleton one. As it happens, on the one hand, commitment on the contractual level is not observed in equilibrium. "Pre-approved" applications are deployed only off the equilibrium path to restrict the players' beliefs so that the many equilibria in the standard version of Hellwig's game are ruled out. On the other, the equilibrium insurance menu is introduced under commitment on the policy level.

Needless to say, it is not simply the interaction between menus and Hellwig's game that drives our result. Cross-subsidization of net income between contracts has been considered also by Maskin and Tirole [19] under a game-theoretic structure similar to Hellwig's but a much more general interpretation of contractual arrangements. This paper identified the set of equilibrium allocations that would emerge if the latter are actually mechanisms: specifications of a game-form to be played between two parties, the set of possible actions for each, and an allocation for each pair of strategies. Even though the authors' main focus was on signalling, they established that the set of equilibrium outcomes remains essentially the same under screening as long as the out-of-equilibrium actions and beliefs are left unrestricted. Much in agrement with Hellwig's intuition, this set is rich to an extent that renders its IIE subset negligible.

By comparison, the present paper demonstrates the necessity of restricting the market participants' out-of-equilibrium beliefs in order to arrive at the given IIE outcome. Our augmented version of Hellwig's game is an example of a mechanism that restricts the out-of-equilibrium beliefs appropriately. It does so by relying heavily on the notion of endogenous commitment and, in this sense, attests to the important role of what are called public actions in Myerson [22]. These are enforceable decisions individual players can publicly-commit themselves to carry out, even if they may turn out ex-post to be harmful to themselves or others. Myerson assigns to the set of public actions center-stage in establishing the existence of neutral optima. The latter form the smallest class of incentive-compatible allocations that are attainable as sequential equilibria of the game in which the informed party proposes mechanisms and satisfy four fundamental axioms of mechanism selection. As we argue in the sequel, our equilibrium allocation is the unique neutral optimum for the insurance economy under study.

The rest of the paper is organized as follows. The next section presents the market for insurance provision in the context of Hellwig's three-stage game. It revisits important results which are commonly-used in applications of Hellwig's model but have been shown rather heuristically in the literature. In particular, we identify their strategic underpinnings and show how, viewed under the light of contractual commitment, they set the stage for our main result. This is presented in Section 3 which analyzes how the interplay between commitment on the contractual and policy levels leads in fact to efficient insurance provision. In Section 4, we discuss and interpret our findings further, in particular vis a vis relevant ones in the literature. Section 5 concludes. It is followed by an Appendix containing the analytical version of our arguments. By contrast, wherever possible, the main text presents the economic intuition behind our claims via graphical examples of cases in which they are true.

2 The Simple Model Revisited

To parsimoniously describe adverse selection in the market for insurance provision, imagine that, after having inferred as much as possible from observable characteristics, the insurance firms have grouped a continuum of customers into two classes of otherwise identical individuals.[1] Across these, the agents differ only in the probability of having an accident, which is known by no one else but the agent herself. For the low-risk class, which contains a fraction $\lambda \in (0, 1)$ of the population, this probability is p_L. The high-risk class includes the remainder of individuals whose accident probability is p_H, with $0 < p_L < p_H < 1$.

Each individual is endowed with wealth $W \in \mathbb{R}_{++}$, to be reduced by the amount $d \in (0, W)$ if she suffers an accident. She may insure herself against this event by accepting an insurance contract $\mathbf{a} = (a_0, a_1) \in \mathbb{R}^2_+$. That is, by paying a premium a_0 if no accident occurs (state $s = 0$) in exchange for receiving the net indemnity a_1 otherwise (state $s = 1$). Having entered this agreement with an insurance supplier, her state-contingent wealth is given uniquely by the vector $\mathbf{w} = (w_0, w_1) = (W - a_0, W - d + a_1)$, a transfer of wealth across states at the premium rate $\frac{\mathrm{d}w_1}{\mathrm{d}w_0} = -\frac{\mathrm{d}a_0}{\mathrm{d}a_1}$. Her preferences over such vectors (equivalently, over the respective contracts) admit an expected utility representation with an identical for all consumers, strictly-increasing, strictly-concave, twice continuously-differentiable Bernoulli utility function $u : \mathbb{R}_{++} \mapsto \mathbb{R}$. For an agent of risk-type $h \in \{L, H\}$, the preference relation will be denoted by $\succsim_h$, its representation being $U_h(\mathbf{w}) = (1 - p_h) u(w_0) + p_h u(w_1)$.

On the supply side of the market, insurance is provided by risk-neutral firms which maximize expected profits: $\Pi_p(\mathbf{a}) = (1 - p) a_0 - p a_1$, when the typical insurance contract is sold to a pool of customers whose average accident probability is $p \in [0, 1]$. It will be convenient to use the particular notation $\Pi_h(\cdot)$ and $\Pi_M(\cdot)$ whenever this probability is, respectively, p_h or the population average, $\overline{p} = \lambda p_L + (1 - \lambda) p_H$. These firms are supposed to have adequate financial resources to be willing and able to supply any number of insurance contracts they think profitable.

In fact, they may supply any collection of contracts that is expected to deliver aggregate profits, even if some of its members might be loss-making in expectation. As will be apparent in the sequel, in the market under study, the relevant collections of this kind are binary and will be referred to henceforth as insurance menus. Unless otherwise stated, the typical one $\{\mathbf{a}_L, \mathbf{a}_H\}$ is separating (the subscript indicating the respective risk-class the contract is meant for) with single contracts corresponding to trivial menus $\{\mathbf{a}, \mathbf{a}\}$. The latter will be referred to as pooling policies if meant to be bought by customers of either risk-class.[2] Needless to say, designing insurance provision in this

[1] The continuum hypothesis is standard in models of this type. It allows us to invoke the strong law of large numbers and claim that an insurance supplier whose policy will serve both types of customers can expect, with virtual certainty, the composition of risk-types in its client pool to be identical to that in the population.

[2] Formally (see Step 5 of our RSW analysis in the Appendix), the distinction between pooling policies and sepa-

way entails the usual incentive-compatibility and individual-rationality constraints:

$$U_h(\mathbf{w}_h) - U_h(\mathbf{w}_{h'}) \geq 0 \tag{1}$$

$$U_h(\mathbf{w}) \geq \overline{u}_h \equiv U_h(W, W-d) \qquad h, h' \in \{L, H\} \tag{2}$$

The insurance market is taken to be competitive in that there is free entry and exit. In equilibrium, therefore, we may observe only menus that expect at least zero aggregate profits. As a consequence, with respect to pooling policies, the admissible space consists of contracts $\mathbf{a} \in \mathbb{R}_+^2$ that satisfy

$$\begin{aligned} U_h(\mathbf{w}) &\geq \overline{u}_h \qquad h, h' \in \{L, H\} \\ \Pi_M(\mathbf{a}) &\geq 0 \end{aligned}$$

Moreover, any menu may be supplied if expected to be demanded (i.e. at least one of its contracts is expected to be bought given that customers choose insurance contracts to maximize their expected utility) and profitable. The workings of this market will be modeled by means of the three-stage game in Hellwig [14]. At stage 1, the insurance companies offer menus of contracts. At stage 2, customers choose contracts from these menus to apply for, each being allowed to apply for only one contract. At stage 3, the firms may reject whatever applications they have received at stage 2.

To make predictions, we will use the notion of sequential (equivalently, with only two risk-types, perfect Bayesian) equilibrium. We seek that is a vector of strategies - one for the firms and one for each type of customer - and a vector of beliefs - at each information set in the game tree - such that the strategies are optimal at each point (sequential rationality) given that the beliefs are (fully) consistent. Under this notion and using the terms "honoring" (or "not withdrawing") a contract to mean that none of its applications is rejected at stage 3, an equilibrium insurance menu is such that (a) in equilibrium, each of its constituent contracts is honored at stage 3 and chosen by at least one risk-class of customers, and (b) there is no other admissible menu that, if offered alongside the one in question, would expect strictly positive profits.

Even though by now standard in the pertinent literature, this definition hinges upon the market participants' beliefs about the profitability of insurance menus. And these beliefs are unambiguous only under full information, in which case the equilibrium set is a singleton, the strictly-separating menu $\{\mathbf{a}_L^F, \mathbf{a}_H^F\}$ where $\mathbf{a}_h^F$ maximizes the expected utility of risk-type h amongst the contracts that break even when demanded exclusively by this type.[3] Otherwise, under adverse selection, these beliefs depend fundamentally upon two defining features of the model: the type of separating

rating menus is that the latter entail at least one strict inequality in (1). If both inequalities bind, the menu will be referred to as strictly separating.

[3] Let $FO_p^k = \{\mathbf{a} \in \mathbb{R}_+^2 : \Pi_p(\mathbf{a}) = k\}$ be the level set of expected profits for some pair $(p, k) \in [0, 1] \times \mathbb{R}_+$. In the (a_0, a_1)-space, this is a line of slope $\frac{\mathrm{d}a_1}{\mathrm{d}a_0} = -\frac{1-p}{p}$ with $k = 0$ defining that through the trivial contract $\mathbf{a} = \mathbf{0}$, the endowment point $(W, W-d)$ in the (w_0, w_1)-space. When the accident probability in question is, respectively, p_h or $\overline{p}$, the latter line will be referred to as the fair-odds line of risk-type h (FO_h^*) or of the market (FO_M^*). In the (w_0, w_1)-space, $\mathbf{a}_h^F = \arg\max_{\mathbf{a} \in \mathbb{R}_+^2 : \Pi_h(\mathbf{a}) \geq 0} U_h(\mathbf{w})$ corresponds to the point of tangency between the indifference curve of risk-type h and the line FO_h^*. This coincides with the intersection of the latter with the 45-degree line, the locus of full-insurance $\{\mathbf{w} \in \mathbb{R}_{++}^2 : w_0 = w_1\}$.

menus it admits and the strategies under which insurance policies are marketed. In what follows, we investigate this relation and its implications when all players' strategies are of the Nash-type.

2.1 The Rothschild-Stiglitz Equilibrium

Suppose for now that, at stage 1 of the game described above, we admit only menus that (i) do not involve cross-subsidization, and (ii) constitute binding contractual offers. Formally, the first requirement restricts the admissible set to menus $\{\mathbf{a}_L, \mathbf{a}_H\} \in \mathbb{R}^4_+$ that satisfy (1)-(2) and

$$\Pi_h(\mathbf{a}_h) \geq 0 \qquad h \in \{H, L\} \tag{3}$$

The second requirement, on the other hand, renders common knowledge that being called upon to act at stage 1 comes with an irreversible commitment to the action chosen at that point. Specifically, no part of an insurance menu may be withdrawn at stage 3, irrespective of the risk-class composition of the pool of customers who chose it at stage 2. This is an exogenous restriction which renders the third stage of Hellwig's game obsolete. It reduces it to a two-stage game in which, at stage 1, the uninformed insurance providers make contractual offers while, at stage 2, the informed customers choose amongst them.[4]

It is easy to see that this version of the game leads to exactly the same equilibrium outcome as the analysis in Rothschild and Stiglitz [28]. First of all, the RSW allocation is the only equilibrium candidate allocation, as the following two results show formally. To see intuitively first why it is not possible to have pooling policies in the equilibrium set, notice that a hypothetical pooling equilibrium policy $\mathbf{a}^*$ ought to just break-even in expectation. To do so, however, it must involve cross-subsidization, expecting losses on the high-risk customers to be matched exactly by expected profits from the low-risk ones.[5] Yet, the very fact that strictly positive profits are extracted by the low-risk type allows for the existence of another contract $\mathbf{a}^1_L$ which delivers strictly positive profits if accepted only by low-risk agents and is such that $\mathbf{a}^1_L \succ_L \mathbf{a}^* \succ_H \mathbf{a}^1_L$. This is shown formally by Proposition 1 and graphically by any point in the interior of the shaded areas in Figures 1-2. In the presence of $\mathbf{a}^*$, therefore, $\mathbf{a}^1_L$ will attract away only the low-risk customers. Clearly, offering it at stage 1 is a strictly-profitable deviation given that the pooling policy is also on offer.

Proposition 1 *Let* $\mathbf{a}$ *be admissible as a pooling policy. Then, it cannot be sustained as an equilibrium allocation in the two-stage game described above.*

[4]In terms of interpreting the third stage of our game by means of Grossman's insight, insurance contracts here can be introduced in the market only via sending out "pre-approved" application forms. Recall that we take the term to mean that it is common knowledge amongst all market participants that any customer who files such an application is guaranteed, in a way that is enforceable whatever her risk-type, delivery of the respective contract.

[5]Given free entry, we ought to have $\Pi_M(\mathbf{a}^*) \geq 0$. Yet, this cannot be a strict inequality. For if $\Pi_M(\mathbf{a}^*) = \epsilon > 0$, we may consider the contract $\hat{\mathbf{a}} = \mathbf{a}^* - (1, -1)\frac{\epsilon}{2}$ which is such that $\hat{\mathbf{a}} \succ_h \mathbf{a}^*$ by either h (it provides strictly more income in either state of the world). In the contingency, therefore, in which $\mathbf{a}^*$ and $\hat{\mathbf{a}}$ are the only policies on offer, the latter contract would attract the entire population of customers and, as a pooling policy itself, would expect profits $\Pi_M(\hat{\mathbf{a}}) = \Pi_M(\mathbf{a}^*) - \frac{\epsilon}{2} > 0$. It constitutes, that is, a profitable deviation, contradicting part (b) of the definition for $\mathbf{a}^*$ to be a market equilibrium. To arrive at the claim in the text, notice that $p_H > p_L$ requires $\Pi_H(\mathbf{a}^*) < \Pi_L(\mathbf{a}^*)$. Clearly, $\Pi_M(\mathbf{a}^*) = 0$ only if $\Pi_H(\mathbf{a}^*) < 0 < \Pi_L(\mathbf{a}^*)$.

Proof. See Appendix C. ■

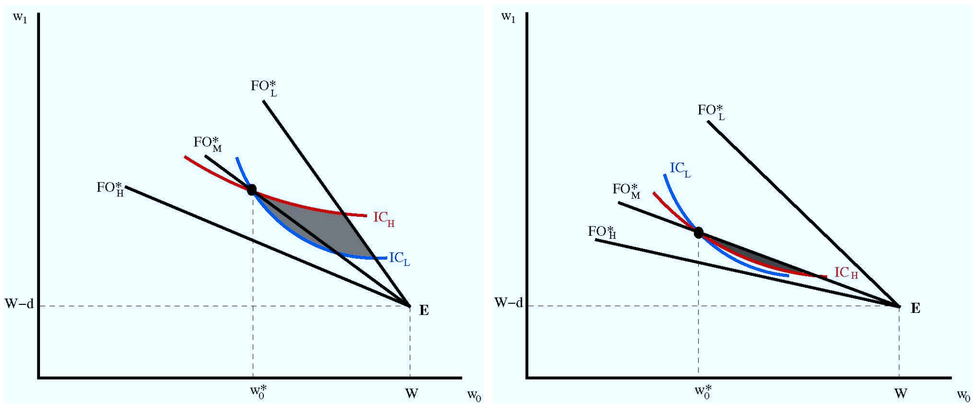

Figure 1: Deviations against pooling policies

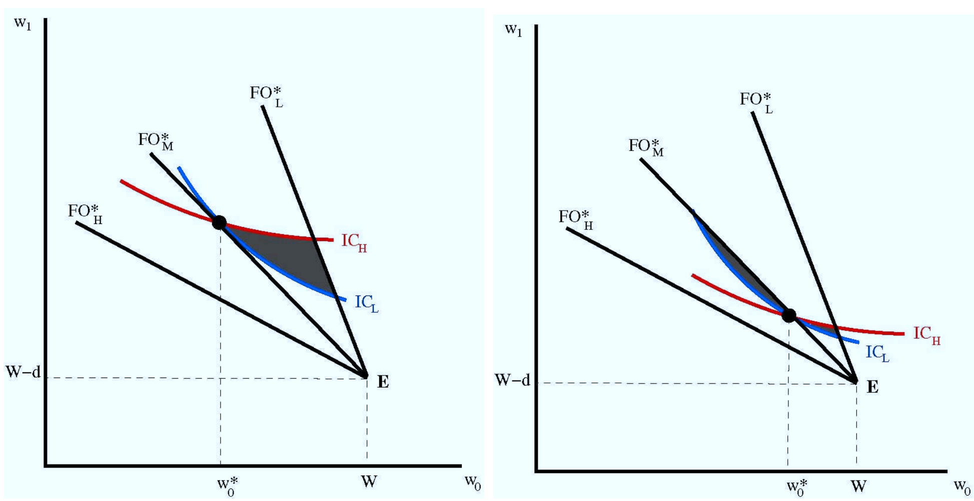

Figure 2: Deviations against pooling policies

An equilibrium policy, therefore, cannot be but a separating menu. Amongst the admissible ones, though, the only legitimate candidate is what will be henceforth referred to as the Rothschild-Stiglitz (RS) menu and denoted by $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$. Its corresponding income allocation $\{\mathbf{w}_L^{**}, \mathbf{w}_H^{**}\}$, the so called Rothschild-Stiglitz-Wilson (RSW) allocation, solves the problem

$$\max_{(\mathbf{w}_L, \mathbf{w}_H) \in \mathbb{R}^4_{++}} U_h(\mathbf{w}_h) \qquad \text{s.t. (1), (2), (3)} \qquad h \in \{L, H\}$$

This is the definition of an RSW allocation relative to zero reservation profits, as it appears in Maskin and Tirole [19]. However, for the economy under study here, it can be identified also as the solution to the more general problem

$$\text{P1}(\mu): \qquad \max_{(\mathbf{a}_L, \mathbf{a}_H) \in \mathcal{A} \times \mathcal{A}} \mu U_L(\mathbf{a}_L) + (1-\mu) U_H(\mathbf{a}_H) \quad \text{s.t. (1), (2), and (3)}$$

where $\mu \in [0,1]$. In what follows the RSW allocation will be denoted by the menu $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$, which is identified in the following result.

Proposition 2 *For any $\mu \in (0,1]$, the solution to P1(μ) is uniquely the menu $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$ given by the following conditions*

(i) $\mathbf{a}_H^{**} = \mathbf{a}_H^{FB}$,

(ii) $\Pi_L(\mathbf{a}_L^{**}) = 0$, *and*

(iii) $U_H(\mathbf{a}_L^{**}) = U_H(\mathbf{a}_H^{FB})$.

This menu solves also the PI(0) problem although so does any incentive compatible menu $\{\mathbf{a}_L, \mathbf{a}_H^{**}\}$ such that $U_L(\mathbf{a}_L) \geq \overline{u}_L$ and $\Pi_L(\mathbf{a}_L) \geq 0$.

Proof. See Appendix B. ■

The RSW allocation plays an important in any competitive market environment with adverse selection because it is the only feasible allocation that maximizes the utility of either risk-type while being also ex-post individually rational. In the terminology of Myerson [22], it is a *safe* allocation, one which would remain feasible even if the firm (the uninformed party) knew the agents' type (the private information of the informed party).

Proposition 3 *Let $\mathbf{a}_L, \mathbf{a}_H$ be separating ($\mathbf{a}_L \succsim_L \mathbf{a}_H \succsim_H \mathbf{a}_L$ with at least one relation strict). If $(\mathbf{a}_L, \mathbf{a}_H) \neq (\mathbf{a}_L^{**}, \mathbf{a}_H^{**})$, then it cannot be sustained as an equilibrium allocation in the two-stage game described above.*

Proof. See Appendix C. ■

To see intuitively why no other separating menu can be an equilibrium, suppose otherwise and let $\{\mathbf{a}_L, \mathbf{a}_H\} \neq \{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$ be one. Observe also that the RSW allocation is unique and maximizes the welfare of the high-risk agents amongst all the separating allocations that are admissible here. It can only be, therefore, $\mathbf{a}_H^{**} \succ_H \mathbf{a}_H$. As shown by Proposition 3, this necessitates the existence of another contract $\mathbf{a}_L^2$ such that $\mathbf{a}_L^2 \succ_L \mathbf{a}_L, \mathbf{a}_H^{**}$ but $\mathbf{a}_H^{**} \succ_H \mathbf{a}_L^2$ and which delivers strictly-positive profits if chosen only by low-risk agents. Consider now a firm offering the menu $\left\{\mathbf{a}_L^2, \mathbf{a}_H^{**}\right\}$. This is separating and attracts either risk-type away from $\{\mathbf{a}_L, \mathbf{a}_H\}$. Doing so, moreover, it breaks-even on the high-risk agents but is strictly profitable on the low-risk ones. Examples of $\mathbf{a}_L^2$ are given by the interior points of the shaded area in either diagram of Figure 3.

Clearly, the RSW menu is the unique equilibrium candidate. Yet, there can be parameter values for which even this is not a viable equilibrium. As Rothchild and Stiglitz pointed out, albeit heuristically, this is bound to happen when there are enough low-risk agents in the population so that the market fair-odds line FO_M^* cuts through the low-risk indifference curve associated with $\mathbf{a}_L^{**}$. Formally, the RSW menu is an equilibrium here if and only if there exists no contract the low-risk type prefers strictly to $\mathbf{a}_L^{**}$ and which delivers zero profits as pooling policy.

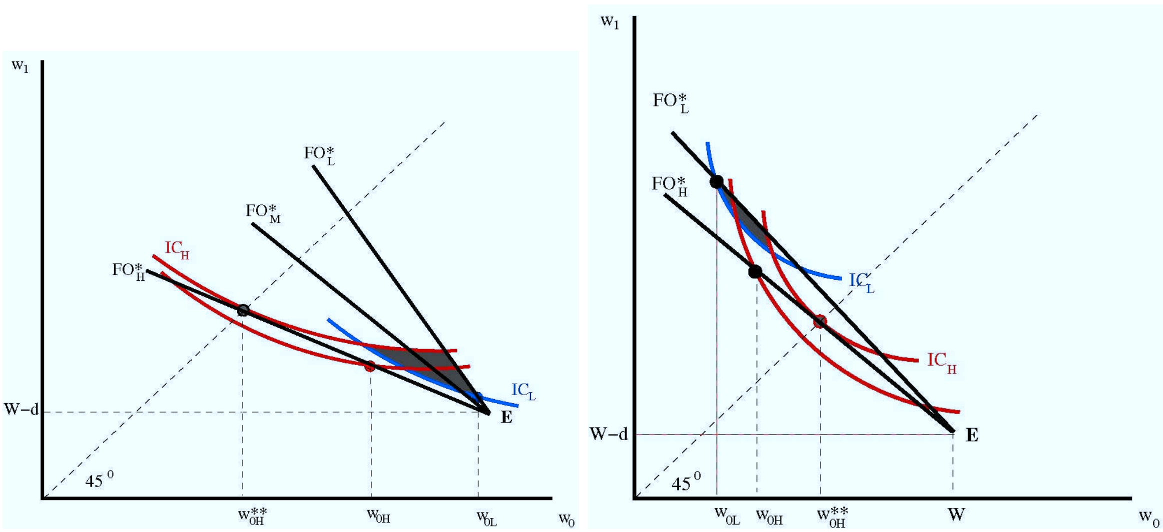

Figure 3: Deviations against a non-RS separating menu

Proposition 4 *Let* $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$ *be the RSW allocation. It is an equilibrium allocation in the two-stage game described above if and only if*

$$\nexists \mathbf{a} \in \mathbb{R}_+^2 : \ \Pi_M(\mathbf{a}) = 0 \ \textit{and} \ \mathbf{a} \succ_L \mathbf{a}_L^{**}$$

Proof. See Appendix C. ■

The contrapositive of the "only if" part of this statement is established by showing that, if there are contracts that expect zero profits as pooling policies and are strictly preferred to $\mathbf{a}_L^{**}$ by the low-risk type, we can construct profitable deviations against the RSW menu. These are contracts $\mathbf{a}^2$ that are strictly profitable as pooling policies and attract at least the low-risk type away ($\mathbf{a}^2 \succ_L \mathbf{a}_L^{**}$). Obviously, if they pull away also the high-risk agents ($\mathbf{a}^2 \succ_H \mathbf{a}_H^{**}$), they are strictly profitable pooling deviations. Otherwise, the high-risk type opts to leave $\mathbf{a}^2$ with only the low-risk agents and, hence, at least as large profits as before (recall the one before the last footnote). Examples of the former case are points in the interior of the shaded area in Figure 4.

2.2 When in doubt, withdraw...

Suppose now that, in the game we just analyzed, the admissibility condition (ii) is altered so as to allow only non-binding contractual offers.[6] This re-enacts the third stage of the original game by rendering two important elements common knowledge. Being called upon to act now at stage 1 caries no commitment to the action chosen at that point. Indeed, a firm will have to decide whether or not to withdraw any part of its current offer at stage 3. It will make that decision, moreover, after having observed the actions of all other firms at stage 1 and after having tried to infer the subsequent choices of all customers at stage 2.

[6] To relate the description once again to Grossman's interpretation of withdrawals, no firm has any "pre-approved" application form at its disposal now. It is common knowledge amongst all market participants that no customer who files an application may be guaranteed delivery of the respective contract.

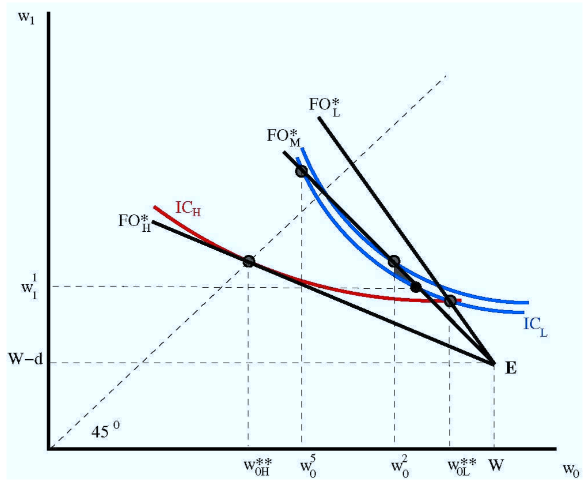

Figure 4: Deviations against the RS menu

To analyze what we will be calling henceforth the standard three-stage game, observe first that the same argument as before precludes any admissible separating menu, but the RS one, from being an equilibrium. The deviant strategy consists now of offering the menu $\left\{\mathbf{a}_L^2, \mathbf{a}_H^{**}\right\}$ at the first stage with the intention to honor it at the third, irrespective of the history of play at that point. Being strictly-separating with $\Pi_L\left(\mathbf{a}_L^2\right) > 0 = \Pi_H\left(\mathbf{a}_H^{**}\right)$, the menu guarantees at least zero profits in any possible contingency. And along the subgame that starts at the beginning of stage 2 and in which this and the original menu are the only ones that have been introduced at stage 1, the former expects strictly-positive profits irrespectively of whether or not the latter is withdrawn at stage 3.

Regarding the admissible pooling policies, reasoning similarly as in the previous version of the game, we conclude that only contracts $\mathbf{a} \in FO_M^*$ are legitimate candidates.[7] And from these, only the ones leaving the low-risk agents at least as well-off as $\mathbf{a}_L^{**}$. This is because, given any $\mathbf{a} \in FO_M^* : \mathbf{a}_L^{**} \succ_L \mathbf{a}$, there are contracts $\mathbf{a}_L^4$ which are strictly-profitable if selected only by low-risk agents and such that $\mathbf{a}_H^{**} \succ_H \mathbf{a}_L^4 \succ_L \mathbf{a}, \mathbf{a}_H^{**}$ (see the shaded area in the left-hand side diagram of Figure 5).[8] Consider then the strategy of offering $\left\{\mathbf{a}_L^4, \mathbf{a}_H^{**}\right\}$ at stage 1 in order to honor it at stage 3, irrespective of the history of play at that point. Being strictly-separating, this menu guarantees at least zero profits in any possible contingency. And in the event in which itself and the

[7]A similar argument to that in footnote 5 applies also here. The deviant strategy now offers $\hat{\mathbf{a}}$ at stage 1 as a pooling policy. Along the subgame that starts at the beginning of stage 2 and in which $\hat{\mathbf{a}}$ and $\mathbf{a}^*$ are the only policies that have been introduced at stage 1, the deviant plan is to honor the former contract at stage 3. In this contingency, the deviant firm expects strictly-positive profits irrespectively of whether or not $\mathbf{a}^*$ gets withdrawn. In any other subgame, the deviant plan is to withdraw $\hat{\mathbf{a}}$ at stage 3 if and only if some contract $\mathbf{a} \in \mathbb{R}_+^2 : \mathbf{a} \succ_L \hat{\mathbf{a}}$ has been introduced at stage 1. Needless to say, in this event, neither $\mathbf{a}^*$ can be honored.

[8]$\mathbf{a}_L^4$ is constructed in the same way as $\mathbf{a}_L$ in Case I in the proof of Proposition 5, once $\mathbf{a}_L^0$ is replaced by $\mathbf{a}_L^{**}$. This gives $\mathbf{a}_L^{**} \succ_h \mathbf{a}_L^4$ for either risk-type. Hence, $\mathbf{a}_H^{**} \succ_H \mathbf{a}_L^4$ given that $\mathbf{a}_H^{**} \sim_H \mathbf{a}_L^{**}$ at the RSW allocation. For the low-risk agents, on the other hand, letting $\Delta = U_L\left(\mathbf{a}_L^{**}\right) - \max\left\{U_L\left(\mathbf{a}_H^{**}\right), U_L\left(\mathbf{a}\right)\right\}$ suffices for $\mathbf{a}_L^4 \succ_L \mathbf{a}_H^{**}, \mathbf{a}$. Finally, since the low-risk agents are under-insured at the RSW allocation, the substitution of $\mathbf{a}_L^0$ with $\mathbf{a}_L^{**}$ works also for the profits, giving $\Pi_L\left(\mathbf{a}_L^4\right) > \Pi_L\left(\mathbf{a}_L^{**}\right) = 0$. Needless to say, there are also here two possible ways to select (κ, ϵ), partitioning the shaded area in the left-hand side diagram of Figure 5 on the basis of whether or not $w_{0L}^4 \geq w_{0L}^{**}$.

pooling policy are the only ones that have been introduced at stage 1, it mounts a strictly-profitable deviation.

In that case, it is strictly-dominant for the low-risk type to select $\mathbf{a}_L^4$ over $\mathbf{a}$. As a result, applications for the pooling policy at stage 2, if there exist any, cannot but come exclusively from high-risk agents. Hence, any insurer offering the pooling policy should expect losses and view its withdrawal as the only sequentially-rational choice at stage 3. Anticipating this, however, and whatever her preference between $\mathbf{a}$ and $\mathbf{a}_H^{**}$, the high-risk type cannot but apply for the latter at stage 2. Nevertheless, the deviant menu is strictly-separating and expects to make profits against the low-risk agents and break even on the high-risk ones.

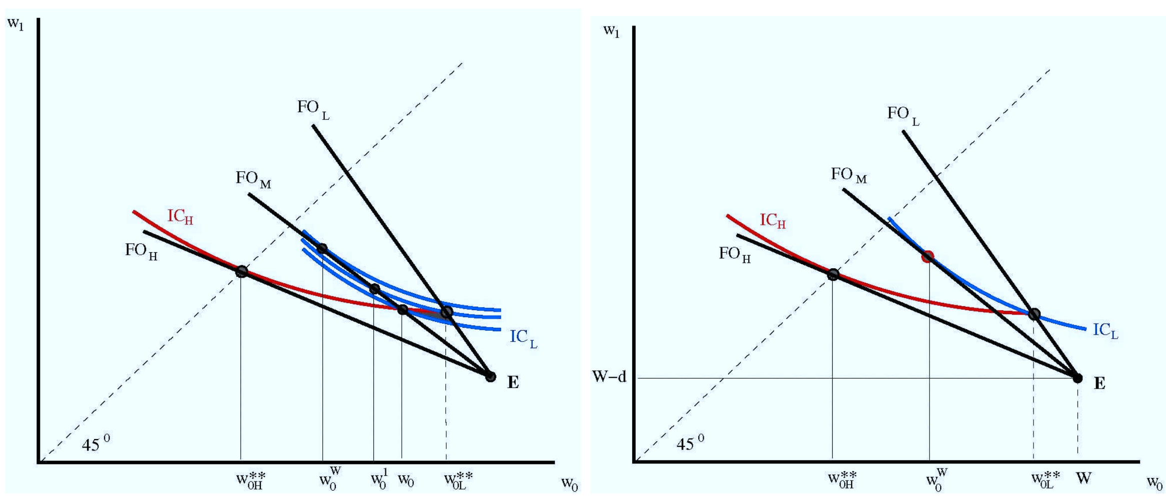

Figure 5: Deviations and Equilibria in Hellwig's game

Evidently, whenever the RS policy is a Nash equilibrium in the previous version of the game, the path of play now evolves essentially in the same way as before. One equilibrium outcome is always that firms offer the RS menu at stage 1, in order to honor it at stage 3, while all agents select the RS contract designed for their type. The only difference is that now an additional equilibrium scenario obtains when there is a unique $\mathbf{a}^1 \in FO_M^* : \mathbf{a}^1 \succsim_L \mathbf{a}_L^{**}$. Of course, by the continuity of the preference $\succsim_L$, it cannot be then but $\mathbf{a}^1 \sim_L \mathbf{a}_L^{**}$ (see the right-hand side diagram of Figure 5). In this case, it is an equilibrium for this contract to be offered as a pooling policy at stage 1 in order to be honored at stage 3. Correspondingly, its suppliers believe that its applicants form a representative sample of the population and indeed every customer is applying for this contract.

As before, the reason why each of these two scenarios can be supported as equilibrium is the very fact that no deviation can mount a credible threat against either unless it pulls away low-risk customers.[9] In both equilibrium outcomes, however, the low-risk type enjoys the welfare she gets

[9] Indeed, no contract can attract only the high-risk type and avoid losses doing so. With respect to the RSW menu, this has been shown in the proof of the "if" part of Proposition 4. Regarding the pooling policy $\mathbf{a}^1$, suppose to the contrary that $\mathbf{a} \in \mathbb{R}_+^2$ gives $\Pi_H(\mathbf{a}) \geq 0$ and $\mathbf{a}^1 \succsim_L \mathbf{a} \succ_H \mathbf{a}^1$. Then if, on the one hand, $\mathbf{a}^1 \succ_H \mathbf{a}_H^{**}$, the separating menu $\{\mathbf{a}^1, \mathbf{a}\}$ Pareto-dominates the RSW allocation (recall that $\mathbf{a}^1 \sim_L \mathbf{a}_L^{**}$) while satisfying the constraints of the efficiency problem the latter solves (observe that $p_L < p_H$ implies that $p_L < \overline{p}$ and, thus, $\Pi_M(\mathbf{a}^1) \leq \Pi_L(\mathbf{a}^1)$, the

under the RSW allocation. It follows then that she cannot be offered a strictly-better outcome by any deviant menu that is designed to be strictly-profitable while separating. Given, moreover, that the Nash equilibrium does exist in the previous version of the game, the only deviant contracts that can attract the low-risk agents and be strictly-profitable as pooling policies are those whose pools of applicants exhibit a fraction of low-risk agents higher than λ.[10]

But against these contracts, withdrawing the equilibrium policy is both credible and sufficient a threat. To see this for the pooling policy $\mathbf{a}^1$, let the deviant firm believe that honoring its policy at stage 3 will be profitable, a consistent belief only if the policy-selection strategies at stage 2 are such that a larger than λ fraction of its customers are low-risk. Given these strategies, however, a smaller than λ fraction of those selecting $\mathbf{a}^1$ are low-risk. As a result, the firms offering the latter policy ought to expect losses and plan to withdraw it at stage 3. Yet, anticipating this at stage 2, none of the customers should apply for $\mathbf{a}^1$. Their only sequentially-rational choice is to select the deviant contract, rendering it pooling with an average quality of applicants exactly equal to the population one. Which contradicts, of course, the deviant suppliers' original belief that honoring their policy at stage 3 will be profitable. Needless to say, the same scenario supports also the RS menu as equilibrium. As this delivers zero profits in every possible contingency, its withdrawal is not necessitated by the deviation but suffices as a credible threat against it.

The equilibrium predictions are different, however, and dramatically so, whenever the previous version of the game has no Nash equilibrium in pure strategies. As we know already, this obtains when there are contracts that expect positive (resp. zero) profits as pooling policies and are weakly (resp. strictly) preferred by the low-risk type to her RS contract. In this case, the RS menu and any pooling policy $\mathbf{a}^* \in FO_M^* : \mathbf{a}^* \succsim_L \mathbf{a}_L^{**}$ can be sustained as pure-strategy sequential equilibria. These are the pooling contracts on the segment between $\mathbf{a}^1$ and $\mathbf{a}^5$ in Figure 4. Recall that none of them were equilibrium in the previous version because, for each one, there existed deviant contracts certain to make strictly-positive profits (in some cases as pooling policies, in others servicing only the low-risk type) *in the presence* of the original policy. Yet, policies can be withdrawn now at stage 3 and this, being common knowledge amongst the players, renders the actual profitability of these deviations dependent upon the players' *beliefs* about their profitability.

Once again, these policies can be supported as equilibria because no deviation can mount a credible threat against them unless it pulls away low-risk customers. Given, however, that the low-risk type is at least as well-off as under the RSW allocation, she cannot be offered a strictly-better outcome by any deviant menu that is designed to be a strictly-profitable separating policy. In this

inequality being strict as long as $\mathbf{a}^1 \neq \mathbf{0}$; recall also that $\Pi_M\left(\mathbf{a}^1\right) = 0$). If, on the other hand, $\mathbf{a}_H^{**} \succsim_H \mathbf{a}^1$, the contradiction becomes that the separating (recall that $\mathbf{a}^1 \sim_L \mathbf{a}_L^{**} \succ_L \mathbf{a}_H^{**}$) menu $\left\{\mathbf{a}^1, \mathbf{a}_H^{**}\right\}$ solves the RSW problem $\forall \mu\,[0,1]$ even though it is not the RSW allocation.

[10]For $\hat{\mathbf{a}} \in \mathbb{R}_{++}^2$, let $\hat{p}^* \in (0,1)$ be given by $\Pi_{\hat{p}^*}(\hat{\mathbf{a}}) = 0$. Since $\frac{\mathrm{d}}{\mathrm{d}p}\left(\frac{1-p}{p}\right) < 0$, we have $(p - \hat{p}^*)\,\Pi_p(\hat{\mathbf{a}}) < 0$ $\forall p \in (0,1) \setminus \{\hat{p}^*\}$ so that $\hat{P} = [0, \hat{p}^*]$ is the set of average accident probabilities across its customers that allow $\hat{\mathbf{a}}$ to avoid losses. Observe now that, the previous version of the game having a Nash equilibrium in pure strategies, it cannot be $\hat{\mathbf{a}} \succ_L \mathbf{a}_L^{**}$ unless $\overline{p} \notin \hat{P}$. As long as $\hat{\mathbf{a}} \succ_L \mathbf{a}_L^{**}$, therefore, the former contract may avoid losses only if $\overline{p} > \hat{p} = \hat{\lambda} p_L + \left(1 - \hat{\lambda}\right) p_H$, where $\hat{\lambda}$ is the share of low-risk customers in its pool of applicants. As $p_L < p_H$, however, the last inequality is equivalent to $\hat{\lambda} > \lambda$.

version of the game, on the equilibrium path, all agents apply for the equilibrium pooling policy or the RS contract designed for their type. Threats of deviations can only arise from pooling contracts but none is introduced at stage 1 because everyone believes that, were it to be introduced, it would be loss-making and, hence, withdrawn at stage 3. Off the equilibrium path, the very fact that a deviant contract may be withdrawn at stage 3 turns itself into a self-fulfilling prophecy if a firm decides to offer it at stage 1. In this contingency, the equilibrium strategy may entail two different prescriptions. One works only against some deviations but does so robustly in a sense to be made precise shortly. The other may be deployed against any deviant contract without, however, the latter property.

The first scenario requires that also the equilibrium policy is withdrawn and may be used against deviations that lie above the line FO_M^* (e.g. the interior points of the lower shaded area in the right-hand side diagram of Figure 2). These contracts are strictly-profitable as pooling policies only if a larger than λ fraction of their customers are low-risk. For this reason, they cannot be honored at stage 3 if the equilibrium policy is withdrawn. Otherwise, they would be selected by either type at stage 2 and become loss-making, their fraction of low-risk customers being then exactly λ (see Section C.2.1 in the Appendix). This equilibrium scenario is described also in Hellwig [14] (pp. 323), his focus being explicitly on fending off deviations that are potentially-profitable against the Wilson pooling policy, the contract that maximizes the welfare of the low-risk type along FO_M^* (depicted on the left-hand side diagram in Figure 2).[11]

Yet, the equilibrium set includes also other pooling contracts on FO_M^* as well as the RS menu, which can be challenged also by deviations on or below FO_M^*. Such deviations are depicted by the interior points of the shaded areas in Figure 4 and in the right-hand side diagram og Figure 1 as well as of the upper shaded area in the diagram on the right-hand side of Figure 2 and of the shaded area on the left-hand side diagram of Figure 1 that lies below the line FO_M^*. These cases call for the second strategic scenario, which is based on an equilibrium-sustaining argument that is equally straightforward although perhaps not as intuitive.

It rests entirely on the fact that the notion of sequential equilibrium puts rather limited constraints on the beliefs players may entertain on information sets off the equilibrium path. As before, any deviant contract, if introduced at stage 1, will be withdrawn at stage 3 because its suppliers expect it to be loss-making given their belief about the average quality of its applicants. Now, however, the equilibrium policy will not be withdrawn in this off-equilibrium event. Threatening to do so is without bite because the deviant contract can be profitable as a stand-alone pooling policy, even if a smaller than λ fraction of its applicants are low-risk. It cannot be profitable, though, if this fraction is too small (in particular, it makes losses on the high-risk type) and the corresponding beliefs of the deviant suppliers are what the equilibrium rests upon (see Section C.2.2 in the Appendix).

[11] This is the contract $\mathbf{a}^W \in FO_M^* : \mathbf{a}^W \succsim_L \mathbf{a}\ \forall \mathbf{a} \in FO_M^*$ and features prominently in Wilson [31]. Notice that no strictly-profitable pooling policy may attract the low-risk agents away from $\mathbf{a}^W$, unless it lies above FO_M^*. This is because, together, $\Pi_M(\hat{\mathbf{a}}) > 0$ and $\hat{\mathbf{a}} \succsim_L \mathbf{a}^W$ imply $\exists \mathbf{a} \in FO_M^* : \mathbf{a} \succ_L \mathbf{a}^W$ (recall the opening observation in Section 4 in the Appendix), an absurdity.

2.3 Equilibrium Selection

Re-introducing the third stage of the game, so that the insurance companies may withdraw their policies if they so wish, leads to a dramatic reversal of the results. Existence of a Nash equilibrium in pure strategies is no longer an issue. If anything, there are multiple equilibria whenever admissible pooling policies Pareto-dominate the RS menu. In fact, the issue now becomes that of equilibrium selection as the Wilson contract strictly Pareto-dominates all other equilibrium allocations.

With this in mind, Hellwig viewed the Wilson policy as the most plausible outcome, being the only equilibrium to survive the stability criterion of Kohlberg and Mertens [16]. Even though Hellwig's claim can be easily substantiated, it was not in his paper; an omission that has misled later scholars, working on applications of this model, into the view that it can be supported by the intuitive criterion of Cho and Kreps [3]. Of course, the criterion is indeed an interpretation of stability and, admittedly, the most straightforward one. In the game under study, however, it lacks the power to single out the Wilson policy.

The Intuitive Criterion

To restrict the out-of-equilibrium beliefs in a way that justifies rejecting a given sequential equilibrium, the intuitive criterion rests upon two integral conditions. The first identifies the sender's types that do not have any incentive to send an out-of-equilibrium message. These types should strictly prefer the equilibrium outcome to anything else they might get out of the receiver's sequentially-rational response to the message, given that the receiver might have any belief with support amongst the types that are allowed to send it. The second condition selects, out of the remaining types of sender, those that do have incentives to send the out-of-equilibrium message. For these types, the equilibrium must be strictly worse than any outcome they may get when the receiver responds to the message optimally, the receiver's belief being again any belief with support amongst the types that are allowed to send it (excluding, of course, the types for whom the first condition applies).

Designed for signaling games, the intuitive criterion may be deployed over the sub-game that begins at stage 2, an out-of-equilibrium sub-path on which, alongside the equilibrium policy, a deviant one has been introduced at stage 1. In this two-stage subgame, the customers move first, selecting the insurance policy they wish to apply for. Inducing the firms' beliefs about the average quality of applications, these choices are signals the informed players send to the uninformed. Following the receipt of these signals, a firm chooses whether or not to honor the policy it has on offer.

In the signalling sub-game, let the two policies introduced at stage 1 be a sequential equilibrium of the overall game and one of its potentially-profitable deviations, a contract that attracts at least the low-risk type away in its presence. We will establish that, applying the intuitive criterion, we may dismiss a non-trivial subset of pooling equilibria as unreasonable. We will also show, however, that this kind of reasoning rejects neither the also non-trivial remaining subset of pooling equilibria nor the RS policy.

With respect to the first claim, let the two policies be a pooling equilibrium, other than the

Wilson one, and a deviant contract which (i) lies below the fair-odds market line, and (ii) is strictly-better (resp. -worse) for the low-risk (resp. high-risk) agents.[12] As we already know, in this case, the fact that the deviation meets condition (i) means that the sequential-equilibrium strategy has the equilibrium contract being honored in the signaling subgame so that the equilibrium allocation corresponds to being insured under the equilibrium contract. By the very choice of deviation, therefore, the high-risk type strictly prefers this outcome over even the best-case scenario that might follow her application to the deviant policy (the event in which the latter is honored at stage 3). Of course, the same claim cannot be made for the low-risk agents.

The high-risk type being the only one with strong incentive to not send the out-of-equilibrium message, the two conditions of the intuitive criterion sort here the two risk-types so that the deviant suppliers should believe that, with probability one, applications for their policy originate from low-risk agents. As a result, they should expect their policy to be strictly profitable and, thus, plan to honor it at stage 3. Applying, therefore, the Cho-Kreps criterion in this example, we are led to regard intuitively unreasonable that the deviant contract gets withdrawn at stage 3 when the equilibrium one is not. In other words, we are led to characterize intuitively unreasonable the very premise upon which our sequential equilibrium argument was based.

Clearly, any of the pooling equilibria that is susceptible to potentially-profitable deviations that satisfy conditions (i)-(ii) is rejected by the intuitive criterion. This is the case for the entire part of the segment between $\mathbf{a}^1$ and $\mathbf{a}^5$ in Figure 4 that lies above $\mathbf{a}^W$. In our quest to single out the latter contract, however, we are still left with the remainder of the segment as well as the RS menu. Each of these may be challenged only by deviations that Pareto-dominate the equilibrium outcome or lie above the line FO_M^*, in which case the sequential equilibrium strategy requires that the equilibrium policy is withdrawn at stage 3.[13] They are all, hence, susceptible only to deviant contracts that, if honored at stage 3, Pareto-dominate the equilibrium outcome. And this renders the intuitive criterion impotent.

When the Pareto-dominance is strict, in the sense that both risk-types are strictly worse off under the equilibrium outcome, the criterion fails because there is no type without strong incentive to send the out-of-equilibrium message. Technically speaking, no type meets the criterion's first condition, which means that also no type meets the second. Intuitively, we may reason as follows. As a pooling policy, the deviant contract avoids losses only if the fraction of its customers that are low-risk does not fall below some cutoff. It will be honored, therefore, as long as its suppliers entertain the corresponding beliefs, which they may well do since no risk-type (in particular, the low-risk) is excluded from applying to their policy.

The deviant policy being honored, though, is an outcome that both risk-types prefer strictly to the given equilibrium. As a consequence, there is no intuitive restriction we may put on the beliefs

[12] Examples are points in the interior of the shaded area below the line FO_M^* in the left-hand side diagram of Figure 1 and of the area that lies between the two indifference curves and below FO_M^* in its right-hand side diagram.

[13] Deviations with the former property are the points in the interior of the shaded area of Figure 4 and of the upper shaded area in the right-hand side diagram of Figure 2. With respect to the latter property, consider the points in the interior of the lower shaded area in the latter diagram, of the shaded area in the left-hand side diagram of the same figure, or of the area delimited by the low-risk indifference curve and the lines FO_M^* and FO_L^* in Figure 4.

of the deviant suppliers regarding the average quality of the applications they receive. Precisely because either risk-type (in particular, the high-risk) has reason to aspire to their policy, we cannot rule out that the fraction of their applicants who are low-risk is in fact below the cutoff. It is also possible, therefore, that they will not honor their policy at the end, leaving whoever chose it at the endowment point; a prospect unpleasant enough to induce both types to stay at the incumbent equilibrium.[14]

This argument needs but a slight modification when the deviant contract Pareto-dominates the equilibrium outcome, albeit not strictly. In the game under study, this obtains when the low-risk agents are indifferent between their equilibrium and the deviant contract. In this case, one type does lack strong incentives to send the out-of-equilibrium message; it (weakly) satisfies the first condition.[15] Being, however, the low-risk type, the intuitive criterion has again no bite. If anything, intuition should have now the deviant suppliers believe that their applications come exclusively from high-risk agents. They should expect, therefore, their policy to make losses and plan to withdraw it at stage 3. Anticipating this, in turn, their applicants should expect to be left at the endowment point. As before, the feasibility of this prospect is enough to justify that both types apply for the equilibrium policy even though a Pareto-dominant one is available.

Divinity

Given the impotence of the intuitive criterion in selecting the Pareto-preferred equilibrium, another obvious recourse is interpreting stability as divinity in the sense of Banks and Sobel [1]. This is a criterion whose real force comes into play precisely when the first intuitive condition fails to identify types with no incentive to send the out-of-equilibrium message. Facing situations, such as the ones described above, where both risk-types wish to defect from the current equilibrium, divinity guides the receiver's beliefs by placing more weight on the type more likely to do so.

Inevitably, this entails utility comparisons across types and, hence, considerable loss of generality for our study.[16] More importantly perhaps, adverse selection becomes most pressing an economic

[14]The worst case scenario for an agent who applies for the deviant policy is to be left at the endowment point. This is an outcome that cannot be strictly preferred to the equilibrium one by either risk-type. Relative to the equilibrium outcome, the endowment is strictly worse for either risk-type whenever the equilibrium strategy prescribes that the equilibrium contract should be honored against the given deviation. It is as good as the equilibrium outcome whenever the equilibrium contract is withdrawn.

[15]Such deviations are depicted by points on the low-risk indifference curve boundary of the shaded area in Figure 4 and of the lower shaded area in the diagram on the right-hand side of Figure 2. In the strict sense of the intuitive criterion as presented in Cho and Kreps [3] (Section IV.3), even in this case, the low-risk type fails to meet the first condition. Their definition depicts either of the criterion's conditions as strict preference. Yet, one could consider relaxing the first to a weak preference (as, in fact, the authors themselves do in Section IV.5). In the game under study, however, even this cannot render the criterion useful. There is no type that satisfies (even weakly) the second condition. Even if there were, actually, it could only be the high-risk and the withdrawal of the deviant policy would again be the only intuitive outcome.

[16]Upon receipt of a message, our receiver has only two pure responses available, withdraw (W) or honor (NW) the policy on offer. We may depict, therefore, her mixed strategy by the probability $r \in [0, 1]$ with which she withdraws her policy at stage 3. Then, in terms of the divinity presentation in Cho and Kreps [3] (Section IV.4), for either

issue exactly if high-risk agents gain more than low-risk ones under a socially-desirable policy change. And in this case, by placing more likelihood on the high-risk type, divinity as well precludes the Pareto-dominant alternative from mounting a successful challenge to the current equilibrium. In fact, in the only case when utility gains here can be unambiguously compared across the two risk-types (the selection problem in the paragraph preceding the last), divinity leads to exactly the same conclusion as the intuitive criterion.[17]

Stability

In our attempt to single out the Wilson policy as the only reasonable equilibrium, we have yet to deal with a non-trivial set of pooling equilibria as well as with the RS menu. These share a common feature: they are all susceptible to deviant contracts that are strictly better, at least for the low-risk type, and lie below the market fair-odds line. And against deviations of this kind, the sequential equilibrium obtains only via the following strategic scenario (Section C.2.2). The deviant contract is deemed loss-making and gets withdrawn at stage 3. By contrast, the equilibrium one is honored because it is believed to be avoiding losses. Anticipating this at stage 2, all agents apply for the equilibrium contract with probability one.

To support this scenario, we need to construct a sequence of vanishing trembles $\left\{r_L^k, r_H^k\right\}_{k\in\mathbb{N}} \in (0,1)$ with the intended interpretation that an agent of risk type h applies to the equilibrium and deviant contracts with probability $\sigma_h^k = 1 - r_h^k$ and $1 - \sigma_h^k$, respectively. This ought to evolve in such a way that, at least along a subsequence, the deviant suppliers believe that the ratio of low- to high-risk amongst their applicants does not exceed the quantity $\hat{\lambda}^* = \frac{1-\hat{p}^*}{\hat{p}^*}$ (the probability as defined in footnote XX). And since strategies ought to be sequentially-consistent, this ratio is given by $\frac{1-\sigma_L^k}{1-\sigma_H^k}$, the relative frequencies with which the two risk-types apply for the deviant contract.

Let us perturb, however, the game by assigning to each risk-type an independent probability of accidentally implementing a fully-randomized strategy, instead of the one she is supposed to play. More precisely, consider a strategy profile $(\widetilde{q}_L, \widetilde{q}_H) \in (0,1)^2$ and a mixture $(\epsilon_L, \epsilon_H) \in (0,1]^2$ to mean that an agent of risk-type h, whose strategy in the original game was to apply for the equilibrium and deviant policies with probability σ_h and $1 - \sigma_h$ respectively, does so in the perturbed game

risk-type h, $D_h^0 = \{r_h\}$ and $D_h = [0, r_h)$, with r_h being the deviant suppliers' mixed response which corresponds to an expected utility for h equal to the utility she derives from her current equilibrium contract. Clearly, to identify the relative sizes of these sets across h, we need to rank the probabilities r_h. The same requirement arises with respect to the the divinity characterization in Banks and Sobel [1] (Section 3). Following an application for the deviant policy, to construct the beliefs of its suppliers that are consistent with no withdrawal - the set $\Gamma(0)$ - we have to compare the schedules $\overline{\mu}(h, r_h) = [0,1]$ and $\overline{\mu}(h, r) = 0$ (1), if $r > r_h$ ($r < r_h$), across h.

[17] In this case, $D_L = \varnothing$ and $D_L^0 = \{NW\} \subset D_H = [0, r_H)$ with $r_H > 0$. According to Criterion D1, therefore, no pooling application may come from low-risk agents. Observe also that, having only two types of sender in this game, the Criteria D1 and D2 coincide. Regarding the exposition in Banks and Sobel [1], we have now $\overline{\mu}(L, 0) = [0,1]$ and $\overline{\mu}(L, r) = 0$ for any $r > 0$, while $\overline{\mu}(H, r_H) = [0,1]$ and $\overline{\mu}(H, r) = 0$ (1) for $r > r_H$ ($r < r_H$). Clearly, $\forall r \in [0,1]$, $\overline{\mu}(L, r) = 1$ implies $\overline{\mu}(H, r) = 1$ while the opposite direction is not true. Observe also that this is the only case in which another refinement, neologism proofness (Farrell [9]), may be deployed here. Given, though, that only the high-risk type strictly prefers the deviant outcome, this concept, as presented by Banks and Sobel (Section 5), also suggests that deviant applications ought to come exclusively from this type.

with probability $\widetilde{\sigma}_h = (1-\epsilon_h)\,\sigma_h + \epsilon_h \widetilde{q}_h$ and $1-\widetilde{\sigma}_h$, respectively. Define then any closed subset of the set of equilibria of the original game to be prestable if $\forall \epsilon_0 \in \mathbb{R}_{++}$ $\exists \epsilon \in \mathbb{R}_{++}$ such that $\forall\,(\widetilde{q}_L, \widetilde{q}_H) \in (0,1)^2$ and $\forall\,(\epsilon_L, \epsilon_H) \in (0,\epsilon)^2$ the perturbed game has at least one equilibrium in the ϵ_0-neighborhood of this subset. Equilibrium stability, in the sense of Kohlberg and Mertens [16], is an identifying feature of the minimal prestable sets.[18]

Given this characterization, the sequential equilibrium scenario under study cannot be stable unless it remains a sequential equilibrium also under perturbations. Yet, as shown in Section C.2.3, we may construct perturbations that are arbitrarily close to the original game but for which no sequence of trembles can meet the condition $\lim_{k\to\infty} \frac{1-\widetilde{\sigma}_L^k}{1-\widetilde{\sigma}_H^k} \leq \hat{\lambda}^*$. Intuitively, the original equilibrium hinges crucially upon forcing the deviant suppliers to be pessimistic enough regarding the average quality of the applications they would have received had they not withdrawn their policy. Perturbing, however, the strategies by which the customers select contracts at stage 2, translates under the sequential-equilibrium reasoning to perturbing the deviant suppliers' belief about the average quality of their applicants. And although the mixtures in these perturbations need to stay close to the original strategic profile, the additional range in the deviant suppliers' beliefs granted by the arbitrary introduction of the randomized profile precludes them from being pessimistic enough. As a result, they find it now optimal to honor their policy at stage 3, a response anything but close to the one they undertook in the original game.[19]

Needless to say, this strategic instability manifests itself also when the scenario under consideration is deployed to support the Wilson policy. Yet, the defining characteristic of this policy is that it is susceptible only to deviations above the market fair-odds line. As a consequence, it can always be supported also by the sequential equilibrium scenario in which both itself as well as the deviant contract are withdrawn at stage 3. Being then indifferent between selecting either policy at stage 2, an agent of risk-type h applies for the Wilson policy with probability $\sigma_h \in [0,1) : 1 \leq \frac{1-\sigma_L}{1-\sigma_H} \leq \hat{\lambda}^*$ (Section C.2.1).

And this is a set of strategic profiles that contains a stable subset. Specifically, as we show in Section C.2.4, every profile in the set $\left\{(\sigma_L, \sigma_H) \in [0,1)^2 : 1 < \frac{1-\sigma_L}{1-\sigma_H} < \frac{1+\hat{\lambda}^*}{2}\right\}$ remains a sequential equilibrium under arbitrary perturbations. Which also means, of course, that each of these perturbations has some equilibrium arbitrarily close to the end points of the set. Its closure, therefore, is prestable. In fact, it is stable because it is minimally prestable. Indeed, with respect to any profile such that $\frac{1+\hat{\lambda}^*}{2} < \frac{1-\sigma_L}{1-\sigma_H} \leq \hat{\lambda}^*$, there are perturbations that are arbitrarily close to the original game

[18] This definition of stability is in the spirit of Section 5.6 in Myerson R.B. *Game Theory: Analysis of Conflict*, Harvard University Press (1997). Of course, analogous perturbations should be considered also regarding the insurers' strategies. In the subgame under study, these players can be indexed by $i \in \{\text{incumbent}, \text{deviant}\}$. Its pure strategies being to withdraw or honor its policy at stage 3, let σ_i be the probability that the ith firm chooses the latter. In the perturbed game, this becomes $\widetilde{\sigma}_i = (1-\epsilon_i)\,\sigma_i + \epsilon_i \widetilde{q}_i$ where the mixture $\epsilon_i \in (0,1]$ and the randomized strategy $\widetilde{q}_i \in (0,1)$ are chosen arbitrarily and are independent from those of other firms and customers.

[19] Recall the preceding footnote. To show instability, we establish that some strategy σ_i, substantially different than the equilibrium σ_i^* in the original game, is optimal for the ith firm against the perturbed selections of customers. Of course, in the perturbed game, the firm's actual strategy is restricted to be $\widetilde{\sigma}_i$. Nevertheless, this can be arbitrarily close to σ_i by appropriate choice of ϵ_i and $\widetilde{q}_i$.

and for which the deviant contract will not be withdrawn.

Endogenous Commitment

As the preceding discussion suggests, to single out the Pareto-optimal amongst the many equilibria of the standard Hellwig game, the notion of stability must be deployed in its pure sense. This is arguably too abstract a refinement, especially when it comes to applications of Hellwig's model. An equally successful but more intuitive one is to restrict the out-of-equilibrium beliefs directly via contractual commitment. By allowing, that is, insurance suppliers to publicly pre-commit, if they so wish, at stage 1 upon honoring their contracts at stage 3. This works as equilibrium selection devise because rendering commitment upon delivering on a contract endogenous is useful here only to deviant suppliers.

Indeed none of the equilibrium contracts identified in Section 2.2 would ever be introduced via "pre-approved" application forms because each faces deviations against which it survives only via the threat of its withdrawal. The Wilson contract aside, however, each is susceptible also to deviant contracts that lie below FO_M^* and offer strict welfare improvements for either risk-type. And against such deviations, the equilibrium scenario is based entirely on offsetting the customers' preference by the belief that the deviant contracts will be withdrawn at stage 3. Yet, this belief is no longer in the support of their reasonable beliefs if the deviations are introduced through "pre-approved" application forms. If their suppliers pre-commit upon honoring them, the deviant contracts will be chosen by all customers at stage 2 and turn into strictly-profitable pooling deviations.

3 Efficient Insurance Provision

Enabling insurance firms not only to withdraw their contracts at stage 3 but also commit at stage 1 upon not withdrawing them singles out the Wilson policy, the only stable equilibrium outcome of the standard Hellwig game. As a notion, therefore, endogenous commitment on the contractual level is powerful in delivering uniqueness of equilibrium in pure strategies. And its power increases even more, along the efficiency dimension, when one allows also insurance suppliers to subsidize net income across contracts.

The latter structural change enlarges the space of admissible insurance menus $\{\mathbf{a}_L, \mathbf{a}_H\}$ into consisting of ones that satisfy (1)-(2) and $\widetilde{\lambda}\Pi_L(\mathbf{a}_L) + \left(1 - \widetilde{\lambda}\right)\Pi_H(\mathbf{a}_H) \geq 0$, where $\widetilde{\lambda}$ is the average ratio of low- to high-risk customers amongst the applicant pool. In equilibrium, of course, the menu must serve both types of customers and the belief of its provider about the average quality of applicants cannot but coincide with the population average ($\widetilde{\lambda} = \lambda$). This, along with the act that contractual commitment is now endogenous, restricts dramatically our predictions regarding the outcome of the Hellwig game. There is now a unique sequential equilibrium which Pareto-dominates even the most desirable pooling equilibrium. In fact, the equilibrium allocation solves the IIE(μ) problem

$$\text{P2}(\mu): \qquad \max_{(\mathbf{w}_L,\mathbf{w}_H)\in\mathbb{R}^4_{++}} \mu U_L(\mathbf{w}_L) + (1-\mu)U_H(\mathbf{w}_H) \qquad \text{s.t. (1)-(2)} \quad \text{and}$$

$$\lambda\Pi_L(\mathbf{a}_L) + (1-\lambda)\Pi_H(\mathbf{a}_H) \quad \geq \quad 0 \tag{4}$$

when the weight is placed entirely upon the welfare of the low-risk type ($\mu = 1$).

These are in fact the IIE allocations relative to zero reservation profits, in the sense of Maskin and Tirole [19]. In the current setting, they can be identified by the following result. Since for a given parameter μ the solution to the problem P2(μ) is unique, it will be called in particular the IIE(μ) allocation. An allocation will be referred to as IIE in general if it is the IIE(μ) allocation for some $\mu \in [0,1]$.

Claim 1 *For any $\mu \in [0,1]$ the IIE(μ) optimum $\{\mathbf{a}^*_L, \mathbf{a}^*_H\}$ is unique. Moreover, it satisfies*

$$\lambda\Pi_L(\mathbf{a}^*_L) + (1-\lambda)\Pi_H(\mathbf{a}^*_H) = 0$$

while being such that

(i) $U_H(\mathbf{w}^*_L) = U_H(\mathbf{w}^*_H) = u(w^*_{0H}) = U_L(\mathbf{w}^*_H) < U_L(\mathbf{w}^*_L)$ *if* $\lambda < \mu$,

(ii) $U_L(\mathbf{w}^*_H) = U_L(\mathbf{w}^*_L) = u(w^*_{0L}) = U_H(\mathbf{w}^*_L) < U_H(\mathbf{w}^*_H)$ *if* $\lambda > \mu$*, and*

(iii) $\mathbf{a}^*_h = \mathbf{a}^*$ *with* $U_h(\mathbf{w}^*) = u(w^*_0)$ *for either* h*, if* $\lambda = \mu$.

Furthermore, for either h, the solution mapping $\mathbf{a}^*_h(\mu) : [0,1] \mapsto \mathbb{R}^2_+$ consists of the following two restrictions

(a) a bijective function $[0,\lambda] \mapsto \mathbb{R}^2_+ \setminus \{\mathbf{a}^{**}_h\}$*, and*

(b) a function $(\lambda,1] \mapsto \mathbb{R}^2_+$ *that is either bijective into* $\mathbb{R}^2_+ \setminus \{\mathbf{a}^{**}_h\}$ *or the constant* $(\lambda,1] \mapsto \{\mathbf{a}^{**}_h\}$.

Proof. See Appendix B. ■

It will be important also to relate the RSW allocation to the concept of interim incentive efficiency. As indicated by the following result, in the current setting, this relation is particularly strong.

Claim 2 *The RSW allocation is IIE only if it is IIE(μ) for all $\mu \in (\lambda, 1]$.*

Proof. See Appendix B. ■

We will establish that the IIE(1) allocation is the unique equilibrium in the augmented Hellwig game in two steps. First, we will show that the outcome in question, referred to henceforth as the IIE(1) allocation, is the only candidate equilibrium allocation. Then, we will argue that the unique menu which delivers this allocation is a sequential equilibrium of the augmented three-stage game. To establish that only the IIE(1) may be an equilibrium allocation, it suffices to show that, as long as the corresponding allocation $\{\mathbf{w}^*_L, \mathbf{w}^*_H\}$ of a candidate equilibrium menu $\{\mathbf{a}^*_L, \mathbf{a}^*_H\}$ does not solve the IIE(1) problem, we can construct a profitable deviation. This is based upon the fact that we can always find another menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ whose allocation $\left\{\mathbf{w}_L^0, \mathbf{w}_H^0\right\}$ is IIE for some $\mu^0 \in (0,1]$ and which constitutes a welfare increase (resp. decrease) for the low-risk (resp. high-risk) agents ($\mathbf{w}_L^0 \succ_L \mathbf{w}_L^*$ but $\mathbf{w}_H^* \succsim_H \mathbf{w}_H^0$).

Suppose then that some firm introduces this menu at stage 1 soliciting applications for $\mathbf{a}_L^0$ via "pre-approved" forms. In this event, being guaranteed a strictly-better outcome, the low-risk customers will leave the equilibrium menu with only high-risk potential applicants on whom it does not expect but losses. It follows that the equilibrium strategy cannot but withdraw the contract $\mathbf{a}_H^*$ at stage 3. Anticipating this at stage 2, the high-risk type cannot but also opt for the deviant menu; albeit, for the contract $\mathbf{a}_H^0$ since the menu is separating (being IIE).

Of course, the deviation just described does not offer a clear incentive to potential challengers of the hypothetical equilibrium. It does attract the low-risk type away but expects zero not strictly positive profits in doing so. Nonetheless, as shown by Propositions 5-6 below, a strictly profitable deviation obtains by replacing one of the contracts in $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ with another in a way that generates a strictly-profitable menu while maintaining the two crucial elements of a successful challenge: the new menu continues to be separating and, relative to the equilibrium one, strictly-better for the low-risk customers. As it turns out, which element of $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ we need to replace depends on whether or not the hypothetical equilibrium menu is itself IIE and, if it is for some $\mu^* \in [0,1)$, on whether or not $\mu^* \geq \lambda$.

Proposition 5 *Let $\{\mathbf{a}_L^*, \mathbf{a}_H^*\}$ be IIE(μ) optimal for some $\mu \in [0,1)$. There exists then another menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ that is*

(i) separating: $\hat{\mathbf{a}}_h \succsim_h \hat{\mathbf{a}}_{h'}$ for $h, h' \in \{L, H\}$,

(ii) strictly profitable: $\lambda \Pi_L(\hat{\mathbf{a}}_L) + (1-\lambda)\Pi_H(\hat{\mathbf{a}}_H) > 0$, and

(iii) such that $\hat{\mathbf{a}}_L \succ_L \mathbf{a}_L^*$ but $\mathbf{a}_H^* \succ_H \hat{\mathbf{a}}_H$.

Proof. See Appendix C. ■

If $\{\mathbf{w}_L^*, \mathbf{w}_H^*\}$ is IIE for some $\mu^* \in [\lambda, 1)$, the menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ can be chosen so that $\left\{\mathbf{w}_L^0, \mathbf{w}_H^0\right\}$ is IIE for some $\mu^0 \in (\mu, 1]$. There exists, moreover, a contract $\hat{\mathbf{a}}_L$ which is strictly preferred to $\mathbf{a}_L^*$ by the low-risk type, expects more profits than $\mathbf{a}_L^0$ when chosen by this type, and sorts strictly the types in conjunction with $\mathbf{a}_H^0$ ($\hat{\mathbf{a}}_L \succ_L \mathbf{a}_H^0 \succ_H \hat{\mathbf{a}}_L$). Examples are points in the interior of the shaded area in the left-hand side diagram of Figure 6. Here, either of the two IIE allocations offer full insurance to the high-risk customers and under-insurance to the low-risk ones. In addition, both menus leave the high-risk type indifferent between the two constituent contracts while both expect to exactly break even if selected by a representative sample of the population of customers. By contrast, the menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\} = \left\{\hat{\mathbf{a}}_L, \mathbf{a}_H^0\right\}$ expects strictly positive profits.

In all other respects, the scenario remains as described before: the deviant strategy is to offer the menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ at stage 1 being pre-committed upon honoring $\hat{\mathbf{a}}_L$ at stage 3. This applies

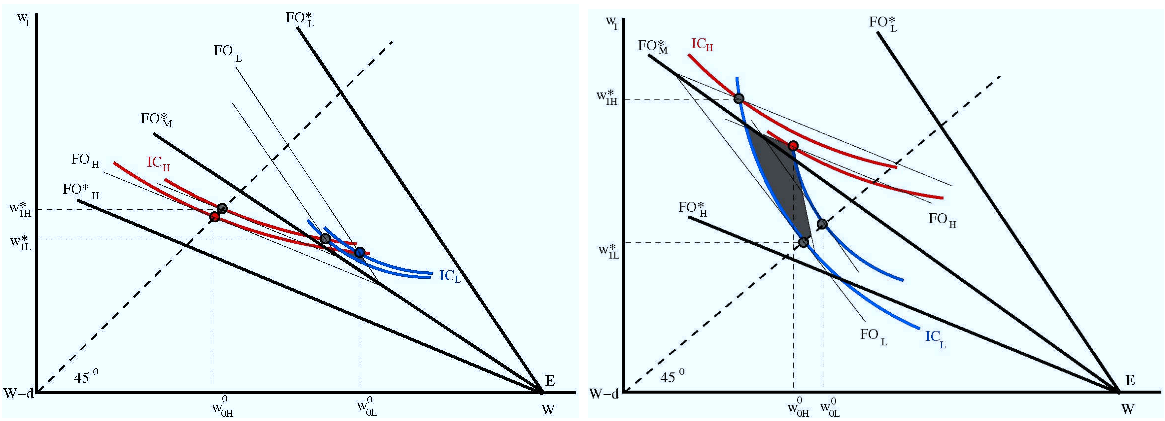

Figure 6: Deviations against IIE policies

also when $\{\mathbf{w}_L^*, \mathbf{w}_H^*\}$ is IIE but for some $\mu^* \in [0, \lambda)$. In this case, $\{\mathbf{a}_L^0, \mathbf{a}_H^0\}$ can be constructed so that its allocation is IIE with $\mu^0 \in (\mu, \lambda)$. There exists then a contract $\hat{\mathbf{a}}_H$ which is strictly worse than $\mathbf{a}_H^*$ for the high-risk customers, expects smaller losses when chosen only by them, and sorts strictly the types in conjunction with $\mathbf{a}_L^0$. This is depicted by points in the interior of the shaded area in the right-hand side diagram of Figure 6. Here, the two IIE allocations offer full insurance to the low-risk customers and over-insurance to the high-risk ones. They both leave also the former risk-type indifferent between their elements while both expect to break even if selected by a representative sample of the population. By contrast, the menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\} = \{\mathbf{a}_L^0, \hat{\mathbf{a}}_H\}$ expects strictly positive profits and the deviant strategy enters it at stage 1 soliciting applications for $\mathbf{a}_L^0$ with "pre-approved" forms.

Proposition 6 *Suppose that the menu $\{\mathbf{a}_L^*, \mathbf{a}_H^*\}$ is feasible but not IIE(μ) optimal for any $\mu \in [0, 1]$. There exists then another menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ that is*

(i) separating: $\hat{\mathbf{a}}_h \succsim_h \hat{\mathbf{a}}_{h'}$ for $h, h' \in \{L, H\}$,

(ii) strictly profitable: $\lambda \Pi_L(\hat{\mathbf{a}}_L) + (1-\lambda)\Pi_H(\hat{\mathbf{a}}_H) > 0$, and

(iii) such that $\hat{\mathbf{a}}_L \succ_L \mathbf{a}_L^*$.

Proof. See Appendix C. ■

Offering a strictly-profitable separating menu under "pre-approved" applications for the low-risk contract is a present and insurmountable challenge also when the hypothetical equilibrium menu does not correspond to an IIE allocation. The only difference is that now the menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ may constitute a welfare improvement even also for the high-risk customers. Situations of this kind are depicted in Figure 7 with respect to the RS menu and the Wilson contract. As we know already, the latter is the only equilibrium pooling candidate under endogenous contractual commitment. Regarding the former, the left-hand side of the figure depicts an example in which the RS menu is not interim incentive efficient even though it is an equilibrium in the Rothschild-Stiglitz setting.

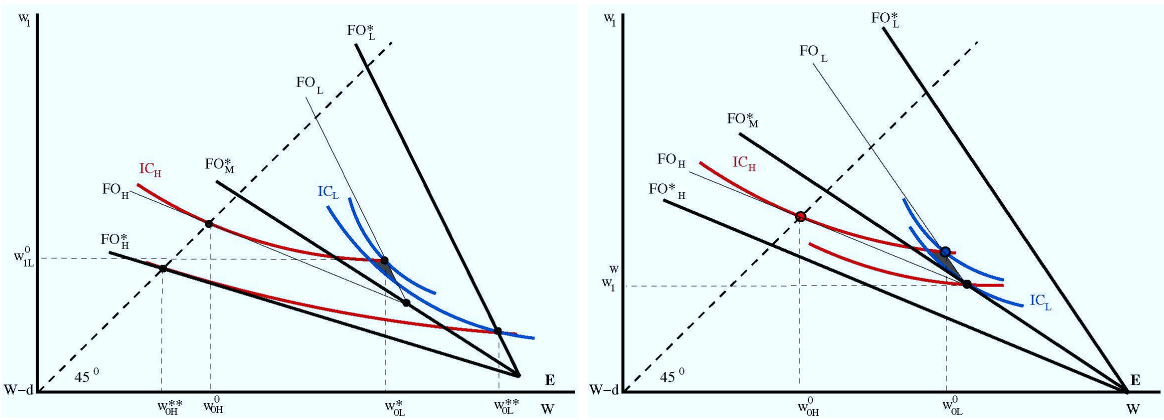

Figure 7: Deviations against non-IIE policies

It remains to show that the unique equilibrium candidate, the menu $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ that corresponds to the IIE(1) allocation, is indeed an equilibrium. This follows from its very characteristic features, being uniquely the IIE(1) optimum.

Proposition 7 *Let the menu* $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ *be the optimum of the IIE(1) problem. Then, it is the equilibrium allocation of the game supported by the following strategy. The firm offers* $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ *at stage 1 with commitment only upon the menu itself, and the plan to withdraw it at stage 3 only if another firm has offered at stage 1 a menu* $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ *such that* $\hat{\mathbf{a}}_L \succ_L \mathbf{a}_L^1$*, with commitment on* $\hat{\mathbf{a}}_L$ *or a plan that includes not withdrawing it at stage 3.*

Proof. Observe first that commitment upon the menu is necessitated by the requirement that the equilibrium strategy profile ought to be subgame perfect. In particular, by the need for subgame perfection along the equilibrium path: along the subgames that begin at stage 2 following the histories in which at least one firm has offered $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ at stage 1 with each of the other firms offering the same menu or being inactive. This is because, in these subgames, the equilibrium strategies of the low and high risk-types must be such that they self-select into $\mathbf{a}_L^1$ and $\mathbf{a}_H^1$, respectively. Yet, as $\Pi_H\left(\mathbf{a}_H^1\right) < 0 < \Pi_L\left(\mathbf{a}_L^1\right)$, it is then optimal for the firm to withdraw $\mathbf{a}_H^1$. And, against the latter strategy, the best response of the high-risk customers is to choose $\mathbf{a}_L^1$. Which renders it, however, a pooling and, as such, loss-making policy.

Regarding the non-equilibrium subgames, recall first that $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ is such that $\mathbf{a}_H^1 \sim_H \mathbf{a}_L^1$ and $\Pi_H\left(\mathbf{a}_H^1\right) < 0 = \lambda\Pi_L\left(\mathbf{a}_L^1\right) + (1-\lambda)\,\Pi_H\left(\mathbf{a}_H^1\right)$. This means that no menu that satisfies (1)-(2) is able to attract away only the high-risk type and avoid losses doing so. For it would have to be some $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ such that $\mathbf{a}_L^1 \succsim_L \hat{\mathbf{a}}_L \succsim_L \hat{\mathbf{a}}_H \succ_H \mathbf{a}_H^1$ and $\Pi_H\left(\hat{\mathbf{a}}_H\right) \geq 0$. Since $\hat{\mathbf{a}}_H \succ_H \mathbf{a}_H^1 \sim_H \mathbf{a}_L^1$, however, this means that the menu $\left\{\mathbf{a}_L^1, \hat{\mathbf{a}}_H\right\}$ is separating. Moreover, $\Pi_H\left(\hat{\mathbf{a}}_H\right) \geq 0 > \Pi_H\left(\mathbf{a}_H^1\right)$ means that $\lambda\Pi_L\left(\mathbf{a}_L^1\right) + (1-\lambda)\,\Pi_H\left(\hat{\mathbf{a}}_H\right) > 0$. In other words, $\left\{\mathbf{a}_L^1, \hat{\mathbf{a}}_H\right\}$ would also solve the IIE(1) problem, an absurdity since its solution is unique.

The potentially profitable deviations, therefore, involve menus that attract away the low-risk type. Recall though that $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ is the unique $\arg\max_{\left(\mathbf{a}_L^1, \mathbf{a}_H^1\right)} U_L\left(\mathbf{a}_L\right)$ s.t. (1), (2), and (4). Moreover, it satisfies the latter constraint as an equality. By uniqueness, therefore, any menu $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ that satisfies (1), (2), and $\hat{\mathbf{a}}_L \succ_L \mathbf{a}_L^1$ must be such that $\lambda\Pi_L(\hat{\mathbf{a}}_L) + (1-\lambda)\Pi_H(\hat{\mathbf{a}}_H) < 0$. Let then $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ be the deviant menu under the plan that does not withdraw $\hat{\mathbf{a}}_L$ at stage 3. If the firm does withdraw $\{\mathbf{a}_L^1, \mathbf{a}_H^1\}$ at stage 3, anticipating the withdrawal, the high-risk customers cannot but opt also for $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ at stage 2. Yet, this turns the deviant menu into one that serves both types, rendering it loss-making. ■

With respect to the formal argument above, notice that the need for commitment upon the menu arises only along the equilibrium path and because the Nash equilibrium needs to be subgame perfect. Without commitment upon the menu, the strategy described in Proposition 7 remains NE without, however, being subgame perfect and, thus, perfect Beysian. With respect to the non-equilibrium subgames, moreover, commitment upon the menu is not necessary at all. For the contract $\mathbf{a}_H^1$ needs to be withdrawn along with $\mathbf{a}_L^1$ only against deviations $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\}$ such that $\mathbf{a}_H^1 \succ_H \hat{\mathbf{a}}_H$. Yet, in such cases, the firm would withdraw it at stage 3 voluntarily. For otherwise the high-risk would select it at stage 2 since $\mathbf{a}_L^1 \sim_H \mathbf{a}_H^1$, but $\Pi_H(\mathbf{a}_L^1) < \Pi_H(\mathbf{a}_H^1) \leq 0$ (Lemma 7).

4 Discussion and Related Literature

Evidently, contractual endogenous commitment - the fact that insurance firms may choose at stage 1 to pre-commit upon honoring a given contract at stage 3 - plays exclusively here the role of restricting the out-of-equilibrium beliefs so that only the IIE(1) menu may be supported as sequential equilibrium. For as we saw, on the equilibrium path, this menu must be introduced at stage 1 without "pre-approved" application forms on either of its two contracts. Nonetheless, some commitment has to be present also in equilibrium; more precisely, on the policy level. As will become apparent by what follows, the IIE(1) menu needs to be introduced at stage 1 as an insurance policy, carrying the binding promise that none of its constituent contracts may be withdrawn unless the menu itself is.

Our equilibrium outcome has been known in the literature (see Crocker and Snow [4]) as the Miyazaki-Wilson allocation. It was established by Miyazaki [21] as the unique equilibrium in a labor market with adverse selection (due to two types of workers in terms of marginal productivity schedules) and firms possessing Wilson foresight. In Wilson [31], it is assumed that each firm correctly anticipates which policies already offered by other firms will become unprofitable as a consequence of any changes in its own offer. It expects then their withdrawal and calculates the profitability of its new offer accordingly. For the insurance provision problem under investigation here, this kind of firm behavior supports always an equilibrium which, with only two risk-types, is almost always unique. Depending on the primitives of the economy, it entails either the RS menu or the Wilson contract - apart from the knife-edge case in which $\mathbf{a}^W \sim_L \mathbf{a}_L^{**}$ and both are valid.

Of course, being able to adjust its current actions according to their effect upon the future choices of its opponents, Wilson's typical firm is not restricted to Nash strategies. And it is the extent of the subsequent complexity in firms' interactions that delivers equilibrium uniqueness.

This becomes evident in Hellwig [14] which could be viewed as an attempt to reconcile anticipatory and Nash-type behavior. The three-stage game permits some anticipation of future reactions but the resulting flexibility in firms' behavior is nowhere near that envisioned by Wilson. Requiring, in addition, sequential rationality and consistent beliefs leads to a rich superset of equilibrium outcomes.

The latter observation is of importance when comparing Miyazaki's result with ours. Both studies regard firms as sophisticated enough to aggregate profits across contracts within douplet menus. And both deliver the IIE(1) allocation as the unique outcome, albeit of a Wilson equilibrium in one but sequential in the other.[20] We obtain it from Hellwig's game when commitment on insurance promises becomes endogenous. This is what keeps our equilibrium set singleton. It also adds a realistic component in the dynamics of insurance provision. Our analysis treats both the practice of requiring applications (which sellers regularly include as part of the insurance transaction) as well as the option to refrain from it (which is also commonly observed) as rational strategies for identifying high-risk buyers and enticing low-risk ones.

As a strategic element, the application process was first studied in Grossman [12]. It induced the high-risk buyers to conceal their identity by mimicking the low-risk choice when offered a separating contractual arrangement. Given that competition imposes zero aggregate profits on menus, the high-risk contract is necessarily loss-making and the firm has a clear incentive to avoid its delivery. Having sorted its customers with a separating menu, it can do so by rejecting applications known to be coming from the high-risk type. It will deliver instead her RS contract, the full-information allocation a high-risk customer can guarantee herself simply by announcing her type. Being able to foresee this, high-risk customers cannot but dissemble their preferences, turning the low-risk contract into a loss-making pooling policy. For the insurance economy under consideration here, we are led back to the Wilson equilibrium even if firms can subsidize net income across contracts.

Now, of course, also the high-risk customers engage in non-Nash strategic behavior, anticipating the effect of their current choices on the sellers' future reactions. This notwithstanding, the strategic dimension of the application process remains at work even when the underlying structure is game-theoretic. Interpreting the rejection of an application as the withdrawal of the respective contractual offer at stage 3, we cannot but conclude that no separating menu can be sustained as Nash equilibrium, unless it gets introduced as a policy. Yet, this is now a result of signalling rather than preference dissembling. The high-risk customers ought to be served on the equilibrium path. If they apply, however, for the high-risk contract at stage 2, the firm can infer their type at stage 3.

[20] There is another difference between the two studies, the underlying economic problem. In a labor market, it is natural to interpret contractual agreements as points in the wage-effort space and take effort as affecting firms' profits through the marginal productivity of labor. This schedule differs across worker-types but it may do so isomorphically-enough for the IIE(1) allocation to be actually first-best. As shown by his example, under certain parameter values of Miyazaki's model, it may be efficient even under full information. This cannot happen in our standard model of an insurance market. Taking the accident probabilities as given exogenously, independent of one's contractual choice, the iso-profits are always linear. More importantly, they have a particular conal shape between the risk-types which, in conjunction with the downward-sloping indifference curves, precludes the IIE(1) allocation from ever solving the full-information efficiency problem.

In the signalling subgame, therefore, its optimal response is to withdraw this contract. And, this being a subgame reached in equilibrium, refusing service must be in the firm's overall strategy.

Grossman presented his insight mainly as a critique against Miyazaki's thesis which identified a given menu of wage-effort contracts with the internal wage structure of a particular firm. It viewed subsequently free exit from the market as sanctioning the withdrawal of entire menus, but not of only an individual contract from a menu. This is admittedly too strong an assumption regarding the market for insurance provision. Here, firms conventionally require customers to apply for particular contracts on a personal basis and they can do so independently of their practice on other elements in their menus. For this reason, the withdrawal of individual insurance contracts ought to be part of an environment with menus. And, as we saw in the preceding paragraph, it ought then to preclude separating contractual arrangements in equilibrium if the underlying structure is Hellwig's three-stage game.

In this case, the set of equilibrium outcomes would still be much richer than the one Grossman imagined, even with two risk-types. In fact, it would coincide with the one in the standard Hellwig model because the analysis of Section 2.2 applies even when firms are allowed to cross-subsidize net income within menus. Specifically, sequential rationality would allow for deviations by separating menus to be neutralized by the perception that the composition of their pools of applicants would be such that they are deemed loss-making and withdrawn. However, in the light of the latter part of the preceding section, this depends crucially upon the firms' being unable to commit on the contractual or policy level.

Under endogenous commitment, a dramatic reversal takes place: the equilibrium cannot but entail a separating contractual arrangement. As follows immediately from our analysis, if firms may choose whether to commit but only on individual contracts, the equilibrium is uniquely the RS menu whenever this is the IIE(1) allocation; otherwise, an equilibrium in pure strategies does not exist. If, in addition, they can introduce menus as policies, the equilibrium is uniquely and always the IIE(1) policy. For it should be clear from the preceding discussion that, on the equilibrium path, the firm ought to condition itself to not withdraw an individual contract from the equilibrium menu unless it withdraws the latter all together. Yet, this is now a matter of strategic choice, not exogenous restriction. It is the firm's optimal response to the equilibrium strategy of the high-risk type. The latter selects the high-risk contract from the IIE(1) menu only if this has been introduced as a policy; otherwise, it opts for the low-risk contract.

Given this epexegesis regarding the strategic underpinnings of our equilibrium outcome, we may turn our attention to its properties and compare it with equilibria in the pertinent literature. In doing so, our principal aim is to provide a convincing account for the central message of the present paper. Namely, under a simple theoretical structure, the augmented version of Hellwig's three-stage game, the forces of market competition should converge upon a single insurance allocation, the most desirable out of those that are efficient under adverse selection. To this end, it is best to first fix ideas about what efficiency ought to mean in the economic environment under investigation.

The standard efficiency concept in economics is the Pareto criterion, mainly due to its obvious appeal when information is complete (no individual has information, about her preferences, endowments, or productive capacity, which is not known by all other individuals). By definition, whenever a given allocation is Pareto-inefficient, there exists another feasible allocation which improves the individuals' welfare unambiguously (in the sense that certainly some individual will be made better off and, equally certainly, no individual will be made worse off). All it takes, therefore, to achieve an unambiguously better economic outcome is for a good (and benevolent) enough outsider to identify and suggest this alternative. And even in the absence of such a welfare economist or social planner, an argument often known as Coase's Theorem suggests that we should still expect to move towards the Pareto-dominant allocation, as long as the costs of bargaining amongst individuals are insignificant. For if bargaining is costless, any of the individuals who will be made better off under the new outcome has a clear incentive to propose the reallocation while no one else has reason to object.

The strength of endorsing Pareto efficiency this way lies in anonymity: to justify a departure from Pareto-inefficient outcomes, there is no need for weighted distributions of gains and losses amongst individuals because no one looses. Its weakness is that it leaves open the question of who is to find the Pareto-improving allocation, an outside planner or members of the economy. It entails, that is, a normative and a positive justification, respectively. Of course, this distinction does not matter under complete information because, without loss of generality, we may assume that the planner knows everything individuals know, which is everything known (indeed, nothing precludes us from anointing any individual as planner).

The distinction is important, though, for economies with incomplete information. In these economies, the individual members have different private information at the time when choices are made. As a result, their decisions and the subsequent outcome depend upon the state of the individuals' information. What matters, therefore, is the decision rule or mechanism, the specification of how decisions are determined as a function of the individuals' information. When the comparison is between mechanisms, however, the normative and positive interpretation of the Pareto criterion may no longer be in agreement. Indeed, the former might admit decision rules the latter would not allow. For it could well be that individuals would unanimously agree to substitute one decision rule with another even though an outside planner could not have identified the new rule as Pareto-improving.

Yet, the role of an outside planner is precisely what an economic theorist assumes when it comes to mechanism design and implementation. To ensure, therefore, that our normative view of Pareto efficiency is not contradicted by that of the individuals in the economy under study, we cannot but disregard a decision rule if it depends upon information individuals hold privately and do not want to reveal. We have to restrict attention, that is, to incentive-compatible decision rules, mechanisms that incentivise each individual to report her private information honestly given that everyone else does the same.

Within the class of incentive-compatible mechanisms, the resulting Pareto-optimal allocations are the ones that achieve incentive efficiency. Albeit stemming from an intuitive requirement, this criterion is subject to when decision rules come up for welfare evaluation because what is optimal for an individual depends crucially on what information she possesses at the time. And, for economies like the market for insurance under consideration here, where at the time she is called upon to act each buyer knows only her private information (her own probability of incurring an income loss), the relevant evaluation stage is the interim one. Indeed, IIE is the appropriate criterion since there cannot be unanimous agreement to depart from an IIE decision rule if some individual knows just her own private information (see Theorem 1 and the subsequent discussion in Holmstrom and Myerson [15]).

In the present setting, a (degenerate) decision rule is nothing but a (douplet) menu of contracts. This describes completely, in each state of the world, how income is allocated between customers and firms per customer-type. As a result, on the one hand, since only one individual (the customers) is informed, condition (1) defines the incentive-compatible allocations. On the other, as competition ensures that firms cannot extract social surplus, the objective function for Pareto-optimality is given by that of the IIE problem, where the weights depend only on the type of the informed party precisely because interim efficiency is the relevant concept.[21] In fact, the two optimality problems of the preceding section define, respectively, the RSW and interim incentive efficient allocations in Maskin and Tirole [19], the reservation allocation being the null trade.[22]

Section 7 of that seminal study considered a three-stage game similar to the one we do, but under a significant generalization of what is meant by contractual arrangement. It assumed that at least two uniformed parties (UP) begin by simultaneously proposing contracts to one informed (IP). A contract, though, is actually a mechanism; it specifies a game form to be played between the two parties, the set of possible actions for each, and an allocation for each pair of their strategies. Following the proposal stage, the IP responds at stage 2. If she accepts a proposal, that game is played out and each party receives the respective outcome at stage 3. Otherwise, each gets its reservation payoff; a contingency that, in a modification, gets replaced by another game in which the two parties alternate in making proposals.

Under this three-stage game, the ensuing set of equilibrium outcomes is very large, even in our simple economy. Since any IIE allocation meets (4) with equality and whatever the value of λ, it does satisfy condition (iv) of Maskin and Tirole's Proposition 7. Any allocation, therefore, is an equilibrium one as long as it satisfies (1), (2), and (4), the latter with equality (see their Proposition 12). It is supported as such by a strategy which prescribes that, following a strictly-profitable deviation by another UP, the outcome of the equilibrium mechanism would be an allocation that

[21] Technically speaking, this claim needs more general decision rules, mapping the type-space $\{H, L\}$ to possibly random allocations (probability distributions on the feasible set A). With respect to such mechanisms, the functions $U_h(\cdot)$ are linear and, consequently, the constraint set is convex as required. For the economy under study, however, the claim does apply even when only deterministic allocations are considered because, as shown by our analysis in the Appendix, the solutions to the IIE problem are of this kind. These were described also in Crocker and Snow [4] (see their Theorem 1). Yet, the present formulation is more general and our analysis more complete as we follow the Khun-Tucker approach. Given that the constraint set is not convex when attention is restricted to deterministic allocations, it is not immediate that the Lagrange conditions are necessary and sufficient for optimality.

[22] To be exact, we refer to their respective counterparts when only deterministic allocations are considered. This is, however, without loss of generality (recall the preceding footnote). We introduce also, with (2), individual-rationality constraints for the customers but this is again without consequence as they do not matter in either problem.

all IP types prefer strictly to that of the deviant.

Yet, this allocation is offered by the equilibrium strategy as a latent threat, not to be delivered necessarily in equilibrium. And this is important in explaining why the Maskin-Tirole result seems so at odds with ours. It indicates fundamental dependence on mechanisms that are much more general that the ones in the present paper, even when attention is restricted to deterministic allocations. Both papers focus on incentive-compatible allocations; hence, on mechanisms in which truthful revelation is a Nash equilibrium (it is in the interest of each informed player to report her type honestly given that everyone else does the same). Nevertheless, seen as mechanisms, all of the games in the two preceding sections stay within the realm of direct revelation. They do not allow equilibrium strategies with latent contracts, to be offered in some off-equilibrium contingency but never implemented in equilibrium.

Of course, this is not the only difference between the two studies. Another emerges in the light of the modified game where mechanism design gets influenced also by the IP. In this case, the equilibrium set shrinks to the outcomes that satisfy the constraints of the IIE problem and (weakly) Pareto-dominate the RSW allocation (see Propositions 13 and 6 in Maskin and Tirole [19]). Equivalently, to the set of equilibrium allocations when the original game entails signalling rather than screening, the IP being now the one to propose mechanisms. And, under this perspective, the distinction between the two papers is drawn even sharper. Our augmented version of the Hellwig game leads to a unique equilibrium allocation with such properties that this game ought to be singled out by the IP under any reasonable theory of mechanism selection. It requires, however, that commitment on insurance provision is endogenous, both on the contractual and the policy level. And, within the Maskin-Tirole approach, the IP cannot exploit this element, even when she is able to stir the process towards a unique equilibrium.

When customers are the ones acting at stage 1, the game form restricts itself to the signalling subgame, the signal being now to suggest a particular contractual arrangement rather than select one already on offer. Adjusting, hence, our analysis in Section 2.2, it is easy to see that any admissible menu with the requisite dominance property, be it separating or trivial, may be supported as sequential equilibrium. It can be guarded against any deviation by the perception that the composition of the pool of customers who suggest the deviant menu renders it loss-making and, thus, precludes any firm from accepting it at stage 2. This logic was deployed above to import Grossman's insight into a version of Hellwig's game that was standard, apart from the fact that firms could subsidize net income across contracts. Under signalling, however, it produces a rich set of separating equilibrium allocations because, standing on the receiving end of insurance proposals, firms are no longer able to sort customers at will.

They can do so only with the consent of the low-risk type and by using the intuitive criterion, a combination powerful enough to admit only one equilibrium outcome. Given any equilibrium menu, separating or trivial, low-risk customers can signal their type by suggesting a contract which makes them (resp. the high-risk) strictly better (resp. worse) off and is strictly profitable when sold only to the low-risk type. More importantly, under the intuitive criterion, their communication is credible since firms interpret it as originating exclusively from this type. The lone survivor is the

RS menu, the only allocation the low-risk type cannot improve upon unilaterally without violating (3) for $h = L$. This can be verified using diagrammatic examples from the preceding section. It is also immediate from Proposition 7 in Maskin and Tirole [19]: their condition (ii) is met since our IP has only two types while the boundary of the feasibility set does not matter.

Hence, when the IP selects mechanisms in the Maskin-Tirole general context, she cannot guarantee herself the IIE(1) allocation apart from a special case (when the RS menu is an IIE and, consequently by Claim 2, the IIE(1) allocation). Yet, the latter is the IP's only reasonable choice when she is called upon to propose mechanisms. This follows from Myerson [22]. In this seminal investigation of mechanism design, an informed party, the principal, plays essentially the same signalling game as above but for the generalization that the other parties, the subordinates, may also be informed. Myerson identified a subset of incentive compatible allocations, the core allocations, and characterized a subset of these, the neutral optima. These are sequential equilibrium outcomes of the game and form the smallest class of allocations satisfying four fundamental axioms of mechanism selection.

Intuitively, if an allocation is not core, there must exist another incentive compatible allocation that would be (i) strictly preferred by some of the principal's types and (ii) implementable given the information revealed by its selection, provided that all the principal's types who prefer the new allocation are expected to propose it. In the paper, the second property is defined as the new allocation being conditionally incentive compatible for the subordinates. Here, however, it can be characterized more simply since the subordinates are uninformed. Suppose that the UP expects the new allocation to be proposed only if the IP's type falls in a particular subset of her type-space. Then the UP should accept it even when he knows that the IP's type lies in this subset.

In the simple insurance economy under investigation here, free entry and exit ensures that firms will acquiesce to a feasible allocation $\{\mathbf{a}_L, \mathbf{a}_H\}$ (separating or trivial) as long as it is incentive compatible and satisfies (4), if it is selected by both risk types, or the relevant condition in (3), otherwise. Within the realm of these restrictions, the IIE(1) allocation is the unique selection of the low-risk type and, by satisfying (4) and $\Pi_L(\mathbf{a}_L^*) > 0$, implementable if $\mathbf{a}_h^* \succ_h \mathbf{a}_h$ for either h or $\mathbf{a}_L^* \succ_h \mathbf{a}_L$ but $\mathbf{a}_H \succsim_H \mathbf{a}_H^*$. In the latter case, moreover, the new proposal comes exclusively from the high-risk type and implementability is ruled out as $\mathbf{a}_H \succ_H \mathbf{a}_H^* \succ_H \mathbf{a}_H^{**}$ necessitates that $\Pi_H(\mathbf{a}_H) < 0$ (see the argument preceding Step 1 of our IIE analysis in the Appendix). Clearly, the IIE(1) is the only core allocation; hence, the unique neutral optimum.

5 Concluding Remarks

In this sense, one may conclude that the present paper leads back to the issue Rothschild and Stiglitz raised originally, albeit under a different perspective. Our result suggests that the lack of efficient outcomes in competitive markets under adverse selection may not be due to the presence of private but rather due to the absence of public information. More precisely, due to the lack of institutions that guarantee the enforcement of two kinds of public commitments by insurance suppliers: to deliver on contracts their customers have applied to via "pre-approved" forms and to abide by insurance promises themselves have marketed as policies.

Our analysis rests upon augmenting contractual admissibility along two dimensions, an accounting and a strategic. The former allows firms to subsidize their net income across contractual offers via the deployment of menus. The latter has firms choosing the extent to which they commit upon their offers. In either of its two forms, the endogeneity of commitment allows the suppliers' promises to play the same strategic role as the public actions do in Myerson [22]. Indeed, be it on the contractual or the policy level, an insurance offer with commitment is a decision which a firm can publicly commit itself to carry out even if it may turn out ex-post to be harmful to itself.

For the insurance economy under consideration here, the interaction between these two dimensions of contactual admissibility renders the IIE(1) allocation the unique sequential equilibrium outcome of a simple direct revelation mechanism that has been used extensively in the literature on applications of contract theory. The game-theoretic structure of Hellwig's model addresses the issue of existence of market equilibrium in pure strategies in a way that is both simple and realistic. Its standard version, however, admits multiple equilibria of which only the RS allocation may be incentive efficient.[23] And, when this is not the case, the selection of the Pareto-optimal equilibrium calls for the exact specification of the equilibrium strategies because it requires the deployment of stability in its technical sense.

As we saw, endogenous contractual commitment can be used as an alternative selection method. Yet, even though intuitively straightforward, this restricts the out-of-equilibrium beliefs of market participants to an extent that precludes the existence of equilibrium in pure strategies if admissibility is augmented also along the accounting dimension. Existence of equilibrium but also uniqueness as well as Pareto-efficiency are restored when the strategic dimension of admissibility allows endogenous commitment also on the policy level. In this sense, our equilibrium outcome demands the simultaneous application of two facets of endogenous commitment in a way that is probably too difficult to establish via real-world market institutions. More realistic settings, such as that in Guerrieri et al. [13], might be viewed as approximating it via the unique equilibrium of another mechanism involving market imperfections. Under this view, the present paper outlines the benchmark mechanism for such approximations.

References

[1] Banks J.S. and J. Sobel (1987): "Equilibrium Selection in Signalling Games," *Econometrica*, 55:647-61.

[2] Bisin A. and P. Gottardi (2006): "Efficient Competitive Equilibria with Adverse Selection," *Journal of Political Economy*, 114:485-516.

[23]This statement requires the following qualification. There are cases in which the pooling equilibria include the allocation that offers full-insurance to both risk-types, the only incentive efficient pooling allocation (it solves the IIE problem for $\mu = \lambda$). This occurs whenever the contract $\mathbf{a}^5$ in Figure 4 lies on or above the 45-degree line. Yet, as we saw in Section 2.3, this equilibrium cannot survive the intuitive criterion.

[3] Cho I.K. and D. Kreps (1987): "Signaling Games and Stable Equilibria," *Quarterly Journal of Economics*, 102:179-221.

[4] Crocker K.J. and A. Snow (1985): "The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information," *Journal of Public Economics*, **26**:207-19.

[5] Dubey P. and J. Geanakoplos (2002): "Competitive Pooling: Rothschild-Stiglitz Reconsidered," *Quarterly Journal of Economics*, 117:1529-70.

[6] Dubey P., J. Geanakoplos, and M. Shubik (2005): "Default and Punishment in General Equilibrium," *Econometrica*, 73:1-37.

[7] de la Fuente A., (2000): "Mathematical Methods and Models for Economists," Cambridge University Press.

[8] Engers M. and L. Fernadez (1987): "Market Equilibrium with Hidden Knowledge and Self-selection," *Econometrica*, 55:425-39.

[9] Farrell J. (1993): "Meaning and Credibility in Cheap-Talk Games," *Games and Economic Behavior*, 5:514-31.

[10] Gale D. (1996): "Equilibria and Pareto Optima of Markets with Adverse Selection," *Economic Theory*, 7:207-35.

[11] Gale D. (1992): "A Walrasian Theory of Markets with Adverse Selection," *The Review of Economic Studies*, 59:229-55.

[12] Grossman H.I. (1979): "Adverse selection, Dissembling, and Competitive Equilibrium," *The Bell Journal of Economics*, 10:336-43.

[13] Guerrieri V., R. Shimer, and R. Wright (2010): "Adverse selection in Competitive Search Equilibrium," *Econometrica*, 78:1823-62.

[14] Hellwig M. (1987): "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection," *European Economic Review*, 31:319-25.

[15] Holmstrom B. and R.B. Myerson (1983): "Efficient and Durable Decision Rules with Incomplete Information," *Econometrica*, 51:1799-819.

[16] Kohlberg E. and J.F. Mertens (1986): "On the Strategic Stability of Equilibria," *Econometrica*, 54:1003-37.

[17] Martin A. (2009): "A Model of Colateral, Investment, and Adverse Selection" *Journal of Economic Theory*, 144:1572-88.

[18] Martin A. (2007): "On Rothschild-Stiglitz as Competitive Pooling," *Economic Theory*, 3:371-86.

[19] Maskin E. and J. Tirole (1992): "The Princpal-Agent Relationship with an Informed Principal, II: Common Values," *Econometrica*, 60:1-42.

[20] Mimra W. and A. Wambach (2011): "A Game-Theoretic Foundation for the Wilson Equilibrium in Competitive Insurance Markets with Adverse Selection," CESIFO Working Paper No. 3412.

[21] Miyazaki H. (1977): "The Rat Race and Internal Labor Markets," *The Bell Journal of Economics*, 8:394-418.

[22] Myerson R.B. (1983): "Mechanism Design by an Informed Principal," *Econometrica*, 51:1767-97.

[23] Netzer N. and F. Scheuer (2013): "A Game Theoretic Approach of Competitive Equilibria with Adverse Selection," *International Economic Review*, forthcoming.

[24] Picard P. (2009): "Participating Insurance Contracts and the Rothschild-Stiglitz Equilibrium Puzzle," preprint.

[25] Prescott E.C. and R.M. Townsend (1984): "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, 52:21-46.

[26] Prescott E.C. and R.M. Townsend (1984): "General Competitive Equilibria Analysis in an Economy with Private Information," *International Economic Review*, 25:1-20.

[27] Riley J.G. (1979): "Informational Equilibrium," *Econometrica*, 47:331-59.

[28] Rothschild M. and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90:629-49.

[29] Rustichini A. and P. Siconolfi (2008): "General Equilibrium in Economies with Adverse Selection," *Economic Theory*, 37:1-29.

[30] von Siemens F. and M. Kosfeld (2009): "Negative Externalities and Equilibrium Existence in Competitive Markets with Adverse Selection," IZA DP No. 4125.

[31] Wilson C. (1977): "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 60:167-207.

Appendices

A Preliminaries

Lemma 1 *Consider an open interval $I \subseteq \mathbb{R}$ and let the functions $F, G : I \mapsto \mathbb{R}$ be differentiable on I, the respective derivative functions being $f, g : I \mapsto \mathbb{R}$. Consider also the function $\Delta : I^4 \mapsto \mathbb{R}$*

given by

$$\Delta(x_1,x_2,x_3,x_4) = F(x_4) - F(x_3) + G(x_2) - G(x_1)$$

If $\{x_i^*\}_{i=1}^4 \in I$ *are such that* $x_1^* < x_2^*$ *and* $(x_3^*, x_4^*) = k(x_1^*, x_2^*)$ *for some* $k \in \mathbb{R}^*$*, then*

$$\Delta(x_1^*, x_2^*, x_3^*, x_4^*) = [kf(k\theta) + g(\theta)](x_2^* - x_1^*)$$

for some $\theta \in (x_1^*, x_2^*)$*.*

Proof. By the fundamental theorem of calculus, we have

$$F(x_4^*) - F(x_3^*) = \int_{x_3^*}^{x_4^*} f(z)\,\mathrm{d}z = \int_{kx_1^*}^{kx_2^*} f(z)\,\mathrm{d}z = k\int_{x_1^*}^{x_2^*} f(kt)\,\mathrm{d}t$$

Therefore,

$$\Delta(x_1^*, x_2^*, x_3^*, x_4^*) = \int_{x_1^*}^{x_2^*} [kf(kt) + g(t)]\,\mathrm{d}t = H(x_2^*) - H(x_1^*)$$

where $H : I \mapsto \mathbb{R}$ is defined by $H(t) = F(kt) + G(t)$. The claim is an immediate consequence of the mean value theorem. ■

Lemma 2 *Let the contracts* $\mathbf{a}^0$ *and* $\mathbf{a}^1$ *be such that* $\mathbf{a}^1 = \mathbf{a}^0 + (\kappa, 1)\epsilon$ *for some* $\kappa, \epsilon \in \mathbb{R}^*$*. Then, the corresponding income allocations* $\mathbf{w}^j = \left(w_0^j, w_1^j\right) = \left(W - a_0^j, W - d + a_1^j\right)$ $(j = 0, 1)$ *are such that*

$$U_h\left(\mathbf{w}^1\right) - U_h\left(\mathbf{w}^0\right) = \begin{cases} \left[p_h u'\left(w_1^0 + \widetilde{\epsilon}\right) - \kappa(1-p_h)u'\left(w_0^0 - \kappa\widetilde{\epsilon}\right)\right]\epsilon & \text{if } \epsilon > 0 \\ \left[p_h u'\left(w_1^0 - \widetilde{\epsilon}\right) - \kappa(1-p_h)u'\left(w_0^0 + \kappa\widetilde{\epsilon}\right)\right]\epsilon & \text{if } \epsilon < 0 \end{cases}$$

for some $\widetilde{\epsilon} \in (0, |\epsilon|)$ *and any* $h = L, H$*.*

Proof. Let first $\epsilon > 0$. Since $\mathbf{w}^1 = \mathbf{w}^0 - (\kappa, -1)\epsilon$, we have

$$\begin{aligned} U_h\left(\mathbf{w}^1\right) - U_h\left(\mathbf{w}^0\right) &= (1-p_h)\left[u\left(w_0^1\right) - u\left(w_0^0\right)\right] + p_h\left[u\left(w_1^1\right) - u\left(w_1^0\right)\right] \\ &= (1-p_h)\left[u\left(w_0^0 - \kappa\epsilon\right) - u\left(w_0^0\right)\right] + p_h\left[u\left(w_1^0 + \epsilon\right) - u\left(w_1^0\right)\right] \end{aligned}$$

The required result follows immediately from applying the preceding lemma with $\mathbf{x}^* = (0, \epsilon)$, $k = -\kappa$, $F(z) = (1-p_h)u\left(w_0^0 + z\right)$, and $\widetilde{F}(z) = p_h u\left(w_1^0 + z\right)$.
When $\epsilon < 0$, on the other hand, we may write $\mathbf{w}^1 = \mathbf{w}^0 + (\kappa, -1)|\epsilon|$. In this case,

$$\begin{aligned} U_h\left(\mathbf{w}^1\right) - U_h\left(\mathbf{w}^0\right) &= (1-p_h)\left[u\left(w_0^0 + \kappa|\epsilon|\right) - u\left(w_0^0\right)\right] + p_h\left[u\left(w_1^0 - |\epsilon|\right) - u\left(w_1^0\right)\right] \\ &= -(1-p_h)\left[u\left(w_0^0\right) - u\left(w_0^0 + \kappa|\epsilon|\right)\right] - p_h\left[u\left(w_1^0\right) - u\left(w_1^0 - |\epsilon|\right)\right] \end{aligned}$$

and we may apply Lemma 1 as before but for $\mathbf{x}^* = (-|\epsilon|, 0)$. ■

Lemma 3 *Let the contracts* $\mathbf{a}^0$ *and* $\mathbf{a}^1$ *be such that* $\mathbf{a}^1 = \mathbf{a}^0 + (\kappa, 1)\epsilon$ *for some* $\kappa, \epsilon \in \mathbb{R}^*$*. Suppose also that one of the following conditions holds* *(i)* $\kappa \in \left(0, \frac{p_h}{1-p_h}\right]$, $\epsilon > 0$, *and* $w_0^1 \geq w_1^1$.

(ii) $\kappa \in \left[\frac{p_h}{1-p_h}, \infty\right)$, $\epsilon < 0$, $w_0^1 \leq w_1^1$.

where $\left(w_0^j, w_1^j\right) = \left(W - a_0^j, W - d + a_1^j\right)$ *denotes the corresponding income allocation for* $j = 0, 1$. *Then,* $\mathbf{a}^1 \succ_h \mathbf{a}^0$ *for* $h = H, L$.

Proof. Consider first condition (i). Since $\kappa \leq \frac{p_h}{1-p_h}$, by the preceding lemma, we ought to have $U_h\left(\mathbf{w}^1\right) - U_h\left(\mathbf{w}^0\right) \geq p_h\left[u'\left(w_1^0 + \widetilde{\epsilon}\right) - u'\left(w_0^0 - \kappa\widetilde{\epsilon}\right)\right]\epsilon$ for some $\widetilde{\epsilon} \in (0, \epsilon)$. Observe, moreover, that $w_1^0 + \widetilde{\epsilon} < w_1^0 + \epsilon = w_1^1 \leq w_0^1 = w_0^0 - \kappa\epsilon < w_0^0 - \kappa\widetilde{\epsilon}$. The claim now follows immediately due to risk-aversion $(u''(\cdot) < 0)$ and $\epsilon > 0$.

Under condition (ii) the argument is equally straightforward. In this case, $U_h\left(\mathbf{w}^1\right) - U_h\left(\mathbf{w}^0\right) \geq p_h\left[u'\left(w_1^0 - \widetilde{\epsilon}\right) - u'\left(w_0^0 + \kappa\widetilde{\epsilon}\right)\right]\epsilon$ for some $\widetilde{\epsilon} \in (0, |\epsilon|)$. Yet now, $w_1^0 - \widetilde{\epsilon} > w_1^0 - |\epsilon| = w_1^1 \geq w_0^1 = w_0^0 + \kappa|\epsilon| > w_0^0 + \kappa\widetilde{\epsilon}$ and $\epsilon < 0$. ■

For the next two results, we denote by

$$I_h(\mathbf{a}) = \frac{\mathrm{d}a_1}{\mathrm{d}a_0} = -\frac{\frac{\partial U_h(W-a_0)}{\partial w_0}\frac{\mathrm{d}w_0}{\mathrm{d}a_0}}{\frac{\partial U_h(W-d+a_1)}{\partial w_1}\frac{\mathrm{d}w_1}{\mathrm{d}a_1}} = \left(\frac{1-p_h}{p_h}\right)\frac{u_h'(W-a_0)}{u_h'(W-d+a_1)} > 0 \qquad h = H, L$$

the slope of the indifference curve of risk-type h at an arbitrary contract point $\mathbf{a} \in \mathbb{R}_+^2$ in the (a_0, a_1)-space. Of course, we have $I_H(\mathbf{a}) < I_L(\mathbf{a})$.

Lemma 4 *Taking* $h, h' \in \{L, H\}$ *with* $h \neq h'$. *For any* $\kappa \in (I_H(\mathbf{a}_h), I_L(\mathbf{a}_h))$ *there exists* $\mathbf{a}_h^0 = \mathbf{a}_h + (1, \kappa)\epsilon$ *with* $\epsilon < 0$ *if* $h = L$ *(resp.* $\epsilon > 0$ *if* $h = H$*) such that* $\mathbf{a}_h^0 \succ_h \mathbf{a}_h \succ_{h'} \mathbf{a}_h^0$.

Proof. Suppose first that $h = L$ and consider the contract $\mathbf{a}_L^0 = \mathbf{a}_L + (1, \kappa)\epsilon$ for some $\kappa \in (I_H(\mathbf{a}_L), I_L(\mathbf{a}_L))$ and $\epsilon < 0$. Define also $\Delta_{h''} = |\kappa - I_{''}(\mathbf{a}_L)|$ for the arbitrary $h'' \in \{L, H\}$. By Lemma 2, we get

$$\begin{aligned}
U_h\left(\mathbf{w}_L^0\right) - U_h(\mathbf{w}_L) &= \left[p_h u_h'(w_{1L} - \widetilde{\epsilon}) - \kappa^{-1}(1-p_h)u_h'\left(w_{0L} + \kappa^{-1}\widetilde{\epsilon}\right)\right]\kappa\epsilon \\
&= \left[\kappa - \frac{(1-p_h)u_h'\left(w_{0L} + \kappa^{-1}\widetilde{\epsilon}\right)}{p_h u_h'(w_{1L} - \widetilde{\epsilon})}\right]p_h u_h'(w_{1L} - \widetilde{\epsilon})\epsilon \\
&= \left[\kappa - I_h\left(\mathbf{a}_L^* - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right)\right]p_h u_h'(w_{1L} - \widetilde{\epsilon})\epsilon
\end{aligned}$$

for some $\widetilde{\epsilon} \in (0, \kappa|\epsilon|)$. Yet, the function $I_h(\cdot)$ is continuous and $\lim_{\widetilde{\epsilon}\to 0} I_h\left(\mathbf{a}_L - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right) = I_h(\mathbf{a}_L)$. For small enough $|\epsilon|$ (and, subsequently, $\widetilde{\epsilon}$), therefore, $\left|I_h\left(\mathbf{a}_L - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right) - I_h(\mathbf{a}_L)\right| < \min\{\Delta_L, \Delta_H\}$. But then,

$$\begin{aligned}
U_L\left(\mathbf{w}_L^0\right) - U_L(\mathbf{w}_L) &= \left[I_L(\mathbf{a}_L) - \Delta_L - I_L\left(\mathbf{a}_L - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right)\right]p_L u_L'(w_{1L} - \widetilde{\epsilon})\epsilon \\
&= -\left[I_L\left(\mathbf{a}_L - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right) - I_L(\mathbf{a}_L) + \Delta_L\right]p_L u_L'(w_{1L} - \widetilde{\epsilon})\epsilon > 0 \\
U_H\left(\mathbf{w}_L^0\right) - U_H(\mathbf{w}_L) &= \left[I_H(\mathbf{a}_L) - I_H\left(\mathbf{a}_L - \left(\kappa^{-1}, 1\right)\widetilde{\epsilon}\right) + \Delta_H\right]p_H u_H'(w_{1L} - \widetilde{\epsilon})\epsilon < 0
\end{aligned}$$

as required. Clearly, if $\{\mathbf{a}_L, \mathbf{a}_H\}$ is separating, we have $U_L\left(\mathbf{w}_L^0\right) > U_L(\mathbf{w}_L) \geq U_L(\mathbf{w}_H)$ and $U_H\left(\mathbf{w}_L^0\right) < U_H(\mathbf{w}_L) \leq U_H(\mathbf{w}_H)$. When $h = H$, the argument is identical but for $\epsilon > 0$. ■

Lemma 5 *Let $h, h' \in \{L, H\}$ with $h \neq h'$. Suppose also that $\mathbf{a}^0$ and $\mathbf{a}^*$ are such that $\mathbf{a}^0 \succ_{h'} \mathbf{a}^*$. Then, there exists $\mathbf{a} = \mathbf{a}^0 + (1, \kappa)\,\epsilon$, with either $\epsilon > 0$ and $\kappa \in \left(-\infty, \min_{h\in\{L,H\}} I_h\left(\mathbf{a}^0\right)\right)$ or $\epsilon < 0$ and $\kappa \in \left(\max_{h\in\{L,H\}} I_h\left(\mathbf{a}^0\right), \infty\right)$, such that $\mathbf{a}^0 \succ_{h'} \mathbf{a} \succ_{h'} \mathbf{a}^*$ and $\mathbf{a}^0 \succ_h \mathbf{a}$.*

Proof. Let $\Delta = U_{h'}\left(\mathbf{w}^0\right) - U_{h'}\left(\mathbf{w}^*\right) > 0$ and consider the contract $\mathbf{a} = \mathbf{a}^0 + (1, \kappa)\,\epsilon$ first for some $\kappa \in \left(-\infty, \min_{h\in\{L,H\}} I_h\left(\mathbf{a}^0\right)\right)$ and $\epsilon > 0$. Let also $\Delta_h = I_h\left(\mathbf{a}^0\right) - \kappa$ for $h \in \{L, H\}$. By Lemma 2 we have

$$
\begin{aligned}
U_h\left(\mathbf{w}\right) - U_h\left(\mathbf{w}^0\right) &= \left[p_h u\left(w_1^0 + \hat{\epsilon}_h\right) - \kappa^{-1}\left(1 - p_h\right) u'\left(w_0^0 - \kappa^{-1}\hat{\epsilon}_h\right)\right]\kappa\epsilon \\
&= \left[\kappa - \frac{\left(1 - p_h\right) u'\left(w_0^0 - \kappa^{-1}\hat{\epsilon}_h\right)}{p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)}\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon \\
&= \left[\kappa - I_h\left(\mathbf{a}^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon \\
&= \left[I_h\left(\mathbf{a}^0\right) - \Delta_h - I_h\left(\mathbf{a}^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon
\end{aligned}
$$

for some $\hat{\epsilon}_h \in (0, \kappa\epsilon)$. Yet, the function $I_h(\cdot)$ is continuous and $\lim_{\hat{\epsilon}_h \to 0} I_h\left(\mathbf{a}^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) = I_h\left(\mathbf{a}^0\right)$. For small enough ϵ (and, subsequently, $\hat{\epsilon}_h$), therefore, $|I_h\left(\mathbf{a}^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) - I_h\left(\mathbf{a}^0\right)| < \min\{\Delta_L, \Delta_H\}$ and, consequently, $U_h\left(\mathbf{w}\right) < U_h\left(\mathbf{w}^0\right)$ for either h. Observe, however, that the last quantity above vanishes as $\epsilon \to 0$. Consequently, with respect to the risk-type h' and for sufficiently small ϵ, we can guarantee that $U_{h'}\left(\mathbf{w}\right) - U_{h'}\left(\mathbf{w}^0\right) \in (-\Delta, 0)$ or $U_{h'}\left(\mathbf{w}\right) > U_{h'}\left(\mathbf{w}^*\right)$.
Take now $\mathbf{a}$ as before but for some $\kappa \in \left(\max_{h\in\{L,H\}} I_h\left(\mathbf{a}^0\right), \infty\right)$ and $\epsilon < 0$. Let also $\Delta_h = \kappa - I_h\left(\mathbf{a}^0\right)$ for $h \in \{L, H\}$. In this case, Lemma 2 gives

$$
\begin{aligned}
U_h\left(\mathbf{w}\right) - U_h\left(\mathbf{w}^0\right) &= \left[p_h u'\left(w_1^0 - \hat{\epsilon}_h\right) - \kappa^{-1}\left(1 - p_h\right) u'\left(w_0^0 + \kappa^{-1}\hat{\epsilon}_h\right)\right]\kappa\epsilon \\
&= \left[\kappa - \frac{\left(1 - p_h\right) u'\left(w_0^0 - \kappa^{-1}\hat{\epsilon}_h\right)}{p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)}\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon \\
&= \left[\kappa - I_h\left(\mathbf{a}^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon \\
&= \left[I_h\left(\mathbf{a}^0\right) + \Delta_h - I_h\left(\mathbf{a}^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_1^0 + \hat{\epsilon}_h\right)\epsilon < 0
\end{aligned}
$$

for some $\hat{\epsilon}_h \in (0, \kappa|\epsilon|)$ and sufficiently small $|\epsilon|$ to ensure that $|I_h\left(\mathbf{a}^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) - I_h\left(\mathbf{a}^0\right)| < \min\{\Delta_L, \Delta_H\}$. Regarding the risk-type h', though, a small enough $|\epsilon|$ guarantees also that $U_{h'}\left(\mathbf{w}^0\right) - U_{h'}\left(\mathbf{w}\right) \in (-\Delta, 0)$ and, thus, $U_{h'}\left(\mathbf{w}\right) > U_{h'}\left(\mathbf{w}^*\right)$ as required. ■

With respect to the following two results, it should be pointed out that, in an optimization problem such as the IIE, $\mathcal{D}$ and $\mathcal{M}$ being independent is meant to say that none of the parameters enters the constraints.

Claim 3 *For $n, m \in \mathbb{N}^*$, let $\mathcal{D} \subseteq \mathbb{R}^n$ and $\mathcal{M} \subseteq \mathbb{R}^m$ be, respectively, independent sets of choice and parameter vectors such that the problem $\max_{\mathbf{x}\in\mathcal{D}} f\left(\mathbf{x}, \mu\right)$ is well-defined for the function $f : \mathcal{D} \times \mathcal{M} \mapsto \mathbb{R}$. Let also $V : \mathcal{M} \mapsto \mathbb{R}$ be the associated value function $V\left(\mu\right) = \max_{\mathbf{x}\in\mathcal{D}} f\left(\mathbf{x}, \mu\right)$ and suppose that $\mathcal{M}$ is convex. Then V is [resp. strictly] convex if $f\left(\mathbf{x}, \cdot\right)$ is [resp. strictly] convex.*

Proof. In what follows, given an arbitrary $\mu \in \mathcal{M}$, we let $\mathbf{x}\left(\mu\right)$ denote an optimum for the optimization problem. Given also arbitrary $k \in [0, 1]$ and $\mu', \mu'' \in \mathcal{M}$, we adopt the notation

$\mu^k = k\mu' + (1-k)\,\mu''$.

To establish the result, take any $\mu', \mu'' \in M$. Since also $\mu^k \in \mathcal{M}$, for any $k \in [0,1]$, we ought to have

$$\begin{aligned} V\left(\mu^k\right) = f\left(\mathbf{x}\left(\mu^k\right), \mu^k\right) &\leq kf\left(\mathbf{x}\left(\mu^k\right), \mu'\right) + (1-k)\, f\left(\mathbf{x}\left(\mu^k\right), \mu''\right) \\ &\leq kf\left(\mathbf{x}\left(\mu'\right), \mu'\right) + (1-k)\, f\left(\mathbf{x}\left(\mu''\right), \mu''\right) \\ &= kV\left(\mu'\right) + (1-k)\, V\left(\mu''\right) \end{aligned}$$

Here, the second inequality results from the optimality of $\mathbf{x}(\mu')$ and $\mathbf{x}(\mu'')$ while the first is due to $f\left(\mathbf{x}\left(\mu^k\right), \cdot\right)$ being convex. The latter inequality is moreover strict when $f\left(\mathbf{x}\left(\mu^k\right), \cdot\right)$ is strictly convex. ■

Lemma 6 *Let $\mathcal{D} \subseteq \mathbb{R}^n$ be a choice set such that the problem $\max_{\mathbf{x}\in\mathcal{D}} \mu f(\mathbf{x}) + (1-\mu)\, g(\mathbf{x})$ is well-defined for the functions $f, g : \mathcal{D} \mapsto \mathbb{R}$ and the parameter $\mu \in \mathcal{M} \equiv [0,1]$. Suppose also that the sets $\mathcal{D}$ and $\mathcal{M}$ are independent and let $S : \mathcal{M} \mapsto \mathcal{D}$ be the solution mapping. For any function $\mathbf{x}^* : \mathcal{M} \mapsto S(\mu)$, we have*

$$f\left(\mathbf{x}^*\left(\mu_1\right)\right) - f\left(\mathbf{x}^*\left(\mu_2\right)\right) \geq 0 \geq g\left(\mathbf{x}^*\left(\mu_1\right)\right) - g\left(\mathbf{x}^*\left(\mu_2\right)\right) \quad \forall \mu_1, \mu_2 \in [0,1] : \mu_1 > \mu_2$$

with both inequalities strict if $S(\mu) \cap S(\mu') = \varnothing$.

Proof. Since the choice set remains the same irrespectively of the value of the parameter, taking $\mathbf{x}^*(\mu_1)$ and $\mathbf{x}^*(\mu_2)$ to be optimal in the respective problems means that

$$\mu_1\left[f\left(\mathbf{x}^*\left(\mu_1\right)\right) - f\left(\mathbf{x}^*\left(\mu_2\right)\right)\right] + (1-\mu_1)\left[g\left(\mathbf{x}^*\left(\mu_1\right)\right) - g\left(\mathbf{x}^*\left(\mu_2\right)\right)\right] \geq 0 \tag{5}$$

$$\mu_2\left[f\left(\mathbf{x}^*\left(\mu_2\right)\right) - f\left(\mathbf{x}^*\left(\mu_1\right)\right)\right] + (1-\mu_2)\left[g\left(\mathbf{x}^*\left(\mu_2\right)\right) - g\left(\mathbf{x}^*\left(\mu_1\right)\right)\right] \geq 0 \tag{6}$$

That is,

$$(\mu_1 - \mu_2)\left[f\left(\mathbf{x}^*\left(\mu_1\right)\right) - f\left(\mathbf{x}^*\left(\mu_2\right)\right) - \left(g\left(\mathbf{x}^*\left(\mu_1\right)\right) - g\left(\mathbf{x}^*\left(\mu_2\right)\right)\right)\right] \geq 0 \tag{7}$$

and $\mu_1 > \mu_2$ requires that

$$f\left(\mathbf{x}^*\left(\mu_1\right)\right) - f\left(\mathbf{x}^*\left(\mu_2\right)\right) \geq g\left(\mathbf{x}^*\left(\mu_1\right)\right) - g\left(\mathbf{x}^*\left(\mu_2\right)\right) \tag{8}$$

Since the result holds trivially if both sides of (8) are zero, we need to examine the following two cases.

(i) Exactly one side of (8) is zero. Letting $g(\mathbf{x}^*(\mu_1)) = g(\mathbf{x}^*(\mu_2))$, requires that $f(\mathbf{x}^*(\mu_1)) > f(\mathbf{x}^*(\mu_2))$; otherwise, (5) is violated given that $\mu_1 > 0$. Similarly, if $f(\mathbf{x}^*(\mu_1)) = f(\mathbf{x}^*(\mu_2))$, it must be $g(\mathbf{x}^*(\mu_1)) \leq g(\mathbf{x}^*(\mu_2))$; otherwise, (6) is violated given that $\mu_2 < 1$.

(ii) Neither side of (8) is zero. If $\mu_1, \mu_2 \in (0,1)$, the quantities on either side of (8) cannot be of the same sign; otherwise, one of (5)-(6) will be violated. Moreover, $\mu_1 = 1$ requires that $f(\mathbf{x}^*(\mu_1)) > f(\mathbf{x}^*(\mu_2))$ in (5) and then it must be $(1-\mu_2)\left[g(\mathbf{x}^*(\mu_1)) - g(\mathbf{x}^*(\mu_2))\right] < 0$ from (6). Furthermore, if $\mu_2 = 0$, we ought to have $g(\mathbf{x}^*(\mu_1)) < g(\mathbf{x}^*(\mu_2))$ in (6) and, hence, $\mu_1\left[f(\mathbf{x}^*(\mu_1)) - f(\mathbf{x}^*(\mu_2))\right] >$

0 by (5).

Now, if $S(\mu) \cap S(\mu') = \varnothing$, then $\mathbf{x}^*(\mu_1)$ [resp. $\mathbf{x}^*(\mu_2)$] is not optimal when the parameter is μ_2 [resp. μ_1]. In this case, all four inequalities (5)-(8) are strict and the result follows from two observations. Neither side of $f(\mathbf{x}^*(\mu_1)) - f(\mathbf{x}^*(\mu_2)) > g(\mathbf{x}^*(\mu_1)) - g(\mathbf{x}^*(\mu_2))$ can be zero; otherwise, one of the strict versions of (5)-(6) will be violated. For the same reason, moreover, the two sides cannot be of the same sign. Needless to say, the larger [resp. smaller] one must be strictly positive [resp. strictly negative]. ■

B Efficiency

Proof of Proposition 2

Consider PI(μ) for some $\mu \in [0,1]$ and let β_h, γ_h, and δ_h be, respectively, the Lagrangean multipliers on the constraints (1)-(3) for risk-type h. The Kuhn-Tucker first-order conditions for an optimum are given by

$$
\begin{aligned}
(\mu + \beta_L^{**} + \gamma_L^{**}) \frac{\partial U_L(\mathbf{w}_L^{**})}{\partial a_{0L}} &= \beta_H^{**} \frac{\partial U_H(\mathbf{w}_L^{**})}{\partial a_{0L}} - \delta_L^{**}(1 - p_L) && (9) \\
(\mu + \beta_L^{**} + \gamma_L^{**}) \frac{\partial U_L(\mathbf{w}_L^{**})}{\partial a_{1L}} &= \beta_H^{**} \frac{\partial U_H(\mathbf{w}_L^{**})}{\partial a_{1L}} + \delta_L^{**} p_L && (10) \\
(1 - \mu + \beta_H^{**} + \gamma_H^{**}) \frac{\partial U_H(\mathbf{w}_H^{**})}{\partial a_{0H}} &= \beta_L^{**} \frac{\partial U_L(\mathbf{w}_H^{**})}{\partial a_{0H}} - \delta_H^{**}(1 - p_H) && (11) \\
(1 - \mu + \beta_H^{**} + \gamma_H^{**}) \frac{\partial U_H(\mathbf{w}_H^{**})}{\partial a_{0H}} &= \beta_L^{**} \frac{\partial U_L(\mathbf{w}_H^{**})}{\partial a_{0H}} + \delta_H^{**} p_H && (12) \\
\beta_h^{**}(U_h(\mathbf{w}_h^{**}) - U_h(\mathbf{w}_{h'}^{**})) &= 0 \qquad h, h' \in \{H, L\} && (13) \\
\gamma_h^{**}(U_h(\mathbf{w}_h^{**}) - \overline{u}_h) &= 0 \qquad h \in \{H, L\} && (14) \\
\delta_h^{**} \Pi_h(\mathbf{a}_h^{**}) &= 0 \qquad h \in \{H, L\} && (15) \\
\beta_h^{**}, \gamma_h^{**}, \delta_h &\geq 0 \qquad h \in \{H, L\} && (16)
\end{aligned}
$$

along with (1), (2), and (3). Here, for either h, we have

$$
\begin{aligned}
\frac{\partial U_h(\mathbf{w}_h)}{\partial a_{0h}} &= (1 - p_h) u'(w_{0h}) \frac{\mathrm{d}w_{0h}}{\mathrm{d}a_{0h}} = -(1 - p_h) u'(w_{0h}) && (17) \\
\frac{\partial U_h(\mathbf{w}_h)}{\partial a_{1h}} &= p_h u'(w_{1h}) \frac{\mathrm{d}w_{1h}}{\mathrm{d}a_{1h}} = p_h u'(w_{1h}) && (18)
\end{aligned}
$$

and it is trivial to check that (9)-(10) give

$$
\begin{aligned}
& (\mu + \beta_L^{**} + \gamma_L^{**}) p_L (1 - p_L) \left[u'(w_{1L}^{**}) - u'(w_{0L}^{**})\right] && (19) \\
= \; & \beta_H^{**} \left[p_H (1 - p_L) u'(w_{1L}^{**}) - p_L (1 - p_H) u'(w_{0L}^{**})\right]
\end{aligned}
$$

while (11)-(12) imply that

$$
\begin{aligned}
& (1 - \mu + \beta_H^{**} + \gamma_H^{**}) p_H (1 - p_H) \left[u'(w_{1H}^{**}) - u'(w_{0H}^{**})\right] && (20) \\
= \; & \beta_L^{**} \left[p_L (1 - p_H) u'(w_{1H}^{**}) - p_H (1 - p_L) u'(w_{0H}^{**})\right]
\end{aligned}
$$

Our analysis will proceed through a series of observations regarding the characteristics of an RSW(μ)-optimal allocation.

Step 1

(i). If the low-risk profit constraint binds at the optimum, the low-risk type cannot be fully insured. We will establish the contrapositive statement, arguing ad absurdum. Let, thus, $w_{1L}^{**} = w_{0L}^{**}$. Then, $U_H(\mathbf{w}_L^{**}) = u(w_{0L}^{**})$ and the high-risk incentive constraint would read $U_H(\mathbf{w}_H^{**}) \geq u(w_{0L}^{**})$. Which cannot be, however, if $\Pi_L(\mathbf{a}_L^{**}) = 0$ because

$$\begin{aligned} U_H(\mathbf{w}_H^{**}) = (1-p_H)\,u(w_{0H}^{**}) + p_H u(w_{1H}^{**}) &\leq u((1-p_H)\,w_{0H}^{**} + p_H w_{1H}^{**}) \\ &= u(W - (1-p_H)\,a_{0H}^{**} + p_H(a_{1H}^{**} - d)) \\ &\leq u(W - p_H d) \\ &< u(W - p_L d) \\ &= u(W - (1-p_L)\,a_{0L}^{**} + p_L(a_{1L}^{**} - d)) \\ &= u((1-p_L)\,w_{0L}^{**} + p_L w_{1L}^{**}) = u(w_{0L}^{**}) \end{aligned}$$

The first inequality here is due to $u(\cdot)$ being everywhere strictly-concave (the binding case obtaining only if $w_{0H}^{**} = w_{1H}^{**}$). The third inequality follows from $p_L < p_H$ and $u(\cdot)$ being strictly-increasing. The second and fourth equalities use that $\mathbf{w}_h^{**} = (W - a_{0h}^{**}, W - d + a_{1h}^{**})$. The second inequality follows from $u'(\cdot) > 0$ since $-[(1-p_H)\,a_{0H}^{**} - p_H a_{1H}^{**}] = -\Pi_H(\mathbf{a}_H^{**}) \leq 0$, by the respective condition in (3). The latter condition, which binds by assumption for the low-risk, is responsible also for the second equality.

(ii). The high-risk profit constraint must bind $\forall \mu \in [0,1]$. The low-risk one must do so if $\mu > 0$. Suppose first that $\Pi_L(\mathbf{a}_L^{**}) > 0$. Then $\delta_L^{**} = 0$ by the corresponding condition in (15), and (9)-(10) read

$$\begin{aligned} (\mu + \beta_L^{**} + \gamma_L^{**})\,(1-p_L)\,u'(w_{0L}^{**}) &= \beta_H^{**}\,(1-p_H)\,u'(w_{0L}^{**}) \\ (\mu + \beta_L^{**} + \gamma_L^{**})\,p_L u'(w_{1L}^{**}) &= \beta_H^{**} p_H u'(w_{1L}^{**}) \end{aligned}$$

Recall, however, that $u'(\cdot) > 0$ and $p_h \in (0,1)$ for either h. It follows that, along with the non-negativity conditions in (16), $\mu > 0$ requires that neither side in either equation is zero. But then one equation may be divided by the other to give $\frac{1-p_L}{p_L} = \frac{1-p_H}{p_H}$, a contradiction. This establishes that $\Pi_L(\mathbf{a}_L^{**}) > 0$ only if $\mu = 0$.

And a trivially similar argument, using the respective condition in (15) and (11)-(12), shows that $\Pi_H(\mathbf{a}_H^{**}) > 0$ only if $\mu = 1$. If $\mu = 1$, however, then $\Pi_L(\mathbf{a}_L^{**}) = 0$ from above and the low-risk cannot be fully-insured (Step 1(i)). Which requires, in turn, that $\beta_H^{**} > 0$. For if $\beta_H^{**} = 0$, since $\mu > 0$, $\beta_L^{**}, \gamma_L^{**} \geq 0$, and $p_L < 1$, (19) gives $u'(w_{1L}^{**}) = u'(w_{0L}^{**})$; equivalently, $w_{1L}^{**} = w_{0L}^{**}$ (under the strict concavity of $u(\cdot)$), a contradiction. Even when $\mu = 1$, therefore, it must be $\beta_H^{**} > 0$ and the argument in the preceding paragraph applies again for the respective condition in (15) and (11)-(12) to rule out that $\Pi_H(\mathbf{a}_H^{**}) > 0$.

Step 2.

We have shown that both profit constraints in (3) cannot but bind at the optimum, as long as

$\mu \in (0,1]$. When $\mu = 0$, the profit constraint on the low-risk type does not have to bind at the optimum but we may take it to be so without significant loss of generality. In this case, we restrict attention to the allocation $\{\mathbf{w}_L^{**}, \mathbf{w}_H^{**}\}$ we derive below, replacing (3) with its binding version:

$$\Pi_h(\mathbf{a}_h) = 0 \qquad h \in \{H, L\} \tag{21}$$

Of course, this is but one of many optima when $\mu = 0$. For then, any allocation $\{\mathbf{w}_L, \mathbf{w}_H^{**}\}$ is optimal, as long as it satisfies the constraints. Yet, $\{\mathbf{w}_L^{**}, \mathbf{w}_H^{**}\}$ is the only optimum that remains so $\forall \mu \in [0,1]$.

Step 3.

(i). The low-risk type does not get full insurance.

This is simply a re-statement of Step 1(i) under the corresponding condition in (21).

(ii). A pooling allocation cannot be optimal.

For a pooling contract $\mathbf{a} = (a_0, a_1)$ to satisfy the two zero-profit conditions in (21), we must have $a_0 = \left(\frac{p_h}{1-ph}\right) a_1$ for either h. Which, since $p_H > p_L$, means actually that the only admissible pooling contract is the trivial one ($\mathbf{a} = \mathbf{0}$). Yet, as shown by what follows, we can do much better for either type than leaving them at the endowment point.

(iii). At least one incentive-compatibility constraint does not bind at the optimum.

Otherwise, $U_L(\mathbf{w}_L^{**}) = U_L(\mathbf{w}_H^{**})$ and $U_H(\mathbf{w}_H^{**}) = U_H(\mathbf{w}_L^{**})$ together imply that

$$\begin{aligned} p_L\left[u\left(w_{0L}^{**}\right) - u\left(w_{0H}^{**}\right) + u\left(w_{1H}^{**}\right) - u\left(w_{1L}^{**}\right)\right] &= u\left(w_{0L}^{**}\right) - u\left(w_{0H}^{**}\right) \\ &= p_H\left[u\left(w_{0L}^{**}\right) - u\left(w_{0H}^{**}\right) + u\left(w_{1H}^{**}\right) - u\left(w_{1L}^{**}\right)\right] \end{aligned}$$

Yet, given $\alpha, \zeta \in \mathbb{R}$ and $\gamma, \delta \in \mathbb{R}^{**}$ with $\gamma \neq \delta$, $\gamma(\alpha + \zeta) = \delta(\alpha + \zeta) = \alpha$ implies $\alpha = \zeta = 0$. It follows, therefore, that the above relations cannot be unless $u(w_{0L}^{**}) = u(w_{0H}^{**})$ and $u(w_{1L}^{**}) = u(w_{1H}^{**})$. Equivalently, unless $w_{0L}^{**} = w_{0H}^{**}$ and $w_{1L}^{**} = w_{1H}^{**}$, an absurd conclusion in light of the preceding step.

(iv). The incentive compatibility constraint of the high-risk type binds if $\mu > 0$.

Suppose the opposite. Then, by the corresponding condition in (13), $\beta_H^{**} = 0$. Which, cannot be, however, if $\mu > 0$. For then (19) requires that the low-risk type gets fully-insured ($w_{1L}^{**} = w_{0L}^{**}$) which is absurd given that her profit constraint binds (Steps 1(i)-(ii)).

(v). The high-risk agents are fully insured.

Let first $\mu > 0$. Steps 3(iii)-(iv) together dictate that $\beta_L^{**} = 0$ in (13). Moreover, the argument that established Step 3(iii) showed also that $\beta_H^{**} > 0$ in this case. And as $1 - \mu, \gamma_H^{**} \geq 0$ while $0 < p_H < 1$, (20) necessitates that $u'(w_{0H}^{**}) = u'(w_{1H}^{**})$. Equivalently, $w_{0H}^{**} = w_{1H}^{**}$ as required.

Let now $\mu = 0$. If $\beta_H^{**} > 0$, it must be $\beta_L^{**} = 0$ (Step 3(iv)) and the result follows again from (20). The same argument applies also for $\beta_H^{**} = 0$ as long as $\beta_L^{**} = 0$. While when $\beta_L^{**} > 0$, $\gamma_L^{**} \geq 0$, $\beta_H^{**} = 0$, and $0 < p_L < 1$ force (19) to require that the low-risk type gets full insurance, which is absurd (Step 3(i)).

(vi). The high-risk individual-rationality constraint is slack ($\gamma_H^{**} = 0$).

Recall the preceding observation and Step 1(i). Since the high-risk profit constraint binds and this type is fully insured, we have

$$\begin{aligned}
\overline{u}_H &= (1-p_H)\,u(W) + p_H u(W-d) \\
&< u((1-p_H)W + p_H(W-d)) \\
&= u(W - p_H d) = u\left(W - (1-p_H)\,a_{0H}^{**} + p_H\left(a_{1H}^{**} - d\right)\right) \\
&= u\left((1-p_H)\,w_{0H}^{**} + p_H w_{1H}^{**}\right) = u\left(w_{0H}^{**}\right) = U_H\left(\mathbf{w}_H^{**}\right)
\end{aligned}$$

(vii). The contract for the high-risk customers is what they would get under prefect information. Under perfect information, each risk-type is offered the full-insurance contract that meets the respective condition in (21). Geometrically, this the intersection point of the 45-degree and the fair-odds line through the endowment point, FO_h^*. Clearly, $\mathbf{w}_H^{**} = \mathbf{w}_H^F$ by Step (v) and the respective condition in (21).

(viii). The low-risk customers are underinsured ($w_{0L}^{**} > w_{1L}^{**}$).
We have just established that the optimal contract for the high-risk type involves full insurance. Hence, $U_H\left(\mathbf{w}_H^{**}\right) = u\left(w_{0H}^{**}\right) = U_L\left(\mathbf{w}_H^{**}\right)$ and the two incentive compatibility constraints can be put together as $U_H\left(\mathbf{w}_L^{**}\right) \le u\left(w_{0H}^{**}\right) \le U_L\left(\mathbf{w}_L^{**}\right)$. The inequality between the first and the last quantities requires that $(p_H - p_L)\left[u\left(w_{1L}^{**}\right) - u\left(w_{0L}^{**}\right)\right] \le 0$. Given that $p_H > p_L$ and $u(\cdot)$ is strictly-increasing, this necessitates that $w_{1L}^{**} \le w_{0L}^{**}$. In fact, $w_{1L}^{**} < w_{0L}^{**}$ since equality has been ruled out (Step 3(i)).

(ix). The low-risk individual-rationality constraint is slack ($\gamma_L^{**} = 0$).
This is immediate once the preceding observation is combined with Lemma 3(i), applied for $h = L$ with $\widetilde{\mathbf{w}}$ and $\mathbf{w}$ being, respectively, the endowment point and $\mathbf{w}_L^{**}$. Recall that either of the latter two points is on the line FO_L^*. Thus, $\mathbf{a}_L^{**} = \mathbf{0} + \left(a_{0L}^{**}, a_{1L}^{**}\right) = \mathbf{0} + \left(\frac{p_L}{1-p_L}, 1\right) a_{1L}^{**}$ with $a_{1L}^{**} > 0$.

(x). $\mathbf{a}_L^{**}$ is given by the intersection of FO_L^* and the locus curve $U_H(\mathbf{w}) = u\left(w_{0H}^F\right)$.
The contract offered to the low-risk type is given by the binding incentive-compatibility constraint of the high-risk type, $U_H\left(\mathbf{w}_L^{**}\right) = u\left(w_{0H}^F\right)$, in conjunction with the corresponding equation in (21).

Step 4.

It remains to verify that the rank condition of the Khun-Tucker theorem (see, for instance, Th. 1.18, Ch. 7 in de la Fuente [7]) is satisfied. To this end, notice that the only inequality constraint that does not bind at the suggested optimum is the low-risk incentive compatibility one. The rank condition, therefore, requires here that the matrix

$$\begin{bmatrix}
\nabla_{(\mathbf{w}_L,\mathbf{w}_H)} U_L(\mathbf{w}_L) - U_L(\mathbf{w}_H) \\
\nabla_{(\mathbf{w}_L,\mathbf{w}_H)} U_H(\mathbf{w}_H) - U_H(\mathbf{w}_L) \\
\nabla_{(\mathbf{w}_L,\mathbf{w}_H)} \Pi_L(\mathbf{a}_L) \\
\nabla_{(\mathbf{w}_L,\mathbf{w}_H)} \Pi_H(\mathbf{a}_H)
\end{bmatrix}
=
\begin{bmatrix}
\frac{\partial U_L(\mathbf{w}_L)}{\partial w_{0L}} & \frac{\partial U_L(\mathbf{w}_L)}{\partial w_{1L}} & -\frac{\partial U_L(\mathbf{w}_H)}{\partial w_{0H}} & -\frac{\partial U_L(\mathbf{w}_H)}{\partial w_{1H}} \\
-\frac{\partial U_H(\mathbf{w}_L)}{\partial w_{0L}} & -\frac{\partial U_H(\mathbf{w}_L)}{\partial w_{1L}} & \frac{\partial U_H(\mathbf{w}_H)}{\partial w_{0H}} & \frac{\partial U_H(\mathbf{w}_H)}{\partial w_{1H}} \\
-(1-p_L) & -p_L & 0 & 0 \\
0 & 0 & -(1-p_H) & -p_H
\end{bmatrix}$$

has rank at least 3 at $\left(\mathbf{w}_L^{**}, \mathbf{w}_H^{**}\right)$. But this is obvious. In fact, this matrix has rank at least 3 everywhere, for its last three columns are linearly independent. Indeed, no linear combination of them can be zero at the third entry given that $p_L > 0$.

To complete the proof, we should point that the uniqueness of $\left(\mathbf{w}_L^{**}, \mathbf{w}_H^{**}\right)$ as the optimum is obvious (subject of course to the qualification in Step 2) while its sufficiency follows from the fact that the RSW(μ) problem has a (strictly) concave objective and a convex feasible set (recall footnote XX in the main text and see, for instance, Theorems 1.6, 1.14, and 1.19, Ch. 7 in de la Fuente [7]).

Step 5.

The claim now follows by the corresponding condition in (21) and Step 3(x) [resp. 3(v)] for $h = L$ [resp. $h = H$]. ■

Proposition 8 *The IIE optimum* $\{\mathbf{w}_h^* = (W - a_{0h}^*, W - d + a_{1h}^*) : h = L, H\}$ *is given by letting*

$$[\lambda (1 - p_L) + (1 - \lambda)(1 - p_H)]\, a_0^* = [\lambda p_L + (1 - \lambda)\, p_H]\, a_1^* \tag{22}$$

$$(1 - p_h)(a_{0h}^* - a_0^*) = p_h (a_{1h}^* - a_1^*) \qquad h = H, L \tag{23}$$

and obtaining the unknowns a_{1L}^**,* a_{1H}^**, and* a_0^* *by one of the following sets of conditions*

(i) $U_H(\mathbf{w}_H^*) = U_H(\mathbf{w}_L^*)$*,* $w_{0H}^* = w_{1H}^*$*, and*

$$\mu p_L (1 - p_L)\left[u'(w_{1L}^*) - u'(w_{0L}^*)\right] = \beta_H^* \left[\begin{array}{c} p_H (1 - p_L)\, u'(w_{1L}^*) \\ -p_L (1 - p_H)\, u'(w_{0L}^*) \end{array}\right] \tag{24}$$

$$\lambda (1 - \mu + \beta_H^*)\, u'(w_{0H}^*) = (1 - \lambda)\left[\mu - \beta_H^* \left(\frac{1 - p_H}{1 - p_L}\right)\right] u'(w_{0L}^*) \tag{25}$$

if $\lambda < \mu$*,*

(ii) $U_L(\mathbf{w}_L^*) = U_L(\mathbf{w}_H^*)$*,* $w_{0L}^* = w_{1L}^*$*, and*

$$(1 - \mu)\, p_H (1 - p_H)\left[u'(w_{1H}^*) - u'(w_{0H}^*)\right] = \beta_L^* \left[\begin{array}{c} p_L (1 - p_H)\, u'(w_{1H}^*) \\ -p_H (1 - p_L)\, u'(w_{0H}^*) \end{array}\right] \tag{26}$$

$$(1 - \lambda)(\mu + \beta_L^*)\, u'(w_{0L}^*) = \lambda \left[1 - \mu - \beta_L^* \left(\frac{1 - p_L}{1 - p_H}\right)\right] u'(w_{0H}^*) \tag{27}$$

if $\lambda > \mu$*, or*

(iii) $a_1^* = d - a_0^*$ *and* $(a_{0h}^*, a_{1h}^*) = (a_0^*, a_1^*)$ *for either* h*, if* $\lambda = \mu$*.*

Proof. Let β_h, γ_h, and δ be, respectively, the Lagrangean multipliers on the incentive-compatibility and individual rationality constraints of risk-type h and the feasible constraint in (4). The Kuhn-

Tucker first-order conditions for an optimum are given by

$$\left(\mu+\beta_L^*+\gamma_L^*\right)\frac{\partial U_L\left(\mathbf{w}_L^*\right)}{\partial a_{0L}} = \beta_H^*\frac{\partial U_H\left(\mathbf{w}_L^*\right)}{\partial a_{0L}}-\delta^*\lambda\left(1-p_L\right) \quad (28)$$

$$\left(\mu+\beta_L^*+\gamma_L^*\right)\frac{\partial U_L\left(\mathbf{w}_L^*\right)}{\partial a_{1L}} = \beta_H^*\frac{\partial U_H\left(\mathbf{w}_L^*\right)}{\partial a_{1L}}+\delta^*\lambda p_L \quad (29)$$

$$\left(1-\mu+\beta_H^*+\gamma_H^*\right)\frac{\partial U_H\left(\mathbf{w}_H^*\right)}{\partial a_{0H}} = \beta_L^*\frac{\partial U_L\left(\mathbf{w}_H^*\right)}{\partial a_{0H}}-\delta^*\left(1-\lambda\right)\left(1-p_H\right) \quad (30)$$

$$\left(1-\mu+\beta_H^*+\gamma_H^*\right)\frac{\partial U_H\left(\mathbf{w}_H^*\right)}{\partial a_{1H}} = \beta_L^*\frac{\partial U_L\left(\mathbf{w}_H^*\right)}{\partial a_{1H}}+\delta^*\left(1-\lambda\right)p_H \quad (31)$$

$$\beta_h^*\left(U_h\left(\mathbf{w}_h^*\right)-U_h\left(\mathbf{w}_{h'}^*\right)\right) = 0 \qquad h,h'\in\{H,L\} \quad (32)$$

$$\gamma_h^*\left(U_h\left(\mathbf{w}_h^*\right)-\overline{u}_h\right) = 0 \qquad h\in\{H,L\} \quad (33)$$

$$\delta^*\left[\lambda\Pi_L\left(\mathbf{a}_L^*\right)+\left(1-\lambda\right)\Pi_H\left(\mathbf{a}_H^*\right)\right] = 0 \quad (34)$$

$$\beta_h^*,\gamma_h^*,\delta^* \geq 0 \qquad h\in\{H,L\} \quad (35)$$

along with (1), (2), and (4). And because of (17)-(18) in the proof of Proposition 2, here also one equation from the pairs (28)-(29) and (30)-(31) can be replaced, respectively, by

$$\begin{aligned}&\left(\mu+\beta_L^*+\gamma_L^*\right)p_L\left(1-p_L\right)\left[u'\left(w_{1L}^*\right)-u'\left(w_{0L}^*\right)\right]\\ =\;&\beta_H^*\left[p_H\left(1-p_L\right)u'\left(w_{1L}^*\right)-p_L\left(1-p_H\right)u'\left(w_{0L}^*\right)\right]\end{aligned} \quad (36)$$

and

$$\begin{aligned}&\left(1-\mu+\beta_H^*+\gamma_H^*\right)p_H\left(1-p_H\right)\left[u'\left(w_{1H}^*\right)-u'\left(w_{0H}^*\right)\right]\\ =\;&\beta_L^*\left[p_L\left(1-p_H\right)u'\left(w_{1H}^*\right)-p_H\left(1-p_L\right)u'\left(w_{0H}^*\right)\right]\end{aligned} \quad (37)$$

As before, our analysis will proceed through a series of observations regarding the characteristics of an IIE(μ) optimal allocation.

Step 1

(i). The profit constraint binds at the optimum $\forall\mu\in[0,1]$.

Suppose that the constraint in (4) is slack at the optimum. Then $\delta^*=0$ by (34) and, given (17)-(18), (28)-(29) read

$$\begin{aligned}\left(\mu+\beta_L^*+\gamma_L^*\right)\left(1-p_L\right)u'\left(w_{0L}^*\right) &= \beta_H^*\left(1-p_H\right)u'\left(w_{0L}^*\right)\\ \left(\mu+\beta_L^*+\gamma_L^*\right)p_L u'\left(w_{1L}^*\right) &= \beta_H^* p_H u'\left(w_{1L}^*\right)\end{aligned}$$

Yet $u'(\cdot)>0$ and $p_h\in(0,1)$ for either h. Hence, in conjunction with the non-negativity conditions in (35), $\mu>0$ requires that neither side in either equation is zero. But then one equation may be divided by the other to give $\frac{1-p_L}{p_L}=\frac{1-p_H}{p_H}$, a contradiction.

This establishes that the constraint in (4) may be slack only if $\mu=0$. However, a trivially similar argument, using now (30)-(31), shows that the constraint in (4) may be slack only if $\mu=1$. Clearly, the constraint in question cannot be slack for any μ.

(ii). Searching for the IIE(μ) optimum, it is without loss of generality to restrict attention to menus $\{\mathbf{a}_L,\mathbf{a}_H\}$ *such that*

$$\Pi_H\left(\mathbf{a}_H\right)\leq 0\leq\Pi_L\left(\mathbf{a}_L\right) \quad (38)$$

As it will turn out, for any $\mu \in [0,1]$, the IIE(μ) optimal contract for the high-risk type $\mathbf{a}_H^*$ is such that $U_H(\mathbf{w}_H^*) \geq u(W - p_H d)$ - the inequality being strict unless it coincides with the RSW-optimal high-risk contract $\mathbf{a}_H^{**}$. Yet, $u(W - p_H d) = U_H(\mathbf{w}_H^{**})$ (recall Step 3(vi) in the proof of Proposition 2), while $\mathbf{a}_H^{**} \in \arg\max_{\mathbf{a}\in\mathbb{R}_+^2 : \Pi_H(\mathbf{a})\geq 0} U_H(\mathbf{a})$ and $\Pi_H(\mathbf{a}_H^{**}) = 0$. Clearly, it cannot be $\Pi_H(\mathbf{a}_H^*) > 0$. It is without loss of generality, therefore, to restrict attention to high-risk contracts that satisfy $\Pi_H(\mathbf{a}_H) \leq 0$. The claim now follows because, by the preceding step, it is also without loss of generality if we restrict attention to the case in which (4) holds as an equality constraint:

$$\lambda \Pi_L(\mathbf{a}_L) + (1-\lambda)\Pi_H(\mathbf{a}_H) = 0 \tag{39}$$

(iii). A menu $\{\mathbf{a}_L, \mathbf{a}_H\}$ satisfying (38)-(39), corresponds uniquely to a contract $\mathbf{a} \in FO_M^*$ such that

$$\begin{aligned} a_0 &\geq 0 && (40)\\ [\lambda(1-p_L) + (1-\lambda)(1-p_H)]\, a_0 &= [\lambda p_L + (1-\lambda) p_H]\, a_1 && (41)\\ (1-p_h)(a_{0h} - a_0) &= p_h (a_{1h} - a_1) \qquad h = H, L && (42) \end{aligned}$$

Let $\mathbf{a}$ be defined (uniquely) by (42), as the intersection of the two fair-odds lines FO_h through the members of the given menu. As pooling policy, it expects profits

$$\begin{aligned} \Pi_M(\mathbf{a}) &= \lambda[(1-p_L)a_0 - p_L a_1] + (1-\lambda)[(1-p_H)a_0 - p_H a_1]\\ &= \lambda[(1-p_L)a_0 - p_L a_1] + (1-\lambda)[(1-p_H)a_0 - p_H a_1]\\ &+ \lambda[(1-p_L)(a_{0L} - a_0) - p_L(a_{1L} - a_1)]\\ &+ (1-\lambda)[(1-p_H)(a_{0H} - a_0) - p_H(a_{1H} - a_1)]\\ &= \lambda[(1-p_L)a_{0L} - p_L a_{1L}] + (1-\lambda)[(1-p_H)a_{0H} - p_H a_{1H}]\\ &= \lambda \Pi_L(\mathbf{a}_L) + (1-\lambda)\Pi_H(\mathbf{a}_H) \end{aligned}$$

Given (39), therefore, it must be $\Pi_M(\mathbf{a}) = 0$, which is just another way to write (41). This equation being, moreover, satisfied also by the endowment point - the trivial contract $(0,0)$ - $\mathbf{a}$ cannot but lie on the market fair-odds line FO_M^*. Finally, notice that

$$\begin{aligned} \left(\frac{1-p_L}{p_L}\right) a_0 &= \frac{1}{p_L}[(1-p_L)a_{0L} - p_L a_{1L} + p_L a_1] = \frac{1}{p_L}[\Pi_L(\mathbf{a}_L) + p_L a_1]\\ &\geq a_1 = \frac{1}{p_H}[(1-p_H)a_0 - \Pi_H(\mathbf{a}_H)] \geq \left(\frac{1-p_H}{p_H}\right) a_0 \end{aligned}$$

where the first and third equalities are due to the respective relations in (42) while the two inequalities follow from the respective sides of (38). Given this, (40) follows immediately since $p_H > p_L$.

Step 2.

In what follows, we study the IIE(μ) problem after having substituted (38)-(39) by the equivalent four-relations system in (40)-(42). Of course, for the menu $\{\mathbf{a}_L, \mathbf{a}_H\}$ to satisfy the two conditions in (42), we can only consider movements along the same slope as the corresponding fair-odds lines FO_h, i.e. contract changes of the form $\mathrm{d}a_{1h} = \left(\frac{1-p_h}{p_h}\right) \mathrm{d}a_{0h}$. This means that the solution to the

IIE(μ) problem can be fully characterized in terms of the quantities a_0^*, a_{0L}^*, and a_{0H}^*. For the three equations in (41)-(42) above can be equivalently written as

$$a_1^* = \left(\frac{1}{\lambda p_L + (1-\lambda) p_H} - 1\right) a_0^* \tag{43}$$

$$a_{1L}^* = \frac{1}{p_L}\left[(1-p_L) a_{0L}^* - \frac{(1-\lambda)(p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0^*\right] \tag{44}$$

$$a_{1H}^* = \frac{1}{p_H}\left[(1-p_H) a_{0H}^* + \frac{\lambda (p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0^*\right] \tag{45}$$

Here, the first equality is simply a re-arrangement of (41) and, given this, the other two equalities follow from the respective ones in (42).

<u>*Step 3.*</u>

With no loss of generality, therefore, the feasibility constraint of the IIE(μ) problem can be replaced by (40) and (43)-(45). Given the latter three equations, we may write

$$\begin{aligned}
w_{0L} &= W - \left(\frac{1}{1-p_L}\right)\left(p_L a_{1L} + \frac{(1-\lambda)(p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0\right) \\
w_{1L} &= W - d + a_{1L} \\
w_{0H} &= W - \left(\frac{1}{1-p_H}\right)\left(p_H a_{1H} - \frac{\lambda (p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0\right) \\
w_{1H} &= W - d + a_{1H}
\end{aligned}$$

We may treat, therefore, a_{1L}, a_{1H}, and a_0 as our free choice variables and regard (1)-(2) and (40) as our inequality constraints. Letting β be the Lagrangean multiplier on the latter, the Kuhn-Tucker first-order conditions are given by (40), (43)-(45), (1)-(2), any one from the pair (28)-(29), any one from the pair (30)-(31), (32)-(33), (36)-(37), and

$$\begin{aligned}
\lambda\left[1 - \mu + \beta_H^* + \gamma_H^* - \beta_L^*\left(\frac{1-p_L}{1-p_H}\right)\right] u'\left(w_{0H}^*\right) &\; - \\
(1-\lambda)\left[\mu + \beta_L^* + \gamma_L^* - \beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] u'\left(w_{0L}^*\right) &= -\frac{\beta^*\left[\lambda p_L + (1-\lambda) p_H\right]}{p_H - p_L} \\
\beta^* a_0^* &= 0, \quad \beta^* \geq 0
\end{aligned} \tag{46}$$

(i). Regarding this new formulation, it is the slackness of (40) that distinguishes the IIE(μ) and RSW(μ) problems.

For if $a_0^* = 0$, (41) requires that also $a_1^* = 0$ and, hence, (42) reduces to (21). In which case, the optimum is the RSW allocation.

<u>*Step 4.*</u>

Hence, in what follows, we take $a_0^* > 0$. As a consequence, $\beta^* = 0$ and (46) can be re-written as

$$\begin{aligned}
&\lambda\left[1 - \mu + \beta_H^* + \gamma_H^* - \beta_L^*\left(\frac{1-p_L}{1-p_H}\right)\right] u'\left(w_{0H}^*\right) \\
&= (1-\lambda)\left[\mu + \beta_L^* + \gamma_L^* - \beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] u'\left(w_{0L}^*\right)
\end{aligned} \tag{47}$$

(i). At least one of the two incentive constraints in (1) binds at the optimum

To see this, suppose that they are both slack so that $\beta_L^* = 0 = \beta_H^*$ by the complementary slackness conditions in (13). By the last equation above, this cannot be if $\mu = \gamma_L^* = 0$ or $1-\mu = \gamma_H^* = 0$. For then, we would have, respectively, $u'(w_{1H}^*) = 0$ or $u'(w_{1L}^*) = 0$; either an absurd conclusion given that $u(\cdot)$ is everywhere strictly monotone. It can only be, therefore, $\mu + \gamma_L^*, 1-\mu+\gamma_H^* > 0$ and (36)-(37) dictate that both risk-types ought to be fully-insured. But then, $U_h(\mathbf{w}_{h'}^*) = u(w_{0h'}^*)$ for $h, h' \in \{H, L\}$ and, thus, $U_L(\mathbf{w}_L^*) \geq U_L(\mathbf{w}_H^*) = U_H(\mathbf{w}_H^*) \geq U_H(\mathbf{w}_L^*) = U_L(\mathbf{w}_L^*)$; another absurd conclusion if both incentive constraints in (1) are slack.

Step 5.

The preceding observation allows for an exhaustive investigation of the IIE(μ) problem by examining the following three cases. In the first two, we ignore the possibility that $\beta_L^* = 0 = \beta_H^*$. As we have just established, this cannot be if $\mu = \gamma_L^* = 0$ or $1-\mu = \gamma_H^* = 0$. And when $\mu+\gamma_L^*, 1-\mu+\gamma_H^* > 0$, it requires that both incentive constraints bind, exactly the situation investigated by the third case below.

(A). Only the high-risk incentive constraint binds at the optimum.

(i). The high-risk [resp. low-risk] type is fully [resp. under] insured at the optimum.

Since only the high-risk incentive constraint binds, we ought to have $\beta_L^* = 0 < \beta_H^*$. By (37), therefore, it must be $w_{0H}^* = w_{1H}^*$. Which implies, in turn, under- or full-insurance for the low-risk type: $w_{0L}^* \geq w_{1L}^*$ (recall Step 3(viii) in the proof of Proposition 2). In fact, it suffices here to consider only strict inequality: $w_{0L}^* > w_{1L}^*$. For, as we have already seen above, if both types are full-insured both incentive constraints in (1) bind, a situation examined in 5(C) below.

(ii). The high-risk individual rationality constraint in (2) *is slack (and, thus,* $\gamma_H^* = 0$*).*

At the optimal high-risk income point (w_{0H}^*, w_{1H}^*), we have now

$$\begin{aligned}
\overline{u}_H &= (1-p_H)\, u(W) + p_H u(W-d) < u\left((1-p_H) W + p_H (W-d)\right) \\
&= u(W - p_H d) < u\left(W - p_H d + \frac{\lambda (p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0^*\right) \\
&= u\left((1-p_H) w_{0H}^* + p_H w_{1H}^*\right) = u(w_{0H}^*) = U_H(\mathbf{w}_H^*)
\end{aligned}$$

where the first inequality is due to risk-aversion while the second follows from $a_0^* > 0$, $p_H > p_L$, and non-satiation.

(iii). The low-risk individual rationality constraint in (2) *is slack (and, thus,* $\gamma_L^* = 0$*) for any* $\mu > 0$.

Let $\mu > 0$ and consider the IIE(μ') and RSW(μ') problems for some $\mu' \in (0, \mu)$. Since $\mu' > 0$, the RSW(μ') optimum is given uniquely by $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$, at which point the low-risk individual rationality constraint is slack (recall Step 3(ix) in the proof of Proposition 2). Denoting, therefore, by $\mathbf{a}_L^*(\mu')$ the low-risk optimal contract in the IIE(μ') problem, we ought to have $U_L(\mathbf{w}_L^*(\mu')) \geq U_L(\mathbf{w}_L^{**}) > \overline{u}_L$, the first inequality since $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$ is feasible in the IIE(μ) problem for all $\mu \in [0,1]$. The claim now follows as it must be $U_L(\mathbf{w}_L^*(\mu)) \geq U_L(\mathbf{w}_L^*(\mu'))$ by Lemma 6.

(iv). Given the above, the optimum is fully characterized here by (43)-(44),

$$a_{0H}^* = d - a_{1H}^* \tag{48}$$

$$a_{1H}^* = \frac{\lambda (p_H - p_L)}{\lambda p_L + (1-\lambda) p_H} a_0^* + (1-p_H)\, d \tag{49}$$

and

$$\mu p_L (1-p_L)\left[u'(w_{1L}^*) - u'(w_{0L}^*)\right] = \beta_H^* \left[\begin{array}{c} p_H (1-p_L) u'(w_{1L}^*) \\ -p_L (1-p_H) u'(w_{0L}^*) \end{array}\right] \tag{50}$$

$$\lambda (1-\mu+\beta_H^*) u'(w_{0H}^*) = (1-\lambda)\left[\mu - \beta_H^* \left(\frac{1-p_H}{1-p_L}\right)\right] u'(w_{0L}^*) \tag{51}$$

$$U_H(\mathbf{w}_L^*) = u(w_{0H}^*)$$

This is a system of seven equations in the seven unknowns (a_0^*, a_1^*, a_{0L}^*, a_{1L}^*, a_{0H}^*, a_{1H}^*, and β_H^*). Here (49) is due to (48) and (45), while (50) and (51) are how (36) and (47) read, respectively.

(v). This solution is valid only if $\lambda < \mu$.

Recall Step A(i). The solution offers full insurance to the high-risk agents for whom the incentive constraint binds. That is,

$$U_H(w_{0H}^*) = u(w_{0H}^*) = U_H(\mathbf{w}_L^*) = u(w_{0L}^*) + p_H\left[u(w_{1L}^*) - u(w_{0L}^*)\right]$$

and, as under-insurance is offered to the low-risk type ($w_{1L}^* < w_{0L}^*$), it must be $u(w_{0L}^*) > u(w_{0H}^*)$; equivalently, $u'(w_{0H}^*) > u'(w_{0L}^*)$ due to risk-aversion. But then, the right-hand side of (51) ought to give

$$\begin{aligned}(1-\lambda)\left[\mu - \beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] u'(w_{0L}^*) &< \left[\mu(1-\lambda) + \lambda\beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] u'(w_{0L}^*) \\ &\leq \left[\mu(1-\lambda) + \lambda\beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] u'(w_{0H}^*) \\ &\leq (\mu(1-\lambda) + \lambda\beta_H^*) u'(w_{0H}^*)\end{aligned}$$

where the first inequality is due to $\beta_H^*, u'(\cdot) > 0$ and $p_h < 1$, the second one is because $u'(w_{0H}^*) > u'(w_{0L}^*)$, and the last inequality follows from $p_H > p_L$. Observe now that $\lambda \geq \mu$ implies $\mu(1-\lambda) \leq \lambda(1-\mu)$, forcing the last inequality above to lead to a negation of (51).

(B). Only the low-risk incentive constraint binds at the optimum.

(i). The low-risk [resp. high-risk] type is fully [resp. over] insured at the optimum.

Since only the low-risk incentive constraint is binding, we ought to have $\beta_L^* > 0 = \beta_H^*$. By (36), therefore, it must be $w_{0L}^* = w_{1L}^*$. But then $U_L(\mathbf{w}_L^*) = u(w_{0L}^*) = U_H(\mathbf{w}_L^*)$ and the two incentive compatibility constraints can be put together as $U_H(\mathbf{w}_H^*) \geq u(w_{0L}^*) \geq U_L(\mathbf{w}_H^*)$. And the inequality between the first and the last quantities requires that $(p_H - p_L)\left[u(w_{1H}^*) - u(w_{0H}^*)\right] \geq 0$. Equivalently, $w_{0H}^* \leq w_{1H}^*$ given that $p_H > p_L$ and $u(\cdot)$ being strictly-increasing. As before, it suffices to consider only the case $w_{0H}^* < w_{1H}^*$. For if both types are fully insured, then either constraint in (1) binds and this situation examined in 5(C) below.

(ii). Both individual rationality constraints in (2) are slack (and, thus, $\gamma_H^* = 0 = \gamma_L^*$).

Observe first that

$$\begin{aligned}u(W - p_L d) &= u((1-p_L) w_{0L}^* + p_L w_{1L}^* - (1-p_L) a_0 - p_L a_1) \\ &< u((1-p_L) w_{0L}^* + p_L w_{1L}^*) = u(w_{0L}^*) = U_L(\mathbf{w}_L^*) = U_H(\mathbf{w}_L^*) < U_H(\mathbf{w}_H^*)\end{aligned}$$

where the first equality follows from (42) for $h = L$ while the first inequality is due to $p_H > p_L$ and non-satiation. The claim now follows since, by risk-aversion,

$$\begin{aligned}\overline{u}_h = (1-p_h)\,u(W) + p_h u(W-d) &< u((1-p_h)\,W + p_h(W-d)) \\ &= u(W - p_h d)\end{aligned}$$

for either h while $u(W - p_H d) < u(W - p_L d)$ by $p_H > p_L$ and non-satiation.

(iii). In this case, therefore, the optimum is given by (43), (45),

$$a_{0L}^* = d - a_{1L}^* \tag{52}$$

$$a_{1L}^* = (1-p_L)\,d - \frac{(1-\lambda)(p_H - p_L)}{\lambda p_L + (1-\lambda)\,p_H} a_0^* \tag{53}$$

and

$$(1-\mu)\,p_H(1-p_H)\left[u'(w_{1H}^*) - u'(w_{0H}^*)\right] = \beta_L^* \begin{bmatrix} p_L(1-p_H)\,u'(w_{1H}^*) \\ -p_H(1-p_L)\,u'(w_{0H}^*) \end{bmatrix} \tag{54}$$

$$(1-\lambda)(\mu + \beta_L^*)\,u'(w_{0L}^*) = \lambda\left[1 - \mu - \beta_L^*\left(\frac{1-p_L}{1-p_H}\right)\right] u'(w_{0H}^*) \tag{55}$$

$$U_L(\mathbf{w}_H^*) = u(w_{0L}^*)$$

giving again a system of seven equations in the seven unknowns: a_0^*, a_1^*, a_{0L}^*, a_{1L}^*, a_{0H}^*, a_{1H}^*, and β_L^*. Here (53) is due to (52) and (44), whereas (54) and (55) are simply how (37) and (47) read, respectively.

(v). The solution is valid only if $\lambda > \mu$.

Recall Step B(i). Since full insurance is offered to the low-risk type and under-insurance to the high-risk while the low-risk incentive constraint is binding, a trivial adaptation of the argument in the first part of Step A(v) concludes that $u(w_{0L}^*) < u(w_{0H}^*)$. Equivalently, $u'(w_{0H}^*) < u'(w_{0L}^*)$, by risk-aversion, and the left-hand side of (55) gives

$$(1-\lambda)(\mu + \beta_L^*)\,u'(w_{0L}^*) \geq (1-\lambda)(\mu + \beta_L^*)\,u'(w_{0H}^*) > [\mu(1-\lambda) - \lambda\beta_L^*]\,u'(w_{0H}^*)$$

Here, the last inequality is due to $\beta_L^*, u'(\cdot) > 0$ which, along with $p_H > p_L$, requires also that the right-hand side of (55) gives

$$\lambda\left[1 - \mu - \beta_L^*\left(\frac{1-p_L}{1-p_H}\right)\right] u'(w_{0H}^*) \leq \lambda(1 - \mu - \beta_L^*)\,u'(w_{0H}^*)$$

Observe now that $\lambda \leq \mu$ is equivalent to $\lambda(1-\mu) \leq \mu(1-\lambda)$. Clearly, (55) cannot hold if $\mu \geq \lambda$.

(C). Both incentive constraints in (1) bind at the optimum.

(i). The optimum is a pooling contract offering full insurance. Recall Step 3(iii) in the proof of Proposition 2. If both incentive constraints are binding, we ought to have $\mathbf{w}_L^* = \mathbf{w}_H^*$ so that the optimum is a pooling contract, denoted henceforth by $\mathbf{a}^* = (a_0^*, a_1^*)$. To specify it, observe that, taking $w_{sh}^* = w_s^*$ for $(s,h) \in \{0,1\} \times \{L,H\}$, the first-order conditions read now

$$(\mu + \beta_L^* + \gamma_L^*)\,p_L(1-p_L)\left[u'(w_1^*) - u'(w_0^*)\right] = \beta_H^* \begin{bmatrix} p_H(1-p_L)\,u'(w_1^*) \\ -p_L(1-p_H)\,u'(w_0^*) \end{bmatrix} \tag{56}$$

$$(1 - \mu + \beta_H^* + \gamma_H^*)\,p_H(1-p_H)\left[u'(w_1^*) - u'(w_0^*)\right] = \beta_L^* \begin{bmatrix} p_L(1-p_H)\,u'(w_1^*) \\ -p_H(1-p_L)\,u'(w_0^*) \end{bmatrix} \tag{57}$$

$$\lambda\left[1-\mu+\beta_H^*+\gamma_H^*-\beta_L^*\left(\frac{1-p_L}{1-p_H}\right)\right]=(1-\lambda)\left[\mu+\beta_L^*+\gamma_L^*-\beta_H^*\left(\frac{1-p_H}{1-p_L}\right)\right] \tag{58}$$

the last equation because $u(\cdot)$ is everywhere strictly monotone. It is trivial, however, to verify that these three equations together give

$$(p_H-p_L)\left[\lambda\beta_L^*p_L+(1-\lambda)\beta_H^*p_H\right]u'(w_1^*)=0$$

Equivalently, $\lambda\beta_L^*p_L+(1-\lambda)\beta_H^*p_H=0$ which, each term of the sum on the left-hand side being non-negative, can be only if $\beta_L^*=\beta_H^*=0$. Yet, the system (56)-(57) gives then

$$\left[(\mu+\gamma_L^*)(1-p_L)+(1-\mu+\gamma_H^*)(1-p_H)\right]\left[u'(w_1^*)-u'(w_0^*)\right]=0$$

which (at least one term of the sum in the first pair of brackets on the left-hand side being strictly positive) cannot be unless $u'(w_1^*)-u'(w_0^*)$. Equivalently, $w_0^*=w_1^*$ as required.
(ii). The optimum is given by the full insurance contact that satisfies (41).
As it has to provide full insurance, the optimal pooling contract must be such that $a_0^*=d-a_1^*$. This along with (41) defines a system of two equations in the two unknowns a_0^* and a_1^*.
(iii). Both individual rationality constraints in (2) *are slack (and, thus,* $\gamma_H^*=0=\gamma_L^*$*).*
Recall Step B(ii). As either risk-type gets full insurance, in this case, we have

$$\begin{aligned}\overline{u}_h<u(W-p_hd) &= u\left((1-p_h)w_0^*+p_hw_1^*-(1-p_h)a_0^*-p_ha_1^*\right)\\ &< u\left((1-p_h)w_0^*+p_hw_1^*\right)=u(w_{0h}^*)=U_h(\mathbf{w}^*)\end{aligned}$$

(iv). The solution is valid only if $\lambda=\mu$.
This is immediate. For the condition is but another way of writing (58) when $\beta_h^*=0=\gamma_h^*$ for either h.
Step 6.
The preceding analysis leads to a complete characterization of the IIE allocations. Specifically, for a given parameter-pair $(\lambda,\mu)\in(0,1)\times[0,1]$, the contrapositives of the observations in 5.A.(v), 5.B.(v), and 5.C.(iv) above leaves one as the unique possibility. If $\lambda<\mu$, for instance, our analysis rules out the cases in 5.B and 5.C, leaving 5.A (i) as the only one relevant.
Step 7.
It remains to verify that the rank condition of the Khun-Tucker theorem is satisfied, for each of the three solutions in 5(A)-(C) above. To this end, notice that, in either of the optima for 5(A)-(B), there is exactly one slack inequality constraint, the respective incentive constraint. The rank condition of the Khun-Tucker theorem, therefore, requires that the matrix

$$\begin{aligned}D(\mathbf{w}_L^*,\mathbf{w}_H^*) &= \begin{bmatrix}\nabla_{(\mathbf{w}_L,\mathbf{w}_H)}U_L(\mathbf{w}_L^*)-U_L(\mathbf{w}_H^*)\\ \nabla_{(\mathbf{w}_L,\mathbf{w}_H)}U_H(\mathbf{w}_H^*)-U_H(\mathbf{w}_L^*)\\ \nabla_{(\mathbf{w}_L,\mathbf{w}_H)}\lambda\Pi_L(\mathbf{a}_L^*)+(1-\lambda)\Pi_H(\mathbf{a}_H^*)\end{bmatrix}\\ &= -\begin{bmatrix}-\frac{\partial U_L(\mathbf{w}_L^*)}{\partial w_{0L}} & -\frac{\partial U_L(\mathbf{w}_L^*)}{\partial w_{1L}} & \frac{\partial U_L(\mathbf{w}_H^*)}{\partial w_{0H}} & \frac{\partial U_L(\mathbf{w}_H^*)}{\partial w_{1H}}\\ \frac{\partial U_H(\mathbf{w}_L^*)}{\partial w_{0L}} & \frac{\partial U_H(\mathbf{w}_L^*)}{\partial w_{1L}} & -\frac{\partial U_H(\mathbf{w}_H^*)}{\partial w_{0H}} & -\frac{\partial U_H(\mathbf{w}_H^*)}{\partial w_{1H}}\\ \lambda(1-p_L) & \lambda p_L & (1-\lambda)(1-p_H) & (1-\lambda)p_H\end{bmatrix}\end{aligned}$$

has rank at least 2. As in the RSW problem, this is trivial to verify. Take, for instance, a linear combination $(\zeta_0, \zeta_1) \in \mathbb{R}^2 \setminus \{\mathbf{0}\}$ of the first two columns. For this to be zero at the first two entries, we ought to have $\sum_{s=0,1} \zeta_s \frac{\partial U_h(\mathbf{w}_L)}{\partial w_{sL}} = 0$ for $h = L, H$. Equivalently, $\frac{\partial U_L(\mathbf{w}_L)}{\partial w_{0L}} / \frac{\partial U_H(\mathbf{w}_L)}{\partial w_{0L}} = \frac{\partial U_L(\mathbf{w}_L)}{\partial w_{1L}} / \frac{\partial U_H(\mathbf{w}_L)}{\partial w_{1L}}$ which, under the given utility specification, reads $\frac{p_L}{p_H} = \frac{1-p_L}{1-p_H}$. Yet, $p_H > p_L$ requires that $\frac{p_L}{p_H} < 1 < \frac{1-p_L}{1-p_H}$.
In 5(C), on the other hand, all three inequality constraints bind at the optimum and the rank condition requires now that the matrix

$$D(\mathbf{w}^*) = \begin{bmatrix} (1-p_L)\, u'(w_0^*) & p_L u'(w_1^*) & -(1-p_L)\, u'(w_0^*) & -p_L u'(w_1^*) \\ -(1-p_H)\, u'(w_0^*) & -p_H u'(w_1^*) & (1-p_H)\, u'(w_0^*) & p_H u'(w_1^*) \\ -\lambda(1-p_L) & -\lambda p_L & -(1-\lambda)(1-p_H) & -(1-\lambda)\, p_H \end{bmatrix}$$

has rank 3. Which is indeed the case since the 3x3 submatrix formed by the first two and the last column is non-singular. By adding the second column to the first and the last, its determinant is the same as that of the matrix

$$\begin{bmatrix} u'(w_0^*) & p_L u'(w_1^*) & 0 \\ -u'(w_0^*) & -p_H u'(w_1^*) & 0 \\ -\lambda & -\lambda p_L & -p_H + \lambda(p_H - p_L) \end{bmatrix}$$

which is given by $[p_H - \lambda(p_H - p_L)]\,(p_H - p_L)\, u'(w_0^*)\, u'(w_1^*) > 0$.
Step 8.
To complete the proof, we should point that the set of first-order conditions (28)-(35) also suffices for an optimum because the IIE problem has a (strictly) concave objective and a convex feasible set (recall footnote XX in the main text and see, for instance, Theorems 1.6, 1.14, and 1.19, Ch. 7 in de la Fuente [7]). By contrast, the first-order conditions in Step 3 remain only necessary.[24] ■

Proof of Claim 2

Recall Proposition 2 and Step 5 in the proof of Proposition 8. Obviously, the RSW allocation can be optimal for an IIE(μ) problem only if $\mu \in (\lambda, 1]$. Suppose then that it is the IIE(μ) optimum for some $\mu \in (\lambda, 1]$. Clearly, it must satisfy either of the two sets of first-order conditions (9)-(14) and (28)-(33). Which requires that $\delta_L^{**}\delta_H^{**} \neq 0$ with $\delta_L^{**}/\delta_H^{**} = \lambda/(1-\lambda)$, in order to have $\beta_h^* = \beta_h^{**}$ along with $\gamma_h^* = \gamma_h^{**}$ for either h, and $\delta_L^{**}/\lambda = \delta^* = \delta_H^{**}/(1-\lambda)$. But then, since the RSW allocation satisfies (9)-(14) for all $\mu \in (\lambda, 1]$, it must satisfy (28)-(33) as well for all $\mu \in (\lambda, 1]$. The result follows given that, for any $\mu \in (\lambda, 1]$, either profit constraint in the RSW problem binds at the RSW allocation and so does, thus, also the profit constraint of the IIE(μ) problem. ■

Proof of Claim 1

The first part of statement, the one concerning uniqueness, is obvious. Since $u(\cdot)$ is everywhere strictly concave, for any $\mu \in [0,1]$, the IIE(μ) problem is about maximizing a strictly concave

[24] Indeed, (28)-(34) defines a system of nine equations in nine unknowns. Yet, thirteen equations are specified in Step 3 while only three additional unknowns are introduced.

objective on a convex set (recall footnote XX in the main text). The second part of the statement follows immediately from the proof of Proposition 8 - in particular, from Steps 1(i), 5.A.(i)-(iii), 5.B.(i)-(ii), and 5.C.(i)-(iii).

Regarding now the last part of the statement, that the solution mapping $\mathbf{a}_h^*(\mu) : [0,1] \mapsto \mathbb{R}_+^2$ is a function follows immediately from the fact that, for any $\mu \in [0,1]$, the IIE(μ) optimum is unique. To show that this function is in fact bijective on $[0,\lambda)$ suppose otherwise. Let, therefore, a_{sh}^* with $(s,h) \in \{0,1\} \times \{L,H\}$ be optimal for $\mu, \mu' \in [0,\lambda)$ with $\mu \neq \mu'$ and recall Step 5.B in the proof of Proposition 8. There must be $\beta_L^*, \hat{\beta}_L \in \mathbb{R}_{++}$ such that the seven equations in Proposition 8(ii) hold, in conjunction with the following two

$$\begin{aligned} (1-\mu')\, p_H (1-p_H) \left[u'(w_{1H}^*) - u'(w_{0H}^*)\right] &= \hat{\beta}_L \begin{bmatrix} p_L(1-p_H)\, u'(w_{1H}^*) \\ -p_H(1-p_L)\, u'(w_{0H}^*) \end{bmatrix} \\ (1-\lambda)\left(\mu' + \hat{\beta}_L\right) u'(w_{0L}^*) &= \lambda \left[1 - \mu' - \hat{\beta}_L \left(\frac{1-p_L}{1-p_H}\right)\right] u'(w_{0H}^*) \end{aligned}$$

Recall also that $w_{0H}^* < w_{1H}^*$. Clearly, $u'(w_{1H}^*) < u'(w_{0H}^*)$ and $p_L(1-p_H)\, u'(w_{1H}^*) < p_H(1-p_L)\, u'(w_{0H}^*)$. Notice also that, since $\beta_L^*, \hat{\beta}_L > 0$, from the second equation above and (55) we ought to have, respectively, $1-\mu'-\hat{\beta}_L\left(\frac{1-p_L}{1-p_H}\right), 1-\mu-\beta_L^*\left(\frac{1-p_L}{1-p_H}\right) > 0$. That is, neither side of the two equations above nor of (54) or (55) may be zero. It follows, therefore, that

$$\frac{1-\mu}{1-\mu'} = \frac{\beta_L^*}{\hat{\beta}_L}, \qquad \frac{\mu+\beta_L^*}{\mu'+\hat{\beta}_L} = \frac{1-\mu-\beta_L^*\left(\frac{1-p_L}{1-p_H}\right)}{1-\mu'-\hat{\beta}_L\left(\frac{1-p_L}{1-p_H}\right)}$$

Yet, together these two equations lead to the contradiction $(\mu, \beta_L^*) = \left(\mu', \hat{\beta}_L\right)$. When the solution mapping is restricted on $(\lambda, 1]$, the argument is trivially similar for the case in which the RSW allocation is not optimal. If the RSW allocation is optimal, on the other hand, it is so for all $\mu \in (\lambda, 1]$ (Claim 2). ■

Lemma 7 *If $\{\mathbf{a}_L^*, \mathbf{a}_H^*\}$ is the IIE(μ) optimum for some $\mu \in (\lambda, 1]$, then $\Pi_H(\mathbf{a}_L^*) < \Pi_H(\mathbf{a}_H^*)$.*

Proof. Suppose to the contrary that $\Pi_H(\mathbf{a}_L^*) \geq \Pi_H(\mathbf{a}_H^*)$. Then, for any $\kappa \in (0,1)$, the contract $\mathbf{a}_H^\kappa = \kappa \mathbf{a}_L^* + (1-\kappa)\mathbf{a}_H^*$ would give $\Pi_H\left(\mathbf{a}_H^k\right) = \kappa \Pi_H(\mathbf{a}_L^*) + (1-\kappa)\Pi_H(\mathbf{a}_H^*) \geq \Pi_H(\mathbf{a}_H^*)$. Due to risk-aversion, moreover, since $\mathbf{a}_L^* \sim_H \mathbf{a}_H^*$, it must be $U_H(\mathbf{a}_H^\kappa) > \kappa U_H(\mathbf{a}_L^*) + (1-\kappa) U_H(\mathbf{a}_H^*) = U_H(\mathbf{a}_L^*)$ again for all $\kappa \in (0,1)$. Similarly, as $U_L(\mathbf{a}_L^*) > U_L(\mathbf{a}_H^*)$, we ought to have $U_L(\mathbf{a}_H^\kappa) > U_L(\mathbf{a}_H^*)$ for all $\kappa \in (0,1)$. Yet, given this, since $\lim_{\kappa \to 0} U_L\left(\mathbf{a}_H^0\right) = U_L(\mathbf{a}_H^*)$, continuity implies that letting $\Delta = U_L(\mathbf{a}_L^*) - U_L(\mathbf{a}_H^*)$ there exists $\overline{\kappa} \in (0,1)$ s.t. $U_L(\mathbf{a}_L^*) > U_L(\mathbf{a}_H^\kappa)$ for all $\kappa \in (0, \overline{\kappa})$.

Putting the above observations together, for any $\kappa \in (0,\overline{\kappa})$, we have $\mathbf{a}_L^* \succ_L \mathbf{a}_H^\kappa \succ_H \mathbf{a}_L^*$. But then the menu $\{\mathbf{a}_L^*, \mathbf{a}_H^\kappa\}$ is separating. As it also satisfies the profit and individual rationality constraints by construction, it is feasible in the PII(μ) problem. But this is absurd. For if $\mu \in (\lambda, 1)$, it means that $\{\mathbf{a}_L^*, \mathbf{a}_H^*\}$ cannot be optimal. If $\mu = 1$, on the other hand, it contradicts the fact that the IIE(1) optimum is unique. ■

C Equilibrium without cross-subsidization

C.1 The Rothschild-Stiglitz game

Proof of Proposition 1

Consider the contract $\mathbf{a}_L^1 = \mathbf{a} + (1,\kappa)\,\epsilon$ for some $\kappa \in (I_H(\mathbf{a}^{**}), I_L(\mathbf{a}))$ and $\epsilon < 0$. By Lemma 4, for small enough $|\epsilon|$, we ought to have $\mathbf{a}_L^1 \succ_L \mathbf{a} \succ_H \mathbf{a}_L^1$. Hence, in the presence of $\mathbf{a}$, the new contract attracts only the low-risk customers. And by doing so, it delivers profits

$$\begin{aligned}\Pi_L\left(\mathbf{a}_L^1\right) &= (1-p_L)(a_0+\epsilon) - p_L(a_1+\kappa\epsilon) = \Pi_L(\mathbf{a}) + [1-p_L(1+\kappa)]\,\epsilon \\ &= \Pi_M(\mathbf{a}) + [\Pi_L(\mathbf{a}) - \Pi_M(\mathbf{a})] + [1-p_L(1+\kappa)]\,\epsilon\end{aligned}$$

But $\Pi_L(\mathbf{a}) > \Pi_M(\mathbf{a})$ and, thus, $1+\kappa > \frac{1}{p_L}\left(1+\frac{1}{\epsilon}[\Pi_L(\mathbf{a}) - \Pi_M(\mathbf{a})]\right)$ if $|\epsilon|$ is sufficiently small. Equivalently, $\Pi_L\left(\mathbf{a}_L^1\right) > \Pi_M(\mathbf{a}) \geq 0$, the last inequality due to $\mathbf{a}$ being an admissible pooling policy. Clearly, $\mathbf{a}_L^1$ is a profitable deviation. ■

Proof of Proposition 3

Let $\{\mathbf{a}_L, \mathbf{a}_H\}$ be an equilibrium separating policy and observe first that it must be $\Pi_L(\mathbf{a}_L) = 0$. For if $\Pi_L(\mathbf{a}_L) = \delta > 0$, we could consider the contract $\mathbf{a}_L^1 = \mathbf{a}_L + (1,\kappa_L)\,\epsilon_L$ for some $\kappa_L \in (I_H(\mathbf{a}_L), I_L(\mathbf{a}_L))$ and $\epsilon_L < 0$. And this would give

$$\Pi_L\left(\mathbf{a}_L^0\right) = (1-p_L)(a_{0L}+\epsilon_L) - p_L(a_{1L}+\kappa_L\epsilon_L) = \Pi_L(\mathbf{a}_L) + [1-p_L(1+\kappa_L)]\,\epsilon_L = \delta + [1-p_L(1+\kappa_L)]\,\epsilon_L$$

and sufficiently small $|\epsilon_L|$ ensures that $|1-p_L(1+\kappa_L)||\epsilon_L| < \delta$. That is, $\Pi_L\left(\mathbf{a}_L^0\right) > 0$. By Lemma 4, moreover, for small enough $|\epsilon_L|$, we also have $\mathbf{a}_L^1 \succ_L \mathbf{a}_L \succ_H \mathbf{a}_L^1$ or $\mathbf{a}_L^1 \succ_L \mathbf{a}_H \succ_H \mathbf{a}_L^1$ since $\{\mathbf{a}_L, \mathbf{a}_H\}$ is separating. Clearly, the new menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H\right\}$ is a separating policy that attracts the low-risk type away from $\{\mathbf{a}_L, \mathbf{a}_H\}$ and makes strictly positive profits. Which is, of course, absurd given that the latter policy is supposed to be an equilibrium one. A trivially similar argument produces a contract $\mathbf{a}_H^1 = \mathbf{a}_H + (1,\kappa_H)\,\epsilon_H$ for some $\kappa_H \in (I_H(\mathbf{a}_L), I_L(\mathbf{a}_L))$ and $\epsilon_H > 0$ showing that it must be $\Pi_H(\mathbf{a}_H) = 0$.

Suppose now that $\{\mathbf{a}_L, \mathbf{a}_H\} \neq \{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$. Since $\{\mathbf{a}_L^{**}, \mathbf{a}_H^{**}\}$ is uniquely optimal for the RSW problem for $\mu \in (0,1)$, it must be $\mathbf{a}_h^{**} \succsim_h \mathbf{a}_h$ for either $h \in \{L, H\}$ with strict preference for at least one type. Hence, there are two cases to consider.

Case (i): $\mathbf{a}_H^{**} \succ_H \mathbf{a}_H$ and $\mathbf{a}_L^{**} \sim_L \mathbf{a}_L$.

Suppose first that $w_{0H} > w_{1H}$. Then, $u'(w_{0H}) < u'(w_{1H})$ by risk-aversion and, thus, $I_h(\mathbf{a}_H) < \frac{1-p_h}{p_h}$ for either h. Hence, since $p_L < p_H$ and, thus, also $I_L(\mathbf{a}_H) > I_H(\mathbf{a}_H)$, letting $\kappa_H \in \left(\frac{1-p_L}{p_L}, \infty\right)$ and $\epsilon_H < 0$ allows Lemma 5 to ensure that the contract $\mathbf{a}_H^2 = \mathbf{a}_H^{**} + (1,\kappa_H)\,\epsilon_H$ is such that $\mathbf{a}_H^{**} \succ_L \mathbf{a}_H^2 \succ_H \mathbf{a}_H$ for sufficiently small $|\epsilon_H|$. Recall also that $\mathbf{a}_L \sim_L \mathbf{a}_L^{**} \succ_L \mathbf{a}_H^{**}$ while $\mathbf{a}_H \succsim_H \mathbf{a}_L$. Clearly, the menu $\left\{\mathbf{a}_L, \mathbf{a}_H^2\right\}$ is separating and attracts at least the high-risk type away from $\{\mathbf{a}_L, \mathbf{a}_H\}$. Doing so, moreover, is strictly profitable since $\Pi_H\left(\mathbf{a}_H^2\right) = \Pi_H(\mathbf{a}_H) + (1-p_H-\kappa_H p_H)\,\epsilon_H = (1-p_H-\kappa_H p_H)\,\epsilon_H > 0$, the second equality since $\Pi_H(\mathbf{a}_H) = 0$ as we have shown at the opening step of this proof.

If $w_{0H} < w_{1H}$, on the other hand, then $I_h(\mathbf{a}_H) > \frac{1-p_H}{p_H}$ for either h and the preceding argument applies now for $\kappa_H \in \left(-\infty, \frac{1-p_H}{p_H}\right)$ and sufficiently small $\epsilon_H > 0$. To complete the examination of this case observe that it cannot be $w_{0H} = w_{1H}$. For, as $\Pi_H(\mathbf{a}_H) = 0$, $\mathbf{a}_H$ would lie on the intersection between the FO^*_H and the 45-degree lines. But then it would coincide with $\mathbf{a}^{**}_H$, contradicting the hypothesis that $\mathbf{a}^{**}_H \succ_H \mathbf{a}_H$.

Case (ii): $\mathbf{a}^{**}_L \succ_L \mathbf{a}_L$ and $\mathbf{a}^{**}_H \succsim_H \mathbf{a}_H$.

Suppose first that $w_{0L} > w_{1L}$. Then, $u'(w_{0L}) < u'(w_{1L})$ and, thus, $I_h(\mathbf{a}_L) < \frac{1-p_L}{p_L}$ for either h. For $\kappa_L \in \left(\frac{1-p_L}{p_L}, \infty\right)$ and $\epsilon_L < 0$ with $|\epsilon_L|$ sufficiently small, therefore, the contract $\mathbf{a}^2_L = \mathbf{a}^{**}_L + (1, \kappa_L)\epsilon_L$ gives $\mathbf{a}^{**}_L \succ_H \mathbf{a}^2_L \succ_L \mathbf{a}_L$ (Lemma 5). Recall, moreover, that $\mathbf{a}^{**}_L \succ_L \mathbf{a}^{**}_H$. Letting, therefore, $\Delta = U_L(\mathbf{a}^{**}_L) - U_L(\mathbf{a}^{**}_H)$ since $\lim_{\epsilon_L \to 0} \mathbf{a}^2_L = \mathbf{a}^{**}_L$, by continuity, it must be $U_L(\mathbf{a}^{**}_L) - U_L\left(\mathbf{a}^2_L\right) < \Delta$ and, thus, $\mathbf{a}^2_L \succ_L \mathbf{a}^{**}_H$ for sufficiently small $|\epsilon_L|$. But then, we have $\mathbf{a}^{**}_H \sim_H \mathbf{a}^{**}_L \succ_H \mathbf{a}^2_L \succ_L \mathbf{a}^{**}_H$. That is, the menu $\left\{\mathbf{a}^2_L, \mathbf{a}^{**}_H\right\}$ is separating and attracts at least the low-risk type away from $\{\mathbf{a}_L, \mathbf{a}_H\}$. Moreover, it is strictly profitable since $\Pi_L\left(\mathbf{a}^2_L\right) = \Pi_L(\mathbf{a}^{**}_L) + (1 - p_L - \kappa_L p_L)\epsilon_L = (1 - p_L - \kappa_L p_L)\epsilon_L > 0$. If $w_{0L} < w_{1L}$, on the other hand, then $I_h(\mathbf{a}_H) > \frac{1-p_H}{p_H}$ for either h and, by Lemma 5 again, the preceding argument applies now for $\kappa_L \in \left(-\infty, \frac{1-p_H}{p_H}\right)$ and sufficiently small $\epsilon_L > 0$. To complete the proof, observe that the case $w_{0L} = w_{1L}$ is not possible. For it would imply that

$$\begin{aligned}
U_H(\mathbf{w}_H) = (1-p_H)u(w_{0H}) + p_H u(w_{1H}) &< u((1-p_H)w_{0H} + p_H w_{1H}) \\
&= u(W - (1-p_H)a_{0H} + p_H(a_{1H} - d)) \\
&= u(W - p_H d) \\
&< u(W - p_L d) \\
&= u(W - (1-p_L)a_{0L} + p_L(a_{1L} - d)) \\
&= u((1-p_L)w_{0L} + p_L w_{1L}) = u(w_{0L}) = U_H(\mathbf{w}_L)
\end{aligned}$$

contradicting that $\{\mathbf{a}_L, \mathbf{a}_H\}$ is incentive compatible for the high-risk type. The first inequality above is due to $u(\cdot)$ being everywhere strictly-concave (and it is necessarily strict since as we have seen $w_{0H} \neq w_{1H}$). The second inequality follows from $p_L < p_H$ and $u(\cdot)$ being strictly-increasing. The second and fifth equalities use that $\mathbf{w}_h = (W - a_{0h}, W - d + a_{1h})$ while the third and fourth follow from the fact that $\Pi_h(\mathbf{a}_h) = 0$ for either h as we have shown at the opening step of this proof. ■

Proof of Proposition 4

"Only if." We will argue by contradiction. Suppose, therefore, that $\mathbf{a} \in FO^*_M : \mathbf{a} \succ_L \mathbf{a}^{**}_L$. Since the low-risk individual-rationality constraint does not bind at the RSW allocation, we also have $\mathbf{a} \succ_L \mathbf{a}^{**}_L \succ_L \mathbf{0}$, the latter contract corresponding to the endowment point (also on FO^*_M). By the continuity of the preference relation $\succsim_L$, therefore, there must exist a convex combination of $\mathbf{a}$ and $\mathbf{0}$ such that the low-risk type is indifferent between this new point and her RS contract.[25]

[25] Recall that there is a one-to-one relation between contract and income points. The continuity of the relation $\succsim_L$ derives from continuous preferences over lotteries on wealth vectors. Given this, the existence of the wealth vector in

There must be, therefore, $\mathbf{a}^1 = \pi\mathbf{a}$ for some $\pi \in (0,1)$ such that $\mathbf{a}^1 \in FO_M^*$ and $\mathbf{a}^1 \sim_L \mathbf{a}_L^{**}$. Notice now that, since $\mathbf{a}, \mathbf{a}^1 \in FO_M^*$ and $\pi < 1$, it ought to be $\mathbf{a}^1 = \mathbf{a} - \left(\frac{\overline{p}}{1-\overline{p}}, 1\right)\epsilon$ for some $\epsilon > 0$. By Lemma 2 then the corresponding income points ought to give

$$\begin{aligned} U_L(\mathbf{w}) - U_L(\mathbf{w}^1) &= \left[p_L u'\left(w_1^1 + \epsilon'\right) - \frac{\overline{p}(1-p_L)}{1-\overline{p}} u'\left(w_0^1 - \frac{\overline{p}\epsilon'}{1-\overline{p}}\right)\right]\epsilon \\ &= \left[p_L u'\left(w_1'\right) - \frac{\overline{p}(1-p_L)}{1-\overline{p}} u'\left(w_0'\right)\right]\epsilon \end{aligned}$$

for some $\epsilon' \in (0, \epsilon)$. Here, the second equality is because we let $\mathbf{w}' = \mathbf{w}^1 + \left(-\frac{\overline{p}}{1-\overline{p}}, 1\right)\epsilon'$ which corresponds to the contract $\mathbf{a}' = \mathbf{a}^1 + \left(\frac{\overline{p}}{1-\overline{p}}, 1\right)\epsilon'$ that is also on FO_M^*. Obviously, since $\mathbf{a} \succ_L \mathbf{a}^1$, it cannot be $\frac{\overline{p}}{1-\overline{p}} \geq \frac{p_L u'\left(w_1'\right)}{(1-p_L)u'\left(w_0'\right)}$.

It must be, therefore, $\frac{\overline{p}}{1-\overline{p}} < \frac{p_L u'\left(w_1'\right)}{(1-p_L)u'\left(w_0'\right)}$ and we may consider another contract $\mathbf{a}^2 = \mathbf{a}' + (\kappa, 1)\varepsilon$ with $\kappa \in \left(\frac{\overline{p}}{1-\overline{p}}, \frac{p_L u'\left(w_1'\right)}{(1-p_L)u'\left(w_0'\right)}\right)$ and $\varepsilon > 0$. As a pooling policy, this gives

$$\Pi_M\left(\mathbf{a}^2\right) = \Pi_M\left(\mathbf{a}'\right) + \left[\kappa(1-\overline{p}) - \overline{p}\right]\varepsilon = \left[\kappa(1-\overline{p}) - \overline{p}\right]\varepsilon > 0$$

the second equality since $\mathbf{a}' \in FO_M^*$. For either h, moreover, applying again Lemma 3 successively gives

$$\begin{aligned} U_h\left(\mathbf{w}^2\right) - U_h\left(\mathbf{w}^1\right) &= U_h\left(\mathbf{w}^3\right) - U_h\left(\mathbf{w}'\right) + U_h\left(\mathbf{w}'\right) - U_h\left(\mathbf{w}^1\right) \\ &= \left[p_h u'\left(w_1' + \varepsilon'\right) - \kappa(1-p_h)u'\left(w_0' - \kappa\varepsilon'\right)\right]\varepsilon \\ &\quad + \left[p_h u'\left(w_1^1 + \epsilon''\right) - \frac{\overline{p}(1-p_h)}{1-\overline{p}} u'\left(w_0^1 - \frac{\overline{p}\epsilon''}{1-\overline{p}}\right)\right]\epsilon' \\ &= \left[\kappa^{-1} - I_h\left(\mathbf{a}' + (\kappa,1)\varepsilon'\right)\right] p_h u'\left(w_1' + \varepsilon'\right)\kappa\varepsilon \\ &\quad + \left[p_h u'\left(w_1^1 + \epsilon''\right) - \frac{\overline{p}(1-p_h)}{1-\overline{p}} u'\left(w_0^1 - \frac{\overline{p}\epsilon''}{1-\overline{p}}\right)\right]\epsilon' \end{aligned}$$

for some $(\varepsilon', \epsilon'') \in (0,\varepsilon) \times (0,\epsilon')$. Yet, $\kappa^{-1} > I_L(\mathbf{a}') > I_H(\mathbf{a}')$ and, letting $\Delta = \kappa^{-1} - I_L(\mathbf{a}')$, we may choose ε (and, subsequently, ε') sufficiently small to guarantee that $|I_h(\mathbf{a}' + (\kappa,1)\varepsilon') - I_h(\mathbf{a}')| < \Delta$

question (and, thus, of the corres ponding contract) is a standard result to be found in textbook derivations of the expected utility theorem. See, for instance, Step 3 of Proposition 6.B.3 in Mas-Collel A., Whinston M.D., and J.R. Green, *Microeconomic Theory*, Oxford University Press (1995).

for either h. But then

$$\begin{aligned} U_h\left(\mathbf{w}^2\right)-U_h\left(\mathbf{w}^1\right) &= \left[\Delta+I_L\left(\mathbf{a}'\right)-I_h\left(\mathbf{a}'+(\kappa,1)\,\varepsilon'\right)\right]p_h u'\left(w_1'+\varepsilon'\right)\kappa\varepsilon \\ &\quad+\left[p_h u'\left(w_1^1+\epsilon''\right)-\frac{\overline{p}\left(1-p_h\right)}{1-\overline{p}}u'\left(w_0^1-\frac{\overline{p}\epsilon''}{1-\overline{p}}\right)\right]\epsilon' \\ &> \left[p_h u'\left(w_1^1+\epsilon''\right)-\frac{\overline{p}\left(1-p_h\right)}{1-\overline{p}}u'\left(w_0^1-\frac{\overline{p}\epsilon''}{1-\overline{p}}\right)\right]\epsilon' \\ &> \left[p_h u'\left(w_1^1+\epsilon'\right)-\frac{\overline{p}\left(1-p_h\right)}{1-\overline{p}}u'\left(w_0^1-\frac{\overline{p}\epsilon'}{1-\overline{p}}\right)\right]\epsilon' \\ &\geq \left[p_L u'\left(w_1^1+\epsilon'\right)-\frac{\overline{p}\left(1-p_L\right)}{1-\overline{p}}u'\left(w_0^1-\frac{\overline{p}\epsilon'}{1-\overline{p}}\right)\right]\epsilon' \\ &= \left[U_L\left(\mathbf{w}\right)-U_L\left(\mathbf{w}^1\right)\right]\frac{\epsilon'}{\epsilon}>0 \end{aligned}$$

where the second inequality follows from the fact that $\epsilon''<\epsilon'$ while $u(\cdot)$ is strictly concave and the last one exploits that $p_H>p_L$.

Let now the contract $\mathbf{a}^2$ be offered in the presence of the RSW policy. As $\mathbf{a}^2\succ_L\mathbf{a}^1\sim_L\mathbf{a}_L^{**}\succ_L\mathbf{a}_H^{**}$ (the first two preferences by construction, the last one by the properties of the RSW allocation), the low-risk type is pulled away. If $\mathbf{a}^2\succ_H\mathbf{a}_H^{**}$, this is also the case for the high-risk type so that $\mathbf{a}^2$ becomes a pooling policy, a strictly profitable one. Otherwise, the deviant contract attracts only the low-risk type, delivering even higher expected profits. For since $p_L<\overline{p}<p_H$, an arbitrary contract gives $\Pi_L(\mathbf{a})\geq\Pi_M(\mathbf{a})\geq\Pi_H(\mathbf{a})$, with either inequality strict unless $\mathbf{a}=\mathbf{0}$. In either case, therefore, offering $\mathbf{a}^2$ is a strictly profitable deviation against the RSW policy.

"If." Notice first that there are no separating menus able to steal any risk-type away from her RSW contract since $\mathbf{a}_h^{**}$ maximizes h's welfare amongst the separating menus that are admissible in this game. Moreover, no contract is able to attract only one risk-type away from the RSW policy and avoid losses doing so. For suppose that $\mathbf{a}'$ is designed in this way with respect to the risk-type h. It ought to be then $\Pi_h(\mathbf{a}')\geq 0$ and $\mathbf{a}_{h'}^{**}\succsim_{h'}\mathbf{a}'\succ_h\mathbf{a}_h^{**}$ for $h'\neq h$. As, however, $\mathbf{a}_h^{**}\succsim_h\mathbf{a}_{h'}^{**}$, this would mean that the separating menu $\left\{\mathbf{a}_{h'}^{**},\mathbf{a}'\right\}$ Pareto-dominates the RSW allocation while satisfying the constraints of the efficiency problem the latter solves. Which is absurd.

Hence, credible challenges may come only from contracts intended to attract both types away for the RSW allocation. As such, though, they must be strictly-profitable as pooling policies. Yet, by hypothesis, there are no strictly-profitable pooling contracts that can attract the low-risk customers. For if $\mathbf{a}^1\in\mathbb{R}_+^2$ is s.t. $\Pi_M\left(\mathbf{a}^1\right)=\epsilon>0$ and $\mathbf{a}^1\succsim_L\mathbf{a}_L^{**}$, the contract $\mathbf{a}=\mathbf{a}^1-(1,-1)\,\epsilon$ is such that $\Pi_M(\mathbf{a})=\Pi_M\left(\mathbf{a}^1\right)-\epsilon=0$ and $\mathbf{a}\succ_h\mathbf{a}^1$ by either h (it offers more income in either state of the world). Yet, then we would have $\Pi_M(\mathbf{a})=0$ and $\mathbf{a}\succ_L\mathbf{a}_L^{**}$ which cannot be by hypothesis. ■

C.2 The Hellwig Game

C.2.1 Deviations above FO_M^*

We will show that the following is a sequential equilibrium scenario. Both the equilibrium and the deviant policies get withdrawn at stage 3. Being then indifferent between applying for either policy, an agent of risk-type h applies for the former with probability

$$\sigma_h \in [0,1): \quad 1 \leq \frac{1-\sigma_L}{1-\sigma_H} \leq \hat{\lambda}^* \qquad \text{where } \hat{\lambda}^* = \frac{1-\hat{p}^*}{\hat{p}^*} \tag{59}$$

To support this, consider a sequence of trembles $\left\{q_L^k, q_H^k\right\}_{k\in\mathbb{N}} \in (0, 1-\sigma_h)^2$ such that $\left(q_L^k, q_H^k\right) \to (0,0)$. The intended interpretation is that, along the sequence, an agent of risk-type h applies for the equilibrium and deviant contracts with probability $\sigma_h^k = \sigma_h + q_h^k$ and $1-\sigma_h^k$, respectively. This is depicted in Figure 8 where the firm's two available end-actions, to honor or withdraw its offer, are given as NW and W, respectively, while its conditional belief as a deviant at the upper node of its information set in by $\frac{\lambda\left(1-\sigma_L^k\right)}{\lambda\left(1-\sigma_L^k\right)+(1-\lambda)\left(1-\sigma_H^k\right)}$. Hence, throughout the deviant information set, the average accident probability is $\hat{p}_k = \frac{\lambda\left(1-\sigma_L^k\right)p_L+(1-\lambda)\left(1-\sigma_H^k\right)p_H}{\lambda\left(1-\sigma_L^k\right)+(1-\lambda)\left(1-\sigma_H^k\right)}$. Equivalently,

$$\frac{1-\hat{p}_k}{\hat{p}_k} = \frac{\hat{\lambda}_k \lambda (1-p_L) + (1-\lambda)(1-p_H)}{\hat{\lambda}_k \lambda p_L + (1-\lambda) p_H} \qquad \text{where } \hat{\lambda}_k = \frac{1-\sigma_L^k}{1-\sigma_H^k} \tag{60}$$

But $\frac{\mathrm{d}}{\mathrm{d}\hat{p}_k}\left(\frac{1-\hat{p}_k}{\hat{p}_k}\right) < 0 < \frac{\partial}{\partial\hat{\lambda}_k}\left(\frac{1-\hat{p}_k}{\hat{p}_k}\right) \ \forall \hat{p}_k, \hat{\lambda}_k \in \mathbb{R}_{++}$ while $\frac{1-\hat{p}_k}{\hat{p}_k} = \frac{1-\overline{p}}{\overline{p}}$ for $\hat{\lambda}_k = 1$. These expressions apply also for $\overline{p}_k$, the average accident probability on the equilibrium information set, once $\hat{\lambda}_k$ is replaced by $\overline{\lambda}_k = \sigma_L^k/\sigma_H^k$.

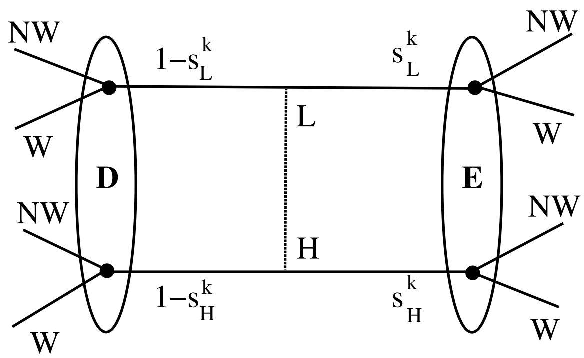

Figure 8: The Signalling Subgame

Let now $q_L^k < q_H^k \ \forall k$ and observe that $\sigma_L^k = \sigma_L + q_L^k < \sigma_L + q_H^k \leq \sigma_H + q_H^k = \sigma_H^k$, the second inequality due to (59). That is, $\overline{\lambda}_k < 1$ or $\overline{p}_k > \overline{p}$ along the sequence and, being on FO_M^*, the equilibrium contract is loss-making as a pooling poling and should get withdrawn at stage 3. If it is the RS policy, on the other hand, it cannot but break even and we may actually impose its withdrawal without loss of generality. Regarding the deviant policy, we have $\hat{\lambda} \equiv \lim_{k\to\infty} \hat{\lambda}_k = \frac{1-\sigma_L}{1-\sigma_H} \leq \hat{\lambda}^*$. Equivalently, $\hat{p} \equiv \lim_{k\to\infty} \hat{p}_k \geq \hat{p}^*$ so that, taking a subsequence if necessary, this policy will not make profits and should be withdrawn as well.

To complete the analysis, some additional observations are in order. It is immediate from the above analysis that the scenario in question cannot be supported if the inequality (59) is violated. For $\frac{1-\sigma_L}{1-\sigma_H} < 1$ and $\frac{1-\sigma_L}{1-\sigma_H} > \hat{\lambda}^*$ require, respectively, that the equilibrium and the deviant contract are honored at stage 3. It is equally immediate, however, that the scenario can be supported for

$\sigma_L = \sigma_H = 1$. One needs only to define $\sigma_h^k = 1 - r_h^k$, for vanishing trembles $\left\{r_L^k, r_H^k\right\}_{k\in\mathbb{N}} \in (0,1)^2$ such that $r_L^k > r_H^k$ $\forall k$.

Notice also that it is impossible to sustain a pooling equilibrium under a scenario in which the deviant policy is not withdrawn at stage 3. For this to happen, the deviant policy must be profitable. Being above FO_M^*, though, the maximum average accident probability that allows it to be profitable is $\hat{p}^* < \overline{p}$. Hence, the deviant is profitable only if $\hat{\lambda}^* > 1$. Equivalently, only if $\overline{\lambda} \equiv \lim_{k\to\infty} \overline{\lambda}_k < 1$ or $\lim_{k\to\infty} \overline{p}_k > \overline{p}$. Yet, the equilibrium contract is on FO_M^* and this inequality forces it to make losses as a pooling policy and precipitates its withdrawal at stage 3. Anticipating this at stage 2, all agents cannot but apply for the deviant policy. But then, $\hat{p} = \overline{p}$ and the deviation delivers losses.

Other sequential equilibria
It remains to consider the scenario in which the deviant policy is withdrawn at stage 3, the equilibrium one is not, and it is strictly dominant for either risk-type to apply for the latter at stage 2 with probability one. As sequential equilibrium, this can be supported by a sequence of trembles $\left\{\left(r_L^k, r_H^k\right)\right\}_{k\in\mathbb{N}} \in (0,1)^2$ such that $\left(r_L^k, r_H^k\right) \to (0,0)$ and $r_L^k/r_H^k \to 0$. Along the sequence now, an agent of risk-type h applies for the equilibrium and deviant contracts with probability $1 - r_h^k$ and r_h^k, respectively, and the preceding formulae apply with $\sigma_h = 1$ and $q_h^k = -r_h^k$ for either h. Hence, $\hat{\lambda}_k = r_L^k/r_H^k \to 0$ and $\lim_{k\to\infty} \frac{1-\hat{p}_k}{\hat{p}_k} = \frac{1-p_H}{p_H} \leq \frac{1-\hat{p}^*}{\hat{p}^*}$, the equality by (60) while the inequality due to the fact that $\Pi_H(\hat{\mathbf{a}}) \leq 0$ by (38) and, hence, $\hat{p}^* \leq p_H$. In other words, $\hat{p} \geq \hat{p}^*$ and, taking a subsequence if necessary, the deviant policy is not expected to be profitable and will be withdrawn at stage 3. Regarding the equilibrium policy, taking a further subsequence if necessary, $\hat{\lambda}_k < 1$ necessitates that $\overline{\lambda}_k > 1$ everywhere along the subsequence. Hence, $\lim_{k\to\infty} \overline{p}_k \geq \overline{p}$ and the equilibrium contract is expected to be profitable and be honored at stage 3. If it is the RS policy, honoring it at stage 3 can be imposed without loss of generality.

C.2.2 Deviations on/below FO_M^*

The argument of the preceding paragraph can be applied again to support the second sequential equilibrium scenario. Notice, moreover, that there exist deviations below the line FO_M^* against which this is the *only* equilibrium. This follows immediately from two observations.
First, against a deviant contract $\hat{\mathbf{a}}$ that lies below FO_M^*, no equilibrium scenario can have the equilibrium policy withdrawn at stage 3. For, in this case, as long as the deviant suppliers honor their policy, all agents of either risk-type cannot but apply for it at stage 2. Yet, the deviant contract lying below FO_M^*, it must be $\overline{p} < \hat{p}^*$ and, thus, $\Pi_M(\hat{\mathbf{a}}) > 0$. If the entire population of agents, therefore, applies for $\hat{\mathbf{a}}$, the strategy of offering it at stage 1 and honoring it at stage 3 is indeed a profitable deviation.
Second, there are contracts such as $\hat{\mathbf{a}}$ above which must be withdrawn at stage 3 in any sequential equilibrium scenario. To see this, recall first that, being susceptible to deviations such as $\hat{\mathbf{a}}$, the equilibrium policy $\mathbf{a}^*$ cannot be the Wilson one. Hence, we can always choose $\hat{\mathbf{a}}$ such that either

(i) $\hat{\mathbf{a}} \succ_h \mathbf{a}^*$ for either risk-type h or (ii) $\hat{\mathbf{a}} \succ_L \mathbf{a}^* \succ_H \hat{\mathbf{a}}$.[26] Suppose now that $\hat{\mathbf{a}}$ is honored at stage 3 on the equilibrium path. Anticipating this at stage 2, either (i) all agents of either risk-type or (ii) only the low-risk ones would find it strictly-dominant to apply for $\hat{\mathbf{a}}$. In each case, the beliefs of its suppliers is such that $\hat{p} \leq \overline{p} < \hat{p}^*$. As before, the strategy of offering $\hat{\mathbf{a}}$ at stage 1 and honoring it at stage 3 would be a profitable deviation.

When the equilibrium policy is the RS one, these arguments remain valid (recall the analysis in Section 4). The claim becomes now that, under no equilibrium scenario, $\mathbf{a}_L^*$ may be withdrawn at stage 3. If it is, honoring their own policy becomes optimal for the deviant suppliers. And against these strategies, the low-risk agents would find it optimal to apply for the deviant contract so that again $\hat{p} \leq \overline{p} < \hat{p}^*$.

C.2.3 All equilibria, but the Wilson policy, are unstable

The equilibria in question are the RS policy and the contracts on FO_M^* that Pareto-dominate it, but for the Wilson one. All of them are subject to potentially-profitable deviations that lie below FO_M^* and do satisfy one of the cases (i)-(ii) in the preceding section. And against these deviations, any sequential equilibrium requires that the deviant policy is withdrawn at stage 3, the equilibrium one is not, and all agents apply at stage 2 for the latter with probability one. Moreover, this will have to be supported by some vanishing sequence of trembles $\left\{r_L^k, r_H^k\right\}_{k\in\mathbb{N}} \in (0,1)^2$ along which an agent of risk-type h applies for the equilibrium and deviant policies with probability $\sigma_h^k = 1 - r_h^k$ and $1 - \sigma_h^k$, respectively (Section C.2.1).

Consider now a perturbation $\{\epsilon_h, \widetilde{r}_h\}_{h=H,L} \in (0,1)^4$. Under the equilibrium strategy profile in question, an agent of risk-type h applies for the equilibrium and deviant contracts in the perturbed game with probability $\widetilde{\sigma}_h^k = (1-\epsilon_h)\,\sigma_h^k + \epsilon_h \widetilde{r}_h$ and $1 - \widetilde{\sigma}_h^k$, respectively. To support this also on the perturbed game, the deviant contract has to be withdrawn at stage 3 at least in the limit.

As $r_H^k \to 0$, however, $\exists\,(n, k_n'') \in \mathbb{N}^2 \setminus \{(0,0)\} : r_H^k < [4n\,(1-\epsilon_H)]^{-1}\ \forall k > k_n''$. Letting, therefore, $k^* = \max\{k', k''\}$ while choosing $n \in (1,\infty)$ sufficiently large, $\epsilon_L\,(1-\widetilde{r}_L) = \hat{\lambda}^*/n$, and $\epsilon_H\,(1-\widetilde{r}_H) = (4n)^{-1}$, we get

$$\begin{aligned} 1-\widetilde{\sigma}_H^k &\equiv (1-\epsilon_H)\,r_H^k + \epsilon_H\,(1-\widetilde{r}_H) < \frac{1}{4n} + \frac{1}{4n} = \frac{1}{2n} = \frac{\epsilon_L\,(1-\widetilde{r}_L)}{2\hat{\lambda}^*} \\ &< \frac{(1-\epsilon_L)\,r_L^k + \epsilon_L\,(1-\widetilde{r}_L)}{2\hat{\lambda}^*} = \frac{1-\widetilde{\sigma}_L^k}{2\hat{\lambda}^*} \qquad \forall k > k^* \end{aligned}$$

That is, $\lim_{k\to\infty} \hat{\lambda}_k \geq 2\hat{\lambda}^* > \hat{\lambda}^*$ or $\hat{p} < \hat{p}^*$. Hence, the deviant policy is believed to be profit-making and will not be withdrawn at stage 3, a contradiction of the desired sequential equilibrium scenario. To complete the argument, observe that the sequence of trembles we considered was arbitrary. It follows, therefore, that the desired scenario cannot be supported as sequential equilibrium. Notice

[26] Examples of $\hat{\mathbf{a}}$ were constructed in Section 1. Graphically, case (i) is depicted by points in the interiors of the shaded area in the right-hand side diagram of Figure 1 or of the upper-left shaded area in the right-hand side of Figure 2. Case (ii) corresponds to points in the interior of the shaded area in the left-hand side of Figure 1 that lies below the line FO_M^*.

also that, being free to choose the randomized profile $\{\widetilde{r}_L, \widetilde{r}_H\}$ arbitrarily, our construction allows the mixture $\{\epsilon_L, \epsilon_H\}$ to be arbitrarily close to the original game.

C.2.4 The Wilson policy is stable

Against the Wilson policy, a potentially-profitable deviation cannot but lie above the line FO_M^*. And against this, the sequential equilibrium scenario has both contracts withdrawn at stage 3. Being then indifferent between applying for either policy, an agent of risk-type h applies for the Wilson contract with probability $\sigma_h \in [0,1)$ such that (59) is satisfied. We will show that this scenario can be supported also under perturbations as long as we restrict (59) to

$$1 < \frac{1-\sigma_L}{1-\sigma_H} < \frac{1+\hat{\lambda}^*}{2} \tag{61}$$

Consider the sequence of trembles $\left\{q_L^k, q_H^k\right\}_{k\in\mathbb{N}} \in (0, 1-\sigma_L) \times (0, 1-\sigma_H)$, with $\lim_{k\to\infty}\left(q_L^k, q_H^k\right) = (0,0)$ and $q_H^k > q_L^k$ for all k, along which an agent of risk-type h applies for the Wilson and the deviant contracts with probability $\sigma_h^k = \sigma_h + q_h^k$ and $1-\sigma_h^k$, respectively. It suffices to establish that such trembles continue to support the given scenario even when the game undergoes an arbitrary perturbation. To this end, consider a randomized profile $(\widetilde{q}_L, \widetilde{q}_H) \in (0,1)^2$ and the perturbation that arises under this profile and a mixture $(\epsilon_L, \epsilon_H) \in (0,\epsilon)^2$ with $\epsilon < \min\left\{\frac{\sigma_H - \sigma_L}{2+\sigma_H}, 1-\sigma_H - \frac{1+\sigma_H-2\sigma_L}{\hat{\lambda}^*}\right\}$. Under this perturbation, and along the given sequence of trembles, an agent of risk-type h applies for the Wilson and the deviant contract with probability $\widetilde{\sigma}_h^k = (1-\epsilon_h)\,\sigma_h^k + \epsilon_h \widetilde{q}_h$ and $1-\widetilde{\sigma}_h^k$, respectively.

To replicate the sequential equilibrium argument, it is enough that the two required conditions, for the Wilson and the deviant contract to be both withdrawn, are met. The former requirement is indeed satisfied since

$$\begin{aligned}
\widetilde{\sigma}_H^k = (1-\epsilon_H)\left(\sigma_H + q_H^k\right) + \epsilon_H\widetilde{q}_H \quad &> \quad (1-\epsilon)\left(\sigma_H + q_H^k\right) + \epsilon_H \widetilde{q}_H \\
&> \quad (1-\epsilon)\left(\sigma_H + q_H^k\right) \\
&> \quad \sigma_L + q_L^k + \epsilon \\
&> \quad (1-\epsilon_L)\left(\sigma_L + q_L^k\right) + \epsilon\widetilde{q}_L > (1-\epsilon_L)\left(\sigma_L + q_L^k\right) + \epsilon_L\widetilde{q}_L = \widetilde{\sigma}_L^k
\end{aligned}$$

Here, the first and last inequalities follow from the fact that $\epsilon_L, \epsilon_H < \epsilon$ whereas the second and the one before the last inequality use the boundedness conditions $\epsilon_H\widetilde{q}_H > 0$ and $\widetilde{q}_L < 1$, respectively. The remaining inequality uses the first upper bound of ϵ. Specifically, as $q_L^k < q_H^k < 1$, we have $\epsilon < \frac{\sigma_H-\sigma_L}{2+\sigma_H} < \frac{\sigma_H-\sigma_L}{1+\sigma_H+q_H^k} < \frac{\sigma_H+q_H^k-\left(\sigma_L+q_L^k\right)}{1+\sigma_H+q_H^k}$.

It remains to verify that the other required condition, for the deviant contract to be withdrawn, is also met. To this end, observe that, given our assumptions, we have

$$\begin{aligned}
\epsilon_L\left(\sigma_L^k - \widetilde{q}_L\right) = \epsilon_L\left(\sigma_L + q_L^k - \widetilde{q}_L\right) \quad &< \quad \epsilon_L\left(\sigma_L + q_L^k\right) < \epsilon\left(\sigma_L + q_L^k\right) < \epsilon\left(\sigma_H + q_H^k\right) \\
&= \quad \epsilon\sigma_H^k \\
&< \quad \epsilon\left(1+\sigma_H^k\right) < \sigma_H^k - \sigma_L^k
\end{aligned} \tag{62}$$

where the third inequality uses that $\sigma_H > \sigma_L$ - by the left-hand side of (61) - and $q_H^k > q_L^k$ while the last inequality follows from $\epsilon < \frac{\sigma_H^k - \sigma_L^k}{1+\sigma_H^k}$ (as established at the very end of the preceding paragraph). In addition,

$$1 - \sigma_H^k + \epsilon_H \left(\sigma_H^k - \widetilde{q}_H\right) > 1 - \sigma_H^k - \epsilon_H \widetilde{q}_H > 1 - \sigma_H^k - \epsilon \widetilde{q}_H > 1 - \sigma_H^k - \epsilon \tag{63}$$

the first two inequalities following from $\epsilon_H s_H > 0$ and $\epsilon_H < \epsilon$, respectively, while the last one using the fact that $\widetilde{q}_H < 1$. Regarding the last quantity above, moreover, we have

$$1 - \sigma_H^k - \epsilon = 1 - \left(\sigma_H + q_H^k\right) - \epsilon > 1 - \sigma_H - \frac{1-\sigma_L}{\hat{\lambda}^*} - \epsilon > 1 - \sigma_H - \frac{1+\sigma_H - 2\sigma_L}{\hat{\lambda}^*} - \epsilon > 0 \tag{64}$$

Here, the first inequality obtains because, choosing a subsequence of $\left\{q_H^k\right\}_{k\in\mathbb{N}}$ if necessary, we can guarantee that $q_H^k < \frac{1-\sigma_L}{\hat{\lambda}^*}$ for all k. The second inequality, on the other hand, follows from $\sigma_H > \sigma_L$ while the last one uses the second upper bound of ϵ. Putting now the results (62)-(64) together, we get

$$\frac{1-\widetilde{\sigma}_L^k}{1-\widetilde{\sigma}_H^k} = \frac{1-\sigma_L^k + \epsilon_L\left(\sigma_L^k - \widetilde{q}_L\right)}{1-\sigma_H^k + \epsilon_H\left(\sigma_H^k - \widetilde{q}_H\right)} < \frac{1+\sigma_H^k - 2\sigma_L^k}{1-\sigma_H^k - \epsilon}$$

For the deviant contract to be withdrawn, it suffices that the last ratio above doesn't exceed $\hat{\lambda}^\star$. Which is indeed the case, at least along a subsequence of $\left\{q_L^k, q_H^k\right\}_{k\in\mathbb{N}}$, since

$$\lim_{k\to\infty} \frac{1+\sigma_H^k - 2\sigma_L^k}{1-\sigma_H^k - \epsilon} = \frac{1+\sigma_H - 2\sigma_L}{1-\sigma_H - \epsilon} < \hat{\lambda}^*$$

by the second upper bound of ϵ. Observe finally that, as $\hat{\lambda}^* > 1$, (61) is indeed a restriction of (59). Moreover, either upper bound of ϵ is well-defined: they are both positive since $\sigma_H > \sigma_L$ while $1 + \sigma_H - 2\sigma_L < \hat{\lambda}^*(1-\sigma_H)$ is in fact the right-hand side of (61). □

The unstable parts

Let now $\frac{1-\sigma_L}{1-\sigma_H} \in \left(\frac{1+\hat{\lambda}^*}{2}, \hat{\lambda}^*\right]$. In this case, $\frac{1-\sigma_L}{1-\sigma_H} = \frac{1+\hat{\lambda}^*+\delta}{2}$ for some $\delta \in \left(0, \frac{\hat{\lambda}^*-1}{2}\right]$, and we may consider the trembles $\widetilde{q}_L = \sigma_H$, $\widetilde{q}_H = \frac{1}{q}\left[(1+q)\sigma_H - \sigma_L\right]$ for some $q \in \left(\frac{\sigma_H - \sigma_L}{1-\sigma_H}, +\infty\right)$, and the mixtures $\epsilon_L = 1/n$ and $\epsilon_H = q/n^2$ for some $n \in \mathbb{R}_{++}$.[27] Then,

$$\begin{aligned}\epsilon_L\left(\sigma_L - \widetilde{q}_L\right) - \hat{\lambda}^*\epsilon_H\left(\sigma_H - \widetilde{q}_H\right) &= \left(\frac{\hat{\lambda}^*}{n^2} - \frac{1}{n}\right)\left(\sigma_H - \sigma_L\right) \\ &> \sigma_H - \sigma_L = \left(\hat{\lambda}^* + \delta\right)\left(1-\sigma_H\right) - \left(1-\sigma_L\right)\end{aligned}$$

where the second equality is just another way of writing our initial condition while the inequality is due to the fact that $\hat{\lambda}^* > 1$ and, thus, $\hat{\lambda}^* > \frac{n+1}{n}$ as long as n is taken to be sufficiently large. Under this perturbation, therefore,

$$\lim_{k\to\infty} \frac{1-\widetilde{\sigma}_L^k}{1-\widetilde{\sigma}_H^k} = \frac{1-\sigma_L + \epsilon_L\left(\sigma_L - \widetilde{q}_L\right)}{1-\sigma_H + \epsilon_H\left(\sigma_H - \widetilde{q}_H\right)} > \hat{\lambda}^* + \frac{\delta\left(1-\sigma_H\right)}{1-\sigma_H + \epsilon_H\left(\sigma_H - \widetilde{q}_H\right)} > \hat{\lambda}^*$$

[27] Here, the lower bound on q is introduced to ensure that $\widetilde{q}_H < 1$.

which is the desired contradiction: at least along a subsequence, the deviant policy is believed to be strictly profitable and will not be withdrawn at stage 3.

With respect to the second equilibrium scenario given in Section C.2.1, recall that it had the deviant policy withdrawn at stage 3, the equilibrium policy honored, and all agents apply at stage 2 for the latter with probability one. We have already established, though, in Section C.2.3 that this is not stable. That argument applies also for the case in which the first scenario above obtains with either risk-type applying for the Wilson contract with probability one.

D Equilibrium with cross-subsidization

Proof of Proposition 5

It will be instructive to distinguish the case $\mu \in [0, \lambda)$ from the one in which $\mu \in [\lambda, 1)$.

Case I. $\mu \in [\lambda, 1)$

Since $\mu < 1$, let $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ be the IIE(μ) optimum for some $\mu^0 \in (\mu, 1]$ and observe the following

1. $U_L\left(\mathbf{w}_L^0\right) > U_L\left(\mathbf{w}_L^*\right)$ but $U_H\left(\mathbf{w}_H^0\right) < U_H\left(\mathbf{w}_H^*\right)$.

This is because, the mapping from the parameter μ to the IIE optimum being bijective (Claim 1), the strict version of Lemma 6 applies; hence, moving to the new allocation must be strictly beneficial [resp. detrimental] for the low-risk [resp. high-risk] type given that $\mu^0 > \mu$.

2. $U_L\left(\mathbf{w}_L^0\right) > U_L\left(\mathbf{w}_H^0\right)$ while $U_H\left(\mathbf{w}_L^0\right) = U_H\left(\mathbf{w}_H^0\right)$ and $U_L\left(\mathbf{w}_L^*\right) \geq U_L\left(\mathbf{w}_H^*\right)$ while $U_H\left(\mathbf{w}_L^*\right) = U_H\left(\mathbf{w}_H^*\right)$.

This is due to either allocation being IIE with the respective parameters satisfying $\mu^0 > \lambda$ and $\mu \geq \lambda$ (Claim 1).

3. $\mathbf{a}_L^0 = \mathbf{a}_L^* - (\epsilon_{0L}, \epsilon_{1L})$ and $\mathbf{a}_H^0 = \mathbf{a}_H^* - (-1, 1)\,\epsilon_H$ for some $(\epsilon_{0L}, \epsilon_{1L}, \epsilon_H) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++}$.

For the first part, just let $\epsilon_{sL} = a_{sL}^* - a_{sL}^0$ for $s = 0, 1$ and observe that, the corresponding wealth allocation being $\mathbf{w}_L^0 = (w_{0L}^* + \epsilon_{0L}, w_{1L}^* - \epsilon_{1L})$, the only restriction $U_L\left(\mathbf{w}_L^0\right) > U_L\left(\mathbf{w}_L^*\right)$ imposes is that it cannot be $(\epsilon_{0L}, -\epsilon_{1L}) \leq \mathbf{0}$. Regarding the second part of the present observation, for the same reason which supports the preceding observation, the high-risk agents ought to be fully-insured in either allocation. That is, it must be $w_{0H}^0 = w_{1H}^0$ and $w_{0H}^* = w_{1H}^*$, or $a_{0H}^* = d - a_{1H}^*$ and $a_{0H}^0 = d - a_{1H}^0$ for the corresponding contracts. Yet, the last two equations are equivalent to $\mathbf{a}_H^0 = \mathbf{a}_H^* - (-1, 1)\,\epsilon_H$ for some $\epsilon_H \in \mathbb{R}$. And since $\mathbf{a}_H^0 \neq \mathbf{a}_H^*$, it cannot but be $\epsilon_H \neq 0$. In fact, it must be $\epsilon_H > 0$. For if $\epsilon_H < 0$, we may write $\mathbf{a}_H^0 = \mathbf{a}_H^* + (-1, 1)\,|\epsilon_H|$ and appeal to Lemma 2 for some $\widetilde{\epsilon}_H \in (0, |\epsilon_H|)$ such that

$$\begin{aligned}
U_H\left(\mathbf{w}_H^0\right) - U_H\left(\mathbf{w}_H^*\right) &= \left[p_H u'\left(w_{1H}^* + \widetilde{\epsilon}_H\right) + (1 - p_H)\, u'\left(w_{0H}^* + \widetilde{\epsilon}_H\right)\right] |\epsilon_H| \\
&= u'\left(w_{0H}^* + \widetilde{\epsilon}_H\right) |\epsilon_H| > 0
\end{aligned}$$

an absurd conclusion, however, in light of observation I.

4. $\lambda\left[(1 - p_L)\,\epsilon_{0L} - p_L \epsilon_{1L}\right] = (1 - \lambda)\,\epsilon_H$

This follow immediately from the preceding observation and the fact that, either allocation being IIE, it must be $\lambda \Pi_L\left(\mathbf{a}_L^*\right) + (1 - \lambda)\,\Pi_H\left(\mathbf{a}_H^*\right) = 0 = \lambda \Pi_L\left(\mathbf{a}_L^0\right) + (1 - \lambda)\,\Pi_H\left(\mathbf{a}_H^0\right)$. Intuitively, moving from the original to new allocation, the gain by making smaller losses from the high-risk type should be exactly offset by the fall in the profit collected from the low-risk one.

5. $U_L\left(\mathbf{w}_H^0\right) < U_L\left(\mathbf{w}_H^*\right)$.

This is because, since $\mathbf{a}_H^* = \mathbf{a}_H^0 + (-1,1)\,\epsilon_H$ with $\epsilon_H > 0$, Lemma 2 guarantees the existence of some $\widetilde{\epsilon}_L \in (0, \epsilon_H)$ such that

$$\begin{aligned} U_L\left(\mathbf{w}_H^*\right) - U_L\left(\mathbf{w}_H^0\right) &= \left[p_L u'\left(w_{1H}^0 + \widetilde{\epsilon}_L\right) + (1-p_L)\, u'\left(w_{0H}^0 + \widetilde{\epsilon}_L\right)\right]\epsilon_H \\ &= u'\left(w_{0H}^0 + \widetilde{\epsilon}_L\right)\epsilon_H > 0 \end{aligned}$$

Putting now our observations together, there exists a menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$, given by $\mathbf{a}_L^0 = \mathbf{a}_L^* - (\epsilon_{0L}, \epsilon_{1L})$ and $\mathbf{a}_H^0 = \mathbf{a}_H^* - (-1,1)\,\epsilon_H$ for some $(\epsilon_{0L}, \epsilon_{1L}, \epsilon_H) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_{++}$ with $(\epsilon_{0L}, \epsilon_{1L}) \leq \mathbf{0}$ ruled out, which is IIE(μ^0) optimal for some $\mu^0 \in (\mu, 1]$ and satisfies

$$\begin{aligned} U_L\left(\mathbf{w}_H^0\right) < U_L\left(\mathbf{w}_H^*\right) \leq U_L\left(\mathbf{w}_L^*\right) &< U_L\left(\mathbf{w}_L^0\right) \\ U_H\left(\mathbf{w}_L^0\right) = U_H\left(\mathbf{w}_H^0\right) &< U_H\left(\mathbf{w}_H^*\right) \\ \lambda\left[(1-p_L)\,\epsilon_{0L} - p_L\epsilon_{1L}\right] &= (1-\lambda)\,\epsilon_H \end{aligned}$$

Let now $\Delta = U_L\left(\mathbf{w}_L^0\right) - U_L\left(\mathbf{w}_L^*\right)$ and consider the contract $\hat{\mathbf{a}}_L = \mathbf{a}_L^0 + (1,\kappa)\,\epsilon$ for some $\kappa \in \left(0, I_H\left(\mathbf{a}_L^0\right)\right)$ and $\epsilon > 0$. Let also $\Delta_h = I_h\left(\mathbf{a}_L^0\right) - \kappa$ for $h = L, H$. Using Lemma 2 once again, the corresponding wealth allocations give

$$\begin{aligned} U_h\left(\hat{\mathbf{w}}_L\right) - U_h\left(\mathbf{w}_L^0\right) &= \left[p_h u'\left(w_{1L}^0 + \hat{\epsilon}_h\right) - \kappa^{-1}(1-p_h)\, u'\left(w_{0L}^0 - \kappa^{-1}\hat{\epsilon}_h\right)\right]\kappa\epsilon \\ &= \left[\kappa - \frac{(1-p_h)\, u'\left(w_{0L}^0 - \kappa^{-1}\hat{\epsilon}_h\right)}{p_h u'\left(w_{1L}^0 + \hat{\epsilon}_h\right)}\right] p_h u'\left(w_{1L}^0 + \hat{\epsilon}_h\right)\epsilon \\ &= \left[\kappa - I_h\left(\mathbf{a}_L^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_{1L}^0 + \hat{\epsilon}_h\right)\epsilon \\ &= \left[I_h\left(\mathbf{a}_L^0\right) - \Delta_h - I_h\left(\mathbf{a}_L^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_{1L}^0 + \hat{\epsilon}_h\right)\epsilon \end{aligned}$$

for some $\hat{\epsilon}_h \in (0, \kappa\epsilon)$. Yet, the function $I_h(\cdot)$ is continuous and $\lim_{\hat{\epsilon}_h \to 0} I_h\left(\mathbf{a}_L^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) = I_h\left(\mathbf{a}_L^0\right)$. For small enough ϵ (and, subsequently, $\hat{\epsilon}_h$), therefore, $|I_h\left(\mathbf{a}_L^0 + \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) - I_h\left(\mathbf{a}_L^0\right)| < \min\left\{\Delta_L, \Delta_H\right\}$ and, consequently, $U_h\left(\hat{\mathbf{w}}_L\right) < U_h\left(\mathbf{w}_L^0\right)$ for either h. With respect to the low-risk type, however, observe that the last quantity above vanishes as $\epsilon \to 0$. Consequently, for sufficiently small ϵ, we can guarantee that $U_L\left(\hat{\mathbf{w}}_L\right) - U_L\left(\mathbf{w}_L^0\right) \in (-\Delta, 0)$ or $U_L\left(\hat{\mathbf{w}}_L\right) > U_L\left(\mathbf{w}_L^*\right)$. But then, we have $U_L\left(\mathbf{w}_L\right) > U_L\left(\mathbf{w}_L^*\right) > U_L\left(\mathbf{w}_H^*\right) > U_L\left(\mathbf{w}_H^0\right)$ and $U_H\left(\mathbf{w}_L\right) < U_H\left(\mathbf{w}_L^0\right) = U_H\left(\mathbf{w}_H^0\right) < U_H\left(\mathbf{w}_H^*\right)$.

Setting, thus, $\left\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\right\} = \left\{\hat{\mathbf{a}}_L, \mathbf{a}_H^*\right\}$ establishes parts (i) and (iii) of the lemma for the case in which $\mu \in [\lambda, 1)$. It remains to show that the latter menu is expected to be strictly profitable when accepted by both types. To this end, notice that, compared to $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ which is IIE and, thus, breaks even across the two-types, the new menu makes exactly the same expected loss from the high-risk agents. It expects, however, more profits from the low-risk since

$$\begin{aligned} \Pi_L\left(\hat{\mathbf{a}}_L\right) - \Pi_L\left(\mathbf{a}_L^0\right) = \left(\frac{1-p_L}{p_L} - \kappa\right) p_L\epsilon &> \left[\frac{1-p_L}{p_L} - I_L\left(\mathbf{a}_L^0\right)\right] p_L\epsilon \\ &= (1-p_L)\left[1 - \frac{u'\left(w_{0L}^0\right)}{u'\left(w_{1L}^0\right)}\right]\epsilon > 0 \end{aligned}$$

Here, the first inequality is since $\kappa < I_H\left(\mathbf{a}_L^0\right) < I_L\left(\mathbf{a}_L^0\right)$. The second inequality follows from risk-aversion and the fact that, the allocation $\left\{\mathbf{w}_L^0, \mathbf{w}_H^0\right\}$ being IIE(μ^0) with $\mu^0 > \lambda$, it must give $w_{0L}^0 > w_{1L}^0$ (Claim 1).

Case II. $\mu \in [0, \lambda)$

Letting now $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ be the IIE(μ^0) optimum for some $\mu^0 \in (\mu, \lambda]$, it is trivial to adopt the steps of reasoning we followed in the previous case to conclude now that $\mathbf{a}_L^0 = \mathbf{a}_L^* + (-1, 1)\,\epsilon_L$ and $\mathbf{a}_H^0 = \mathbf{a}_H^* - (\epsilon_{0H}, \epsilon_{1H})$ with $\epsilon_L > 0$ and $\epsilon_{0H}, \epsilon_{1H} \in \mathbb{R}$ but $(\epsilon_{0H}, -\epsilon_{1H}) \geq \mathbf{0}$ not possible. Moreover, the corresponding wealth allocations $\mathbf{w}_L^0 = \mathbf{w}_L^* + (1, -1)\,\epsilon_L$ and $\mathbf{w}_H^0 = \mathbf{w}_H^* + (\epsilon_{0H}, -\epsilon_{1H})$ give now

$$\begin{aligned} U_L\left(\mathbf{w}_L^*\right) &< U_L\left(\mathbf{w}_L^0\right) = U_L\left(\mathbf{w}_H^0\right) \\ U_H\left(\mathbf{w}_L^0\right) &< U_H\left(\mathbf{w}_H^0\right) < U_H\left(\mathbf{w}_H^*\right) \\ \lambda\epsilon_L &= (1-\lambda)\left[(1-p_H)\,\epsilon_{0H} - p_H\epsilon_{1H}\right] \end{aligned}$$

Here, $U_L\left(\mathbf{w}_H^0\right) = U_L\left(\mathbf{w}_L^0\right)$ and $U_H\left(\mathbf{w}_H^0\right) > U_H\left(\mathbf{w}_L^0\right)$ are because of IIE(μ^0) optimality with $\mu^0 < \lambda$ (Claim 1) while $U_L\left(\mathbf{w}_L^0\right) > U_L\left(\mathbf{w}_L^*\right)$ and $U_H\left(\mathbf{w}_H^0\right) < U_H\left(\mathbf{w}_H^*\right)$ are by construction.

Let now $\Delta = U_H\left(\mathbf{w}_H^0\right) - U_H\left(\mathbf{w}_L^0\right)$ and consider the contract $\mathbf{a}_H = \mathbf{a}_H^0 - (1, \kappa)\,\epsilon$ for some $\kappa \in \left(I_L\left(\mathbf{a}_H^0\right), +\infty\right)$ and $\epsilon > 0$. Let also $\Delta_h = \kappa - I_h\left(\mathbf{a}_H^0\right)$ for $h = L, H$. By Lemma 2 again, it must be

$$\begin{aligned} U_h\left(\mathbf{w}_H^0\right) - U_h\left(\mathbf{w}_H\right) &= \left[p_h u'\left(w_{1H}^0 - \hat{\epsilon}_h\right) - \kappa^{-1}(1-p_h)\,u'\left(w_{0H}^0 + \kappa^{-1}\hat{\epsilon}_h\right)\right]\kappa\epsilon \\ &= \left[\kappa - I_h\left(\mathbf{a}_H^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_{1H}^0 - \hat{\epsilon}_h\right)\epsilon \\ &= \left[I_h\left(\mathbf{a}_H^0\right) + \Delta_h - I_h\left(\mathbf{a}_H^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right)\right] p_h u'\left(w_{1H}^0 - \hat{\epsilon}_h\right)\epsilon > 0 \end{aligned}$$

for some $\hat{\epsilon}_h \in (0, \kappa\epsilon)$ and sufficiently small ϵ to ensure that $|I_h\left(\mathbf{a}_H^0 - \left(\kappa^{-1}, 1\right)\hat{\epsilon}_h\right) - I_h\left(\mathbf{a}_H^0\right)| < \min\{\Delta_L, \Delta_H\}$ for either h. Regarding the high-risk type, though, observe that, for sufficiently small ϵ, we can also guarantee that $U_H\left(\mathbf{w}_H\right) - U_H\left(\mathbf{w}_H^0\right) \in (-\Delta, 0)$; consequently, $U_H\left(\mathbf{w}_H\right) > U_H\left(\mathbf{w}_L^0\right)$. Clearly, $U_L\left(\mathbf{w}_L^0\right) = U_L\left(\mathbf{w}_H^0\right) > U_L\left(\mathbf{w}_H\right), U_L\left(\mathbf{w}_L^*\right)$ and $U_H\left(\mathbf{w}_L^0\right) < U_H\left(\mathbf{w}_H\right) < U_H\left(\mathbf{w}_H^0\right) < U_H\left(\mathbf{w}_H^*\right)$. Setting, thus, $\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\} = \left\{\mathbf{a}_L^0, \hat{\mathbf{a}}_H\right\}$ establishes parts (i) and (iii) of the lemma in this case. Part (ii), on the other hand, follows from the fact the menu in question makes exactly the same expected profit as the IIE menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ from the low-risk agents but smaller losses from the high-risk. Indeed,

$$\begin{aligned} \Pi_H\left(\hat{\mathbf{a}}_H\right) - \Pi_H\left(\mathbf{a}_H^0\right) = \left(\kappa - \frac{1-p_H}{p_H}\right) p_H\epsilon &> \left(I_H\left(\mathbf{a}_H^0\right) - \frac{1-p_H}{p_H}\right) p_H\epsilon \\ &= (1-p_H)\left[\frac{u'\left(w_{0H}^0\right)}{u'\left(w_{1H}^0\right)} - 1\right]\epsilon > 0 \end{aligned}$$

where the first inequality follows from the choice of κ whereas the last one obtains because the allocation $\left\{\mathbf{w}_L^0, \mathbf{w}_H^0\right\}$, being IIE($\mu^0$) with $\mu^0 < \lambda$, over-insures the high-risk agents (Claim 1). ■

Proof of Proposition 6

In what follows, let $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ denote the IIE(1) optimum with $\left\{\mathbf{w}_L^1, \mathbf{w}_H^1\right\}$ the corresponding wealth allocation. Being feasible, the menu $\{\mathbf{a}_L^*, \mathbf{a}_H^*\}$ satisfies the constraints of the IIE(1) problem. For the corresponding wealth allocation, therefore, we ought to have: $U_L\left(\mathbf{w}_L^1\right) > U_L\left(\mathbf{w}_L^*\right)$. Given this, we may proceed as in the $\mu \in [\lambda, 1)$ case in the proof of Lemma 5, albeit with $\left\{\mathbf{a}_L^1, \mathbf{a}_H^1\right\}$ as our starting menu.

Reasoning in the same way as there, it is trivial to verify that, for small enough $\epsilon_H > 0$, there exists a menu $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$, defined by $\mathbf{a}_H^0 = \mathbf{a}_H^1 - (1,-1)\,\epsilon_H$ and $\mathbf{a}_L^0 = \mathbf{a}_L^1 + (\epsilon_{0L}, \epsilon_{1L})$ with $\epsilon_{0L}, \epsilon_{1L} \in \mathbb{R}$ but $(\epsilon_{0L}, -\epsilon_{1L}) \leq \mathbf{0}$ not possible, which is optimal for the IIE(μ^0) problem for some $\mu^0 \in (\lambda, 1)$ and whose wealth allocation, $\mathbf{w}_L^0 = \left(w_{0L}^1 - \epsilon_{0L}, w_{1L}^1 + \epsilon_{1L}\right)$ and $\mathbf{w}_H^0 = \mathbf{w}_H^1 + (1,1)\,\epsilon_H$, satisfies

$$
\begin{aligned}
U_L\left(\mathbf{w}_H^0\right) < U_L\left(\mathbf{w}_L^0\right) &< U_L\left(\mathbf{w}_L^1\right) \\
U_L\left(\mathbf{w}_H^*\right) \leq U_L\left(\mathbf{w}_L^*\right) &< U_L\left(\mathbf{w}_L^0\right) \\
U_H\left(\mathbf{w}_L^0\right) &= U_H\left(\mathbf{w}_H^0\right) \\
\lambda\left[(1-p_L)\,\epsilon_{0L} - p_L\epsilon_{1L}\right] &= (1-\lambda)\,\epsilon_H
\end{aligned}
$$

Regarding these relations, $\mu^0 > \lambda$ is ensured by the continuity of the IIE problem in the parameter μ: we can choose μ^0 to be arbitrarily close to 1 so as to guarantee that $1 - \mu^0 < 1 - \lambda$. Given $\mu^0 > \lambda$ then, the incentive constraint of the low-risk [resp. high-risk] type does not [resp.does] bind at the optimum. That is, $U_L\left(\mathbf{w}_H^0\right) < U_L\left(\mathbf{w}_L^0\right)$ and $U_H\left(\mathbf{w}_H^0\right) = U_H\left(\mathbf{w}_L^0\right)$. The remaining relations are by construction or immediate, apart from $U_L\left(\mathbf{w}_L^*\right) < U_L\left(\mathbf{w}_L^0\right)$.

To verify the latter, let $\Delta_1 = U_L\left(\mathbf{w}_L^1\right) - U_L\left(\mathbf{w}_L^*\right)$ and notice that, as a mapping $V : [0,1] \mapsto \mathbb{R}$, the value function of the IIE(μ) problem is convex on $[0,1]$ (Lemma 3), and, thus, continuous at any $\mu \in (0,1)$ by a well-known result (see, for instance, Ch.6 Th. 2.14 in [7]). Observe also that, since $\mu^0 > \lambda$, it must be $U_L\left(\mathbf{w}_L^0\right) \neq U_H\left(\mathbf{w}_H^0\right)$. Otherwise, we would have $U_L\left(\mathbf{w}_L^0\right) = U_H\left(\mathbf{w}_H^0\right) = U_H\left(\mathbf{w}_L^0\right)$ and $U_L\left(\mathbf{w}_L^0\right) = U_H\left(\mathbf{w}_L^0\right)$ cannot be unless $\mathbf{w}_L^0$ provides full insurance, which is absurd since $\mu^0 \neq \lambda$. Given these two observations, taking μ^0 sufficiently close to 1 guarantees that

$$
\max\left\{\left(1-\mu^0\right)\left|U_L\left(\mathbf{w}_L^0\right) - U_H\left(\mathbf{w}_H^0\right)\right|,\ \left|V(1) - V\left(\mu^0\right)\right|\right\} < \Delta_1/2
$$

But then we have indeed

$$
\begin{aligned}
U_L\left(\mathbf{w}_L^0\right) &> \mu^0 U_L\left(\mathbf{w}_L^0\right) + \left(1-\mu^0\right) U_H\left(\mathbf{w}_H^0\right) - \frac{\Delta_1}{2} \\
&= V\left(\mu^0\right) - \frac{\Delta_1}{2} > V(1) - \Delta_1 = U_L\left(\mathbf{w}_L^1\right) - \Delta_1 = U_L\left(\mathbf{w}_L^*\right)
\end{aligned}
$$

Define now $\Delta = U_L\left(\mathbf{w}_L^0\right) - \max\left\{U_L\left(\mathbf{w}_H^0\right), U_L\left(\mathbf{w}_L^*\right)\right\}$ and proceed to construct the contract $\hat{\mathbf{a}}_L$ exactly as in the $\mu \in [\lambda, 1)$ case in the proof of Lemma 5. Here, $U_L\left(\hat{\mathbf{w}}_L\right) - U_L\left(\mathbf{w}_L^0\right) \in (-\Delta, 0)$ implies that $U_L\left(\hat{\mathbf{a}}_L\right) > \max\left\{U_L\left(\mathbf{w}_H^0\right), U_L\left(\mathbf{w}_L^*\right)\right\}$. Setting $\left\{\hat{\mathbf{a}}_L, \hat{\mathbf{a}}_H\right\} = \left\{\hat{\mathbf{a}}_L, \mathbf{a}_H^0\right\}$ gives a separating menu ($\hat{\mathbf{a}}_L \succ_L \mathbf{a}_H^0 \sim_H \mathbf{a}_L^0 \succ_H \hat{\mathbf{a}}_L$) such that $\hat{\mathbf{a}}_L \succ_L \mathbf{a}_L^*$. Moreover, compared to $\left\{\mathbf{a}_L^0, \mathbf{a}_H^0\right\}$ which is IIE and, thus, breaks even across the two-types, the new menu expects exactly the same loss from the high-risk agents but more profit from the low-risk ones. ■