Optimal Mechanism under Adverse Selection: The Canonical Insurance Problem

Theodoros M. Diasakos∗ Kostas Koufopoulos†

September 16, 2014

Abstract

This paper revisits the problem of adverse selection in the insurance market of Rothschild and Stiglitz [34]. We extend the game-theoretic structure in Hellwig [20] to a mechanism in which the Miyazaki-Wilson-Spence (MWS) allocation is (i) the only candidate subgame-prefect Nash-equilibrium allocation, and (ii) supported always by a perfect Bayesian equilibrium. As is well-known, the MWS allocation is the unique incentive-efficient and individually rational maximizer of the utility of the most profitable type. In fact, given that the informed player has only two types, it is the unique neutral optimum in the sense of Myerson [28].

Keywords: Insurance Market, Adverse Selection, Interim Incentive Efficiency, Neutral Optimum.

JEL Classification Numbers: D86

∗School of Economics & Finance, University of St Andrews, U.K. (corresponding author at td31@st-andrews.ac.uk).
†Dept. of Banking and Financial Management, University of Pireaus, Greece. We have benefited greatly from comments and suggestions by Aris Boukouras, Tasos Dosis, Dino Gerardi, Dominik Grafenhofer, Peter Hammond, Martin Hellwig, Miltos Makris, Motty Perry, Heracles Polemarchakis, Philip Reny, Paolo Siconolfi, and Yiannis Vailakis. Diasakos acknowledges also financial support by the Scottish Institute for Research in Economics. Needless to say, any errors are ours.
1 Introduction

An old yet still open issue in applied micro-economic theory concerns the nature of competitive equilibrium in insurance markets under adverse selection. This dates back at least to the seminal contribution of Rothschild and Stiglitz [34], and has led to the emergence of a significant body of literature which exhibits renewed momentum in the wake of the recent financial crisis. The main conclusions of this literature can be summarized as follows. If the market has an equilibrium, then in equilibrium, insurance deductibles are used as a sorting device to transmit information from the informed side of the market (the insurance customers) to the uninformed one (the insurance firms). However, such a market may not have an equilibrium at all, and when it does, the typical equilibrium allocation will not be interim incentive efficient (IIE) in the sense of Maskin and Tirole [25] (i.e., Pareto-optimal in the interim sense subject to being incentive compatible for the customers and individually rational for the firms).

Restricting attention to the setting in Rothschild and Stiglitz [34] (what we will be referring to henceforth as the “canonical” insurance problem), the present paper proposes a game-theoretic mechanism in which the conjunction of Bertrand competition amongst the firms, on the one hand, and Nash-type strategic interaction between all market participants, on the other, delivers the strongest of results. Namely, the so-called in the literature Miyazaki-Wilson-Spence (MWS) allocation emerges always as the unique equilibrium outcome. As is well-known, this allocation entails the separation of customers across insurance contracts, may involve cross-subsidization, and is always IIE. Being in fact the most preferred point on the IIE frontier for the most profitable type of customer, it is the most desirable allocation (in the sense of neutral optimality, a notion to be described precisely in the sequel).

More specifically, we consider the game in Hellwig [20] but expand the firms’ strategy space along two dimensions. Our firms can subsidize their net income across contracts by offering menus - in fact, from their perspective, menus of contracts are the very objects of trade. Our firms can also publicly commit at stage 1, if they so wish, to not withdraw a given menu at stage 3 irrespectively of the history of play at that point. And together, these two expansions render ours a special case of the game-theoretic structure in Maskin and Tirole [25]. This seminal study extended Hellwig’s game to allow for cross-subsidization between contracts, along with a much more general interpretation of contractual arrangements. The latter are (finite) mechanisms: specifications of a game-form to be played, the set of possible actions for the players, and an allocation for each profile of strategies.

In this sense, the innovation in the present study is about endogenous commitment: the firm’s choice of whether or not to pre-commit upon the delivery of a given menu. This builds upon the insight in Grossman [18] where it was claimed that, in order to realistically interpret the last two stages in Hellwig’s game (especially within the context of insurance or credit provision), one ought to imagine customers submitting applications for insurance and firms deciding whether or not to approve the applications they receive. We do envision nevertheless an additional element in this process: our firms may choose to send their customers “pre-approved” application forms. And in our view, this is an insight of equal realistic importance, especially given what by now has become
common practice in the markets for standardized insurance or credit provision.\(^1\) It readily admits, moreover, a game-theoretic interpretation, as a *public action* in the terminology of Myerson \cite{myerson1981game}: an enforceable decision a player can publicly-commit herself to carry out, even if it may turn out ex-post to be harmful to her or others.

Needless to say, the underlying issue being so fundamental (and not only for insurance economics), there is a large and important literature on competitive equilibria under adverse selection. Its principal aim has been to propose mechanisms, along with their 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 categories, based upon the extent to which the mechanism allows the players’ behavior to be strategic.

One class of models (see, amongst others, Prescott and Townsend \cite{prescott1972insurance}-\cite{prescott1973insurance}, Rustichini and Siconolfi \cite{rustichini2002insurance}, Dubey and Geanakoplos \cite{dubey2003insurance}, Dubey et al. \cite{dubey2004insurance}, and Guerrieri et al. \cite{guerrieri2007insurance}) has focused on Walrasian mechanisms, the central message being 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 of equilibrium, some studies have introduced rationing or suppressed the requirement that firms are profit-maximizing while imposing at the same time quantity constraints on trade.\(^2\) 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.\(^3\)

Another perspective has been to look at mechanisms in which competitive equilibrium is sup-

---

\(^1\) Many a reader will have at some point received in the mail application forms, especially from credit card companies, that name the recipient “pre-approved” or “pre-selected” for a particular product. Many a reader will have also filled in and mailed back some of these forms only to find out that the small print one typically misses at the bottom of the form renders the terms “pre-approved” and “pre-selected” devoid of any meaning other than ”pre-screened as marketing target.” In sharp contrast, our mechanism demands explicitly that either term is legally binding, and prescribes implicitly the required supporting institutional substructure as a necessary condition for efficient insurance provision in the presence of adverse selection.

\(^2\) A notable exception is Bisin and Gottardi \cite{bisin2008insurance} where, instead of being constrained, the Walrasian mechanism is enhanced with the implicit presence of institutions that monitor trade appropriately. Restricting attention also to the canonical insurance economy, this paper shows that the RSW 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 RSW allocation is not IIE, however, markets for consumption rights are required. The ensuing Arrow-Lindahl equilibria internalize the consumption externality due to adverse selection. In fact, by varying the endowment of consumption rights, they trace the entire IIE frontier but for the MWS point. The latter can be obtained only as the limit of a sequence of equilibria.

\(^3\) In its pure form, the Walrasian approach was initiated by Prescott and Townsend \cite{prescott1972insurance}-\cite{prescott1973insurance} and revisited recently by Rustichini and Siconolfi \cite{rustichini2002insurance}. Under rationing, the refinement criteria range from subgame perfection (Guerrieri et al. \cite{guerrieri2007insurance}) to the universal divinity of Banks and Sobel \cite{banks1982market} (Gale \cite{gale1980market}) or the stability of Kohlberg and Mertens \cite{kohlberg1980stability} (Gale \cite{gale1980stability}). The latter notion has been deployed also under quantity constraints (Dubey and Geanakoplos \cite{dubey2003insurance}, Dubey et al. \cite{dubey2004insurance}) but seems to be more binding in that environment. As shown in Martin \cite{martin2006market}, 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.
ported by strategic behavior, and this has produced two separate lines of approach. In the early models (see Wilson [38], Miyazaki [27], Grossman [18], Riley [33], but also Engers and Fernandez [14]), some of the players exhibit strategic behavior which is not of the Nash-type. However, more recent studies (see Netzer and Scheuer [29], Mimra and Wambach [26], Dosis [11] but also Picard [30], von Siemens and Kosfeld [36], Inderst and Wambach [22], Wambach [37], as well as Asheim and Nilssen [1]) have been built upon the game-theoretic foundation in Hellwig [20] or its generalization in Maskin and Tirole [25]. They are thus directly related to the present paper and will be discussed in some detail in the sequel. Yet, there are two fundamental differences setting apart these papers from the present that should be noted on the outset.

On the one hand, and apart from extremely special cases, the equilibrium outcome is not uniquely the MWS allocation. In fact, whenever the RSW allocation is not IIE, a multiplicity of contractual arrangements can be supported as equilibria - with the equilibrium set typically being rich enough to render its IIE subset (let alone the MWS singleton) null. On the other hand, and more importantly perhaps, each and every one of these studies proposes changes in the specification of Hellwig’s game along either of two dimensions: restricting the set of permissible contracts, or altering the timing - by anticipating the stage at which firms may withdraw their contracts to precede that at which the customers choose while at the same time imposing structure on the way in which contractual offers can be withdrawn. The unifying theme, therefore, of this recent literature is to propose interesting ways in which strategic interaction in markets with adverse selection can be formalized game-theoretically. However, this seems to be done in a way that obscures rather than facilitates our understanding of the real issue at hand, the properties of competition in markets with adverse selection.

In sharp contrast, but in line with the approach in Rothschild and Stiglitz [34] as well as Hellwig [20], the present study remains within the realm of a basic and well-known environment, Bertrand competition. In fact, deploying neither exogenously-imposed restrictions on the set of contracts the firms may offer, nor any structure in the way or changes in the timing under which these offers are made, our mechanism embeds either of the two seminal studies. And by doing so, it sheds light in important aspects of Bertrand competition under asymmetric information that seem to have been ignored by the two papers, causing some confusion in the subsequent literature.

Formally speaking, the Hellwig and Rothschild-Stiglitz models can be viewed, respectively, as examples of Bertrand competition with or without continuous payoffs for the firms. From this perspective, the issue with existence of equilibrium in Rothschild and Stiglitz [34] is something that was not immediate at the time when even Hellwig [20] was written, but certainly is now. From a more intuitive perspective, however, our notion of endogenous commitment allows the Hellwig and Rothschild-Stiglitz models to be viewed, respectively, as examples of Bertrand competition with or without exit. Indeed, the two models result directly from our mechanism if one restricts attention to singleton menus that can be offered, respectively, only without or only under public
commitment. And the latter distinction in particular shows that, contrary to what seems to be the common belief in the literature, the restriction either model imposed on firm competition under asymmetric information went far beyond not allowing firms to cross-subsidize between risk-types.

This emerges from our mechanism if one continues to restrict firms into making offers, respectively, only without or only under public commitment, but allows them in either case to do so with non-singleton menus if they so wish. As we establish in the sequel, in the latter case, the absence of the possibility for exit on the one hand restricts the set of candidate equilibria to allocations that are safe in the sense of Holmstrom and Myerson [21] (i.e., in the current context, to menus that avoid losses irrespectively of the type of customer who chooses them). The free-entry component of Bertrand competition on the other requires that only IIE allocations may be supported as equilibria. In other words, free entry without exit renders the RSW allocation (the only allocation that is both safe and IIE) the unique equilibrium candidate. Yet, the RSW allocation is IIE in the canonical insurance problem if and only if it coincides with the MWS allocation.

It follows, therefore, that the real issue with firms’ competition under asymmetric information in Rothschild and Stiglitz [34] and Hellwig [20] has to do, respectively, solely with the existence and solely with the efficiency of the equilibrium allocation. Equally clearly, a fundamental desideratum regarding Bertrand competition under asymmetric information is to allow firms not only the choice to exit but also the ability to commit to not exit.

To this end, the present paper highlights the role of endogenous public commitment, along with that of the required supporting institutional substructure. Needless to say, the idea of market environments that allow for commitment as endogenously-emerging strategic choice exists already in the literature (see, for instance, Caruana and Einav [7] but also Dosis [11], Netzer and Scheuer [?], as well as Bernheim and Whinston [5] or Baker et al. [?]). Yet, the distinguishing feature of our mechanism is that the implicitly assumed commitment technology to facilitate this strategic choice is required off rather than on the equilibrium path.

Considering the equilibrium path of a game as the model’s predicted behaviour, the fact that the strategic option to commit is not chosen along this path goes a long way in absolving us as modelers from the need of a real motivational...

---

5 And to clarify another point that has been confusing for the canonical insurance problem literature, the RSW allocation coincides with the MWS one on a strict subset of the space of primitive parameters for which the equilibrium exists in Rothschild and Stiglitz [34].

6 Caruana and Einav [7] consider a game in which players can change their previously announced actions but at a cost. A similar idea is to be found also in Netzer and Scheuer [?], where Hellwig’s game is altered so that the firms may withdraw their contracts only at a cost (and at a stage preceding the one at which the customers choose). By contrast, adopting the notion an early version (see Diasakos and Koufopoulos [10]) of the present paper introduced as contractual commitment, Dosis [11] views binding contractual offers as costless strategic choices. These are, however, at the disposal of the informed rather than the uninformed party, in a signalling adaptation of Hellwig’s game which restricts the firms into making only binding offers. Baker et al. [?] allow firms to choose between contracts that are binding, because they are based on objectively-verifiable even though poor performance measures, and contracts that are not, as they are written upon subjective but non-verifiable indicators. Quite similarly, Bernheim and Whinston [5] analyse binding versus non-binding contractual agreements as the choice between complete and incomplete contracts. Yet, as models of endogenous commitment, none of these two relates to the canonical insurance problem, where one assumes (amongst many other things) contractual completeness in order to focus on the issue of adverse selection.
story for the commitment technology in question.

Yet, as a built-in feature, our mechanism carries also a non-strategic component of commitment which concerns all paths of play: it does not allow firms to withdraw a given contract from a menu unless they do so for the menu as a whole. The motivation for this type of commitment is none other than a benchmark for mechanism optimality. Considering the adverse selection problem at hand, our mechanism can be viewed as demonstrating the operational content of the abstract argument in Myerson [28] for the existence of the set of neutral optima, an argument that depends fundamentally upon the deployment of public actions to restrict the out-of-equilibrium beliefs appropriately. The notion of neutral optima was introduced in Holmstrom and Myerson [21] as the smallest class of incentive-compatible allocations that satisfy four fundamental axioms of mechanism selection. As it turns out, in the canonical insurance problem, the set of neutral optima is the singleton consisting of the MWS allocation. As this coincides with the equilibrium outcome of our mechanism, it renders the present study a rather unique in the current literature application of fundamental theoretical insight.7

The balance of the paper is structured as follows. The next section introduces the canonical insurance problem and recalls some results concerning the efficiency of the feasible allocations that will be used in the subsequent analysis. Section 3 describes the mechanism, how it always admits the MWS allocation as its unique, strictly speaking perfect-Bayesian albeit essentially sequential equilibrium outcome, as well as how it relates to the issues about competition under adverse selection discussed above. In Section 4, we interpret our findings vis a vis the directly related ones in the literature. Section 5 concludes, being followed by appendices that contain the proofs of the results in the main text as well as supporting technical material.

2 The Canonical Insurance Problem

The most parsimonious model of adverse selection in the market for insurance provision is the following. There is a continuum of unit measure of risk averse agents, each and everyone of whom has the same wealth endowment $W \in \mathbb{R}^{+}$ and faces the idiosyncratic risk of it being reduced by the amount $d \in (0, W)$. Formally, this endowment risk will be depicted by two states of nature: the “accident” and “no accident” state (denoted, respectively, by $s = 1$ and $s = 0$). Either state is meant to be perfectly observable and verifiable by a court of law, while each agent belongs to one of two risk classes. For the low-risk class (whose relevant quantities will be denoted by the subscript $L$), the accident occurs with probability $p_L \in (0, 1)$. For the high-risk class (whose quantities will be denoted by the subscript $H$), it obtains with larger probability $p_H \in (p_L, 1)$. The class to which an agent belongs will be referred to also as her “type,” and taken to be her own private information.

7 Balkenborg and Makris [3] is a notable contribution in the quest for a constructive characterization of the set of neutral optima. When the informed party has only two types and all utilities are quasi-linear in the transfers, their concept of assured allocation is shown to coincide with the MWS allocation, the unique neutral optimum. In terms of implementation, however, they suggest the same game form as in Netzer and Scheuer [?]; hence, a dense equilibrium set in which, even with two types, the neutral optima comprise but a null subset.
The share $\lambda \in (0,1)$ of low risk agents in the population, however, is common knowledge.\footnote{The continuum hypothesis is standard in models of this type. It allows us to invoke some version of the law of large numbers and assume that exactly the shares (i) $p_h$ of risk-type $h \in \{L,H\}$ agents, and (ii) $\lambda p_L + (1-\lambda) p_H$ of the entire population will eventually have an accident. Of course, applying laws of large numbers on a continuum of random variables is not without technical caveats - for some references on the technical complications and their remedies, see for instance footnote 3 in Netzer and Scheuer [29].}

Partitioning the population into these two risk classes is meant to capture all the relevant information about the problem of adverse selection. We imagine, therefore, that we have arrived at this partition after having inferred as much as possible from the agents’ observable characteristics. As a result, apart from being of a particular risk type, all agents are taken to be otherwise identical. More precisely, they all have von Neumann-Morgenstern preferences over wealth lotteries with the same state- and type-independent Bernoulli utility function $u : \mathbb{R}_+ \mapsto \mathbb{R}$. The latter is everywhere strictly-increasing, strictly-concave, and twice continuously-differentiable.

An agent may insure herself against her endowment risk by accepting an insurance contract. The typical one will be denoted by the vector $a = (a_0, a_1) \in \mathbb{R}_+^2$ where $a_0$ is the premium paid in exchange for receiving the net indemnity $a_1$ if an accident occurs. The corresponding wealth vector, denoted by $w(a) = (w_0(a_0), w_1(a_1)) \equiv (W - a_0, W - d + a_1)$, entails a transfer across the two states at the premium rate $d w_1(a_1) / d w_0(a_0) = -a_0 / a_1$. The expected utility of an agent of risk-type $h \in \{L,H\}$ is given by $U_h(w(a)) \equiv (1 - p_h) u(w_0(a_0)) + p_h u(w_1(a_1))$.

The space of \textit{admissible} insurance contracts will be given by $\mathcal{A} \equiv [0,W] \times [0,W] \times [0,W + W - d]$ for some arbitrarily-large, strictly-positive constant $W$. Needless to say, there is a one-to-one relation between this space and that of the corresponding wealth vectors $W \equiv [0,W] \times [0,W + W - d]$. It entails only slight abuse of notation, therefore, to take the agents’ preference relation over lotteries on $W$, denoted by $\succ_h$ for risk-type $h$, as ordering also lotteries on $\mathcal{A}$. And to abuse notation slightly further, we will also take the function $U_h(\cdot)$ as being defined directly on $\mathcal{A}$.

On the supply side of the market, insurance is provided by risk-neutral firms which maximize expected profits. To ensure competition, we will assume that their collection is a set $\mathcal{N} \subset \mathbb{N}$ with $|\mathcal{N}| \geq 2$. The expected profit of the typical firm is given by $\Pi_h(a) = (1 - p_h) a_0 - p_h a_1$, when the typical contract is bought by the entire class of risk-type $h$ customers. The insurance market is characterized also by free entry in the sense that the firms have adequate financial resources to be able to deliver, following the realization of the accident state, any collection of contracts from $\mathcal{A}$ to any subset of the population.

### 2.1 Feasibility and Efficiency

The above data notwithstanding, the collections of contracts that are directly relevant for the problem at hand form a much simpler space under the standard assumption that the choice of contract by an agent is \textit{exclusive}. More precisely, each agent is restricted to choose at most one contract from those that are available in the market when she is called upon to act. Under this condition, and given that the population is partitioned into two risk classes, the relevant collections of contracts are tuples $\{a_h, a_{h'}\}$ where $a_h, a_{h'} \in \mathcal{A}$ with the subscript $h, h' \in \{L,H\}$ indicating the...
risk-class the respective contract is meant for. In what follows, we will be referring to such tuples as insurance *menus*.

In fact, since our analysis will lie entirely within the realm of the direct revelation principle and maintain that each market participant is rational with an outside option available, the space of relevant insurance menus ought to be further restricted to those that are at the same time *incentive compatible* for either risk type and *individually rational* for all market participants. Formally, the typical menu \( m = \{ a_L, a_H \} \) will be said to meet the former requirement if it satisfies the condition

\[
(\text{IC}_h). \quad U_h (a_h) \geq U_h (a_{h'}) \quad h, h' \in \{ L, H \}
\]

which ensures that truthfully revealing one’s risk-type is a Nash equilibrium amongst the customers (it is in one’s best interests to report her type honestly given that everyone else does the same). By contrast, the menu is said to be individually rational for risk-type \( h \) if

\[
(\text{IR}_h). \quad U_h (a_h) \geq \pi_h \equiv U_h (0) \quad h \in \{ L, H \}
\]

reflecting the very fact that each customer enters the insurance market while still in possession of the endowment lottery.\(^9\) As for the firms, since they can guarantee themselves no losses simply by abstaining from market activity, the menu itself should avoid losses, at least for some belief the firm may have about the distribution of risk-types in its customer pool. In other words, the menu is individually rational for the firm if

\[
(\text{IR}_F). \quad \exists p \in [0, 1] : \Pi (m|p) \equiv p \Pi_L (a_L) + (1 - p) \Pi_H (a_H) \geq 0
\]

The space of relevant menus, therefore, is given by\(^10\)

\[
\mathcal{M} = \{ m \in \mathcal{A} \times \mathcal{A} : (\text{IR}_F), (\text{IC}_h) - (\text{IR}_h) \ h \in \{ L, H \} \}
\]

and within this space, the search for optimal menus will be based upon the Pareto criterion. As usual, given two menus \( m = \{ a_L, a_H \} \) and \( m' = \{ a'_L, a'_H \} \), the former will be said to Pareto-dominate the latter if \( U_h (a_h) \geq U_h (a'_{h'}) \) for either \( h \) with at least one inequality being strict. It is well-known, moreover, that it suffices here to search for the so called *(interim)* incentive efficient (IE) allocations: the ones that are Pareto-optimal amongst those that are (i) IC\(_h\) for either \( h \), and (ii) IR\(_F\) for \( p = \lambda \). Formally, the IE frontier consists of those menus that solve the IE(\( \mu \)) problem:\(^11\)

\[
\max_{\{ a_L, a_H \} = m \in \mathbb{R}^2 \times \mathbb{R}^2} \mu U_L (a_L) + (1 - \mu) U_H (a_H) : \quad \Pi (m|\lambda) \geq 0
\]

---

\(^9\)With no loss of generality, we may let \( \pi_h > 0 \) for either \( h \). For as \( U_h (0) = (1 - p_h) u (W) + p_h u (W - d) \), normalizing the customers’ outside option to \( u (W - d) = 0 \) implies that \( U_h (0) = (1 - p_h) u (W) > (1 - p_h) u (W - d) = 0 \).

\(^10\)Notice that \( \mathcal{M} \) includes the trivial menu \( \{ 0^2, 0^2 \} \). In terms of a firm’s action, this reflects the choice of not participating in the market.

\(^11\)Defining Pareto-optimality via the IE(\( \mu \)) problem is based upon a well-known supporting hyperplane argument, which requires in turn that the feasible set is convex. And even though the linear profit constraint is trivially convex, the IC\(_h\) constraints are not. For this reason, it is standard practice in the literature (see, for instance, Myerson [28] or Maskin and Tirole [25]) to assume that allocations are assigned randomly. In the present setting, this refers to distributions on \( \mathcal{A} \times \mathcal{A} \). Letting \( v \) be a typical one, the IC\(_h\) constraint now reads \( \int_{\mathcal{A} \times \mathcal{A}} \{ U_h (a_h) - U_h (a_{h'}) \} \, dv (a_h, a_{h'}) \geq 0 \) and, being linear, it is trivially convex in \( v \). Of course, in the present setting the IE allocations are all deterministic so that the assumption in question can be left in the background as a purely technical caveat.
for some $\mu \in [0,1]$.

In what follows, we will be referring to an allocation as IE if it is the IE($\mu$) for some $\mu \in [0,1]$. Our main focus, however, will be on two allocations that feature prominently in the literature on competitive market environments with adverse selection. The Miyazaki-Wilson-Spence (MWS) allocation solves the IE(1) problem and its importance is obvious - it maximizes the utility of the most profitable risk-type.\textsuperscript{12} The Rothschild-Stiglitz-Wilson (RSW) allocation, on the other hand, is the only allocation that solves the RSW($\mu$) problem:

$$\max_{m \in \mathbb{R}^2} \mu U_L(a_L) + (1 - \mu) U_H(a_H) : \ (IC_h), \Pi_h(a_h) \geq 0 \ h \in \{L, H\}$$

for every $\mu \in [0,1]$. This allocation maximizes the utility of either risk-type while being also \textit{ex-post} individually-rational for the firms. Its importance lies in the latter characteristic which renders it a \textit{safe} allocation in the terminology of Myerson [28]: one the uninformed party would continue to view as individually-rational even if she knew the private information of the informed, whatever this information might be.

To be precise, this is the RSW allocation \textit{relative to the endowment} in the terminology of Maskin and Tirole [25].\textsuperscript{13} In this terminology, the IE($\mu$) optimum for $\mu \in (0,1)$ is interim efficient (and, thus, also weakly interim efficient) \textit{relative to the prior belief} $\{\lambda, 1 - \lambda\}$ - which in what follows will be also the interim belief given that, the uniformed party being the one to make the proposals, no updating of beliefs takes place at the interim stage. In the current setting, the RSW allocation is given by the menu $m^{**} = \{a_L^{**}, a_H^{**}\}$ which offers full-insurance to the high-risk type ($a_{0H}^{**} = d - a_{1H}^{**}$) while being such that $U_H(a_L^{**}) = U_H(a_H^{**})$ and $\Pi_h(a_h^{**}) = 0$ for $h = L, H$.\textsuperscript{14} And rather similarly, the IE($\mu$) allocations are also such that at least one risk-type receives full insurance while her IC constraint is binding. In fact, and to list some properties we will use in our proofs, the IE frontier can be described as follows.

\textbf{Lemma 1} Take an arbitrary $\mu \in [0,1]$. The IE($\mu$) optimum $m^\mu = \{a_L^\mu, a_H^\mu\}$ is unique and such that

(i) $d_{0H}^{\mu} = d - a_{1H}^{\mu}$, $a_{0L}^{\mu} < d - a_{1L}^{\mu}$, $U_H(a_L^{\mu}) = U_H(a_H^{\mu})$ and $U_L(a_H^{\mu}) < U_L(a_L^{\mu})$ if $\lambda < \mu$,

(ii) $d_{0H}^{\mu} > d - a_{1H}^{\mu}$, $a_{0L}^{\mu} = d - a_{1L}^{\mu}$, $U_H(a_L^{\mu}) < U_H(a_H^{\mu})$ and $U_L(a_H^{\mu}) = U_L(a_L^{\mu})$ if $\lambda > \mu$, and

(iii) $a_h^\mu = a_h^\mu$ for $\mu \in \{L, H\}$ with $a_0^{\mu} = d - a_1^{\mu}$ if $\lambda = \mu$.

\textsuperscript{12} As established in Lemma 1(b) below, for any $\mu \in (\lambda, 1]$, we may restrict attention to the choice set $m \in \mathbb{R}^2_{++} \times \mathbb{R}^2_{++}$. And as $p_H > p_L$, any $a \in \mathbb{R}^2_{++}$ gives $\Pi_L(a) > \Pi_H(a)$.

\textsuperscript{13} Recall that the endowment is the point $(W, W - d)$ in the $(w_0, w_1)$-space. It corresponds to the origin $0^2$ in the $(a_0, a_1)$-space.

\textsuperscript{14} In fact, $a_H^\mu$ is the first-best allocation for the high-risk type, the one she would receive if she were to reveal her type. Formally, it is defined as $a_H^\mu = \arg \max_{a \in A: \Pi_H(a) \geq 0} U_H(a)$. Graphically, it is depicted by the intersection of the lines $\{(a_0, a_1) : a_0 + a_1 = d\}$ and $F_{\Omega H}^\mu = \{a \in A : \Pi_H(a) = 0\}$. The former line is the locus of full-insurance $\{(w_0, w_1) \in W : w_0 = w_1\}$ in the $(w_0, w_1)$-space. The latter has slope $da_1/da_0 = -(1 - p_H)/p_H$ (a fair-odds line for the high-risk type) and goes through the endowment contract $0^2$.\textsuperscript{15}
Moreover, $\Pi_M (m^\mu | \lambda) = 0$ and $\Pi_L (a_L^\mu) \geq 0 \geq \Pi_H (a_H^\mu)$ with either inequality strict unless $a_H^\mu = a_H^{**}$. Furthermore, for either $h$, the solution mapping $a_h^*: [0, 1] \mapsto \mathbb{R}^2$ with $a_h^* (\mu) = a_h^\mu$ consists of two functions

(a) $[0, \lambda) \mapsto \mathbb{R}^2 \setminus \{a_h^{**}\}$ which is bijective, and

(b) $(\lambda, 1] \mapsto \mathbb{R}^2_{++}$ which is either the constant $(\lambda, 1] \mapsto \{a_h^{**}\}$ or a bijective function $(\lambda, 1] \mapsto (0, b) \times (0, b)$ for some finite $b > 0$. In fact, for $\mu \in (\lambda, 1]$, $a_h^* (\mu)$ is a continuous function and so is the value function $\mu U_L (a_L^* (\mu)) + (1 - \mu) U_H (a_H^* (\mu))$.

**Proof.** Conditions (i)-(iii) are well-known (see, for instance, Theorem 1 in Crocker and Snow [8]). The remaining properties follow from Lemmas 14, 15, and 18 in Appendix B. ■

Of course, the definitions of the RSW($\mu$) and IE($\mu$) problems do not require that the relevant space of menus is $\mathcal{M}$ per se. Yet, and as opposed to the entirety of the recent literature on the canonical insurance problem, to do so in the sequel comes without loss of generality.\(^{15}\) For in the next section we will restrict attention to the IE allocations that lie arbitrarily close to the MWS one. And for these allocations, assuming that $\mathcal{A}$ is the contractual space is innocuous (see Lemma 1(b) above but also Corollary 1 in Appendix B). More importantly perhaps, the same can be said with respect to the individual rationality constraint of either type of the insurance customers.

**Claim 1** The following statements are true.

(i) $a_h^{**} \succ_h 0$ \quad $\forall h = L, H$

(ii) $a_H^\mu \succ H 0$ \quad $\forall \mu \in [0, 1]$

(iii) $\exists \mu_L \in [0, 1]: a_L^\mu \succ_L 0$ \quad $\forall \mu \in [\mu_L, 1]$

**Proof.** See Appendix B. ■

To complete introducing the efficient allocations that will be focal for the subsequent analysis, we should outline the relation between the RSW allocation and the IE allocations that are arbitrarily close to the MWS - in particular, between the former and the latter. In the current setting, this relation is particularly strong.\(^{16}\)

**Claim 2** The following statements are equivalent.

\(^{15}\)For every $\mu \in [0, 1]$, restricting attention to menus $m \in \mathcal{A} \times \mathcal{A}$ renders the feasible set of the IE($\mu$) problem compact for every $\lambda \in (0, 1)$. The set in question is closed (since the constraints are all weak inequalities between continuous functions) while $\mathcal{A}$ is itself bounded. Equally importantly, the correspondence from the parameter vector $(\mu, \lambda) \in [0, 1] \times (0, 1)$ to the feasible set is not only compact-valued but also continuous: it is in fact constant with respect to $\mu$ while it depends on $\lambda$ only through the profit constraint linearly (hence, it suffices to recall the well-known argument for the continuity of the budget correspondence in the standard consumer theory problem - see, for instance, Lemmata 9.22-9.23 in Sundaram [9]). The objective being also continuous in $(\mu, \lambda)$, Berge’s theorem of the maximum applies. Needless to say, these observations apply also for the RSW($\mu$) problem.

\(^{16}\)XX Netzer and Scheuer [?]
(i) The RSW allocation is $IE$.

(ii) The RSW allocation is $IE(\mu)$ $\forall \mu \in (\lambda, 1]$.

(iii) We have
\[
\frac{\lambda}{1 - \lambda} \leq \frac{(p_H - p_L) u'(w_{1L}^{**}) u'(w_{0L}^{**})}{p_L (1 - p_L) [u'(w_{1L}^{**}) - u'(w_{0L}^{**})] u'(w_{1H}^{**})}
\]  

(1)

Proof. See Appendix B.

As it turns out, the relation between the RSW allocation and those that are $IE(\mu)$ for $\mu \in (\lambda, 1]$ can be thought of as being driven by the contract that lies on the intersection of the fair-odds lines $FO_h = \{ a \in A : \Pi_h(a) = \Pi_h(a_n^H) \}$. That is, on the intersection of the two lines with slope $da_1/da_0 = -(1 - p_h)/p_h$ that go through the $IE(\mu)$ optimal contract for risk-type $h$.

Lemma 2 Take an arbitrary $\mu \in [0, 1]$ and let the $IE(\mu)$ optimal menu $m^\mu = \{a_L^\mu, a_H^\mu\}$ define the contract $a^\mu = (a_0^\mu, a_1^\mu)$ by
\[
(1 - p_h) (a_{0h}^\mu - a_0^\mu) = p_h (a_{1h}^\mu - a_1^\mu) \quad h = H, L
\]

(2)

This contract is unique and such that $a_0^\mu \geq 0$ while $\Pi_M(a^\mu) = \Pi(\{a^\mu, a^\mu\} | \lambda) = 0$.

Proof. See Appendix B.

Graphically (see the left panel of Figure 2), the RSW and $IE(\mu) : \mu \in (\lambda, 1]$ allocations differ by the extent to which $a^\mu$ is away from the endowment point on the market fair-odds line $FO_{M}^E = \{ a \in A : \Pi_M(a) = 0 \}$.\(^\text{17}\) Formally, what distinguishes them is whether or not the constraint $a_0^\mu \geq 0$ binds. For if it does, it cannot but be $a^\mu = 0$. But then (2) reads $\Pi_h(a_h) = 0$ for either $h$ and the $IE(\mu) : \mu \in (\lambda, 1]$ problem coincides with the RSW($\mu$) : $\mu \in (\lambda, 1]$ one.

3 Mechanism and Analysis

The workings of the canonical insurance market will be modeled by means of the three-stage game in Hellwig [20] - albeit, with two modifications. The first refers to the fact that, in the present analysis, the object of trade for the firms will not be the contract but the menu, the typical element in $M$. The second modification gives rise to our actual mechanism. Namely, as a choice decision at stage one, the firm’s offer will now have two dimensions. It will entail on the one hand choosing a menu of contracts from $M$, and on the other the decision of whether or not to offer this menu with commitment upon its delivery.

Formally, our mechanism operates through the following stages.

Stage 0: Nature decides the risk-type $h \in \{L, H\}$ of each customer. This is neither observable nor verifiable by a third party.

\(^{17}\text{The equation } \Pi_M(a^\mu) = 0 \text{ being obviously satisfied by } 0^2, \text{ the contract } a^\mu \text{ cannot but lie on the line } FO_{M}^E.\)
Stage 1: The firms decide simultaneously on their menu offers. Each firm must choose one menu from $\mathcal{M}$ and at the same time whether or not it will offer it with commitment.

Stage 2: The customers choose simultaneously from the collection of menus the firms have offered at stage 1, after observing whether or not a given menu has been offered with commitment. Each customer can choose up to one menu, and exactly one contract from that menu. If she doesn’t choose any menu, she is left to face her endowment lottery while if no menu is chosen by any customer the game ends. If a customer chooses a menu with commitment, she receives her preferred contract from that menu. And if all firms have made or all customers choose menu offers with commitment, the game ends. Otherwise, it continues to the next stage.

Stage 3: After observing all the menu offers of stage 1, including whether each has been made with or without commitment, the firms that made offers without commitment decide simultaneously whether or not to withdraw them. If the respective firm withdraws the menu a customer had chosen at stage 2, the latter is left to face her endowment lottery. Otherwise, she receives her preferred contract from that menu. And, in either case, the game ends.

We will denote the above game by $\Gamma(S_F,S_A,\Pi_{h \in \{L,H\}},U_{h \in \{L,H\}})$. Pure strategy profiles for the firms will be denoted by $s = (s_1,\ldots,s_N) \in S_F$ with the typical firm’s strategy having three components. The first two reflect the action of the $n$th firm at stage 1, which entails the choice of a menu $m_n \in \mathcal{M}$ to be offered along with the decision $c_n \in \mathcal{C} \equiv \{C,NC\}$ of whether or not to commit upon its delivery (with $C$ and $NC$ having the obvious interpretation). Abusing notation slightly, it will facilitate the ensuing exposition to depict the two dimensions of the firm’s action at stage 1 as the restrictions on the respective domain of a single action $s_n^1 \in S_n^1 \equiv \mathcal{M} \times \mathcal{C}$. More precisely, we will depict the typical action-tuple $(m_n,c_n)$ by $(s_{n|\mathcal{M}},s_{n|\mathcal{C}}^1) = s_n^1$ and the typical history to be observed at the beginning of stage 2 as $s^1 \in S^1 = \prod_{n \in N} S_n^1$.

Of course, the type of a customer making a choice at stage 2 not being observable to anyone but herself, for each history $s^1 \in S^1$, the ensuing stages 2 and 3 define a sub-game that involves simultaneous moves. That is, the extensive form $\Gamma$ admits but one proper subgame $\Gamma(s^1)$ starting at stage 2. As a result, the third component of the $n$th firm’s strategy is given by the function $s_n^2 : S^1 \mapsto \{W, NW\}$. This prescribes whether or not to withdraw its menu at stage 3 (with $W$ and $NW$ having again the obvious interpretation). Yet, there is no such decision to be made by those firms who chose to commit at stage 1. Letting, therefore,

$$\mathcal{N}_C(s^1) = \{n \in \mathcal{N} : s_{n|\mathcal{C}}^1 = C\}$$

be the collection of firms who commit upon their menus at stage 1, we will impose the restriction

$$s_n^2(s^1) = NW \quad \forall n \in \mathcal{N}(s^1), \forall s^1 \in S^1$$

---

18 Recall our assumption that every customer is endowed with the trivial contract 0. Alternatively, we may assume the existence of a zeroth firm that always offers the trivial menu \{0,0\} at stage 1, with or without commitment but under the plan to not withdraw it at stage 3 in the latter case.
To complete defining the firms’ strategy space, let $S_S^2 = \{W, NW\}^{S_1}$, $S^2 = \prod_{n \in N} S^2_n$, and $S_F = S^1 \times S^2$. It remains to describe the strategy of the typical customer at the beginning of stage 2. To facilitate notation for this end, in what follows, the contract in the typical menu that is meant for the risk-type $h \in \{L, H\}$ will be denoted also as a function $a : M \mapsto A$, where $a_h = a_h(m)$ whenever $m = (a_L, a_H)$. Given this, the typical risk-type’s strategy will be given by the function $\alpha_h : S^1 \mapsto A$ which prescribes the contract to be chosen in the subgame $\Gamma(s^1)$ for each history $s^1 \in S^1$.\(^{19}\) The space $S_A = A^{S^1}$ of such functions is the strategy space of either risk-type.

### 3.1 Equilibrium

To make predictions, we will use the pure-strategies version of the perfect Bayesian equilibrium in Maskin and Tirole [25]. This is basically a pure-strategies sequential equilibrium but for the following caveat. The firms’ contractual action space in stage 1 being infinite, the requirement for consistency of beliefs between this stage and 2 is not well-defined. As a result, we seek a vector of (pure) strategies (a profile of strategies for the firms and a strategy for each type of customer), and a vector of beliefs at each information set of the typical subgame $\Gamma(s^1)$ such that (i) the strategies are optimal under these beliefs (sequential rationality), and (ii) the belief at the beginning of $\Gamma(s^1)$ is the prior $\{\lambda, 1 - \lambda\}$ while the subsequent beliefs are everywhere (fully) consistent with the strategies.\(^{20}\) Under this notion of equilibrium and using the expression “honoring a menu” to mean that it is not withdrawn at stage 3, an equilibrium menu is such that (a) on the equilibrium path, it is honored at stage 3 and chosen by at least one risk-class of customers, and (b) there is no other menu that, if offered alongside the one in question, would expect strictly positive profits under some (fully) consistent profile of beliefs.

Formally, consider a given strategy profile $(s, A_L, A_H) \in S_F \times S_A \times S_A$. We first define the payoffs for the stage-2 subgame $\Gamma(s^1)$. To this end, let

$$
\mathcal{N}_{NW}(s) = \{n \in N : s^2_n(s^1) = NW\} \supset \mathcal{N}_C(s^1)
$$

$$
\mathcal{M}_{NW}(s) = \{m \in M : m = s^1_n|_M, n \in \mathcal{N}_{NW}(s)\}
$$

be the collection of firms that plan to honor their menus at stage 3 within the subgame $\Gamma(s^1)$ under $s = (s^1, s^2)$, along with the collection of their respective menus. Optimality on the customers’ part

\(^{19}\)By the very definition of our mechanism, this contract is unique. Whenever $\alpha_h(s^1)$ is to be found in more than one menus from $\cup_{n \in N} s^1_n|_M$, we will assume that each of these menus receives the same share of the total applications made for this contract by the risk-type $h$.

\(^{20}\)Needless to say, the latter part of the statement in (ii) entails the existence of strategic profiles that are arbitrarily close to the equilibrium one and assign positive probability to every move on the game tree, such that the belief vectors that satisfy Bayes’ rule for these strictly positive strategy profiles are arbitrarily close to the equilibrium beliefs. We should point out also that, in Maskin and Tirole [25] (see, in particular, their footnote 13), the requirement for consistency of beliefs between the stages 1 and 2 is replaced by that of updating the prior to an interim belief via Bayes rule. In the present study, however, the uniformed party being the one to make the proposals, no updating of beliefs takes place at the interim stage.
requires their best response function to be given by $\alpha_h^*: S^1 \times S^2 \to A$ such that\(^{21}\)

$$\alpha_h^*(s) \in \begin{cases} \arg\max_{m \in M_{NW}(s)} \ U_h(a_h(m)) & \text{if } \exists m \in M_{NW}(s) : U_h(a_h(m)) \geq \bar{u}_L \\ \{0\} & \text{otherwise} \end{cases} \quad (3)$$

for $h = L, H$.

For the firms’ payoff, on the other hand, we need to distinguish between two kinds of subgames that lie within $\Gamma(s^1)$. First, there are those firms that have made offers at stage 1 with commitment, each of these defining a trivial subgame in which there is a singleton decision node for each risk-type. To obtain the payoff for these firms, let

$$M_h(s) = \{ m \in M_{NW}(s) : a_h(m) = \alpha_h^*(s) \}$$

be the collection of menus that are optimal for the risk-type $h$ under the given firms’ strategy profile. For each $n \in \mathcal{N}_C(s^1)$, the payoff is given by

$$\Pi_n^C(s^1_{n|\mathcal{M}}, s_{-n}, \alpha) = \begin{cases} \sum_{h \in \{L, H\}} \lambda_h \Pi_h(\alpha_h^*(s)) & \text{if } a_h(s^1_{n|\mathcal{M}}) = \alpha_h^*(s) \\ 0 & \text{otherwise} \end{cases}$$

where $\lambda_L = \lambda = 1 - \lambda_H$.

The second type of subgames within $\Gamma(s^1)$ refer to the last two stages of the game without commitment. Which is nothing but a simple signalling game in which the informed party sends a signal to the uniformed by choosing a menu (Figure XX). Here, let $\lambda_h: S^1 \to [0, 1]$ be a belief function, in the sense that $\lambda_h(s^1)$ is the probability the $n$th firm assigns to its menu being selected by the risk-type $h$ at stage 2 given the history $s^1$ and the agents’ strategy profile $\alpha = (\alpha_L, \alpha_H)$.

For $n \in \mathcal{N} \setminus \mathcal{N}_C(s^1)$, therefore, the payoff will be given by $\Pi_n^W(s^1_{n|\mathcal{M}}, s_{-n}, \alpha)$ and

$$\Pi_n^W(s^1_{n|\mathcal{M}}, s_{-n}, \alpha) = \begin{cases} \sum_{h \in \{L, H\}} \lambda_h(s^1) \Pi_h(\alpha_h(s^1_{n|\mathcal{M}})) & \text{if } \sum_{h \in \{L, H\}} \lambda_h(s^1) > 0 \\ 0 & \text{otherwise} \end{cases}$$

And optimality requires that the best response is given by the function $s_n^{2*}: S \times S_{-n} \times S_A \to \{W, NW\}$ where $s_n^{2*}(s^1_{n|\mathcal{M}}, s_{-n}, \alpha) = NW$ if and only if $\Pi_n(NW|s^1_{n|\mathcal{M}}, s_{-n}, \alpha) \geq 0$.

Letting finally

$$\Pi_n^{NW}(s^*_{-n}, \alpha^*) = \max_{m \in \mathcal{M}} \Pi_n^{NW}(m, s^*_{-n}, \alpha^*)((s^1_{n|\mathcal{M}}, NC), s^*_{-n}))$$

$$\Pi_n^C(s^*_{-n}, \alpha^*) = \max_{m \in \mathcal{M}} \Pi_n^C(m, s^*_{-n}, \alpha^*)((s^1_{n|\mathcal{M}}, C), s^*_{-n}))$$

$$\Pi_n(s^*_{-n}, \alpha^*) = \max \{ \Pi_n^C(s^*_{-n}, \alpha^*), \Pi_n^{NW}(s^*_{-n}, \alpha^*) \}$$

\(^{21}\)Recall that we restrict attention to incentive-compatible menus. Moreover, we require that out of a given menu a customer may choose but one contract. Hence, there is no loss of generality in assuming that, out of a given menu $m$, the risk-type $h$ chooses the contract $a_h(m)$, the one that is intended for her type. Notice also that no customer will ever choose any contract from a menu if she believes that the latter will be withdrawn. For under our normalization of the function $u$, any such menu is dominated strictly by the endowment lottery (recall footnote 9).
an equilibrium of $\Gamma$ is a profile of strategies $(s^*, \alpha^*)$ and beliefs $\{\lambda_{hn}\}_{n \in \mathcal{N} \setminus \mathcal{N}_C(s^*)}$ such that for each $n \in \mathcal{N}$ we have

$$s_n^* \in \begin{cases} 
\arg \Pi_n (s_{\neg n}^*, \alpha^*) \times \{s_{\neg n}^*\} \times \{\text{NW}\} & \text{if } \Pi_n (s_{\neg n}^*, \alpha^*) \geq 0 \\
\mathcal{M} \times \{\text{NC}\} \times \{W\} \cup \emptyset \times \{C\} \times \{\text{NW}\} & \text{otherwise} 
\end{cases}$$

with

$$s_{n|C}^{1*} = \begin{cases} 
C & \text{if } \Pi_n^C (s_{\neg n}^*, \alpha^*) \geq \Pi_n^{\text{NW}} (s_{\neg n}^*, \alpha^*) \\
\text{NC} & \text{if } \Pi_n^C (s_{\neg n}^*, \alpha^*) < \Pi_n^{\text{NW}} (s_{\neg n}^*, \alpha^*)
\end{cases}$$

while for each $n \in \mathcal{N} \setminus \mathcal{N}_C (s^*)$

$$\lambda_{hn} (s_{1*}) = \begin{cases} 
\frac{\lambda_h}{|M_n(s)|} & \text{if } a_h (s_{1*}^{n|\mathcal{M}}) = \alpha_h^* (s^*) \\
0 & \text{otherwise} 
\end{cases} \quad h \in \{L, H\} \quad (4)$$

3.2 Efficient Insurance Provision

To show that the MWS allocation is the unique equilibrium of our mechanism, we will argue in two steps: by establishing it first as the unique candidate equilibrium outcome of our three-stage game, and by demonstrating subsequently that indeed it can be attained as an equilibrium.

For the first part of the argument, it suffices to show that a profitable deviation exists as long as a given candidate equilibrium menu $m^* = \{a_L^*, a_H^*\}$ does not solve the IE(1) problem. Formally, this hinges upon applying Lemma 3 to establish Proposition 1 below, for the statement of which it will be instructive to define first two particular sets of contracts under the typical firms’ strategic scenario $s \in S_f$. Namely, the collection of contracts that are preferred by the low risk-type out of all the menus that are honoured at stage 3,

$$A_L (s) = \{0\} \cup \arg \max_{m \in \mathcal{M}_{\text{NW}} (s)} U_L (a_L (m))$$

along with the collection of menus that are honoured at stage 3 and include contracts from $A_L (s)$

$$\mathcal{M}_L (s) = \{m \in \mathcal{M}_{\text{NW}} (s) : a_L (m) \in A_L (s)\}$$

Lemma 3 Let $m = \{a_L, a_H\}$ be any menu with $a_H \in \mathbb{R}^2$ and $a_L \in \mathbb{R}^2 \setminus \{a_L^1\}$. There exists a menu $\{\hat{a}_L, \hat{a}_H\} = \hat{m} \in \mathcal{M}$ with the following properties

(i) $\Pi_L (\hat{a}_L) > 0 \geq \Pi_H (\hat{a}_H)$, (ii) $\Pi (\hat{m}|\lambda) > 0$, and (iii) $\hat{a}_L \succ_L a_L$.

Proof. See Appendix C.

Proposition 1 Let $s^* \in S_f$ be part of an equilibrium strategic profile in $\Gamma$. This can be only if $\mathcal{M}_L (s^*) = \{m^1\}$. 

14
Proof. See Appendix C. 

Intuitively, the first part of the argument follows from the fact that, as long as \( m^* \neq m^1 \), by letting \( \mu \in (\lambda, 1) \) be sufficiently close to 1, we can guarantee that the IE(\( \mu \)) optimum constitutes a welfare increase from \( m^* \) for the low-risk type \( (a^\mu_L \succeq_L a^*_L) \). Suppose then that some firm deviates by introducing this IE menu at stage 1 with commitment. In this event, being guaranteed a strictly-better outcome, at stage 2 the low-risk customers cannot but choose the deviant menu over the equilibrium one. Being IE, moreover, the deviant menu avoids losses even in the worst case scenario in which the entirety of the high-risk customers choose it over the equilibrium one as well.

Of course, the deviation just described does not offer a clear incentive to potential challengers of the hypothetical equilibrium menu - in the worst case scenario, it expects zero, not strictly positive profits. Nonetheless, a strictly profitable deviant menu can be constructed by replacing the contract \( a^\mu_L \) by another \( \hat{a}_L \), in a way that the menu \( \hat{m} = \{\hat{a}_L, a^\mu_H\} \) is strictly-profitable even when chosen by the entire population of customers, while maintaining the two crucial elements of a successful challenge: \( \hat{m} \) remains separating and, relative to the equilibrium one, strictly-better for the low-risk customers (Figures 1-2). Equally importantly perhaps, the deviation remains strictly-preferred to the endowment lottery for either risk-type.

The remaining part of our argument consists of guaranteeing that the unique equilibrium candidate, the MWS menu \( m^1 = \{a^1_L, a^1_H\} \), is indeed an equilibrium. To this end, we establish first that, on the equilibrium path, this menu is offered without commitment. Next, we show that no menu is able to attract away from \( m^1 \) only the high-risk type and avoid losses doing so. As a result, the potentially profitable deviations from \( m^1 \) cannot but refer to menus \( \hat{m} \) such that \( a^L_L (\hat{m}) \succ_L a^1_L \) and, hence, \( \Pi (\hat{m} | \lambda) < 0 \). We need to distinguish now between two possible cases.

If the deviant menu \( \hat{m} \) is offered with commitment, to support the desired equilibrium it suffices that the firms offering \( m^1 \) at stage 1 withdraw it at stage 3. This is of course an optimal response on their part the MWS menu will not be selected by any agent of the low-risk type while \( \Pi_H (a^1_H) \leq 0 \). Anticipating this withdrawal, moreover, the high-risk customers cannot but opt also for \( a_H (\hat{m}) \) at
Figure 2: Deviations against the RSW menu (when not IE) and the Wilson pooling contract stage 2. Which turns, however, $\hat{m}$ into a menu that serves both types, rendering it a loss-making deviation.

For the case in which the deviant menu is offered with commitment, it will be instructive to introduce some more notation. For any given menu $m$ and any given belief $\{p, 1 - p\}$, viewing the profit $\Pi(m|p) = p\Pi_L(a_L(m)) + (1 - p)\Pi_H(a_H(m))$ as a $[0, 1] \mapsto \mathbb{R}$ function in $p$, we may define its fixed point $p(m)$ by

$$\Pi(m|p(m)) = 0$$

Needless to say, the relation $\Pi(m|p) - \Pi(m|p(m)) = [p - p(m)] [\Pi_L(a_L(m)) - \Pi_H(a_H(m))]$ defines the range of beliefs for which the menu is deemed non loss-making as follows\textsuperscript{22}

$$\mathcal{P}(m) = \begin{cases} [0, 1] & \text{if } \Pi_L(a_L(m)) = \Pi_H(a_H(m)) \\ [p(m), 1] & \text{if } \Pi_L(a_L(m)) > \Pi_H(a_H(m)) \\ [0, p(m)] & \text{otherwise} \end{cases}$$

Given this notation, $\Pi(\hat{m}|\lambda) < 0$ requires obviously that $\lambda \notin \mathcal{P}(\hat{m})$. Which implies in turn that, when the deviant menu $\hat{m}$ is offered with commitment, there are two subcases to examine: $\mathcal{P}(\hat{m}) = [p(\hat{m}), 1]$ and $\mathcal{P}(\hat{m}) = [0, p(\hat{m})]$; equivalently, $p(\hat{m}) > \lambda$ and $p(\hat{m}) < \lambda$, for some $p(\hat{m}) \in [0, 1]$.

\textsuperscript{22}Recall that $\Pi_L(a_L) \geq \Pi_H(a_H)$ with equality iff the RSW allocation is IE (Claim 1). As a result, it cannot but be

$$\mathcal{P}(m^1) = \begin{cases} [0, 1] & \text{if the RSW allocation is IE} \\ [\lambda, 1] & \text{otherwise} \end{cases}$$

with $\lambda \in \mathcal{P}(m^1)$ in either case.
As it turns out, when \( p(\hat{m}) > \lambda \), there are two possible equilibrium scenarios: one in which either of the menus \( m^1 \) and \( \hat{m} \) gets withdrawn, and another in which the latter is withdrawn but the former is not. To obtain the first scenario as a sequential equilibrium one, consistency of beliefs requires that the risk-type \( h \) applies for \( m^1 \) and \( \hat{m} \) with probability \( \sigma_h \) and \( 1 - \sigma_h \), respectively, where

\[
\sigma_h \in [0, 1]: \quad 1 \leq \frac{1 - \sigma_L}{1 - \sigma_H} \leq \frac{p(\hat{m})(1 - \lambda)}{(1 - p(\hat{m}))}\lambda
\]

More precisely, to verify that beliefs are consistent we consider sequences of trembles \( \{q^k_L, q^k_H\} \in (0, 1 - \sigma_h)^2 \) such that \( (q^k_L, q^k_H) \to (0, 0) \). Along such a sequence, an agent applies for \( m^1 \) and \( \hat{m} \) with probabilities \( \sigma_h^k = \sigma_h + q^k_h \) and \( 1 - \sigma_h^k \), respectively. Moreover, the conditional belief of either firm is given as above but for the fact that now \( \hat{m} \) gets withdrawn while \( m^1 \) does not is supported by letting \( \sigma_h = 1 \) for either type and considering the sequences of trembles \( \{r^k_L, r^k_H\} \in (0, 1)^2 \) such that \( (r^k_L, r^k_H) \to (0, 0) \). Along such a sequence, the agent applies for \( m^1 \) and \( \hat{m} \) with probabilities \( \sigma_h^k = 1 - r^k_h \) and \( 1 - \sigma_h^k \), respectively, while the conditional belief of either firm is given as above but for the fact that now \( \hat{m} \) gets withdrawn while \( m^1 \) does not.

\[
\lambda^k_{L,n} = \frac{\lambda^k L \lambda}{\lambda^k H \lambda + 1 - \lambda}, \quad \lambda^k_{H,n} = \frac{1 - \lambda}{\lambda^k H \lambda + 1 - \lambda}, \quad \lambda^k_n = \begin{cases} 
\frac{1 - \sigma_L - q^k_h}{1 - \sigma_H - q^k_H} & \text{if } n = \hat{n} \\
\frac{\sigma_L + q^k_H}{\sigma_H + q^k_H} & \text{if } n = n^* 
\end{cases}
\]

On the other hand, the scenario in which \( \hat{m} \) gets withdrawn while \( m^1 \) does not is supported by letting \( \sigma_h = 1 \) for either type while the trembles-induced beliefs are governed by 6. Equally obviously, taking \( \sigma_h = 0 \) for either \( h \) in the first scenario above, the preceding observations can be stated formally as follows.

**Theorem 1** The IE(1) optimum is supported as the equilibrium of \( \Gamma \) by the following (symmetric) strategy and belief profiles. Each firm \( n \in N \) chooses the action \( s^1_{n} = (m^1, NC) \) and follows the strategy \( s^2^* : S^1 \to \{W, NW\} \) given by

\[
s^2^* (s^1) = \begin{cases} 
W & \text{if } \exists n \in N \setminus \{n\} : a_L \left( s^1_{n|\Lambda} \right) \succ_L a_L \left( s^1_{n|\Lambda} \right), \quad s^1_{n|C} = C \\
W & \text{if } \exists n \in N \setminus \{n\} : a_L \left( s^1_{n|\Lambda} \right) \succ_L a_L \left( s^1_{n|\Lambda} \right), \quad s^1_{n|C} = NC, \quad \lambda < p \left( s^1_{n|\Lambda} \right) \\
NW & \text{otherwise}
\end{cases}
\]

\(^{23}\)It is noteworthy that the conjunction of the strategies \( \sigma_L = 1 = \sigma_H \) and the belief formation in 6 can be shown to support also the scenario in which either of the menus \( m^1 \) and \( \hat{m} \) gets withdrawn.
or

\[
(b) \quad s_n^2(s^1) = \begin{cases} 
W & \text{if } \exists n \in N \setminus \{n\} : a_L(s^1_n) \succeq_L a_L(s^1_{n(M)}) \text{, } s^1_{n(M)} = C \\
\text{NW} & \text{otherwise}
\end{cases}
\]

Moreover, each customer follows the strategy in (3) while the associated beliefs by the firms are as in (4).

Proof. See Appendix C. ■

3.2.1 The Rothschild-Stiglitz Equilibrium [to be polished]

Suppose, in particular, that at stage 1 of our game the firms are restricted to choosing menus that (i) do not involve cross-subsidization, and (ii) cannot be offered without commitment. Formally, the first requirement restricts the admissible set to menus \( \{a_L, a_H\} \in \mathbb{R}^4_+ \) that satisfy (IC\( h \)) but also \( \Pi_h(a_h) \geq 0 \) for either \( h \). 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.

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 [34]. First of all, as shown by Lemmas 4-5 below, the RSW allocation is the only equilibrium candidate allocation. To see intuitively first why it is not possible to have pooling policies in the equilibrium set, notice that a hypothetical pooling equilibrium policy \( a^* \) 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.\(^{24}\) Yet, the very fact that strictly positive profits are extracted by the low-risk type allows for the existence of another contract \( a_L^1 \) which delivers strictly positive profits if accepted only by low-risk agents and is such that \( a_L^1 \succeq_L a^* \succeq_H a_L^1 \). This is shown formally by Lemma 4 and graphically by any point in the interior of the shaded areas in Figures 3-4. In the presence of \( a^* \), therefore, \( a_L^1 \) 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.

**Lemma 4** The menu \( \{a_L, a_H\} \) is an equilibrium of the RS game only if (i) it is separating, and (ii) \( \Pi_h(a_h) = 0 \) for either \( h \).

---

\(^{24}\)Given free entry, we ought to have \( \Pi_M(a^*) \geq 0 \). Yet, this cannot be a strict inequality. For if \( \Pi_M(a^*) = \epsilon > 0 \), we may consider the contract \( \hat{a} = a^* - (1, -1) \frac{\epsilon}{2} \) which is such that \( \hat{a} \succeq_h a^* \) by either \( h \) (it provides strictly more income in either state of the world). In the contingency, therefore, in which \( a^* \) and \( \hat{a} \) 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 \( \Pi_M(\hat{a}) = \Pi_M(a^*) - \frac{\epsilon}{2} > 0 \). It constitutes, that is, a profitable deviation, contradicting part (b) of the definition for \( a^* \) to be a market equilibrium. To arrive at the claim in the text, notice that \( p_H > p_L \) requires \( \Pi_H(a^*) < \Pi_L(a^*) \). Clearly, \( \Pi_M(a^*) = 0 \) only if \( \Pi_H(a^*) < 0 < \Pi_L(a^*) \).
Figure 3: Deviations against pooling policies

Figure 4: Deviations against pooling policies

**Proof.** See Appendix D. ■

An equilibrium policy, therefore, cannot be but a separating menu. Amongst the admissible ones, though, the only legitimate candidate is the RSW allocation.

**Lemma 5** The RSW menu \( \{a^*_L, a^*_H\} \) is the unique equilibrium candidate of the RS game.

**Proof.** See Appendix D. ■

To see intuitively why no other separating menu can be an equilibrium, suppose otherwise and let \( \{a_L, a_H\} \neq \{a^*_L, a^*_H\} \) 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, \( a^*_H >_H a_H \). This necessitates the existence of another contract \( a^2_L \) such that \( a^2_L >_L a_L, a^*_H \) but \( a^*_H >_H a^2_L \) and which delivers strictly-positive profits if chosen only by low-risk agents. Consider now a firm offering the menu \( \{a^2_L, a^*_H\} \). This is separating and attracts...
either risk-type away from \( \{a_L, a_H\} \). Doing so, moreover, it breaks-even on the high-risk agents but is strictly profitable on the low-risk ones. Examples of \( a_L^2 \) are given by the interior points of the shaded area in either diagram of Figure 5.

![Figure 5: Deviations against a non-RS separating menu](image)

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 \( FO^E_M \) cuts through the low-risk indifference curve associated with \( a_L^{**} \). Formally, the RSW menu is an equilibrium here if and only if there exists no contract the low-risk type prefers strictly to \( a_L^{**} \) and which delivers zero profits as pooling policy.

**Claim 3** Restricting admissibility to singleton or safe menus (i.e., ruling out cross-subsidization), the equilibrium of the RS game exists if and only if \( \nexists a \in A : a \succ_L a_L^{**} \text{ and } \Pi_M(a) \geq 0 \).

**Proof.** See Appendix D.  

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 \( a_L^{**} \) by the low-risk type, we can construct profitable deviations against the RSW menu. These are contracts \( a^2 \) that are strictly profitable as pooling policies and attract at least the low-risk type away \( (a^2 \succ_L a_L^{**}) \). Obviously, if they pull away also the high-risk agents \( (a^2 \succ_H a_H^{**}) \), they are strictly profitable pooling deviations. Otherwise, the high-risk type opts to leave \( a^2 \) 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 6.

Yet, the preceding result hinges crucially upon the very fact that we restrict attention to singleton or safe menus - in other words, upon the fact that we do not allow for cross-subsidization.\(^{25}\)

If we allow for more general menus, the issue of existence coincides with that of efficiency.

\(^{25}\)Since the RSW allocation is IE only if it is IE(1) while a pooling policy cannot be IE(1), Claim 4 implies also
Claim 4  *Allowing for general menus, the equilibrium of the RS game exists if and only if the RSW allocation is IE.*

**Proof.** See Appendix D. ■

Finally, the issue in question has bite.

**Proposition 2** Let $\mathcal{X} \subset \mathcal{U} \times (0, 1)^3$ be the subset of parameters for which the following conditions apply simultaneously.

(i) The RSW allocation is not IE, and

(ii) The RSW menu is the equilibrium of the RS game

Then, $\mathcal{X}$ is non-null.

**Proof.** See Appendix D. ■

4 Discussion and Related Literature [to be re-written]

Known in the literature (see Crocker and Snow [8]) as the Miyazaki-Wilson-Spence allocation, our equilibrium outcome was established first by Miyazaki [27] 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 [38], 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 that the RSW allocation is IE only if it exists as equilibrium in Rothschild and Stiglitz [34]. This is a widely-used assertion which, to the best of our knowledge, has not been shown formally before.
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 [20] 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 IE(1) allocation as the unique outcome, albeit of a Wilson equilibrium in one but sequential in the other. 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 [18]. 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

\footnote{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 IE(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 IE(1) allocation from ever solving the full-information efficiency problem.}
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. 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 ?? 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 IE(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 IE(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 IE(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 loses. 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, IE is the appropriate criterion since there cannot be unanimous agreement to depart from an IE decision rule if some individual knows just her own private information (see Theorem 1 and the subsequent discussion in Holmstrom and Myerson [21]).

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 (??) 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 IE problem, where the weights depend only on the type of the informed party precisely because interim efficiency is the relevant concept. In fact, the two optimality problems of the preceding section define, respectively, the RSW and interim incentive efficient allocations in Maskin and Tirole [25], the reservation allocation being the null trade.

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 IE allocation meets (??) with equality and whatever the value of \( \lambda \), it does satisfy condition (iv) of Maskin and Tirole’s Proposition 7. Any allocation, therefore, is an equilibrium one as long as it satisfies (??), (11), and (??), 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 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 IE problem and (weakly) Pareto-dominate the RSW allocation (see Propositions 13 and 6 in Maskin and Tirole [25]). 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 ??, 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 (??) 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 [25]: 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 IE(1) allocation apart from a special case (when the RS menu is an IE and, consequently by Claim 2, the IE(1) allocation). Yet, the latter is the IP’s only reasonable choice when she is called upon to propose mechanisms. This follows from Myerson [28]. 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 \( \{a_L, a_H\} \) (separating or trivial) as long as it is incentive compatible and satisfies \((??)\), if it is selected by both risk types, or the relevant condition in \((??)\), otherwise. Within the realm of these restrictions, the IE(1) allocation is the unique selection of the low-risk type and, by satisfying \((??)\) and \(\Pi_L(a^*_L) > 0\), implementable if \(a^*_h \succ_h a_h\) for either \(h\) or \(a^*_L \succ_h a_L\) but \(a_H \succ_H a^*_H\). In the latter case, moreover, the new proposal comes exclusively from the high-risk type and implementability is ruled out as \(a_H \succ_H a^*_H \succ_H a^{**}_H\) necessitates that \(\Pi_H(a_H) < 0\). Clearly, the IE(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 [28]. 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 contractual 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. 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. [19], 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


---

27 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 $\alpha_5$ in Figure 5 lies on or above the 45-degree line. Yet, as we saw in Section ??, this equilibrium cannot survive the intuitive criterion.


Appendices

A Preliminaries

Lemma 6 Let the contracts \( a' \) and \( a'' \) be such that \( a'' = a' + (\kappa, 1)\epsilon \) for some \( \kappa, \epsilon \in \mathbb{R}^* \). Then, the corresponding income allocations \( \left(w''_0, w''_1\right) = \left(W - a'_0, W - d + a'_1\right) \) (\( j = 0, 1 \)) are such that

\[
U_h \left(a''\right) - U_h \left(a'\right) = \left\{
\begin{array}{ll}
[phu' (w'_1 + \epsilon) - \kappa (1 - ph) u' (w'_0 - \kappa \epsilon)] & \text{if } \epsilon > 0 \\
[phu' (w'_1 - \epsilon) - \kappa (1 - ph) u' (w'_0 + \kappa \epsilon)] & \text{if } \epsilon < 0
\end{array}
\right.
\]

for some \( \epsilon \in (0, |\epsilon|) \) and either \( h \in \{H, L\} \).

Proof. Observe first that \( w'' = w' - (\kappa, -1)\epsilon \) gives

\[
U_h \left(a''\right) - U_h \left(a'\right) = (1 - ph) [u (w''_0) - u (w'_0)] + ph [u (w''_1) - u (w'_1)]
\]

Define then the function \( H(z) \equiv (1 - ph) u (w'_0 - \kappa z) + ph u (w'_1 + z) \). As \( U_h \left(a''\right) - U_h \left(a'\right) = H(\epsilon) - H(0) \), the claim is an immediate result of the mean value theorem. 

Lemma 7 Let the contracts \( a' \) and \( a'' \) be as in Lemma 6 and suppose that one of the following conditions holds

(i) \( \kappa \in \left(0, \frac{ph}{1 - ph}\right], \epsilon > 0 \), and \( w''_0 \geq w''_1 \)

(ii) \( \kappa \in \left[\frac{ph}{1 - ph}, \infty\right), \epsilon < 0 \), and \( w''_0 \leq w''_1 \)

(iii) \( \kappa = -1, \epsilon > 0 \), and \( w''_0 = w''_1 \)

Then, \( a'' >_h a' \) for either \( h \in \{H, L\} \).

Proof. The claim follows from the preceding lemma and risk-aversion \( (u''(\cdot) < 0) \). This is obvious under (iii) since \( U_h \left(a''\right) - U_h \left(a'\right) = ph \left[u' (w'_0 + \epsilon) - u' (w'_0 - \kappa \epsilon)\right] \). Under (i), on the other hand, \( \kappa \leq \frac{ph}{1 - ph} \) means that \( U_h \left(a''\right) - U_h \left(a'\right) \geq ph \left[u' (w'_0 + \epsilon) - u' (w'_0 - \kappa \epsilon)\right] \). Yet, \( (w''_0, w''_1) = (w'_0, w'_1) - (\kappa, -1)\epsilon \) and, thus, \( w'_1 + \epsilon < w'_1 + \epsilon = w''_1 \leq w''_0 = w'_0 - \kappa \epsilon < w'_0 - \kappa \epsilon \).
Similarly, under (ii), $U_h(a'') - U_h(a') \geq p_h [u'(w'_1 - \bar{\epsilon}) - u'(w'_0 + \bar{\epsilon})] \epsilon$ for some $\bar{\epsilon} \in (0, |\epsilon|)$. And now $(w''_0, w''_1) = (w'_0, w'_1) + (\kappa, -1) |\epsilon|$ implies that $w'_1 - \bar{\epsilon} > w'_1 - |\epsilon| = w''_1 \geq w''_0 = w'_0 + \kappa |\epsilon| > w'_0 + \bar{\epsilon}$.

For the next two results, we will denote by

$$I_h(a) = \frac{\partial U_h(W - a_0)}{\partial w_0} \frac{d w_0}{d a_0} = \left(\frac{1 - p_h}{p_h}\right) \frac{u'(W - a_0)}{u' (W - d + a_1)}$$

the slope of the indifference curve of risk-type $h$ at an arbitrary point in the $(a_0, a_1)$-space. We have of course $0 < I_H(a) < I_L(a)$ for any $a \in A$.

**Lemma 8** Let $h, h' \in \{H, L\}$ with $h \neq h'$. For every $\kappa \in (I_H(a_L), I_L(a_L))$ there exists $\epsilon_0 > 0$ such that, for all $\epsilon \in (-\epsilon_0, 0)$ if $h = L$ [resp. $\epsilon \in (0, \epsilon_0)$ if $h = H$], the contract $a'_h = a_h + (1, \kappa) \epsilon$ gives $a'_h \succ_h a_h \succ_{h'} a'_h$.

**Proof.** Suppose first that $h = L$ and consider the contract $a'_L = a_L + (1, \kappa) \epsilon$ for some $\kappa \in (I_H(a_L), I_L(a_L))$ and $\epsilon < 0$. Define also $\Delta_{h''} = |\kappa - I_{h''}(a_L)|$ for $h'' \in \{H, L\}$. By Lemma 6, we get

$$U_h(a'_L) - U_h(a_L) = \left[ p_h u'_h (w_{1L} - \bar{\epsilon}) - \kappa^{-1} (1 - p_h) u'_h (w_{0L} + \kappa^{-1} \bar{\epsilon}) \right] \kappa \epsilon$$

$$= \left[ \kappa - \frac{(1 - p_h) u'_h (w_{0L} + \kappa^{-1} \bar{\epsilon})}{p_h u'_h (w_{1L} - \bar{\epsilon})} \right] p_h u'_h (w_{1L} - \bar{\epsilon}) \epsilon$$

$$= \left[ \kappa - I_h(a_L - (\kappa^{-1}, 1) \bar{\epsilon}) \right] p_h u'_h (w_{1L} - \bar{\epsilon}) \epsilon$$

for some $\bar{\epsilon} \in (0, |\kappa| |\epsilon|)$. Yet, the function $I_h(x) : \mathbb{R} \mapsto \mathbb{R}_{++}$ with $I_h(x) = I_h(a_L - (\kappa^{-1}, 1) x)$ is continuous and $\lim_{x \to 0} I_h(x) = I_h(a_L)$. There exists, therefore, some sufficiently small $\bar{\epsilon}_0 > 0$ so that $|I_h(a_L - (\kappa^{-1}, 1) \bar{\epsilon}) - I_h(a_L)| < \min \{\Delta_L, \Delta_H\}$ for all $\bar{\epsilon} \in (0, \bar{\epsilon}_0)$. But then, we ought to have

$$U_L(a'_L) - U_L(a_L) = \left[ I_L(a_L) - \Delta_L - I_L(a_L - (\kappa^{-1}, 1) \bar{\epsilon}) \right] p_L u'_L (w_{1L} - \bar{\epsilon}) \epsilon$$

$$= - I_L(a_L - (\kappa^{-1}, 1) \bar{\epsilon}) - I_L(a_L) + \Delta_L \right] p_L u'_L (w_{1L} - \bar{\epsilon}) \epsilon > 0$$

$$U_H(a'_L) - U_H(a_L) = \left[ I_H(a_L) - I_H(a_L - (\kappa^{-1}, 1) \bar{\epsilon}) + \Delta_H \right] p_H u'_H (w_{1L} - \bar{\epsilon}) \epsilon < 0$$

for some $\bar{\epsilon} \in (0, |\kappa| |\epsilon|)$ and all $\epsilon \in (-\bar{\epsilon}_0/\kappa, 0)$. Setting $\epsilon_0 = \bar{\epsilon}_0/\kappa$ gives the result. When $h = H$, the argument is trivially similar. 

**Lemma 9** Let $h, h' \in \{H, L\}$ with $h \neq h'$. Suppose also that $a'$ and $a^*$ are such that $a' \succ_{h'} a^*$. Then, for every $\kappa \in (-\infty, I_H(a'))$ [resp. $\kappa \in (I_L(a'), +\infty)$] there exists $\epsilon_0 > 0$ such that, for all $\epsilon \in (0, \epsilon_0)$ [resp. $\epsilon \in (-\epsilon_0, 0)$], the contract $a = a' + (1, \kappa) \epsilon$ gives $a' \succ_{h'} a \succ_{h'} a^*$ and $a' \succ_h a$.

**Proof.** Let $\Delta = U_{h'}(a') - U_{h'}(a^*) > 0$ and consider the contract $a = a' + (1, \kappa) \epsilon$ first for some $\kappa \in (-\infty, I_H(a'))$ and $\epsilon > 0$. Let also $\Delta_{h''} = I_{h''}(a') - \kappa$ for $h'' \in \{H, L\}$. Clearly, $\Delta_L > \Delta_H > 0$.
while by Lemma 6 we have

\[
U_{h''}(a) - U_{h''}(a') = \left[ p_{h''}u \left( w'_1 + \hat{\epsilon}_{h''} \right) - \kappa^{-1} \left( 1 - p_{h''} \right) u' \left( w'_0 - \kappa^{-1} \hat{\epsilon}_{h''} \right) \right] \kappa \epsilon
\]

\[
= \left[ \kappa - \frac{\left( 1 - p_{h''} \right) u' \left( w'_0 - \kappa^{-1} \hat{\epsilon}_{h''} \right) - \hat{\epsilon}_{h''} \kappa \epsilon}{p_{h''}u' \left( w'_1 + \hat{\epsilon}_{h''} \right)} \right] p_{h''}u' \left( w'_1 + \hat{\epsilon}_{h''} \right) \epsilon
\]

\[
= \left[ \kappa - I_{h''} \left( a' + \left( \kappa^{-1}, 1 \right) \hat{\epsilon}_{h''} \right) \right] p_{h''}u' \left( w'_1 + \hat{\epsilon}_{h''} \right) \epsilon
\]

\[
= \left[ I_{h''} \left( a' \right) - \Delta_{h''} - I_{h''} \left( a' - \left( \kappa^{-1}, 1 \right) \hat{\epsilon}_{h''} \right) \right] p_{h''}u' \left( w'_1 + \hat{\epsilon}_{h''} \right) \epsilon
\]

for some \( \hat{\epsilon}_{h''} \in (0, \kappa \epsilon) \). Yet, the function \( I_{h''}(x) : \mathbb{R} \mapsto \mathbb{R}_{++} \) with \( I_{h''}(x) = I_{h''}(a' + (\kappa^{-1}, 1) x) \) is continuous and \( \lim_{x \to 0} I_{h''}(x) = I_{h''}(a') \). For small enough \( \bar{\epsilon}_{h''} > 0 \), therefore, we ought to have \( |I_{h''}(a' + (\kappa^{-1}, 1) \hat{\epsilon}_{h''}) - I_{h''}(a')| < \Delta_{h''} \) for all \( \hat{\epsilon}_{h''} \in (0, \bar{\epsilon}_{h''}) \). Consequently, \( U_{h''}(a) < U_{h''}(a') \) for any \( \epsilon \in (0, \bar{\epsilon}_{h''}/\kappa) \). Observe, however, that \( U_{h''}(a) - U_{h''}(a') \) itself vanishes as \( \epsilon \to 0 \). Clearly, with respect to the risk-type \( h' \), a sufficiently small \( \bar{\epsilon}_{h'} \) ensures also that \( U_{h'}(a) - U_{h'}(a') \in (-\Delta, 0) \) or \( U_{h'}(a) > U_{h'}(a^*) \). The result now follows by letting \( \epsilon_0 = \min \{ \bar{\epsilon}_{h}, \bar{\epsilon}_{h'} \} / \kappa \). The argument for the case \( \kappa \in (I_L(a'), +\infty) \) is trivially similar. \( \blacksquare \)

**Lemma 10** Let \( \mathcal{D} \subseteq \mathbb{R}^n \) be a choice set such that the problem \( \max_{x \in \mathcal{D}} \mu f(x) + (1 - \mu) g(x) \) 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.\(^{28}\) For any function \( x^* : \mathcal{M} \mapsto S(\mu) \), we have

\[ f(x^*(\mu_1)) - f(x^*(\mu_2)) \geq 0 \geq g(x^*(\mu_1)) - g(x^*(\mu_2)) \quad \forall \mu_1, \mu_2 \in [0, 1] : \mu_1 > \mu_2 \]

with both inequalities strict if \( S(\mu_1) \cap S(\mu_2) = \varnothing \).

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

\[
\mu_1 \left[ f(x^*(\mu_1)) - f(x^*(\mu_2)) \right] + (1 - \mu_1) \left[ g(x^*(\mu_1)) - g(x^*(\mu_2)) \right] \geq 0
\]

\[
\mu_2 \left[ f(x^*(\mu_2)) - f(x^*(\mu_1)) \right] + (1 - \mu_2) \left[ g(x^*(\mu_2)) - g(x^*(\mu_1)) \right] \geq 0
\]

That is,

\[
\left( \mu_1 - \mu_2 \right) \left[ f(x^*(\mu_1)) - f(x^*(\mu_2)) \right] + (1 - \mu_1 - \mu_2) \left[ g(x^*(\mu_1)) - g(x^*(\mu_2)) \right] \geq 0
\]

and \( \mu_1 > \mu_2 \) requires that

\[
f(x^*(\mu_1)) - f(x^*(\mu_2)) \geq g(x^*(\mu_1)) - g(x^*(\mu_2))
\]

The required result holds trivially if at least one side of (10) is zero. Let, thus, neither side of (10) be zero. If \( \mu_1, \mu_2 \in (0, 1) \), the quantities on either side of (10) cannot be of the same sign; otherwise, one of (7)-(8) will be violated. Obviously, the larger [resp. smaller] one must be strictly positive [resp. strictly negative in this case. If \( \mu_1 = 1 \), on the other hand, it must be \( f(x^*(\mu_1)) > f(x^*(\mu_2)) \) in

\(^{28}\) \( \mathcal{D} \) and \( \mathcal{M} \) being independent is meant here to say that the parameter \( \mu \) does not enter any of the constraints.
(7). And as \( \mu_2 < 1 \), it follows from (8) that \( g(x^*(\mu_1)) < g(x^*(\mu_2)) \). Finally, if \( \mu_2 = 0 \), we ought to have \( g(x^*(\mu_1)) < g(x^*(\mu_2)) \) in (8) and, hence, \( f(x^*(\mu_1)) > f(x^*(\mu_2)) \) in (7) given that \( \mu_1 > 0 \).

Now, if \( S(\mu_1) \cap S(\mu_2) = \emptyset \), then \( x^*(\mu_1) \) [resp. \( x^*(\mu_2) \)] is not optimal when the parameter is \( \mu_2 \) [resp. \( \mu_1 \)]. In this case, all four inequalities (7)-(10) are strict and the result follows from two observations. Neither side of \( f(x^*(\mu_1)) - f(x^*(\mu_2)) > g(x^*(\mu_1)) - g(x^*(\mu_2)) \) can be zero; otherwise, one of the strict versions of (7)-(8) will be violated. For the same reason, the two sides cannot be of the same sign. Needless to say, the larger [resp. smaller] one must be strictly positive [resp. strictly negative].

**Lemma 11** Let \( A^-_h = \{a \in A : a_0 = d - a_1, \Pi_h(a) \leq 0 \} \) be the collection of contracts that offer full insurance but are not profitable against the risk-type \( h \in \{H, L\} \). Then,

\[
U_h(a) > \overline{u}_h \quad \forall a \in A^-_h
\]

**Proof.** For any \( h \in \{H, L\} \) and any \( a \in A^-_h \), we have

\[
\overline{u}_h = (1 - p_h)u(W) + p_hu(W - d) < u((1 - p_h)W + p_h(W - d)) = u(W - p_hd) \\
\leq u(W - p Hd - \Pi_h(a)) = u(W - (1 - p_h)a_0 + p_h(a_1 - d)) = u((1 - p_h)w_0 + p_hw_1) = u(w_0) = U_h(a)
\]

The first inequality above is due to \( u(\cdot) \) being everywhere strictly-concave. The second follows from the hypothesis that \( \Pi_h(a) \geq 0 \) in conjunction with the fact that \( u(\cdot) \) is everywhere strictly-increasing. Finally, the last two equalities use that \( a \) offers full insurance.

**Lemma 12** Suppose that the menu \( \{a_L, a_H\} \) is such that the low-risk type is not over-insured \( (w_{0L} \geq w_{1L}) \) while the high-risk type is fully-insured \( (w_{0H} = w_{1H}) \) and her IC constraint binds \( (U_H(a_H) = U_H(a_L)) \). Then, we have

(i) \( a_{0L} - a_{0H} \leq 0 \leq a_{1L} - a_{1H} \), and

(ii) \( \Pi_H(a_H) \leq \Pi_L(a_L) \).

**Proof.** To establish (i), observe first that the high-risk type being fully-insured with the IC constraint binding means that

\[
u(w_{0H}) = U_H(a_H) = U_H(a_L) = (1 - p_H)u(w_{0L}) + p_Hu(w_{1L}) \leq u(w_{0L})
\]

By non-satiation, therefore, we ought to have \( w_{0L} \geq w_{0H} \) or \( a_{0L} \leq a_{0H} \). Similarly,

\[
u(w_{1H}) = U_H(a_H) = U_H(a_L) \geq u(w_{1L})
\]
gives $w_{1L} \geq w_{1H}$ or $a_{1L} \geq a_{1H}$.

Part (ii) now follows since

$$\Pi_H (a_H) = (1 - p_H) a_{0H} - p_H a_{1H} \leq (1 - p_H) a_{0L} - p_H a_{1L} = \Pi_L (a_L)$$

where the first inequality deploys the ranking between $a_{sL}$ and $a_{sH}$ for $s \in \{0, 1\}$ from (i) while the second uses that $p_H > p_L$. ■

We conclude this section with a small digression on the following convex cone

$$C_z = \{ x \in \mathbb{R}^2 : x_2 - z_2 = l (x_1 - z_1) , \ l \in [l_1, l_2] \}$$

where $l_1, l_2 \in \mathbb{R}$ with $l_1 < l_2$ while $z \in \mathbb{R}^2$. This cone is the intersection of two half-planes, the one to the right of the line $\{ x \in \mathbb{R}^2 : (x_2 - z_2) = l_1 (x_1 - z_1) \}$ and the one to the left of $\{ x \in \mathbb{R}^2 : (x_2 - z_2) = l_2 (x_1 - z_1) \}$. It is pointed at $z$ and separated into two half-cones: $C_z^+ = \{ x \in C_z : x_1 \geq z_1 \}$ and $C_z^- = \{ x \in C_z : x_1 \leq z_1 \}$.

**Lemma 13** Let $x, y \in C_z$. Obviously, there must exist $r_1, r_2 \in [l_1, l_2]$ such that $x_2 - z_2 = r_1 (x_1 - z_1)$ and $y_2 - z_2 = r_2 (y_1 - z_1)$. Suppose also that $r_1 > r_2$ and $x_1 \leq y_1$. Then $y_2 - x_2 \leq r_2 (y_1 - x_1)$ [resp. $y_2 - x_2 \geq r_1 (y_1 - x_1)$] implies that $x, y \in C_z^+$ [resp. $x, y \in C_z^-$]. In fact, $x, y \in C_z^+ \setminus \{ z \}$ [resp. $x, y \in C_z^- \setminus \{ z \}$] if the inequality is strict.

**Proof.** We will establish the claim for the $x, y \in C_z^+$ case. To this end, consider that, under $y_2 - z_2 = r_2 (y_1 - z_1)$, the relation $y_2 - x_2 \leq r_2 (y_1 - x_1)$ can be re-written as $z_2 - x_2 + r_2 (y_1 - z_1) \leq r_2 (y_1 - x_1)$. Which is equivalent to $x_2 - z_2 \geq r_2 (x_1 - z_1)$. Since $r_2 < r_1$, however, this can be only if $x_1 \geq z_1$. Which, under $x_1 \leq y_1$, requires in turn that also $y_1 \geq z_1$. Obviously, if $y_2 - x_2 < r_2 (y_1 - x_1)$, we get $z_1 < x_1$. The argument to show that $x, y \in C_z^-$ is trivially similar. ■

**B Efficiency**

**Lemma 14** For any $\mu \in [0, 1]$, the $IE(\mu)$ optimum satisfies the following conditions.

(i) $\Pi (m^\mu | \lambda) = 0$,

(ii) $a_H^\mu \succeq_H a_H^{**}$, and

(iii) $\Pi_L (a_L^\mu) \geq 0 \geq \Pi_H (a_H^\mu)$, with either inequality strict unless $a_H^\mu = a_H^{**}$.

**Proof.** Letting $\beta_h$ and $\delta$ be, respectively, the Lagrangean multipliers on the IC$_h$ and $\Pi_M (m^\mu | \lambda) \geq 0$
If \( \mu \) above can be zero. But then we ought to have \( \delta \mu = 0 \); \( \mu \) then must be \( 0 \), and \( (11)-(12) \) read
\[
(\mu + \beta_1^\mu)u'(w_{0L}^\mu) = \beta_H^1 \frac{1-p_H}{1-p_L} u'(w_{0L}^\mu) + \delta \mu \lambda
\]
(11)
\[
(\mu + \beta_1^\mu)u'(w_{1L}^\mu) = \beta_H^1 \frac{p_H}{p_L} u'(w_{1L}^\mu) + \delta \mu \lambda
\]
(12)
\[
(1 - \mu + \beta_H^1)u'(w_{0H}^\mu) = \beta_H^1 \frac{1-p_H}{1-p_L} u'(w_{0H}^\mu) + \delta \mu (1 - \lambda)
\]
(13)
\[
(1 - \mu + \beta_H^1)u'(w_{1H}^\mu) = \beta_H^1 \frac{p_H}{p_L} u'(w_{1H}^\mu) + \delta \mu (1 - \lambda)
\]
(14)
\[
\beta_H^1 (U_h(a_H^\mu) - U_h(a_h^\mu)) = 0 \quad h, h' \in \{H, L\}
\]
(15)
\[
\delta \mu \left[ \lambda \Pi_L(a_L^\mu) + (1 - \lambda) \Pi_H(a_H^\mu) \right] = 0
\]
(16)
\[
\beta_H^1, \delta \mu \geq 0 \quad h \in \{H, L\}
\]
(17)
along with \( (IC_h) \) for either \( h \) and \( \Pi_M(m^\mu|\lambda) \geq 0 \).

To show (i), suppose that the profit constraint is slack at the optimum. Then it must be \( \delta \mu = 0 \) in (16), and (11)-(12) read
\[
(\mu + \beta_1^\mu) (1 - p_L) u'(w_{0L}^\mu) = \beta_H^1 (1 - p_H) u'(w_{0L}^\mu)
\]
\[
(\mu + \beta_1^\mu) p_L u'(w_{1L}^\mu) = \beta_H^1 p_H u'(w_{1L}^\mu)
\]
If \( \mu > 0 \), however, due to \( u'(\cdot) > 0, p_h \in (0,1) \) for either \( h \), and (17) neither side in either equation above can be zero. But then we ought to have \( \frac{1-p_L}{p_L} = \frac{1-p_H}{p_H} \), a contradiction. Hence, the profit constraint may be slack only if \( \mu = 0 \). Yet, a trivially similar argument, using now (13) and (14), shows that the constraint in question may be slack only if \( \mu = 1 \).

To show next (ii), let first \( \mu = 1 \) and suppose that \( a_{H_L}^{\mu*} \succ_H a_1^H \). As it must be \( a_1^L \succ_L a_1^L \approx_H a_{H_L}^{\mu*} \approx_H a_1^H \approx_L a_1^L \), the menu \( \hat{m} = \{a_1^L, a_{H_L}^{\mu*}\} \) satisfies the IC constraints of the IE(1) problem; hence, there are two cases to consider. If \( \Pi_M(\hat{m} | \lambda) \geq 0 \), the new menu is optimal for the IE(1) problem and, given (i) above, \( \Pi_H(a_{H_L}^{\mu*}) = 0 \) requires that also \( \Pi_L(a_1^L) = 0 \). In this case, however, the IE(1) problem would coincide with the RSW(1) one. Yet, the optimum in the latter problem being unique, this cannot be unless \( a_{H_L}^{\mu*} = a_1^H \) for either \( h \). Which contradicts, of course, our hypothesis that \( a_{H_L}^{\mu*} \succ_H a_1^H \). Suppose then that \( 0 > \Pi_M(\hat{m} | \lambda) = \lambda \Pi_L(a_1^L) + (1 - \lambda) \Pi_H(a_1^H) \). Equivalently, \( \Pi_L(a_1^L) < 0 \) since \( \Pi_L(a_{H_L}^{\mu*}) = 0 \). By (i) again, however, \( \Pi_L(a_1^L) < 0 \) can only be if \( \Pi_H(a_1^H) > 0 \).

That is, \( \Pi_H(a_1^H) > \Pi_L(a_1^L) \). which Lemma 12, however, renders absurd given that the low-risk [resp. high-risk] types get under [resp. full] insurance IE(1) optimum while the IC\(_H\) constraint binds. We just established, therefore, that \( a_1^H \succ_H a_{H_L}^{\mu*} \). Which implies in turn that \( a_{H_L}^{\mu*} \succ_H a_1^H \) also for any \( \mu \in [0,1) \) (Lemma 10).

To establish finally (iii), recall that \( a_{H_L}^{\mu*} \) is the unique solution to the problem \( \max_{a \in \mathbb{R}^2} \Pi_H(a) \geq 0 U_H(a) \), giving \( \Pi_H(a_{H_L}^{\mu*}) = 0 \). Since \( a_{H_L}^{\mu*} \approx_H a_{H_L}^{\mu*} \), it cannot but be \( \Pi_H(a_{H_L}^{\mu*}) \leq 0 \) for all \( \mu \in [0,1] \), with equality only if \( a_{H_L}^{\mu*} = a_1^H \). And given this, it follows from (i) above that \( \Pi_L(a_1^H) \geq 0 \) for all \( \mu \in [0,1] \), again with equality only if \( a_{H_L}^{\mu*} = a_1^H \). \( \blacksquare \)
which is zero by Lemma 14(i). To show next that $a^\mu$ is defined uniquely by (2) is obvious. As a pooling policy, moreover, it expects profits

$$\Pi_M(a^\mu) = \lambda [(1-p_L) a^\mu_L - p_L a^\mu_H] + (1-\lambda) [(1-p_H) a^\mu_H - p_H a^\mu_L]$$

which is zero by Lemma 14(i). To show next that $a^\mu_0 \geq 0$ notice that

$$\left(\frac{1-p_L}{p_L}\right) a_0 = \frac{1}{p_L} [(1-p_L) a_0L - p_L a_1L + p_L a_1] = \frac{1}{p_L} [\Pi_L(a_L) + p_L a_1]$$

$$\geq a_1 = \frac{1}{p_H} [(1-p_H) a_0 - \Pi_H(a_H)] \geq \left(\frac{1-p_H}{p_H}\right) a_0$$

where the first and third equalities are due to the respective relations in (2) while the two inequalities follow from Lemma 14(ii). Given this and $p_H > p_L$, it must be $a^\mu_0 \geq 0$ as required.

Lemma 15 For any $\mu \in [0,1]$, the IE($\mu$) optimum is unique.

Proof. The feasible set of the IE($\mu$) problem is convex in $(a_L, a_H)$ (recall footnote 11 in the main text) for any $\mu \in [0,1]$. Given that the objective is also strictly concave in $(a_L, a_H)$ for any $\mu \in (0,1)$, uniqueness of the optimum follows by a well-known result (see, for instance, Theorem 1.20 in de la Fuente [9]). By the same result, $a^\mu_H$, the high-risk optimal contract for the IE(0) problem is also unique. And given this, the uniqueness of $a^\mu_0$ follows from the fact that the profit and IC_L constraints bind at the IE(0) optimum. To see this, suppose that $\{a_L, a^\mu_H\}$ and $\{\hat{a}_L, a^\mu_H\}$ are distinct optima of the IE(0) problem. Letting $a_L = \hat{a}_L + (\epsilon_0, \epsilon_1)$, $a_L \neq \hat{a}_L$ means that at least one of $\epsilon_0, \epsilon_1$ is non-zero. Without any loss of generality, therefore, we may take $\epsilon_1 \neq 0$ and write $a_L = \hat{a}_L + (\kappa, 1) \epsilon_1$ with $\kappa = \epsilon_0/\epsilon_1$. It is easy to check now that, the profit constraint being binding at either optimum (recall Lemma 14), it must be $\Pi_L(a_L) = \Pi_L(\hat{a}_L)$; equivalently, $\kappa = p_L/(1-p_L)$. In addition, the IC_L constraint being also binding at either optimum, $U_L(a_L) = U_L(\hat{a}_L)$ and, by Lemma 6, there exists $\bar{\epsilon} \in (0,|\epsilon|)$ s.t. $u'(w_{1L} + \bar{\epsilon}) = u'(w_{0L} - \frac{p_H}{1-p_L})$ [resp. $u'(w_{1L} - \bar{\epsilon}) = u'(w_{0L} + \frac{p_H}{1-p_L})$] if $\epsilon > 0$ [resp. if $\epsilon < 0$]. Under risk-aversion, however, $u'$ is strictly-monotone and this implies in turn that $w_{1L} < w_{0L}$ [resp. $w_{1L} > w_{0L}$] if $\epsilon > 0$ [resp. if $\epsilon < 0$]. Which is absurd. For it means that $a_L \succ_L \hat{a}_L$ in either case (Lemma 7). A trivially similar argument establishes uniqueness also when $\mu = 1$.

Proof of Claim 1

To establish (i), observe first that since the high risk-type gets fully-insured at the RSW allocation while her profit constraint binds, that $a^\mu_H \succ_H 0$ follows immediately from Lemma 11. Notice next
that \( \Pi_L(a^*_L) = 0 \) allows us to write \( a^*_L = 0 + (a^*_{0L}, a^*_{1L}) = 0 + \left( \frac{p_L}{1-p_L}, 1 \right) a^*_{1L} \) for some \( a^*_{1L} \in \mathbb{R} \). In fact, it must be \( a^*_{1L} > 0 \). Otherwise, since \( p_H > p_L \) and \( w^*_0 \geq w^*_L \), \( 0 = a^*_L + \left( \frac{p_L}{1-p_L}, 1 \right) |a^*_{1L}| \) means that \( 0 \succ_H a^*_L \) due to Lemma 7(i). But then it ought to be \( 0 \succ_H a^*_L \sim_H a^*_{1H} \), which is absurd. Now, since \( w^*_0 \geq w^*_L \) and \( a^*_L = 0 + \left( \frac{p_L}{1-p_L}, 1 \right) a^*_{1L} \) with \( a^*_{1L} > 0 \), it follows that \( a^*_L \succ_H 0 \), again by Lemma 7(i).

To show next (iii), recall that the IE(1) and RSW(1) problems have a more strict profit constraint. It must be, therefore, \( U_L(a^*_L) \geq U_L(a^*_L) > \bar{u}_L \), the latter inequality since \( a^*_L \succ_H 0 \). The claim now follows from the fact that both \( U_L(\cdot) \) and the solution function \( a^*_H(\mu) : (0,1) \to \mathcal{A} \) are continuous (the latter by Berge’s maximum theorem - recall footnote 15 in the main text). To complete now the proof, note that, given (i) above, part (ii) of the claim follows immediately from Lemma 14(ii). ■

**Corollary 1** The RSW allocation is such that \( a^*_{1h} > 0 \forall (s,h) \in \{0,1\} \times \{L, H\} \).

**Proof.** Suppose to the contrary that \( \Pi_h(a^*_{1h}) = 0 \). We can write \( a^*_{1h} = 0 + \left( \frac{p_h}{1-p_h}, 1 \right) a^*_{1h} \) for all \( h \). It follows then that \( a^*_{1h} > 0 \) for either \( h \). For otherwise, by Lemma 7(i), \( a^*_{1h} < 0 \) would require that \( 0 = a^*_{1h} - \left( \frac{p_h}{1-p_h}, 1 \right) a^*_{1h} \) in conjunction with \( w^*_{0h} \geq w^*_{1h} \) imply that \( 0 \succ_H a^*_{1h} \) for either \( h \). Which contradicts, though, Claim 1(i). The latter claim rules out also the case \( a^*_{1h} = 0 \). For since \( \Pi_h(a^*_{1h}) = 0 \), this could be only if \( a^*_{1h} = 0 \). To complete the proof, it suffices to note that, under \( \Pi_h(a^*_{1h}) = 0, a^*_{1h} > 0 \) implies that also \( a^*_{1h} > 0 \) for either \( h \). ■

**Lemma 16** Take an arbitrary \( \mu \in (\lambda,1) \). The IE(\( \mu \)) optimum gives \( \Pi_H(a^*_{LH}) < \Pi_H(a^*_{LH}) \).

**Proof.** Suppose to the contrary that \( \Pi_H(a^*_{LH}) \geq \Pi_H(a^*_{LH}) \). Then, for any \( \kappa \in (0,1) \), the contract \( a^*_H = \kappa a^*_L + (1-\kappa) a^*_L \) would give \( \Pi_H(a^*_H) = \kappa \Pi_H(a^*_L) + (1-\kappa) \Pi_H(a^*_L) \geq \Pi_H(a^*_L) \). In addition, since \( a^*_L \sim_H a^*_L \), it must be \( U_L(a^*_H) > \kappa U_L(a^*_L) + (1-\kappa) U_H(a^*_H) = U_H(a^*_H) \) by risk-aversion and again for all \( \kappa \in (0,1) \). Similarly, as \( a^*_L \succ_H a^*_L \), we ought to have \( U_L(a^*_L) > U_L(a^*_L) \) for all \( \kappa \in (0,1) \). Yet, since \( \lim_{\kappa \to 0} U_L(a^*_H) = U_L(a^*_H) \), there exists \( \kappa \in (0,1) \) s.t. \( U_L(a^*_H) - U_L(a^*_H) < \Delta \equiv U_L(a^*_L) - U_L(a^*_H) \) for all \( \kappa \in (0,\kappa) \). For \( \kappa \in (0,\kappa) \), therefore, we have \( a^*_L \succ_L a^*_L \succ_H a^*_H \), so that the menu \( \{a^*_L, a^*_H\} \) is separating. As it also satisfies the profit constraint, it is feasible in the IE(\( \mu \)) problem. But this is absurd. For if \( \mu \in (\lambda,1) \), it means that \( \{a^*_L, a^*_H\} \) cannot be optimal. And if \( \mu = 1 \), it contradicts the uniqueness of the IE(1) optimum (recall Lemma 15). ■

**Lemma 17** Take arbitrary \( \mu, \mu' \in [\lambda,1] \) with \( \mu < \mu' \). The respective IE optima give

\[
\Pi_H(a^*_H) - \Pi_H(a^*_H) < \Pi_L(a^*_L) - \Pi_L(a^*_L)
\]

**Proof.** By Lemma 14(i), we ought to have \( \Pi(m^{|\lambda}) = \Pi(m^{|\lambda}) \); hence, it suffices to show that \( \Pi_L(a^*_L) > \Pi_L(a^*_L) \). To this end, suppose to the contrary that \( \Pi_L(a^*_L) \leq \Pi_L(a^*_L) \). Then, for any \( \kappa \in (0,1) \), the contract \( a^*_L = \kappa a^*_L + (1-\kappa) a^*_L \) would give \( \Pi_L(a^*_H) = \kappa \Pi_L(a^*_L) + 

38
(1 − κ) Π_L (a_L^H) ≥ Π_L (a_L^0). In addition, since by Lemma 10 it cannot but be U_L (a_L^0) > U_L (a_L^H), it must be U_L (a_L^0) > κU_L (a_L^0) + (1 − κ) U_L (a_L^H) > U_L (a_L^H) for all κ ∈ (0, 1). Yet, we also have a_L^H ≻ H a_L^0 ≻_L a_L^H, and since lim_{κ→0} U_H (a_L^κ) = U_H (a_L^H), there exists κ ∈ (0, 1) s.t. |U_H (a_L^κ) − U_H (a_L^H)| < Δ ∋ U_H (a_L^κ) − U_H (a_L^H) for all κ ∈ (0, κ). For κ ∈ (0, κ), therefore, we have a_L^H >_L a_L^κ >_L a_L^κ >_H a_L^H so that the menu {a_L^κ, a_L^H} is separating. As it also satisfies the profit constraint, it is feasible in the IE(µ) problem. Which contradicts, however, the optimality of {a_L^κ, a_L^H}.

**Proof of Claim 2**

(i) iff (ii). There is nothing to show for the “if” part. For the “only if,” that the RSW allocation can be IE(µ) optimum only if µ ∈ (λ, 1) follows from the fact that the IC_H constraint binds and the IC_L doesn’t in either allocation while the latter is unique (Lemma 15). Suppose then that the RSW allocation is the IE(µ) optimum for some µ ∈ (λ, 1). Recall also the Kuhn-Tucker first-order conditions for the latter in the proof of Lemma 14. Letting β_h and δ_h be, respectively, the Lagrangean multipliers on the IC_h and Π_h (a_h) ≥ 0 constraints in the RSW problem, the Kuhn-Tucker first-order conditions now are given by

\[
(\mu + \beta_h^{**}) u'(w_{0L}^{**}) = \beta_h^{**} \frac{1 - p_H}{1 - p_L} u'(w_{0L}^{**}) + \delta_h^{**} \tag{18}
\]

\[
(\mu + \beta_L^{**}) u'(w_{1L}^{**}) = \beta_L^{**} \frac{p_H}{p_L} u'(w_{1L}^{**}) + \delta_L^{**} \tag{19}
\]

\[
(1 - \mu + \beta_H^{**}) u'(w_{0H}^{**}) = \beta_H^{**} \frac{1 - p_L}{1 - p_H} u'(w_{0H}^{**}) + \delta_H^{**} \tag{20}
\]

\[
(1 - \mu + \beta_L^{**}) u'(w_{1H}^{**}) = \beta_L^{**} \frac{p_H}{p_L} u'(w_{1H}^{**}) + \delta_L^{**} \tag{21}
\]

\[
\beta_h^{**} (U_h (a_h^{**}) - U_h (a_h^{**})) = 0 \quad h, h' ∈ \{H, L\} \tag{22}
\]

\[
\delta_h^{**} \Pi_h (a_h^{**}) = 0 \quad h ∈ \{H, L\} \tag{23}
\]

\[
\beta_h^{**}, \delta_h ≥ 0 \quad h ∈ \{H, L\} \tag{24}
\]

along with (IC_h) and Π_h (a_h) ≥ 0 for either h. As the two IC_h constraints are the same in the RSW and IE(µ) problems, letting a_h^{**} = a_h^H for each h, it must be β_h^{**} = β_h^{**}. To satisfy then (35)-(38) and (11)-(14) simultaneously, it cannot but be \( \delta_L^{**}/λ = \delta_H^{**}/(1 - λ). \) But then, as the RSW allocation satisfies (35)-(38) for all µ ∈ (λ, 1), it must satisfy also (11)-(14) for all µ ∈ (λ, 1).

(ii) iff (iii). For the “only if” part, take an arbitrary µ ∈ (λ, 1) and recall the first-order conditions for the IE(µ) optimum in the proof of Lemma 14. Any one equation from each of the pairs (11)-(12) and (13)-(14) can be replaced, respectively, by

\[
(\mu + \beta_L^H) p_L (1 - p_L) [u'(w_{1L}^0) - u'(w_{0L}^0)] = \beta_L^H [p_H (1 - p_L) u'(w_{1L}^0) - p_L (1 - p_H) u'(w_{0L}^0)] \tag{25}
\]
and
\[
(1 - \mu + \beta_H^\mu) p_H (1 - p_H) [u'(w_{1H}^\mu) - u'(w_{0H}^\mu)] = \beta_L^\mu [p_L (1 - p_H) u'(w_{1L}^\mu) - p_H (1 - p_L) u'(w_{0L}^\mu)]
\]
(26)

By Lemma 2, moreover, it is without loss of generality to study the IE(\mu) problem after having substituted the binding profit constraint by the equivalent three-relations system in (27)-(29). Of course, for a menu \{a_L, a_H\} to satisfy (2), we can only consider movements along the same slope as on the corresponding fair-odds lines \text{FO}_h: \; da_{1h} = \left(\frac{1-h}{p_h}\right) da_{0h}. Which means that the IE(\mu) optimum can be fully characterized in terms of the quantities \(a_0^\mu, a_{0L}^\mu, \) and \(a_{0H}^\mu\) since the three equations in (27)-(29) can be equivalently written as
\[
a_1^\mu = \left(\frac{1}{\lambda p_L + (1 - \lambda) p_H} - 1\right) a_0^\mu
\]
(27)
\[
a_{1L}^\mu = \frac{1}{p_L} \left[ (1 - p_L) a_{0L}^\mu - \frac{(1 - \lambda)(p_H - p_L) a_0^\mu}{\lambda p_L + (1 - \lambda) p_H} \right]
\]
(28)
\[
a_{1H}^\mu = \frac{1}{p_H} \left[ (1 - p_H) a_{0H}^\mu + \frac{\lambda(p_H - p_L) a_0^\mu}{\lambda p_L + (1 - \lambda) p_H} \right]
\]
(29)

Here, the first equality is simply a re-arrangement of (27). Given this, the other two equalities follow from the respective ones in (2).

We just established that, with no loss of generality, the profit constraint of the IE(\mu) problem can be replaced by (27)-(29). We may write, therefore,
\[
w_{0L}^\mu = W - a_{0L}^\mu
\]
(30)
\[
w_{1L}^\mu = W - d + \frac{1}{p_L} \left[ (1 - p_L) a_{0L}^\mu - \frac{(1 - \lambda)(p_H - p_L) a_0^\mu}{\lambda p_L + (1 - \lambda) p_H} \right]
\]
(31)
\[
w_{0H}^\mu = W - a_{0H}^\mu
\]
(32)
\[
w_{1H}^\mu = W - d + \frac{1}{p_H} \left[ (1 - p_H) a_{0H}^\mu + \frac{\lambda(p_H - p_L) a_0^\mu}{\lambda p_L + (1 - \lambda) p_H} \right]
\]
(33)

and view the IE(\mu) problem as one in which \(a_1^\mu, a_{1L}^\mu, \) and \(a_{1H}^\mu\) for \((s, h) \in \{0, 1\} \times \{L, H\}\) are the choice variables with (27) but also \(a_0^\mu \geq 0\) being the constraints. Given this, the first-order conditions of the IE(\mu) problem are given now by (IC_h), (15), (25)-(26), any one equation from each of the pairs (11)-(12) and (13)-(14) once we have set \(\delta^\mu = 0\), as well as\(^{29}\)
\[
\lambda \left[ 1 - \mu + \beta_H^\mu - \beta_L^\mu \left(\frac{1 - p_L}{1 - p_H}\right) \right] u'(w_{0H}^\mu) - \\
(1 - \lambda) \left[ \mu + \beta_L^\mu - \beta_H^\mu \left(\frac{1 - p_H}{1 - p_L}\right) \right] u'(w_{0L}^\mu) = -\frac{\beta_L^\mu [\lambda p_L + (1 - \lambda) p_H]}{p_H - p_L}
\]
(34)
\[
\beta_L^\mu a_0^\mu = 0, \quad \beta^\mu \geq 0
\]

This is a system of nine equations in the nine unknowns (the six choice variables along with \(\beta_h^\mu\) for \(h \in \{L, H\}\) and \(\beta^\mu\)).

\(^{29}\)To obtain (34), one equates to zero the derivative of the Lagrangean \(\mu (U_L (a_L) + (1 - \mu) U_H (a_H)) + \beta_L (U_L (a_L) - U_L (a_H)) + \beta_H (U_H (a_H) - U_H (a_L)) + \beta a_0\) w.r.t. \(a_0\) using the four equations in (30)-(33).

40
Next, consider the RSW problem. For an arbitrary $\mu \in [0,1]$, letting $\beta^*_h$ and $\delta^*_h$ be, respectively, the Lagrangean multipliers on the constraints $IC_h$ and $\Pi_h (a_h) \geq 0$, the Kuhn-Tucker first-order conditions are given by

\[
(\mu + \beta^*_h) u'(w^{**}_0) = \beta^*_H \frac{1 - p_H}{1 - p_L} u'(w^{**}_0) + \delta^*_L
\]

(35)

\[
(\mu + \beta^*_L) u'(w^{**}_1) = \beta^*_H \frac{p_H}{p_L} u'(w^{**}_1) + \delta^*_L
\]

(36)

\[
(1 - \mu + \beta^*_H) u'(w^{**}_0) = \beta^*_L \frac{1 - p_L}{1 - p_H} u'(w^{**}_0) + \delta^*_H
\]

(37)

\[
(1 - \mu + \beta^*_H) u'(w^{**}_1) = \beta^*_L \frac{p_L}{p_H} u'(w^{**}_1) + \delta^*_H
\]

(38)

\[
\beta^*_h (U_h (a^*_h) - U_h (a_h^*)) = 0 \quad h, h' \in \{H, L\}
\]

(39)

\[
\delta^*_h \Pi_h (a^*_h) = 0 \quad h \in \{H, L\}
\]

(40)

\[
\beta^*_h, \delta_h \geq 0 \quad h \in \{H, L\}
\]

(41)

along with ($IC_h$) and $\Pi_h (a_h) \geq 0$. It is trivial then to check that (35)-(36) give

\[
(\mu + \beta^*_H) p_L (1 - p_L) [u'(w^{**}_1) - u'(w^{**}_0)] = \beta^*_H [p_H (1 - p_L) u'(w^{**}_1) - p_L (1 - p_H) u'(w^{**}_0)]
\]

(42)

We are now in position to develop the main argument. Suppose that the menu $\{a^*_h, a^*_H\}$ is optimal for the IE($\mu$) problem. As $\mu > \lambda$, both the RSW allocation and the IE($\mu$) optimum are such the $IC_L$ constraint is slack so that we ought to have $\beta^*_L = 0 = \beta^*_L$ for either $h$. And as also $w^{**}_0 = w^{**}_1$, (34) reads

\[
(1 - \lambda) \left( \mu - \frac{\beta^*_H p_H}{p_L} \right) u'(w^{**}_1) = \lambda (1 - \mu + \beta^*_H) u'(w^{**}_0) = \frac{\beta^*_H [\lambda p_L + (1 - \lambda) p_H]}{p_H - p_L}
\]

Multiply now both sides of this equation by $A = p_H (1 - p_L) u'(w^{**}_1) - p_L (1 - p_H) u'(w^{**}_0)$. Under (42), we get

\[
\frac{\beta^*_H A [\lambda p_L + (1 - \lambda) p_H]}{p_H - p_L} = (1 - \lambda) \left[ \mu [p_H (1 - p_L) u'(w^{**}_1) - p_L (1 - p_H) u'(w^{**}_0)] \right. \\
- \mu p_H (1 - p_L) [u'(w^{**}_1) - u'(w^{**}_0)] \\
\left. - \lambda [p_H (1 - p_L) u'(w^{**}_1) - p_L (1 - p_H) u'(w^{**}_0)] \right] u'(w^{**}_1) \\
- \lambda \left[ (1 - \mu) (p_H - p_L) u'(w^{**}_1) \right. \\
\left. + p_L [1 - \mu p_L - (1 - \mu) p_H] [u'(w^{**}_1) - u'(w^{**}_0)] \right] u'(w^{**}_0)
\]

(43)

Recall, though, that $w^{**}_1 < w^{**}_0$. Since also $u''(\cdot) < 0$, $p_H > p_L$, and $u'(\cdot) > 0$, it follows that $A > (p_H - p_L) u'(w^{**}_0) > 0$. But then, $\beta^*_H \geq 0$ iff the right-hand side of (43) is non-negative. Yet, the latter condition is equivalent to

\[
0 \leq (\mu (1 - \lambda) (p_H - p_L) u'_L (w^{**}_1) + \lambda p_L [1 - \mu p_L - (1 - \mu) p_H] u'_H (w^{**}_1)) u'_L (w^{**}_0)
\]

\[
- [\lambda p_L (1 - p_H) + \lambda (p_H - p_L) (1 - \mu + \mu p_L)] u'_L (w^{**}_1) u'_H (w^{**}_0)
\]

41
whose right-hand side has the following partial derivative w.r.t. $\mu$

$$
\lambda (1 - p_L) u'_L (w^*_{1L}) u'_H (w^*_{H}) + [(1 - \lambda) u'_L (w^*_{1L}) + \lambda p_L u'_H (w^*_{H})] (p_H - p_L) u'_L (w^*_{0L}) > 0
$$

Clearly, the right-hand side of (43) is non-negative only if it is so for $\mu = 1$. That is, only if

$$
0 \leq (1 - \lambda) (p_H - p_L) u' (w^*_{1L}) u' (w^*_{0L}) - \lambda p_L (1 - p_L) [u' (w^*_{1L}) - u' (w^*_{0L})] u' (w^*_{0H})
$$

Yet, as $1 - p_L > 1 - p_H > 0$ while $u' (w^*_{1L}) > u' (w^*_{0L})$, the last inequality above is but another way to express (1). For the “if” part, it suffices simply to observe that, if (1) is satisfied, then the right-hand side of (43) is non-negative for $\mu = 1$ and, thus, the RSW allocation is optimal in the IE(1) problem. That is, statement (i) of the claim applies and we have shown already that so does (ii). ■

**Lemma 18** Consider an arbitrary $\mu \in [0, 1]$. For either $h \in \{L, H\}$, the IE($\mu$) solution mapping $a^*_h : [0, 1] \mapsto \mathbb{R}^2$ given by $a^*_h (\mu) = a^*_h \mu$ consists of two functions

(i) $[0, \lambda] \mapsto \mathbb{R}^2 \setminus \{a^*_h\}$ which is bijective, and

(ii) $[0, \lambda] \mapsto \mathbb{R}^2_{++}$ which is either the constant $(\lambda, 1) \mapsto \{a^*_h\}$ or a bijective function $(\lambda, 1) \mapsto (0, b) \times (0, b)$ for some finite $b > 0$. In fact, for $\mu \in (\lambda, 1)$, $a^*_h (\mu)$ is a continuous function and so is the value function $\mu U_L (a^*_L (\mu)) + (1 - \mu) U_H (a^*_H (\mu))$.

**Proof.** That the solution mapping is a function follows immediately from the fact that, for any $\mu \in [0, 1]$, the IE($\mu$) optimum is unique (Lemma 15). The remainder of the argument will proceed in steps as follows.

**Step I(a).** Recall the first-order conditions for an IE optimum in the proofs of Lemma 14 but also Claim 2- in particular (25), (26), and (34). Taking again an arbitrary $\mu \in [0, 1]$, we can show that $\beta^\mu_L = 0 = \beta^\mu_H$ necessitates that either IC$\beta$ constraint binds. To see this, notice first that, under (34), we cannot have $\beta^\mu_L = 0 = \beta^\mu_H$ unless $\mu \notin \{0, 1\}$. For if $\mu = 0$ or $\mu = 1$, we ought to have, respectively, $u' (w^\mu_{1L}) = 0$ or $u' (w^\mu_{0L}) = 0$; either an absurd conclusion given that $u(\cdot)$ is everywhere strictly increasing. It can only be, therefore, $\mu, 1 - \mu > 0$ and (25)-(26) dictate that both risk-types ought to be fully-insured. But this means that $U_h (a^\mu_h) = u (w^\mu_{0H}) = U_{h'} (a^\mu_{h'})$ for $h, h' \in \{H, L\}$ with $h \neq h'$. As a result the two IC constraints give, $U_h (a^\mu_h) \geq U_h (a^\mu_{h'}) = U_{h'} (a^\mu_{h'}) \geq U_{h'} (a^\mu_{h'}) = U_h (a^\mu_h)$, which cannot be unless either constraint binds.

**Step I(b).** We can also establish that both IC$\beta$ constraints binding necessitates that the optimal is a pooling contract - and, thus, corresponds to case (iii) in Lemma 1. This is because $U_h (w^\mu_h) = U_{h'} (w^\mu_{h'})$ for either $h, h' \in \{L, H\}$ means that

$$
\begin{align*}
p_L [u (w^\mu_{0L}) - u (w^\mu_{0H}) + u (w^\mu_{1H}) - u (w^\mu_{1L})] & = u (w^\mu_{0L}) - u (w^\mu_{1L}) \\
& = p_L [u (w^\mu_{0L}) - u (w^\mu_{0H}) + u (w^\mu_{1H}) - u (w^\mu_{1L})]
\end{align*}
$$

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 $u (w^\mu_{0L}) = u (w^\mu_{0H})$ and $u (w^\mu_{1L}) = u (w^\mu_{1H})$.
as well as the IE($\mu'$) optimum for $\mu, \mu' \in [0, \lambda]$ with $\mu \neq \mu'$. Since the IC$_H$ constraint ought to be slack in this optimum, it must be $\beta_H^\mu = 0 = \beta_H^\mu$ and the preceding observation necessitates that either of the multipliers $\beta_L^\mu$ and $\beta_L^\mu$ are strictly positive. And of course, the latter two multipliers satisfy, respectively, the conditions (25)-(26) and

$$
(1 - \mu') p_H (1 - p_H) \left[ u'(w_{1H}^\mu) - u'(w_{0H}^\mu) \right] = \beta_L^\mu \left[ p_L (1 - p_H) u'(w_{1H}^\mu) - p_H (1 - p_L) u'(w_{0H}^\mu) \right] \tag{44}
$$

$$
(1 - \lambda) \left( \mu' + \beta_L^\mu \right) u'(w_{0L}^\mu) = \lambda \left[ 1 - \mu' - \beta_L^\mu \left( \frac{1 - p_L}{1 - p_H} \right) \right] u'(w_{0H}^\mu) \tag{45}
$$

Yet, since case (i) in Lemma 1 is the relevant one here, it must be also $w_{0H}^\mu < w_{1H}^\mu$. Which, in conjunction with $p_L < p_H$, imply that $p_L (1 - p_H) u'(w_{1H}^\mu) < p_H (1 - p_L) u'(w_{0H}^\mu)$. Furthermore, since either of $\beta_L^\mu$ and $\beta_L^\mu$ is strictly positive, (26) and (45) imply that so must be either of $1 - \mu' - \beta_L^\mu \left( \frac{1 - p_L}{1 - p_H} \right)$ and $1 - \mu - \beta_L^\mu \left( \frac{1 - p_L}{1 - p_H} \right)$. In other words, neither side of (25)-(26) or (44)-(45) may be zero. It follows, therefore, that $1 - \mu' (1 - \mu') = \beta_L^\mu / \beta_L^\mu$ and

$$
\frac{\mu + \beta_L^\mu}{\mu' + \beta_L^\mu} = \frac{1 - \mu - \beta_L^\mu \left( \frac{1 - p_L}{1 - p_H} \right)}{1 - \mu' - \beta_L^\mu \left( \frac{1 - p_L}{1 - p_H} \right)}
$$

hold simultaneously, leading to the contradiction $(\mu, \beta_L^\mu) = (\mu', \beta_L^\mu')$. When the solution mapping is restricted on $(\lambda, 1)$, the argument is trivially similar for the case in which the RSW allocation is not IE. And when it is, by Claim 2, it remains so for all $\mu \in (\lambda, 1]$ giving a constant solution function.

Step III(a). Turning now to the claim that the restriction of $a_h^*$ on $(\lambda, 1]$ gives strictly positive values, there is nothing to show of course if the RSW allocation is IE (Corollary 1). For the case when it is not, recall that the high-risk type gets full-insurance at the optimum while the IC$_H$ constraint binds. Formally, the former condition means that $U_H(a_H^*) = u(w_{0H}^*) = u((1 - p_H) w_{0L}^* + p_H w_{1H}^*)$ while the latter gives

$$
u((1 - p_H) w_{0H}^* + p_H w_{1H}^*) = U_H(a_L^*) = (1 - p_H) u(w_{0L}^*) + p_H u(w_{1L}^*) < \mu((1 - p_H) w_{0L}^* + p_H w_{1L}^*)
$$

Here, the inequality is due to risk-aversion and, by non-satiation, it requires that $w_{1L}^* - w_{1H}^* > \frac{1 - p_L}{p_H} (w_{0L}^* - w_{0H}^*)$. Recall also that the low-risk type is under-insured at the optimum. That is, $w_{0L}^* > w_{1L}^*$ and, by non-satiation again, the two equalities above can be put together as $u(w_{0H}^*) = (1 - p_H) u(w_{0L}^*) + p_H u(w_{1H}^*)$ to give $u(w_{0H}^*) < u(w_{0L}^*)$ or $w_{0H}^* < w_{0L}^*$. Yet, $w_h = (W - a_{0h}, w - d + a_{1h})$ for either risk-type and our observations can be re-written as $a_{1H}^* - a_{1L}^* < \frac{1 - p_H}{p_H} (a_{0H}^* - a_{0L}^*)$ and $a_{0L}^* < a_{0H}^*$. Recall now Lemmas 2 and 13. Setting $x = a_h^*$, $y = a_{H}^*$, $z = a_{L}^*$, $r_1 = \frac{1 - p_L}{p_L}$, and $r_2 = \frac{1 - p_H}{p_H}$ ensures that Lemma 13 applies so that $a_{0h}^* > a_{0}^*$ for either $h$. By (2)
then, we must also have $a_{th}^* > a_t^*$ for either $h$. The result now follows from $a^* \geq 0$. This completes the argument for the case in which the RSW allocation is not IE. When it is, the claim follows from Corollary 1.

**Step III(b).** To show next that, when $\mu \in (\lambda, 1]$, the solutions are bounded from above, by Claim 2(ii) there is again nothing to show when the RSW allocation is IE. To establish the claim when it is not, recall first Lemma 14(iii). This implies that $a_{0H}^\mu \leq \rho_H a_{1H}^\mu / (1 - \rho_H)$ and $a_{0L}^\mu \leq \rho_L a_{0L}^\mu / (1 - \rho_L)$. By Lemma 12(i), moreover, it must be also $a_{0L}^\mu \leq a_{0H}^\mu$ and, thus, it suffices to show that $a_{1H}^\mu$ is bounded. To this end, notice that, the high-risk being fully insured at the solution, we ought to have $a_{0H}^\mu = d - a_{1H}^\mu$; equivalently, $\Pi_H (a_{1H}^\mu) = (1 - p_H) d - a_{1H}^\mu$. The claim now follows from Lemma 17. Indeed, it cannot be $\Pi_H (a_{1H}^\mu) < \Pi_H (a_{1H}^{\mu'})$ unless $a_{1H}^\mu > a_{1H}^{\mu'}$ for any $\mu, \mu' \in [\lambda, 1]$ with $\mu < \mu'$. As a result, one may choose any $b \in [a_{1H}^\lambda, +\infty)$.

**Step III(c).** To complete the argument, that $a_1^\mu : (\lambda, 1] \mapsto (0, b) \times (0, b)$ is continuous follows from Berge’s theorem of the maximum. And given this, the resulting value function is also continuous since $U_h$ is everywhere continuous for either $h$. ■

### C Equilibrium Analysis

**Proof of Lemma 3**

Let $\{a_1^\mu, a_1^H\}$ be the IE($\mu$) optimum for some $\mu \in (\lambda, 1)$. Since the IE(1) optimum is unique, we ought to have: $U_L (a_1^1) > U_L (a_L)$. In addition,

$$U_L (a_1^\mu) < U_L (a_L^1), \quad U_H (a_1^\mu) = U_H (a_1^H), \quad U_L (a_1^\mu) < U_L (a_L^1)$$

Regarding these relations, all but the last inequality follow from the choice of $\mu$ (Lemma 10). And to verify the last inequality, let $\Delta_1 = U_L (a_1^1) - U_L (a_L)$ and recall that the value $V : [0, 1] \mapsto \mathbb{R}$ and solution $a_{1e(L,H)}^\mu : [0, 1] \mapsto \mathbb{R}$ of the IE($\mu$) problem are continuous at any $\mu \in (\lambda, 1)$ (Lemma 18) whereas $U_{1e(L,H)} (\cdot)$ is continuous everywhere. Observe also that, since $\mu > \lambda$, it must be $U_L (a_1^\mu) \neq U_H (a_1^\mu)$. Otherwise, we would have $U_L (a_1^\mu) = U_H (a_1^H) = U_H (a_1^H)$. Yet, $U_L (a_1^\mu) = U_H (a_1^H)$ cannot be unless $w_1^L$ provides full insurance, which is absurd since $\mu \neq \lambda$.

Taking, therefore, $\mu$ sufficiently close to 1 guarantees that

$$\max \{ (1 - \mu) |U_L (a_1^\mu) - U_H (a_1^H)|, |V (1) - V (\mu)| \} < \Delta_1 / 2$$

and indeed

$$U_L (a_1^\mu) > \mu U_L (a_L^1) + (1 - \mu) U_H (a_H^\mu) - \frac{\Delta_1}{2}$$

$$= V (\mu) - \frac{\Delta_1}{2} > V (1) - \Delta_1 = U_L (a_1^1) - \Delta_1 = U_L (a_L)$$

Consider now the contract $\hat{a}_L = a_L^\mu + (1, \kappa) \epsilon$ for some $\kappa \in (0, I_H (a_L^\mu))$ and $\epsilon > 0$. Since $a_L^\mu > L a_L$, by Lemma 9, a sufficiently small $\epsilon$ gives $a_L^\mu > L \hat{a}_L > L a_L$ and $a_L^\mu > H \hat{a}_L$. And since also $a_H^\mu \sim_H a_L^\mu$, the menu $\hat{m} = \{\hat{a}_L, \hat{a}_H\} = \{\hat{a}_L, a_L^\mu\}$ is IC$_H$, and satisfies part (iii) of the claim. It will be,
moreover, \( IR_h \) for either \( h \) if we choose \( \mu \) sufficiently close to 1 (Lemma 1). And to show that it is also \( IC_L \), recall the actual proof of Lemma 9. To establish that \( \hat{a}_L \succ_L a_L \), we used the fact that \( U_L(\hat{\omega}_L) - U_L(a_L^0) \in (-\Delta, 0) \) for sufficiently small \( \epsilon \) and where \( \Delta = U_L(a_L^0) - U_L(a_L) \). It is trivial to check that the argument remains valid if we define instead \( \Delta = U_L(a_H^0) - U_L(a_L) \) to obtain \( U_L(\hat{a}_L) > \max \{ U_L(a_H^0), U_L(a_L) \} \), again for sufficiently small \( \epsilon \). This establishes part (iii) of the claim as well as that \( \hat{m} \in \mathcal{M} \).

To show next part (ii) of the claim, notice that, compared to \( \{ a_L^\mu, a_H^\mu \} \) which is IE 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

\[
\Pi_L(\hat{a}_L) - \Pi_L(a_L^\mu) = \left( \frac{1 - p_L}{p_L} - \kappa \right) p_L \epsilon \geq \left[ \frac{1 - p_L}{p_L} - \Pi_L(a_L^\mu) \right] p_L \epsilon = (1 - p_L) \left[ 1 - \frac{w_L^\mu(w_L^\epsilon)}{w_L^\mu(\hat{w}_n)} \right] \epsilon > 0
\]

Here, the first inequality is because \( \kappa < I_H(\hat{a}_L^\mu) < I_L(\hat{a}_L^\mu) \). The second follows from risk-aversion and the fact that, \( \{ a_L^\mu, a_H^\mu \} \) being IE(\( \mu \)) with \( \mu > \lambda \), it must be \( w_L^\mu > w_L^\epsilon \) by Lemma 1. The latter lemma guarantees also that part (i) of the claim is valid since \( \Pi_L(\hat{a}_L) > \Pi_L(a_L^\mu) \geq 0 \geq \Pi_H(\hat{a}_H^\mu) \).

\[\blacksquare\]

Proof of Proposition 1

To establish the claim by contradiction, let \( (s^*, \alpha^*) \in S_F \times S_A \) be an equilibrium strategic profile in which \( a_L^* \in A_L(s^*) \) for some \( a_L^* \in A \setminus \{ a_L^1 \} \). By Lemma 3, there exists a menu \( \hat{m} = (\hat{a}_L, \hat{a}_H) \) which is IC as well as IR for either risk-type and such that \( \hat{a}_L \succ_L a_L^*, \Pi_L(\hat{a}_L) > 0 \geq \Pi_H(\hat{a}_H) \), and \( \lambda \Pi_L(\hat{a}_L) + (1 - \lambda) \Pi_H(\hat{a}_H) \geq 0 \). Take then an arbitrary firm \( n \in \mathcal{N} \) and suppose that it deviates to the strategy \( \hat{s}_n \) which entails the action \( \hat{s}_n^1 = (\hat{m}, C) \) at stage 1 (and subsequently the choice \( \hat{s}_n^2(\hat{s}_n^1, \cdot) = NW \) at stage 3, everywhere on \( S_n^1 \)). Letting \( \hat{s} = (\hat{s}_n, s_n^*) \), observe that \( \hat{a}_L \succ_L a_L^* \in A_L(s^*) \) implies that \( \alpha_L^*(\hat{s}) = \hat{a}_L \) and \( \mathcal{M}(\hat{s}) = \{ \hat{m} \} \). Clearly, the firm should expect profits

\[
\Pi_n^C(\hat{m}, s_n^*, \alpha^*) = \lambda_L \Pi_L(\hat{a}_L) + \left\{ \begin{array}{ll} \lambda_H(\hat{\mathcal{M}}_H(\hat{s})) \Pi_H(\hat{a}_H) & \text{if } \hat{a}_H = \alpha_H^*(\hat{s}) \\ 0 & \text{otherwise} \end{array} \right.
\]

and there are two possible cases. If \( \hat{m} \notin \mathcal{M}_H(\hat{s}) \), then \( \hat{a}_H \neq \alpha_H^*(\hat{s}) \) and

\[
\Pi_n^C(\hat{m}, s_n^*, \alpha^*) = \lambda_L \Pi_L(\hat{a}_L) \geq \lambda_L \Pi_L(\hat{a}_L) + \lambda_H \Pi_H(\hat{a}_H) = \lambda_L \Pi(\hat{a}_L) + (1 - \lambda) \Pi_H(\hat{a}_H) > 0
\]

If, on the other hand, \( \hat{m} \in \mathcal{M}_H(\hat{s}) \), then \( |\mathcal{M}_H(\hat{s})| \geq 1 \) and

\[
\Pi_n^C(\hat{m}, s_n^*, \alpha^*) = \lambda_L \Pi_L(\hat{a}_L) + \frac{\lambda_H(\hat{\mathcal{M}}_H(\hat{s}))}{|\mathcal{M}_H(\hat{s})|} \Pi_H(\hat{a}_H)
\]

\[
\geq \frac{1}{|\mathcal{M}_H(\hat{s})|} [\lambda_L \Pi_L(\hat{a}_L) + \lambda_H \Pi_H(\hat{a}_H)] > 0
\]

45
In either case, therefore, $\hat{s}_n$ is a profitable deviation against the profile $(s_{-n}^*, \alpha^*)$. Which is of course absurd.

We have just established that $s^* \in S_F$ being part of an equilibrium strategic profile necessitates that $\mathcal{A}_L (s^*) = \{a^*_L\}$. The claim now follows from the fact that the IE(1) optimum is unique. For this means that any menu $m = \{a^*_L, a_H^*\}$ with $a^*_L \succ_L a_H$ but $a_H \neq a^*_H$ must be either loss-making or violate the high-risk IC constraint, or both. And since $m \in \mathcal{M}_{NW} (s^*)$ requires that the menu avoids losses, it must be $a^*_L \succ_H a_H$. Without any loss of generality, therefore, we may replace the menu in question by the policy $\{a^*_L\}$ and view $\mathcal{M}_{NW} (s^*)$ as consisting of plans to offer at stage 1 and honour at stage 3 either the IE(1) optimum or the pooling policy $\{a^*_L\}$. Yet, the latter leads to the absurdity that firms are honouring loss-making policies in equilibrium. For if $s_{n|M}^* = \{a^*_L\}$ for some $n \in \mathcal{N}$, then

$$
\Pi_{NW}^n (\{a^*_L\}, s_{-n}^*, \alpha^*) = \frac{1}{|\mathcal{M}_{NW} (s^*)|} \left[ \lambda \Pi_L (a^*_L) + (1 - \lambda) \Pi_H (a^*_H) \right] < \frac{1}{|\mathcal{M}_{NW} (s^*)|} \left[ \lambda \Pi_L (a^*_L) + (1 - \lambda) \Pi_H (a^*_H) \right] = 0
$$

the inequality by Lemma 16. $
$

**Proof of Theorem 1**

Recall that, along the equilibrium path, $m^1$ is the only menu that is not withdrawn at stage 3 (Proposition 1). Our argument will proceed in steps as follows.

**Step 1.** Observe first that, on the equilibrium path, there can be no firm offering $m^1$ with commitment. To show this by contradiction, suppose that the equilibrium profile entails some $n \in \mathcal{N}^*$ whose action at stage 1 is $s_n^1 = (m^1, C)$. By Lemma 8, there exist $\epsilon_0 > 0$ and $\kappa \in (I_H (a^*_L), I_L (a^*_L))$ such that, for any $\epsilon \in (-\epsilon_0, 0)$, the contract $\hat{a}_L = a^*_L + (1, \kappa) \epsilon$ gives $\hat{a}_L \succ_L a^*_L \succ_H \hat{a}_L$. And as $a^*_H \sim_H a^*_L$, we have in fact $a^*_H \succ_H a^*_L \succ_H a^*_L$. Of course, the new contract gives

$$
\Pi_L (\hat{a}_L) - \Pi_L (a^*_L) = \left( \frac{1 - p_L}{p_L} - \kappa \right) p_L \epsilon < \left[ \frac{1 - p_L}{p_L} - I_L (a^*_L) \right] p_L \epsilon
$$

$$
= (1 - p_L) \left[ 1 - \frac{u' (w_{1H}^*)}{u' (w_{1L}^*)} \right] \epsilon < 0
$$

Nonetheless, as $\lim_{\epsilon \to 0} \Pi_L (\hat{a}_L) = \Pi_L (a^*_L) > 0$, for sufficiently small $|\epsilon|$ it must be $\Pi_L (\hat{a}_L) > 0$. Consider now that some firm $\hat{n} \in \mathcal{N} \setminus \{n\}$ deviates to the action $\hat{s}_n^1 = (\{\hat{a}_L\}, C)$ at stage 1 (and subsequently the choice $s_n^2 (\hat{s}_n^1, \cdot) = NW$ at stage 3, everywhere on $S_{-n}^1$). Since it is offered with commitment while the equilibrium profile $s_{-n}^*$ does not honour any other menu but $m^1$, the deviant policy must attract only and all agents of the low-risk type. Which leads, however, to the absurdity that the deviation is profitable (for sufficiently small $|\epsilon|$).

**Step 2.** Notice next that no menu is able to attract away from $m^1$ only the high-risk type and avoid losses doing so. To see this, let $\hat{m} \in \mathcal{M}$ be such that $a^*_L \succ_L a_L (\hat{m})$ and $a_H (\hat{m}) \succ_H a^*_H$. Then it must be $a_H (\hat{m}) \succ_H a^*_H$ because $a^*_H \succ_H a^*_H$. For the latter relation, observe that it is trivially true if the IE(1) and RSW allocations coincide. Otherwise, recall that the high-risk agents are fully-insured in
either allocation. That is, \( a^1_{iH} = d - a^1_{iH} \) and \( a^{**}_{iH} = d - a^{**}_{iH} \), which imply that \( a^1_H = a^{**}_H + (-1, \epsilon) \) for some \( \epsilon \in \mathbb{R} \). In fact, it must be \( \epsilon > 0 \) since \( \epsilon = \Pi_H (a^{**}_H) - \Pi_H (a^1_H) \) while \( \Pi_H (a^{**}_H) = 0 > \Pi_H (a^1_H) \) in this case (recall Lemma 1). But then that \( a^1_H > H a^{**}_H \) follows from Lemma 7(iii).

We have indeed, therefore, that \( a_H (\hat{m}) > H a^1_H \). And since \( a^{**}_H = \arg \max_{a \in A^2: \Pi_H (a) \geq 0} U_H (a) \) while \( \Pi_H (a^{**}_H) = 0 \), it must be \( \Pi_H (a_H (\hat{m})) < 0 \). Let now the firm \( \hat{n} \in \mathcal{N} \) deviate against the equilibrium profile \( s^*_\hat{n} \) by offering the menu \( \hat{m} \) at stage 1. Needless to say, it may do so with or without commitment.

If \( \hat{s}^1_n = (\hat{m}, NC) \), there is nothing to show when \( \Pi_L (a_L (\hat{m})) \leq 0 \). For the deviant offer will have to be withdrawn at stage 3 since

\[
\Pi^{NW}_{\hat{n}} (\hat{m}, s^*_{\hat{n}}, \alpha^*) = \lambda_{L\hat{n}} \Pi_L (a_L (\hat{m})) + \lambda_{H\hat{n}} \Pi_H (a_H (\hat{m})) < 0
\]

irrespective of the deviant’s belief \( \lambda_{\hat{n}} \) regarding the distribution of risk-types in its pool of applicants. And when \( \Pi_L (a_L (\hat{m})) > 0 \), we may construct sequential scenarios under which \( \hat{s}^2_n (\hat{s}^1_n, s^1_{\hat{n}}) = W, \hat{s}^2_{\hat{n}} (\hat{s}^1_n, s^1_{\hat{n}}) = NW \), and \( \sigma_L = 1 = \sigma_{H\hat{n}} \) are the optimal responses under the beliefs formation process in (6). To this end, observe that \( \Pi_L (a_L (\hat{m})) > 0 > \Pi_H (a_H (\hat{m})) \) allows us to define the quantity \( \Delta = -\frac{1}{\lambda_{H\hat{n}} (a_L (\hat{m}))} \) > 0. Taking then a sequence in (6) s.t. \( r^k_L / r^k_H < \min \{ \Delta, 1 \} \) ensures that \( \frac{\lambda_{L\hat{n}}}{\lambda_{H\hat{n}}} < -\frac{\Delta (a_L (\hat{m}))}{\Pi_L (a_L (\hat{m}))} \) everywhere along the sequence. But then \( \lambda_{L\hat{n}} \leq \frac{\Pi_L (a_L (\hat{m}))}{\Pi_L (a_L (\hat{m}))} \) and the deviant menu is indeed withdrawn at stage 3 since

\[
\Pi^{NW}_{\hat{n}} (\hat{m}, s^*_{\hat{n}}, \alpha^*) = \lambda_{L\hat{n}} \Pi_L (a_L (\hat{m})) + \lambda_{H\hat{n}} \Pi_H (a_H (\hat{m})) \leq 0
\]

By contrast, since \( \lambda^k_{L\hat{n}} = r^k_L / r^k_H < 1 \), we must have \( \lambda^k_{L\hat{n}} = \frac{1-r^k_L}{1-r^k_H} > 1 \) along the sequence. Equivalently, \( \lambda^k_{L(n^*)} > \lambda \) and \( \lambda^k_{H(n^*)} < 1 - \lambda \) (since \( \lambda^k_{L} \) [resp. \( \lambda^k_{H} \)] is strictly increasing [resp. decreasing] in \( \lambda^k_{n^*} \)). But then it cannot but be \( \lambda_{L(n^*)} \geq \lambda \) and \( \lambda_{H(n^*)} \leq 1 - \lambda \). And since \( \Pi_L (a_L^1) \geq 0 > \Pi_H (a_H^1) \) (Lemma 1), we ought to have

\[
\Pi^{NW}_{n^*} (m^1, \hat{s}_n, s^*_{(n^*, \hat{n})}, \alpha^*) = \lambda_{L\hat{n}} \Pi_L (a_L^1) + \lambda_{H\hat{n}} \Pi_H (a_H^1) \geq \lambda \Pi_L (a_L^1) + (1 - \lambda) \Pi_H (a_H^1) = 0
\]

Hence, it is indeed optimal for \( n^* \) to honour \( m^1 \) at stage 3.

Let now \( \hat{s}^1_n = (\hat{m}, C) \). In this case,

\[
\Pi^{C}_{\hat{n}} (\hat{m}, s^*_{\hat{n}}, \alpha^*) = \begin{cases} 
\lambda_{L\hat{n}} \Pi_L (a_L (\hat{m})) + (1 - \lambda) \Pi_H (a_H (\hat{m})) & \text{if } a_L (\hat{m}) = \alpha^* (\hat{m}, s^*_{\hat{n}}) \\
(1 - \lambda) \Pi_H (a_H (\hat{m})) & \text{otherwise}
\end{cases}
\]

Yet, \( a_L (\hat{m}) = \alpha^* (\hat{m}, s^*_{\hat{n}}) \) obtains only if \( a_L (\hat{m}) \sim L a^1_L \) and, thus, \( |M_L (\hat{m}, s^*_{\hat{n}}) | \geq 1 \). Which, in conjunction with \( \Pi_L (a_L (\hat{m})) > 0 \), necessitates that

\[
\Pi^{C}_{\hat{n}} (\hat{m}, s^*_{\hat{n}}, \alpha^*) \leq \lambda \Pi_L (a_L (\hat{m})) + (1 - \lambda) \Pi_H (a_H (\hat{m}))
\]

By the uniqueness of the IE(1) optimum, however, \( a_L (\hat{m}) \sim L a^1_L \) and \( a_H (\hat{m}) > H a^1_H \) together render the last quantity above negative.
Step 3. Clearly, the potentially profitable deviations cannot but entail menus \( \hat{m} \in \mathcal{M} : a_L(\hat{m}) \succ_L a^1_L \). Which, \( m^1 \) being the unique IE(1) optimum with \( \lambda \Pi_L(a^1_L) + (1 - \lambda) \Pi_H(a^1_H) = 0 \), implies in turn that the menus in question are such that

\[
\lambda \Pi_L(a_L(\hat{m})) + (1 - \lambda) \Pi_H(a_H(\hat{m})) < 0 \tag{46}
\]

Again we need to distinguish between whether or not the deviant offer is made with commitment. Consider first the deviation \( \hat{s}^1_n = (\hat{m}, C) \) for some \( \hat{n} \in \mathcal{N} \). Since \( a_L(\hat{m}) \succ_L a^1_L \), the equilibrium profile described in the theorem dictates that all other firms withdraw \( m^1 \) at stage 3 (an optimal response since \( \Pi_H(a^1_H) \leq 0 \) while the menu will not be selected by any agent of the low-risk type). Anticipating this, the high-risk customers cannot but opt also for \( a_H(\hat{m}) \) at stage 2. Yet, this turns the deviant menu into one that serves both types, rendering it loss-making under (46).

Next, suppose that \( \hat{s}^1_n = (\hat{m}, NC) \). Given (46), it cannot but be \( \lambda \not\in \mathcal{P}(\hat{m}) \). Which implies in turn that there are two subcases to examine: \( \mathcal{P}(\hat{m}) = [p(\hat{m}), 1] \) and \( \mathcal{P}(\hat{m}) = [0, p(\hat{m})] \); equivalently, \( p(\hat{m}) > \lambda \) and \( p(\hat{m}) < \lambda \), for some \( p(\hat{m}) \in [0, 1] \).

(i) \( p(\hat{m}) > \lambda \)

Notice first that, in this case, there is no belief the deviant firm may entertain to render \( \hat{m} \) profitable. For otherwise, we ought to have \( \lambda_{L,\hat{h}} > p(\hat{m}) \). Which requires in turn that \( \lambda_{L,\hat{h}} > p(\hat{m}) \), at least along a subsequence. We ought to have, therefore, \( \lambda_{L,\hat{h}} > \lambda \) or, equivalently, \( \lambda_{n,\hat{h}} > 1 \). In either of the beliefs formations (5)-(6), however, the latter relation can obtain only if \( \lambda_{n,\hat{h}} > 1 \). That is, only if \( \lambda_{L,\hat{n}} < \lambda \) and \( \lambda_{H,\hat{n}} > 1 - \lambda \). But then it cannot but be \( \lambda_{L,\hat{n}} \leq \lambda \) and \( \lambda_{H,\hat{n}} \geq 1 - \lambda \). And since \( \Pi_L(a^1_L) > 0 \geq \Pi_H(a^1_H) \), we ought to have

\[
\Pi_{\hat{n},n^*}^N \left( m^1, \hat{s}_\hat{n}, s^*_n(n^*,\hat{n}), \alpha^* \right) = \lambda_{L,\hat{n}} \Pi_L(a^1_L) + \lambda_{H,\hat{n}} \Pi_H(a^1_H) \\
\leq \lambda \Pi_L(a^1_L) + (1 - \lambda) \Pi_H(a^1_H) = 0
\]

Hence, it is optimal for the firm \( n^* \) to withdraw \( m^1 \) at stage 3. Of course, anticipating this at stage 2, all agents of either risk-type must follow the strategy \( \sigma_h = 0 \). Which means, however, that \( \lambda_{L,\hat{n}} = \lambda - 1 - \lambda_{H,\hat{n}} \), contradicting under (46) our hypothesis that the deviation is profitable. Given this observation, there are but two possible equilibrium scenarios.

(i.a) Both policies are withdrawn

To support a sequential equilibrium scenario in which both the equilibrium and the deviant menus get withdrawn at stage 3, let

\[
\sigma_h \in [0, 1] : \quad 1 - \frac{1 - \sigma_L}{1 - \sigma_H} \leq \frac{p(\hat{m})(1 - \lambda)}{\lambda(1 - p(\hat{m}))} \tag{47}
\]

\[
(q^k_L, q^k_H) \in [0, 1 - \sigma_L] \times [0, 1 - \sigma_H] : \quad q^k_L \leq q^k_H \quad \forall k \in \mathbb{N} \tag{48}
\]

and observe that \( \sigma^k_L = \sigma_L + q^k_L < \sigma_L + q^k_H = \sigma_H + q^k_H = \sigma^k_H \), the second inequality due to the lower bound in (47). Taking, thus, the beliefs formation as in (5) we have \( \lambda_{n,\hat{h}} < 1 \) and, by the same argument as in the preceding paragraph, the equilibrium menu will be withdrawn at stage 3. Regarding the deviant one, since \( \lambda < p(\hat{m}) \), it must be \( \frac{p(\hat{m})(1 - \lambda)}{\lambda(1 - p(\hat{m}))} > 1 \). Moreover,
The deviant menu is withdrawn but the equilibrium one is not

If the RSW allocation is not IE, we have also $\lambda \not\in P(\hat{m})$ and it is optimal to withdraw $\hat{m}$ as loss-making. If $\lambda_{\hat{n}} = p(\hat{m})$, on the other hand, the menu expects exactly zero profits and we may dictate its withdrawn without loss of generality.

As a remark, it should be noted that the scenario in question cannot be supported if (47) is violated. For $\frac{1 - \sigma_L}{1 - \sigma_H} < 1$ and $\frac{1 - \sigma_L}{1 - \sigma_H} > \frac{p(\hat{m})(1 - \lambda)}{\lambda(1 - p(\hat{m}))}$, respectively, the equilibrium and the deviant menu strictly profitable, precluding the respective withdrawals at stage 3. By contrast, the scenario is supported for $\sigma_L = 1 = \sigma_H$ under the beliefs’ formation in (6) with $r_L^k > r_H^k \forall k$ and $r_L^k / r_H^k \rightarrow 1$. Since now the beliefs are such that $\lim_{k \rightarrow \infty} \lambda_{\hat{n}}^k = 1$, we have $\lambda_{\hat{n}} = \lambda$ so that $\lambda_{\hat{n}} \not\in P(\hat{m})$ as required. Moreover, $\lambda_{\hat{n}}^k > 1$ requires that $\lambda_{\hat{n}}^k < 1$. Equivalently, $\lambda_{\hat{n}}^k < 1$. But then it must be $\lambda_{\hat{n}}^k < 1$ and, thus, $\lambda_{\hat{n}}^k \leq \lambda$. Now, if the RSW allocation is IE, on the other hand, then $\Pi_{\hat{n}}^{NW} \left(m^1, s_{\hat{n}}, s_{L(n^*, \hat{\lambda})}^*, \alpha^* \right) = 0$ and we may dictate the withdrawl of $m^1$ without loss of generality.

(i.b) The deviant menu is withdrawn but the equilibrium one is not

In this scenario, $\sigma_L = 1 = \sigma_H$ is strictly dominant for all agents at stage 2. As sequential equilibrium, this can be supported again under the beliefs’ formation in (6) albeit now with $r_L^k > r_H^k \forall k$ and $r_L^k / r_H^k \rightarrow 0$. For then we have $\lim_{k \rightarrow \infty} \lambda_{\hat{n}}^k = 0 \not\in P(\hat{m})$ and the deviant menu will indeed be withdrawn at stage 3. Regarding the equilibrium one, taking a subsequence if necessary so that $\lambda_{\hat{n}}^k < 1$ everywhere, means that $\lambda_{\hat{n}}^k > 1$ or $\lambda_{\hat{n}}^k > \lambda$. As a result, $\lambda \leq \lambda_{\hat{n}}^k$ so that $\lambda_{\hat{n}}^k \in P(m^1)$ and it is optimal to honour $m^1$ at stage 3.

(ii) $p(\hat{m}) < \lambda$

In this case, it is impossible to sustain an equilibrium scenario in which both policies are withdrawn at stage 3. This is because, for the deviant policy to not be profitable, we must have $\lambda(\hat{n}) < p(\hat{m})$. That is, $\lambda_{\hat{n}} < \lambda$ which requires a subsequence of strategic profiles along which $\lambda_{\hat{n}}^k < 1$. This implies in turn that $\lambda_{\hat{n}}^k > 1$ or $\lambda_{\hat{n}}^k > \lambda$. But then it must be $\lambda_{\hat{n}}^k \geq \lambda$, which is absurd as it renders $m^1$ profitable. Given this observation, there is only one possible equilibrium scenario.

The deviant menu is withdrawn but the equilibrium one is not

This can be supported by the same profile as in (i.b), but for the requirement that $r_L^k / r_H^k \rightarrow 1$ so that $\lim_{k \rightarrow \infty} \lambda_{\hat{n}}^k = 1$ and, thus, $\lambda_{\hat{n}} = \lambda \not\in P(\hat{m})$. ■

D Equilibrium in the Rothschild-Stiglitz game

Proof of Lemma 4

To establish (i), we will show first that it cannot be $a_L = a_H$. To this end, suppose that $\{a\}$ is an equilibrium (pooling) policy and consider the contract $a_L^\epsilon = a + (1, \kappa) \epsilon$ for some $\kappa \in (I_H(a), I_L(a))$ and $\epsilon < 0$. By Lemma 8, for small enough $|\epsilon|$, we ought to have $a_L^\epsilon > L a \succ_H a_L^0$ so that, in the presence of $a$, the new contract attracts only the low-risk customers. And by doing so, it delivers
profits

\[ \Pi_L (a_L^0) = (1 - p_L) (a_0 + \epsilon) - p_L (a_1 + \kappa \epsilon) = \Pi_L (a) + [1 - p_L (1 + \kappa)] \epsilon \]

= \[ \Pi_M (a) + [\Pi_L (a) - \Pi_M (a)] + [1 - p_L (1 + \kappa)] \epsilon \]

But \( \Pi_L (a) > \Pi_M (a) \) and, thus, \( 1 + \kappa > \frac{1}{p_L} (1 + \frac{1}{\epsilon} [\Pi_L (a) - \Pi_M (a)]) \) if \(|\epsilon|\) is sufficiently small. Equivalently, \( \Pi_L (a_L^0) > \Pi_M (a) \geq 0 \), the last inequality due to \( a \) being an equilibrium policy. Clearly, \( \{a_L^0\} \) is a profitable (pooling) deviation against \( \{a\} \). Which is absurd. To complete the argument, note that, for any menu \( \{a_L, a_H\} \), it cannot be \( a_L \sim_h a_H \) for either \( h \) unless \( a_L = a_H \).

To show now (ii), let \( \{a_L, a_H\} \) be an equilibrium separating menu. We will establish first that \( \Pi_h (a_h) \leq 0 \) for either \( h \). To do so for \( h = L \), let \( \Pi_L (a_L) = \delta > 0 \) and consider the contract \( a'_L = a_L + (1, \kappa_L) \epsilon_L \) for some \( \kappa_L \in (I_H (a_L), I_L (a_L)) \) and \( \epsilon_L < 0 \). This gives

\[ \Pi_L (a'_L) = (1 - p_L) (a_0 L + \epsilon_L) - p_L (a_1 L + \kappa_L \epsilon_L) = \Pi_L (a_L) + [1 - p_L (1 + \kappa_L)] \epsilon_L \]

\[ = \delta + [1 - p_L (1 + \kappa_L)] \epsilon_L \]

and sufficiently small \( |\epsilon_L| \) ensures that \( |1 - p_L (1 + \kappa_L)| |\epsilon_L| < \delta \) or \( \Pi_L (a'_L) > 0 \). Yet, by Lemma 8, we also have \( a'_L \succ L a_L \succ H a'_L \). Which, since \( a_L \succ L a_H \succ H a_L \), gives \( a'_L \succ L a_H \succ H a'_L \). Clearly, \( \{a'_L, a_H\} \) is another separating menu that attracts the low-risk type away from \( \{a_L, a_H\} \) and makes strictly positive profits doing so. Which contradicts, of course, the hypothesis that the latter menu is an equilibrium one. A trivially similar argument, using a contract \( a'_H = a_H + (1, \kappa_H) \epsilon_H \) for some \( \kappa_H \in (I_H (a_L), I_L (a_L)) \) and sufficiently small \( \epsilon_H > 0 \), necessitates that \( \Pi_H (a_H) \leq 0 \). That \( \Pi_h (a_h) = 0 \) for either \( h \) follows now immediately from the fact that an equilibrium menu cannot be loss-making.

**Proof of Lemma 5**

By Lemma 4, we may restrict our attention to separating menus \( \{a_L, a_H\} \) such that \( \Pi_h (a_h) = 0 \) for either \( h \). Suppose then that \( \{a_L, a_H\} \neq \{a_L^*, a_H^*\} \). Since \( \{a_L, a_H\} \) is feasible in the RSW problem and \( \{a_L^*, a_H^*\} \) uniquely optimal for \( \mu \in (0, 1) \), it must be \( a_L^* \ngeq H a_h \) for either \( h \) with strict preference for at least one type. Hence, there are two cases to consider.

(i). \( a_H^* \succ H a_H \) and \( a_L^* \sim H a_L \)

Suppose first that \( w_{0H} > w_{1H} \). Then, \( u'(w_{0H}) < u'(w_{1H}) \) and, thus, \( I_h (a_H) < \frac{1 - p_L}{p_H} \) for either \( h \). As also \( I_L (a_H) > I_H (a_H) \), for \( \kappa_H \in \left( \frac{1 - p_L}{p_H}, \infty \right) \) and \( \epsilon_H < 0 \) with \( |\epsilon| \) sufficiently small, by Lemma 9, the contract \( a_L^2 = a_H^* - (1, \kappa_H) \epsilon_H \) gives \( a_L^* \succ L a_H^2 \succ H a_H \). And since \( a_L \sim_L a_L^* \succ L a_H^* \) while \( a_H \succ H a_L \), the menu \( \{a_L, a_L^2\} \) is separating and attracts at least the high-risk type away from \( \{a_L, a_H\} \). Doing so, moreover, is strictly profitable since \( \Pi_H (a_H^2) = \Pi_H (a_H) + (1 - p_H - \kappa_H p_H) \epsilon_H = (1 - p_H - \kappa_H p_H) \epsilon_H > 0 \), the second equality because \( \Pi_H (a_H) = 0 \). If \( w_{0H} < w_{1H} \), on the other hand, then \( I_h (a_H) > \frac{1 - p_H}{p_H} \) 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 \( \Pi_H (a_H) = 0 \), \( a_H \) would lie on the intersection between the \( FO_H^* \) and the 45-degree lines. But then it would coincide with \( a_H^* \), contradicting the
hypothesis that \( \mathbf{a}_H^* \succ_H \mathbf{a}_H \).

(ii). \( \mathbf{a}_H^* \succeq_H \mathbf{a}_H \) and \( \mathbf{a}_L^* \succ_L \mathbf{a}_L \)

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 some \( \epsilon^0 > 0 \), therefore, the contract \( \mathbf{a}_L^2 = \mathbf{a}_L^* + (1, \kappa_L) \epsilon_L \) gives \( \mathbf{a}_L^* \succ_H \mathbf{a}_L^2 \succ_L \mathbf{a}_L \) for any \( \epsilon_L \in (-\epsilon^0,0) \) (again by Lemma 9). 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 \( \lim_{\epsilon_L \to 0} \mathbf{a}_L^2 = \mathbf{a}_L^* \), by continuity it must be \( U_L(\mathbf{a}_L^*) - U_L(\mathbf{a}_L^2) < \Delta \) and, thus, \( \mathbf{a}_L^2 \succ_L \mathbf{a}_H^* \) for sufficiently small \( |\epsilon_L| \). But then, we have \( \mathbf{a}_H^* \sim_H \mathbf{a}_L^* \succ_H \mathbf{a}_L^2 \succ_L \mathbf{a}_H^* \). That is, the menu \( \{ \mathbf{a}_L^2, \mathbf{a}_H^* \} \) is separating and attracts at least the low-risk type away from \( \{ \mathbf{a}_L, \mathbf{a}_H \} \). Moreover, it is strictly profitable since \( \Pi_L(\mathbf{a}_L^2) = \Pi_L(\mathbf{a}_L^*) + (1-p_L - \kappa_L p_L) \epsilon_L = 1 - p_L - \kappa_L p_L > 0 \).

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 9 once again, the preceding argument applies now for \( \kappa_L \in (-\infty, \frac{1-p_H}{p_H}) \) 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

\[
U_H(\mathbf{w}_H) = (1-p_H)u(w_{0L}) + p_Hu(w_{1L}) < u(1-p_H)w_{0H} + p_WH
\]

contradicting that \( \{ \mathbf{a}_L, \mathbf{a}_H \} \) is IC 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 shown in case (i), \( w_{0H} \neq w_{1H} \)). The second inequality follows from \( p_H < p_H \) 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 \( \Pi_h(\mathbf{a}_h) = 0 \) for either \( h \).

**Proof of Claim 4**

“Only if.” It is straightforward to establish the contrapositive. For if the RSW menu is not IE, we may consider again the menus \( \{ \mathbf{a}_L, \mathbf{a}_H^* \} \) and \( \{ \mathbf{a}_L^0, \mathbf{a}_H^0 \} \) as in the proof of Lemma 3. Recall that the former menu is separating, strictly profitable, and such that \( \mathbf{a}_L \succ_L \mathbf{a}_L^0 \). The latter, on the other hand, is the IE(\( \mu^0 \)) optimum for some \( \mu^0 \in (\lambda, 1) \). We ought to have, therefore, \( \mathbf{a}_L^0 \succeq \mathbf{a}_H^* \) for either \( h \) (at least one preference being in fact strict). Clearly, \( \{ \mathbf{a}_L, \mathbf{a}_H^0 \} \) is a strictly profitable deviation against the RSW policy.

“If.” Notice first that no contract is able to attract only one risk-type away from the RSW policy and avoid losses doing so. For suppose that \( \{ \mathbf{a}_h^* \} \) is designed to achieve this with respect to the risk-type \( h \). It ought to be then \( \Pi_h(\mathbf{a}_h^*) \geq 0 \) and \( \mathbf{a}_h^* \succeq h' \mathbf{a}_h' \succ h \mathbf{a}_h^* \) for \( h' \neq h \). As, however, \( \mathbf{a}_h^* \succeq h \mathbf{a}_h^* \), this would mean that the separating menu \( \{ \mathbf{a}_h^*, \mathbf{a}_h' \} \) Pareto-dominates the RSW one while being feasible in the RSW problem, for any \( \mu \in [0,1] \). Which is absurd.
Hence, credible challenges to the RSW policy may come only from menus that attract both types away from \( \{a_L^{**}, a_H^{**}\} \) and are strictly-profitable doing so. Yet, there are no such menus if the RSW allocation is IE. ■

Proof of Claim 3

"Only if." We will argue by contradiction. To this end, it is without any loss of generality to begin by the hypothesis that \( \exists a \in A \text{ s.t. } a \succ_L a_L^{**} \) and \( \Pi_M(a) = 0 \). For if \( a^1 \in A \text{ is s.t. } \Pi_M(a^1) = \epsilon > 0 \) and \( a^1 \succ_L a_L^{**} \), the contract in question is given by \( a = a^1 - (1, -1) \epsilon \). Indeed, it is trivial to check that \( \Pi_M(a) = \Pi_M(a^1) - \epsilon = 0 \) and \( a \succ h_a^1 \) for either \( h \).

Let then \( a \succ_L a_L^{**} \) and \( \Pi_M(a) = 0 \). Since the low-risk IR constraint does not bind at the RSW allocation, we also have \( a \succ_L a_L^{**} \succ_L 0 \), the latter contract corresponding to the endowment point (which is also on \( FO^*_M \)). The continuity of the preference relation \( \succ_L \) then requires the existence of a convex combination of \( a \) and \( 0 \) such that the low-risk type is indifferent between this new point and \( a_L^{**} \).

There must be that is \( a^1 = \pi a \) for some \( \pi \in (0, 1) \) such that \( a^1 \in FO^*_M \) and \( a^1 \sim_L a_L^{**} \). Notice now that, since \( a, a^1 \in FO^*_M \) and \( \pi < 1 \), it ought to be \( a^1 = a - \left( \frac{\bar{p}}{1 - \bar{p}}, 1 \right) \epsilon \) for some \( \epsilon > 0 \) where \( \bar{p} = \lambda p_L + (1 - \lambda) p_H \). By Lemma 6 then we ought to have

\[
U_L(w) - U_L(w^1) = \left[ p_L u'(w_1' + \epsilon') - \frac{p_L(1 - p_L)}{1 - \bar{p}} u'(w_0' - \frac{p p'}{1 - \bar{p}}) \right] \epsilon = \left[ p_L u'(w_1') - \frac{p_L(1 - p_L)}{1 - \bar{p}} u'(w_0') \right] \epsilon
\]

for some \( \epsilon' \in (0, \epsilon) \). Here, the second equality follows by letting \( w' = w^1 + \left( -\frac{\bar{p}}{1 - \bar{p}}, 1 \right) \epsilon' \), which corresponds to the contract \( a' = a^1 + \left( \frac{\bar{p}}{1 - \bar{p}}, 1 \right) \epsilon' \) that is also on \( FO^*_M \).

Yet, by construction, \( a \succ_L a_L^{**} \sim_L a^1 \). Hence, it must be \( \frac{\bar{p}}{1 - \bar{p}} < \frac{p_L u'(w_1')}{(1 - p_L) u'(w_0')} = I_L(a')^{-1} \) and we may consider another contract \( a^2 = a' + (\kappa, 1) \epsilon \) with \( \kappa \in \left( \frac{\bar{p}}{1 - \bar{p}}, I_L(a')^{-1} \right) \) and \( \epsilon > 0 \). As a pooling policy, this gives

\[
\Pi_M(a^2) = \Pi_M(a') + [\kappa (1 - \bar{p}) - \bar{p}] \epsilon = [\kappa (1 - \bar{p}) - \bar{p}] \epsilon > 0
\]

\( \text{Recall that there is a one-to-one relation between contract and income points. The continuity of the relation } \succ_L \text{ derives from continuous preferences over lotteries on wealth vectors. Given this, the existence of the wealth vector in question (and, thus, of the corresponding contract) is a standard result to be found in textbook derivations of the expected utility theorem. See, for instance, Step 3 of Proposition 6.B.3 in Mas-Collel A., Whinston M.D., and J.R. Green, Microeconomic Theory, Oxford University Press (1995).} \)
the second equality since \( a' \in FO^*_M \). For either \( h \), moreover, applying again Lemma 6 successively gives

\[
U_h (w^2) - U_h (w^1) = U_h (w^2) - U_h (w') + U_h (w') - U_h (w^1)
\]

\[
= [p_h u' (w'_1 + \varepsilon') - \kappa (1 - p_h) u' (w'_0 - \kappa \varepsilon')] \varepsilon + p_h u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_h)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon''}{1 - \bar{p}} \right) \varepsilon' \\
= [\kappa^{-1} - I_h (a' + (\kappa, 1) \varepsilon')] p_h u' (w'_1 + \varepsilon') \kappa \varepsilon + p_h u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_h)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon''}{1 - \bar{p}} \right) \varepsilon'
\]

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

\[
U_h (w^2) - U_h (w^1) = [\Delta + I_L (a') - I_h (a' + (\kappa, 1) \varepsilon')] p_h u' (w'_1 + \varepsilon') \kappa \varepsilon + p_h u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_h)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon''}{1 - \bar{p}} \right) \varepsilon' \\
> [p_h u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_h)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon''}{1 - \bar{p}} \right) \varepsilon'] \\
> [p_h u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_h)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon'}{1 - \bar{p}} \right) \varepsilon'] \\
\geq [p_L u' (w'_1 + \varepsilon') - \frac{\bar{p} (1 - p_L)}{1 - \bar{p}} u' \left( w'_0 - \frac{\bar{p} \varepsilon''}{1 - \bar{p}} \right) \varepsilon']
\]

\[
= [U_L (w) - U_L (w^1)] \frac{\varepsilon'}{\varepsilon} > 0
\]

Here the second inequality follows from the fact that \( \varepsilon'' < \varepsilon' \) while \( u (\cdot) \) is strictly concave. By contrast, the last inequality exploits that \( p_H > p_L \).

Let now the contract \( a^2 \) be offered in the presence of the RSW policy. As \( a^2 \succ_L a^1 \sim_L a^*_L \succ_L a^*_H \) (the first two preferences by construction, the last one by the properties of the RSW allocation), the low-risk type is pulled away. If \( a^2 \succ_H a^*_H \), this is also the case for the high-risk type so that \( \{ a^2 \} \) becomes a strictly profitable pooling policy - recall (49). Otherwise, it attracts only the low-risk type, delivering even higher expected profits. For since \( p_L < p < p_H \), any contract \( \bar{a} \) gives \( \Pi_L (\bar{a}) \geq \Pi_M (\bar{a}) \geq \Pi_H (\bar{a}) \), with either inequality strict unless \( \bar{a} = \mathbf{0} \). In either case, therefore, \( \{ a^2 \} \) is a strictly profitable deviation against the RSW policy.

If. Recall that, in establishing the “if” part of Claim 4, we showed that credible challenges to the RSW policy may come only from menus that attract both types away from \( \{ a^*_L, a^*_H \} \) and are strictly-profitable doing so. Given that only safe menus are now admissible, no separating menu can attract either risk-type \( h \) away from \( a^*_H \) since the latter maximizes \( h \)‘s welfare amongst all safe separating menus. The only possibility to consider, therefore, is strictly-profitable singleton menus (i.e., pooling contracts). Yet, by hypothesis, there is none that can attract the low-risk customers from the RSW menu. ■
Proof of Proposition 2

Take an arbitrary \( (u, p_L, p_H) \in \mathcal{U} \times (0, 1)^2 \) and consider the function \( a^* : (0, 1) \to \mathcal{A} \) s.t.

\[
a^* (\lambda) = \arg \max_{a \in \mathcal{A} : \Pi_H(a) \geq 0} U_L(a)
\]

Of course, the objective function being everywhere strictly concave (and, thus, also continuous) and the constraint set compact for each \( \lambda \in (0, 1) \), \( a^* (\cdot) \) is indeed a well-defined solution function. It also results in a value function \( U_L(a^*(\cdot)) \) that is everywhere strictly increasing. To see this, recall that \( p_H > p_L \) implies that \( \Pi_L(a) > \Pi_H(a) \) everywhere on \( \mathcal{A} \). And since the constraint above cannot but bind at the optimum, it must be \( \Pi_L(a^*(\lambda)) > 0 > \Pi_H(a^*(\lambda)) \) for any \( \lambda \in (0, 1) \). Which implies in turn that, for any \( \lambda, \lambda' \in (0, 1) : \lambda' < \lambda \), we have

\[
\lambda \Pi_L(a^*(\lambda)) + (1 - \lambda) \Pi_H(a^*(\lambda)) = 0
\]

\[
= \lambda \Pi_L(a^*(\lambda')) + (1 - \lambda) \Pi_H(a^*(\lambda'))
\]

\[
< \lambda \Pi_L(a^*(\lambda')) + (1 - \lambda) \Pi_H(a^*(\lambda'))
\]

That is, \( a^*(\lambda) \neq a^*(\lambda') \) and, thus, \( U_L(a^*(\lambda')) < U_L(a^*(\lambda)) \) given that \( a^*(\cdot) \) is a function.

Let now \( a^*_L : (0, 1) \to \mathcal{A} \) be the function that maps to the IE(1)-optimal contract for the low risk-type in the economy \( E(u, p_L, p_H, \lambda) \). By Lemma 1, this is also well-defined while \( \Pi_L(a^*_L) > 0 > \Pi_H(a^*_H) \). Clearly, the preceding argument can be deployed again to establish that also \( U_L(a^*_L(\cdot)) \) is everywhere strictly increasing.

Next, recall that the RSW allocation \( \{a^*_L, a^*_H\} \) is independent of the parameter \( \lambda \). Hence, \( U_L(a^*(\cdot)) \) being strictly monotone, the value

\[
\lambda_0 \in (0, 1) : U_L(a^*(\lambda_0)) = U_L(a^*_L)
\]

is defined uniquely. By Claim 3, moreover, the RSW menu is the equilibrium of the two-stage game and, thus, condition (ii) of the proposition obtains in the economy \( E(u, p_L, p_H, \lambda_0) \). Notice also that the singleton menu \( \{a^*(\lambda_0)\} \) is feasible in the IE(1) problem for \( E(u, p_L, p_H, \lambda_0) \) but clearly not optimal. In other words, \( a^*_L(\lambda_0) >_L a^*(\lambda_0) \sim_L a^*_H \) and, thus, also condition (i) obtains in the economy \( E(u, p_L, p_H, \lambda_0) \).

Let now \( \Delta = U_L(a^*_L(\lambda_0)) - U_L(a^*(\lambda_0)) \) and consider an arbitrary \( \lambda < \lambda_0 \). Since \( U_L(a^*(\lambda)) < U_L(a^*(\lambda_0)) = U_L(a^*_L) \), the RSW allocation remains the equilibrium of the two-stage game and condition (ii) applies also in the economy \( E(u, p_L, p_H, \lambda) \). Moreover, it must be \( U_L(a^*_L(\lambda)) < U_L(a^*_L(\lambda_0)) \). Yet, being the value function of the IE(1) problem, \( U_L(a^*_L(\cdot)) \) is continuous (Lemma 18). There exists, therefore, some \( \lambda_1 \in [0, \lambda_0) \) such that \( U_L(a^*_L(\lambda_1)) - U_L(a^*_L(\lambda)) < \Delta \) for all \( \lambda \in (\lambda_1, \lambda_0) \). It follows then that

\[
U_L(a^*_L(\lambda)) - U_L(a^*_L) = U_L(a^*_L(\lambda)) - U_L(a^*_L(\lambda_0)) + U_L(a^*_L(\lambda_0)) - U_L(a^*_L) - U_L(a^*_L(\lambda_0)) + \Delta + 0 > 0
\]
for all $\lambda \in (\lambda_1, \lambda_0)$. Clearly, the RSW allocation is not the IE(1) optimum and consequently (recall Claim 2) condition (i) applies in the economy $E (u, p_L, p_H, \lambda)$ for any $\lambda \in (\lambda_1, \lambda_0)$.

We have just shown that both conditions of the theorem obtain in the economy $E (u, p_L, p_H, \lambda)$ for any $\lambda \in (\lambda_1, \lambda_0)$. And as $(u, p_L, p_H)$ was arbitrarily chosen from $\mathcal{U} \times \left\{ (p_1, p_2) \in (0, 1)^2 : p_1 < p_2 \right\}$ while $(\lambda_1, \lambda_0]$ has positive Lebesgue measure on $(0, 1)$, the claim follows. □