Web Appendix for Overconfidence, Insurance and Paternalism.

Alvaro Sandroni* Francesco Squintani†

Abstract

This paper is not self-contained. It is the mathematical appendix of the paper "Overconfidence, Insurance and Paternalism."

1 Mathematical Appendix

Equilibrium Analysis. This section formalizes the graphical equilibrium analysis of Section II. Before presenting the analysis, we formally define locally-competitive equilibrium.

A *locally-competitive equilibrium* is a set of contracts A such that when each contract $\boldsymbol{\alpha} \in A$ is available in the market, (i) no contract $\boldsymbol{\alpha} \in A$ makes strictly negative expected profits, and (ii) there is an $\varepsilon > 0$ such that any contract $\boldsymbol{\alpha}'$ for which $||\boldsymbol{\alpha} - \boldsymbol{\alpha}'|| < \varepsilon$ for any $\boldsymbol{\alpha} \in A$, would not make strictly positive profits.

The first step in the equilibrium analysis shows that overconfident and low-risk agents pool together, and together they separate from high-risk agents. For future reference, we define the marginal rate of substitution associated to contract $\boldsymbol{\alpha}$ and risk p, as:

$$M(\boldsymbol{\alpha}, p) = \frac{(1-p)\,U'(W-\alpha_1)}{pU'(W-d+\alpha_2)}.$$

Proposition 1 *In the unique locally-competitive equilibrium, high-risk individuals choose the contract $\boldsymbol{\alpha}^H = (p_H d, (1-p_H)d)$. Low-risk and overconfident individuals choose the contract $\boldsymbol{\alpha}^{LO}$ that solves the maximization problem*

$$\max_{\boldsymbol{\alpha}} V(W, d; p_L, \boldsymbol{\alpha}), \tag{1}$$

subject to the non-negativity constraint $\boldsymbol{\alpha} \geqq \mathbf{0}$, and to the incentive compatibility and zero-profit conditions:

$$V(W, d; p_H, \boldsymbol{\alpha}^H) \geq V(W, d; p_H, \boldsymbol{\alpha}), \tag{2}$$

$$(1-p_{LO})\,\alpha_1 - p_{LO}\alpha_2 = 0. \tag{3}$$

As long as $\boldsymbol{\alpha}^{LO} > \mathbf{0}$, the insurance price P^{LO} equals p_{LO} and increases in κ.

Proof. Step 1. *In equilibrium, types L and O pool on the same contract $\boldsymbol{\alpha}^{LO}$, type H chooses a different contract $\boldsymbol{\alpha}^H$.*

*Department of Economics, University of Economics, 3708 Locust Walk, Philadelphia, PA 19104 and Kellogg School of Management, MEDS Department, 2001 Sheridan Rd., Evanston, IL 60208, USA email: sandroni@sas.upenn.edu.

†Universita' degli Studi di Brescia. Via S. Faustino, 74/B, 25122 Brescia, Italy. Email: squintan@eco.unibs.it.

For any contract $\boldsymbol{\alpha}$, bought by types H, L and O with probabilities σ_H^α, σ_L^α, and σ_O^α, respectively, let the average risk be:

$$p_\alpha = \frac{p_H \left(\kappa\sigma_O^\alpha + (1-\kappa-\lambda)\,\sigma_H^\alpha\right) + p_L\lambda\sigma_L^\alpha}{\kappa\sigma_O^\alpha + (1-\kappa-\lambda)\,\sigma_H^\alpha + \lambda\sigma_L^\alpha},$$

Consider any equilibrium contract $\boldsymbol{\alpha}$ such that $\sigma_H^\alpha \geq 0$, and $\sigma_L^\alpha + \sigma_O^\alpha > 0$. Hence, $V(W,d;p_L,\boldsymbol{\alpha}) = V(W,d;p_L,\boldsymbol{\beta})$ for any equilibrium contract $\boldsymbol{\beta}$ such that $\sigma_L^\beta + \sigma_O^\beta > 0$, and $V(W,d;p_H,\boldsymbol{\alpha}) \leq V(W,d;p_L,\boldsymbol{\beta})$ for any contract $\boldsymbol{\beta}$ such that $\sigma_H^\beta > 0$. Further, competition requires that $\pi(\boldsymbol{\alpha}) \equiv (1-p_\alpha)\alpha_1 - p_\alpha\alpha_2 = 0$, or else there is a local profitable deviation, by continuity.

Suppose by contradiction that $p_\alpha > p_{LO}$. Because $(1-p_L)/p_L > (1-p_H)/p_H$, it follows that $M(\boldsymbol{\alpha}, p_L) > M(\boldsymbol{\alpha}, p_H)$. Since U is twice differentiable, there is an $\varepsilon > 0$ small enough such that for any $m \in (M(\boldsymbol{\alpha}, p_H), M(\boldsymbol{\alpha}, p_L))$, the contract $\boldsymbol{\alpha} - \varepsilon(1,m)$ is purchased by all type L and O agents but not by type H agents. Hence, $\boldsymbol{\alpha} - \varepsilon(1,m)$ yields expected profit $(1-p_{LO})(\alpha_1 - \varepsilon) - p_{LO}(\alpha_2 - \varepsilon m)$, which is strictly bigger than $\pi(\boldsymbol{\alpha}) = 0$ for ε small enough because $p_\alpha > p_{LO}$. Because $\boldsymbol{\alpha} - \varepsilon(1,m)$ is a local profitable deviation, $\boldsymbol{\alpha}$ cannot be an equilibrium contract.

Because $p_\alpha \leq p_{LO}$ for any equilibrium contract $\boldsymbol{\alpha}$ such that $\sigma_L^\alpha + \sigma_O^\alpha > 0$, it follows that (i) $\sigma_H^\alpha = 0$ whenever $\sigma_L^\alpha + \sigma_O^\alpha > 0$, and that (ii) $p_\alpha = p_{LO}$ for all $\boldsymbol{\alpha}$ such that $\sigma_L^\alpha + \sigma_O^\alpha > 0$. Because $\pi(\boldsymbol{\alpha}) = 0$ for all equilibrium contracts, and $U'' < 0$, there are therefore at most two equilibrium contracts $\boldsymbol{\alpha}, \boldsymbol{\beta}$, with $\boldsymbol{\alpha} > \boldsymbol{\beta}$, such that $\sigma_L^\alpha + \sigma_O^\alpha > 0$ and $\sigma_L^\beta + \sigma_O^\beta > 0$. Because $M(\boldsymbol{\alpha}, p_H) < (1-p_H)/p_H < (1-p_{LO})/p_{LO}$, there is an $\varepsilon > 0$ small enough such that for any $m \in ((1-p_{LO})/p_{LO}, M(\boldsymbol{\alpha}, p_L))$, the contract $\boldsymbol{\alpha} - \varepsilon(1,m)$ is purchased by all type L and O agents but not by type H agents. The profit $\pi(\boldsymbol{\alpha} - \varepsilon(1,m))$ is strictly positive because $m > (1-p_{LO})/p_{LO}$. This concludes that types L and O must pool on the same contract $\boldsymbol{\alpha}^{LO}$. Because type H must separate from types L and O , and U is concave and twice differentiable, type H purchase a single different contract $\boldsymbol{\alpha}^H$ with probability one.

Step 2. *There exists a unique locally-competitive equilibrium, characterized in the statement of Proposition 1.*

By Step 1, if a locally-competitive equilibrium exists, it is a pair of distinct contracts $\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO}$ such that $\boldsymbol{\alpha}^{LO} \in \arg\max_{\boldsymbol{\alpha}} V(W,d;p_L,\boldsymbol{\alpha})$ s.t. $\boldsymbol{\alpha} \geqq \mathbf{0}$, $(1-p_{LO})\,\alpha_1 - p_{LO}\alpha_2 = 0$, $V(W,d;p_H,\boldsymbol{\alpha}^H) \geq V(W,d;p_H,\boldsymbol{\alpha})$; and $\boldsymbol{\alpha}^H \in \arg\max_{\boldsymbol{\alpha}'} V(W,d;p_H,\boldsymbol{\alpha}')$, s.t. $\boldsymbol{\alpha}' \geqq \mathbf{0}$, $(1-p_H)\,\alpha_1' - p_H\alpha_2' = 0$, $V(W,d;p_L,\boldsymbol{\alpha}^{LO}) \geq V(W,d;p_L,\boldsymbol{\alpha}')$. By construction, any other pair of contracts admits local profitable deviations. The contracts $\boldsymbol{\alpha}^H$ and $\boldsymbol{\alpha}^{LO}$ do not admit local deviations $\boldsymbol{\alpha}$ such that, respectively, $\sigma_H^\alpha > 0$, and $\sigma_L^\alpha + \sigma_O^\alpha > 0$. Because $(1-p_{LO})\,\alpha_1^{LO} - p_{LO}\alpha_2^{LO} = 0$, contract $\boldsymbol{\alpha}^{LO}$ has no local deviation $\boldsymbol{\alpha}$ with any distribution σ^α.

Suppose by contradiction that the constraint $V(W,d;p_L,\boldsymbol{\alpha}^{LO}) \geq V(W,d;p_L,\boldsymbol{\alpha}^H)$ binds in the solution of the $\boldsymbol{\alpha}^H$-maximization problem. Because $M(\boldsymbol{\alpha}, p_H) < M(\boldsymbol{\alpha}, p_L)$ for all $\boldsymbol{\alpha}$ and $V(W,d;p_H,\boldsymbol{\alpha}^H) \geq V(W,d;p_H,\boldsymbol{\alpha}^{LO})$, it follows that $\boldsymbol{\alpha}^H > \boldsymbol{\alpha}^{LO}$. But this and $V(W,d;p_L,\boldsymbol{\alpha}^{LO}) = V(W,d;p_L,\boldsymbol{\alpha}^H)$ are incompatible with $(1-p_{LO})\,\alpha_1^{LO} - p_{LO}\alpha_2^{LO} = 0$ and $(1-p_H)\,\alpha_1^H - p_H\alpha_2^H = 0$. Because $V(W,d;p_L,\boldsymbol{\alpha}^{LO}) > V(W,d;p_L,\boldsymbol{\alpha}^H)$, the contract $\boldsymbol{\alpha}^H$ does not admit any local profitable deviations $\boldsymbol{\alpha}$. Because U is twice differentiable and $U'' < 0$, the solution to the $\boldsymbol{\alpha}^H$-maximization problem is $\boldsymbol{\alpha}^H = (p_H d, (1-p_H)d)$. A solution to the $\boldsymbol{\alpha}^{LO}$-maximization problem exists and is unique because $U'' < 0$ and $M(\boldsymbol{\alpha}, p_H) < M(\boldsymbol{\alpha}, p_L)$ for all $\boldsymbol{\alpha}'$.

Finally, we note that, because $p_H > p_L$, $dp_{LO}/d\lambda < 0$ and $dp_{LO}/d\kappa > 0$. By condition (3), the price $P^{LO} = \alpha_1^{LO}/(\alpha_1^{LO} + \alpha_2^{LO})$ equals p_{LO}, and hence it increases in κ. ■

The equilibrium characterization is completed in the Proposition 2 below, which also reports our comparative statics results, and determines perfect-competitive equilibrium existence. For any parameter constellation (W, d, p_H, p_L), the thresholds κ_1 and κ_2, functions of λ, uniquely solve respectively:

$$V(W, d; p_H, \boldsymbol{\alpha}) = U(W - p_H d),\ p_{LO}\alpha_2 = (1 - p_{LO})\,\alpha_1,\ M(\boldsymbol{\alpha}, p_L) = (1 - p_{LO})/p_{LO}; \quad (4)$$
$$M(\mathbf{0}, p_L) = (1 - p_{LO})/p_{LO}. \quad (5)$$

where the variables κ and λ are embedded in the expression $p_{LO} = (\kappa p_L + \lambda p_H)/(\kappa + \lambda)$.

Proposition 2 *The incentive compatibility condition (2) binds if and only if $\kappa < \kappa_1(\lambda)$. For $\kappa_1(\lambda) < \kappa < \kappa_2(\lambda)$, the equilibrium contract $\boldsymbol{\alpha}^{LO}$ satisfies the tangency condition*

$$M(\boldsymbol{\alpha}, p_{LO}) = (1 - p_{LO})/p_{LO}. \quad (6)$$

Hence $V(W, d; p_H, \boldsymbol{\alpha}^{LO}) < V(W, d; p_H, \boldsymbol{\alpha}^H)$, and both $V(W, d; p_L, \boldsymbol{\alpha}^{LO})$ and $V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ decrease in κ and increase in λ, as long as the Relative Risk Aversion coefficient of U is bounded by $(W - d)/d$. For $\kappa > \kappa_2(\lambda)$, low-risk and overconfident individuals are uninsured: $\boldsymbol{\alpha}^{LO} = \mathbf{0}$. The locally-competitive equilibrium $(\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO})$ is also perfectly competitive if and only if $\lambda > \lambda_0(\kappa)$, where the function λ_0 is such that $\lambda_0^{-1} < \kappa_1$.

Proof. Let $\bar{\boldsymbol{\alpha}} = (\bar{\alpha}_1, \bar{\alpha}_2)$ be the contract pinned down by condition (3) and by the binding incentive compatibility condition (2). Differentiating these equations, we obtain:

$$\frac{d\bar{\alpha}_1}{dp_{LO}} = \frac{(\bar{\alpha}_1 + \bar{\alpha}_2)\,p_H U'(W - d + \bar{\alpha}_2)}{\Delta} > 0,\ \frac{d\bar{\alpha}_2}{dp_{LO}} = \frac{(\bar{\alpha}_1 + \bar{\alpha}_2)\,(1 - p_H)U'(W - \bar{\alpha}_1)}{\Delta} > 0, \quad (7)$$

where the quantity $\Delta = (1 - p_{LO})\,p_H U'(W - d + \bar{\alpha}_2) - p_{LO}(1 - p_H)U'(W - \bar{\alpha}_1)$ is positive because $U'' < 0$, $-\bar{\alpha}_1 > -d + \bar{\alpha}_2$ and $p_H > p_{LO}$. Because $dp_{LO}/d\lambda < 0$ and $dp_{LO}/d\kappa > 0$, we obtain that $d\bar{\alpha}_1/d\kappa > 0$, $d\bar{\alpha}_1/d\lambda < 0$, $d\bar{\alpha}_2/d\kappa > 0$, and $d\bar{\alpha}_2/d\lambda < 0$.

Let $\chi = (1 - p_{LO})/p_{LO}$. Because

$$dM(\bar{\boldsymbol{\alpha}}, p_L) = \frac{1 - p_L}{p_L}\left[-\frac{U''(W - \bar{\alpha}_1)}{U'(W - d + \bar{\alpha}_2)}d\bar{\alpha}_1 - \frac{U''(W - d + \bar{\alpha}_2)\,U'(W - \bar{\alpha}_1)}{(U'(W - d + \bar{\alpha}_2))^2}d\bar{\alpha}_2\right],$$

we obtain: $dM(\bar{\boldsymbol{\alpha}}, p_L)/d\kappa > 0$. Because $d\chi/dp_{LO} < 0$ and $dp_{LO}/d\kappa > 0$, we have shown that for any λ, there is a unique threshold κ_1 pinned down by system (4) and that $M(\bar{\boldsymbol{\alpha}}, p_L) > (<)(1 - p_{LO})/p_{LO}$ if and only if $\kappa > (<)\kappa_1(\lambda)$. Because $d\chi/d\lambda < 0, d\chi/dp_{LO} < 0$ and $dp_{LO}/d\lambda < 0$, κ_1 is strictly increasing in λ by the implicit function theorem.

Suppose that $\kappa < \kappa_1(\lambda)$, and that, by contradiction, condition (2) does not bind in equilibrium: $(1 - p_H)\,U(W - \alpha_1^{LO}) + p_H U(W - d + \alpha_2^{LO}) < U(W - p_H d)$. Since $U'' < 0$, and both $\bar{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}^{LO}$ satisfy condition (3), it must be that $\bar{\boldsymbol{\alpha}} < \boldsymbol{\alpha}^{LO}$ and hence that $M(\boldsymbol{\alpha}^{LO}, p_L) < M(\bar{\boldsymbol{\alpha}}, p_L) < (1 - p_{LO})/p_{LO}$. Because U is twice differentiable, there is an $\varepsilon > 0$ small enough such that for any $m \in (M(\boldsymbol{\alpha}^{LO}, p_L), (1 - p_{LO})/p_{LO})$, the contract $\boldsymbol{\alpha}^{LO} + \varepsilon\,(1, m)$ is chosen by type L and O but not by type H, and makes strictly positive profit. This concludes that for $\kappa < \kappa_1(\lambda)$, $\boldsymbol{\alpha}^{LO} = \bar{\boldsymbol{\alpha}}$.

Suppose that $\kappa > \kappa_1(\lambda)$, and hence that $M(\bar{\boldsymbol{\alpha}}, p_L) > (1 - p_{LO})/p_{LO}$. Suppose by contradiction that $\boldsymbol{\alpha}^{LO} = \bar{\boldsymbol{\alpha}}$ in equilibrium. Note that $M(\bar{\boldsymbol{\alpha}}, p_H) < (1 - p_H)/p_H < (1 - p_{LO})/p_{LO}$. Since $U'' < 0$ and U is smooth, for any $\varepsilon > 0$ small enough, and $m \in ((1 - p_{LO})/p_{LO}, M(\bar{\boldsymbol{\alpha}}, p_L))$, the contract $\bar{\boldsymbol{\alpha}} - \varepsilon\,(1, m)$ is chosen only by types L and O, and not by type H, and yields strictly positive profit. This proves that condition (2) does not bind in equilibrium.

Since $dp_{LO}/d\kappa > 0$, for any λ there is a unique threshold $\kappa_2(\lambda)$ such that $M(\mathbf{0}, p_L) > (<)(1 - p_{LO})/p_{LO}$ if and only if $\kappa > (<)\kappa_2(\lambda)$. When $\kappa > \kappa_2(\lambda)$, the constraint $\boldsymbol{\alpha} \geq \mathbf{0}$ binds in equilibrium, whereas when $\kappa_1(\lambda) < \kappa < \kappa_2(\lambda)$, the equilibrium contract $\boldsymbol{\alpha}^{LO}$ is pinned down by condition (3) and by the tangency condition (6). Since $dp_{LO}/d\lambda < 0$, the function κ_2 is increasing in λ.

Low-risk individuals' utility $V(W, d; p_L, \boldsymbol{\alpha}^{LO})$ decreases in p_{LO} –hence decreasing in κ and increasing in λ– by a simple revealed-preference argument. The overconfident agents' utility $V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ decreases in p_{LO} if the insurance coverage $\alpha_1^{LO} + \alpha_2^{LO}$ decreases in p_{LO}, because the marginal rate of substitution $M(\boldsymbol{\alpha}^{LO}, p_H)$ is larger than $M(\boldsymbol{\alpha}^{LO}, p_L)$. Indeed, we differentiate conditions (3) and (6) with respect to the quantity χ, decreasing in p_{LO}, and obtain:

$$\frac{\partial\left(\alpha_1^{LO} + \alpha_2^{LO}\right)}{\partial\chi} = \alpha_1^{LO} - (1+\chi)\, p_L \frac{U'\left(W - d + \alpha_2^{LO}\right) + U''\left(W - d + \alpha_2^{LO}\right)\alpha_2^{LO}}{(1-p_L)\, U''(W - \alpha_1^{LO}) + \chi^2 p_L U''\left(W - d + \alpha_2^{LO}\right)}.$$

This derivative is positive because $U'\left(W - d + \alpha_2^{LO}\right) + U''\left(W - d + \alpha_2^{LO}\right)\alpha_2^{LO} > 0$, which follows by the hypothesis that $-U'(w)\, w/U''(w) < (W-d)/d$.

By construction, the pair $(\boldsymbol{\alpha}^{LO}, \boldsymbol{\alpha}^H)$ is the (unique) perfectly-competitive equilibrium if and only if it does not admit any pooling, possibly large profitable deviation $\boldsymbol{\alpha}$. Hence, it is necessary and sufficient that $V(W, d; p_L, \boldsymbol{\alpha}^{LO}) \geq V(W, d; p_L, \boldsymbol{\beta})$, where $p_{LH} = \lambda p_L + (1-\lambda)p_H$ and

$$\boldsymbol{\beta} = \arg\max_{\boldsymbol{\alpha}} V(W, d; p_L, \boldsymbol{\alpha}) \quad \text{s.t.} \quad p_{LH}\alpha_2 \leq (1 - p_{LH})\alpha_1,\ \boldsymbol{\alpha} \geqq \mathbf{0}. \tag{8}$$

When $\kappa \geq \kappa_1(\lambda)$, condition (2) does not bind in equilibrium. Thus, by revealed preferences, $V(W, d; p_L, \boldsymbol{\alpha}^{LO}) \geq V(W, d; p_L, \boldsymbol{\beta})$ because $p_{LH} \geq p_{LO}$, and hence $\left(\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO}\right)$ is the perfectly-competitive equilibrium.

Suppose that $\kappa < \kappa_1(\lambda)$. The utility $V(W, d; p_L, \boldsymbol{\alpha}^{LO})$ decreases in κ and increases in λ because $dp_{LO}/d\kappa > 0$, $dp_{LO}/d\lambda < 0$ and

$$\frac{\partial V(W, d; p_L, \boldsymbol{\alpha}^{LO})}{\partial p_{LO}} = -\left(1 - p_L\right) U'(W - \alpha_1^{LO}) \frac{\alpha_1^{LO} + \alpha_2^{LO}}{1 - p_{LO}} - p_L U'(W - d + \alpha_2^{LO}) \frac{\alpha_1^{LO} + \alpha_2^{LO}}{p_{LO}} < 0,$$

after substituting in condition (7). By revealed preferences, $V(W, d; p_L, \boldsymbol{\beta})$ increases in λ but it is constant in κ (p_{LH} depends only on λ). Hence, there is a unique strictly-increasing threshold λ_0, function of κ, such that $\left(\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO}\right)$ is a perfectly-competitive equilibrium if and only if $\lambda > \lambda_0(\kappa)$.
■

Policy Recommendations This section proves our policy results 1, 2, 3, and 4. We begin by formally stating and proving Result 2. So, we need to formally define the general mechanism design problem in our model. Because agents differ in actual risk $p \in \{p_H, p_L\}$ and perceived risk $\hat{p} \in \{p_H, p_L\}$, we let the type space be $\Psi = \{p_H, p_L\} \times \{p_H, p_L\}$. The type distribution ρ is easily derived from the parameter κ and λ. An *allocation* is a profile $\boldsymbol{\alpha}^* : \Psi \to \mathbb{R}_+^2$, and $A^* = \mathbb{R}_+^{2\Psi}$ is the set of allocations. An allocation $\boldsymbol{\alpha}^*$ is *incentive compatible* if

$$\hat{V}(\boldsymbol{\psi}, \boldsymbol{\alpha}^*(\boldsymbol{\psi})) \geq \hat{V}(\boldsymbol{\psi}, \boldsymbol{\alpha}^*(\boldsymbol{\psi}')) \text{ for all } (\boldsymbol{\psi}, \boldsymbol{\psi}') \in \Psi^2, \tag{9}$$

where the perceived expected utility of any type $\boldsymbol{\psi} = (p, \hat{p})$ with contract $\boldsymbol{\alpha}$ is $\hat{V}(\boldsymbol{\psi}, \boldsymbol{\alpha}) = V(W, d; \hat{p}, \boldsymbol{\alpha})$. The allocation $\boldsymbol{\alpha}^*$ is *feasible* if $\sum_{\boldsymbol{\psi}\in\Psi} \rho_{\boldsymbol{\psi}} \pi(\boldsymbol{\psi}, \boldsymbol{\alpha}^*(\boldsymbol{\psi})) \leq 0$ where for any type

$\boldsymbol{\psi} = (p, \hat{p})$, the profit of a contract $\boldsymbol{\alpha} \in \mathbb{R}_+^2$ is $\pi^*(\boldsymbol{\psi}, \boldsymbol{\alpha}) = (1-p)\alpha_1 - p\alpha_2$. Because of monotonicity of individuals' utilities, we can restrict attention without loss of generality to *budget-balanced* allocations $\boldsymbol{\alpha}^*$ that satisfy

$$\sum_{\boldsymbol{\psi} \in \Psi} \rho_{\boldsymbol{\psi}} \pi^*(\boldsymbol{\psi}, \boldsymbol{\alpha}^*(\boldsymbol{\psi})) = 0. \tag{10}$$

A mechanism designer implements an allocation $\boldsymbol{\alpha}^*$ on the basis of the information revealed by the agents. Each individual *only knows* her perceived risk $\hat{p}$, and she (maybe mistakenly) believes that her actual risk p coincides with $\hat{p}$. She can only communicate her perceived ability $\hat{p}$ to the mechanism-designer. Hence we restrict attention to allocations $\boldsymbol{\alpha}^*$ that are constant across the actual risk p. We let $\bar{A} = \{\boldsymbol{\alpha}^* \in A^* : \boldsymbol{\alpha}^*(p_L, \hat{p}) = \boldsymbol{\alpha}^*(p_H, \hat{p}), \text{ for any } \hat{p} \in \{p_H, p_L\}\}$. We can now formally restate and prove Result **??**.

Result 2 *Suppose that $\kappa > \kappa_1(\lambda)$. Then there is no allocation $\boldsymbol{\alpha}^* \in \bar{A}$ that improves the expected utility of both high and low risk agents with respect to the equilibrium outcome $\left(\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO}\right)$.*

Proof. Any candidate allocation $\boldsymbol{\alpha}^*$ must satisfy $V(W, d; p_H, \boldsymbol{\alpha}^*(p_H, p_H)) \geq V\left(W, d; p_H, \boldsymbol{\alpha}^H\right)$. In equilibrium $\boldsymbol{\alpha}^H \in \arg\max_{\boldsymbol{\alpha}} V(W, d; p_H, \boldsymbol{\alpha})$ such that $p_H\alpha_2 = (1-p_H)a_1$. Hence, the candidate allocation $\boldsymbol{\alpha}^*$ must satisfy $p_H\alpha_2^*(p_H, p_H) \geq (1-p_H)\alpha_1^*(p_H, p_H)$. The contracts $\boldsymbol{\alpha}^*(p_L, p_L)$ and $\boldsymbol{\alpha}^*(p_H, p_L)$ coincide by construction. By the budget-balance condition (10) this constrains the terms of the contracts

$$p_{LO}\alpha_2^*(p_L, p_L) \leq (1-p_{LO})\,\alpha_1^*(p_L, p_L). \tag{11}$$

But when $\kappa > \kappa_1(\lambda)$, in equilibrium, $\boldsymbol{\alpha}^{LO} = \arg\max_{\boldsymbol{\alpha}} V(W, d; p_L, \boldsymbol{\alpha})$ such that $p_L\alpha_2 = (1-p_L)a_1$, by Proposition 2. Hence, the allocation $\boldsymbol{\alpha}^*$ cannot be better than $\boldsymbol{\alpha}^{LO}$ for agents of type $\boldsymbol{\psi} = (p_L, p_L)$, i.e. $V(W, d; p_L, \boldsymbol{\alpha}^*(p_L, p_L)) < V\left(W, d; p_L, \boldsymbol{\alpha}^{LO}\right)$. ■

Result 1 immediately follows from the proof of Result 2.

Proof of Result 1. For any compulsory insurance contract $\boldsymbol{\beta} > \mathbf{0}$, the associated equilibrium allocation $\boldsymbol{\alpha}^*$ such that $\boldsymbol{\alpha}^*(p_L, p_L) = \boldsymbol{\alpha}^*(p_L, p_H) = \boldsymbol{\beta} + \boldsymbol{\alpha}^{LO}(\boldsymbol{\beta})$ and $\boldsymbol{\alpha}^*(p_H, p_H) = \boldsymbol{\beta} + \boldsymbol{\alpha}^H(\boldsymbol{\beta})$ is budget balanced and incentive compatible. Furthermore, $V(W, d; p_H, \boldsymbol{\alpha}^*(p_H, p_H)) > V\left(W, d; p_H, \boldsymbol{\alpha}^H\right)$ because $\beta_2/\beta_1 = (1-p_{LH})/p_{LH} < (1-p_{LO})/p_{LO}$, and $\alpha_1^*(p_H, p_H) + \alpha_2^*(p_H, p_H) = d$. The proof of Result 2 thus concludes that, for $\kappa > \kappa_1(\lambda)$, $V(W, d; p_L, \boldsymbol{\alpha}^*(p_L, p_L)) < V\left(W, d; p_L, \boldsymbol{\alpha}^{LO}\right)$. ■

In order to prove Result **??**, we first formally describe the equilibrium of our model with overconfident and underconfident agents.

Proposition 3 *In the unique locally-competitive equilibrium, the contract of high-risk and underconfident agents is $\boldsymbol{\alpha}^{HU}$ such that*

$$p_{HU}\alpha_2^{HU} - (1-p_{HU})\,\alpha_1^{HU} = 0, \; M\left(\boldsymbol{\alpha}^{HU}, p_H\right) = (1-p_{HU})/p_{HU} \tag{12}$$

the contract of low-risk and overconfident agents is

$$\begin{aligned} \boldsymbol{\alpha}^{LO} \;&=\; \max_{\boldsymbol{\alpha}} V(W, d; p_L, \boldsymbol{\alpha}) \\ &\quad \textit{s.t.}\;\; \boldsymbol{\alpha} \geqq \mathbf{0}, \;\; p_{LO}\alpha_2 = (1-p_{LO})\,\alpha_1, \;\; V(W, d; p_H, \boldsymbol{\alpha}^{HU}) \geq V(W, d; p_H, \boldsymbol{\alpha}). \end{aligned} \tag{13}$$

Proof. For any contract $\boldsymbol{\alpha}$, let

$$p_\alpha = \frac{p_H\left(\kappa\sigma_O^\alpha + \eta\sigma_H^\alpha\right) + p_L\left(\lambda\sigma_L^\alpha + \upsilon\sigma_U^\alpha\right)}{\kappa\sigma_O^\alpha + \eta\sigma_H^\alpha + \lambda\sigma_L^\alpha + \upsilon\sigma_U^\alpha},$$

where σ_U^α is the probability that U purchases $\boldsymbol{\alpha}$.

Arguments in the proof of Proposition 1 conclude that $p_\alpha \le p_{LO}$ for any equilibrium $\boldsymbol{\alpha}$ such that $\sigma_L^\alpha + \sigma_O^\alpha > 0$. Also, $p_\alpha \le p_{HU}$ for any equilibrium $\boldsymbol{\alpha}$ such that $\sigma_H^\alpha + \sigma_U^\alpha > 0$, or else the contract $\boldsymbol{\alpha} + \varepsilon(1, m)$ with $m > M(\boldsymbol{\alpha}, p_H)$ would be a profitable deviation for $\varepsilon > 0$ small enough. These two results conclude that (i) $\sigma_H^\alpha + \sigma_U^\alpha = 0$ and $p_\alpha = p_{LO}$ whenever $\sigma_L^\alpha + \sigma_O^\alpha > 0$, and (ii) $\sigma_L^\alpha + \sigma_O^\alpha = 0$ and $p_\alpha = p_{HU}$ whenever $\sigma_H^\alpha + \sigma_U^\alpha > 0$. Because U is concave and twice differentiable, types L and O pool on the same contract $\boldsymbol{\alpha}^{LO}$ and types H and U pool on a different contract $\boldsymbol{\alpha}^{HU}$.

Arguments in the proof of Proposition 1, with obvious modifications, conclude that there exists a unique locally-competitive equilibrium such that $\boldsymbol{\alpha}^{LO}$ is as specified in program (13) and $\boldsymbol{\alpha}^{H} \in \arg\max_{\boldsymbol{\alpha}'} V(W, d; p_H, \boldsymbol{\alpha}')$, s.t. $\boldsymbol{\alpha}' \geqq \mathbf{0}$, $(1 - p_H)\,\alpha_1' - p_H \alpha_2' = 0$. Hence $\boldsymbol{\alpha}^{HU}$ is determined by equations (12). ■

We can now prove Result **??**.

Proof of Result ??. The proof of Proposition 2, with obvious modifications, concludes that (i) the incentive compatibility constraint $V(W, d; p_H, \boldsymbol{\alpha}^{HU}) \ge V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ does not bind in equilibrium if and only if $\kappa > \bar{\kappa}\,(\upsilon, \lambda)$ where $\bar{\kappa}$ solves

$$\begin{aligned} M\,(\bar{\boldsymbol{\alpha}}, p_L) &= (1 - p_{LO})/p_{LO}, & (14)\\ p_{LO}\bar{\alpha}_2 &= (1 - p_{LO})\,\bar{\alpha}_1,\ V(W, d; p_H, \boldsymbol{\alpha}^{HU}) = V(W, d; p_H, \bar{\boldsymbol{\alpha}}), & (15)\end{aligned}$$

and that (ii), when $V(W, d; p_H, \boldsymbol{\alpha}^{HU}) > V(W, d; p_H, \boldsymbol{\alpha})$, the locally-competitive equilibrium $(\boldsymbol{\alpha}^{HU}, \boldsymbol{\alpha}^{LO})$ is perfectly competitive.

The general mechanism design problem defined above applies to our model also when $\upsilon > 0$. The proof of Result 2, with obvious modifications, shows that, when $V(W, d; p_H, \boldsymbol{\alpha}^{HU}) > V(W, d; p_H, \boldsymbol{\alpha})$, there is no incentive-compatible budget-balanced mechanism that improves all agents' welfare upon the equilibrium $\left(\boldsymbol{\alpha}^{HU}, \boldsymbol{\alpha}^{LO}\right)$.

We conclude the proof by showing that the function $\bar{\kappa}$ decreases in υ. Differentiating equations 12, we obtain:

$$\begin{aligned} \frac{d\alpha_1^{HU}}{dp_{HU}} &= \Delta\left[p_{HU}\left((1 - p_H)\,U'\left(W - \alpha_1^{HU}\right) + p_H U'\left(W - d + \alpha_2^{HU}\right)\right)\right.\\ &\quad \left. + \left(\alpha_1^{HU} + \alpha_2^{HU}\right) p_H\,(1 - p_{HU})\,U''\left(W - d + \alpha_2^{HU}\right)\right]\\ \frac{d\alpha_2^{HU}}{dp_{HU}} &= \Delta\left[(1 - p_{HU})\left((1 - p_H)\,U'\left(W - \alpha_1^{HU}\right) + p_H U'\left(W - d + \alpha_2^{HU}\right)\right)\right.\\ &\quad \left. - \left(\alpha_1^{HU} + \alpha_2^{HU}\right) p_{HU}\,(1 - p_H)\,U''\left(W - \alpha_1^{HU}\right)\right]\end{aligned}$$

where $\Delta = \left[p_{HU}^2\,(1 - p_H)\,U''\left(W - \alpha_1^{HU}\right) + (1 - p_{HU})^2\,p_H U''\left(W - d + \alpha_2^{HU}\right)\right]^{-1} < 0.$

Differentiating the expression (15), and then substituting for $d\alpha_1^{HU}/dp_{HU}$ and $d\alpha_2^{HU}/dp_{HU}$, we obtain:

$$\begin{aligned} \frac{d\bar{\alpha}_1}{dp_{HU}} &= \frac{p_{LO} p_H U'(W - d + \alpha_2^{HU}) d\alpha_2^{HU} - p_{LO}(1 - p_H) U'(W - \alpha_1^{HU}) d\alpha_1^{HU}}{p_H\,(1 - p_{LO})\,U'(W - d + \bar{\alpha}_2) - p_{LO}(1 - p_H) U'(W - \bar{\alpha}_1)}\\ &\propto \Delta p_H U'(W - d + \alpha_2^{HU})\left[-\left(\alpha_1^{HU} + \alpha_2^{HU}\right) p_{HU}\,(1 - p_H)\,U''\left(W - \alpha_1^{HU}\right)\right.\\ &\quad \left. + (1 - p_{HU})\left((1 - p_H)\,U'\left(W - \alpha_1^{HU}\right) + p_H U'\left(W - d + \alpha_2^{HU}\right)\right)\right]\\ &\quad -\Delta(1 - p_H) U'(W - \alpha_1^{HU})\left[\left(\alpha_1^{HU} + \alpha_2^{HU}\right) p_H\,(1 - p_{HU})\,U''\left(W - d + \alpha_2^{HU}\right)\right.\\ &\quad \left. + p_{HU}\left((1 - p_H)\,U'\left(W - \alpha_1^{HU}\right) + p_H U'\left(W - d + \alpha_2^{HU}\right)\right)\right] \equiv \Psi\end{aligned}$$

because $p_{LO} > 0$, $p_H(1-p_{LO}) > p_{LO}(1-p_H)$ and $U'(W-d+\bar{\alpha}_2) > U'(W-\bar{\alpha}_1)$. Remembering that $\Delta < 0$, and $U'' < 0$,

$$\begin{aligned}\Psi &< \Delta p_H(1-p_{HU})U'(W-d+\alpha_2^{HU})\left((1-p_H)U'\left(W-\alpha_1^{HU}\right)+p_H U'\left(W-d+\alpha_2^{HU}\right)\right)\\ &\quad -\Delta(1-p_H)p_{HU}U'(W-\alpha_1^{HU})\left((1-p_H)U'\left(W-\alpha_1^{HU}\right)+p_H U'\left(W-d+\alpha_2^{HU}\right)\right)\\ &\propto -\left(p_H(1-p_{HU})U'(W-d+\alpha_2^{HU})-p_{HU}(1-p_H)U'(W-\alpha_1^{HU})\right)<0,\end{aligned}$$

because $p_H > p_{HU}$ and $U'(W-d+\alpha_2^{HU}) > U'(W-\alpha_1^{HU})$. Because $dp_{HU}/dv < 0$, we conclude that $d\bar{\alpha}_1/dv > 0$. As $d\bar{\alpha}_2/dp_{HU} = [(1-p_{HU})/p_{HU}][d\bar{\alpha}_1/dp_{HU}]$, we have $d\bar{\alpha}_2/dv > 0$.

Differentiating $M(\bar{\boldsymbol{\alpha}}, p_L)$, we obtain

$$dM(\bar{\boldsymbol{\alpha}}, p_L) = \frac{1-p_L}{p_L}\left[-\frac{U''(W-\bar{\alpha}_1)}{U'(W-d+\bar{\alpha}_2)}d\bar{\alpha}_1 - \frac{U''(W-d+\bar{\alpha}_2)U'(W-\bar{\alpha}_1)}{(U'(W-d+\bar{\alpha}_2))^2}d\bar{\alpha}_2\right].$$

Hence $dM(\bar{\boldsymbol{\alpha}}, p_L)/dv > 0$. Letting $\chi = (1-p_{LO})/p_{LO}$, because $d\chi/dp_{LO} < 0$ and $dp_{LO}/dv = 0$, and because $dM(\bar{\boldsymbol{\alpha}}, p_L)/d\kappa > 0$, $d\chi/dp_{LO} < 0$ and $dp_{LO}/d\kappa > 0$, $\bar{\kappa}$ decreases in v by the implicit function theorem. ■

We conclude by proving result 4.

Proof of Result 4. For any fraction κ, let $\boldsymbol{\alpha}^{LO}(\kappa)$ be the associated contract as calculated in Proposition 2. Suppose that in equilibrium all low-risk and overconfident agents join the program. For q large enough, $\kappa' = (1-q)\kappa < \kappa_2(\lambda)$. By Proposition 2, the low-risk agents' equilibrium utility (and the overconfident agents' perceived utility) $V\left(W, d, p_L, \boldsymbol{\alpha}^{LO}(\kappa')\right)$ decreases in κ' when $\kappa' \leq \kappa_2(\lambda)$, and it is constant in κ' for $\kappa' \geq \kappa_2(\lambda)$. Hence, for c small enough, $V(W-c, d, p_L, \boldsymbol{\alpha}^{LO}(\kappa')) > V\left(W, d, p_L, \boldsymbol{\alpha}^{LO}(\kappa)\right)$. This implies that (i) all low-risk and overconfident agents join the training program, hence verifying our equilibrium imputation, and (ii) in equilibrium low-risk agents benefit from the adoption of voluntary training programs.

The high-risk agents' equilibrium utility $V(W, d; p_H, \boldsymbol{\alpha}^H)$ is constant in κ. Because $c > 0$, they choose not to join training programs. By Proposition 2, when $\kappa \geq \kappa_2(\lambda)$, the equilibrium overconfident agents' utility $V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ is constant in κ. When $\kappa \in [\kappa_1(\lambda), \kappa_2(\lambda)]$, $V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ decreases in κ. For any $\kappa > \kappa_1$, $V(W, d; p_H, \boldsymbol{\alpha}^{LO})$ is smaller than the high-risk agents utility $V\left(W, d, p_H, \boldsymbol{\alpha}^H\right)$. For c small enough, agents who remain overconfident despite participating in the program improve their actual welfare because $V(W-c, d; p_H, \boldsymbol{\alpha}^{LO}(\kappa')) > V(W, d; p_H, \boldsymbol{\alpha}^{LO}(\kappa))$, as $\kappa' < \kappa$. Overconfident agents who change their beliefs improve their actual welfare because $V(W-c, d; p_H, \boldsymbol{\alpha}^H) > V(W, d; p_H, \boldsymbol{\alpha}^{LO}(\kappa))$. ■