ISSN 1755-5361

University of Essex

Department of Economics

Discussion Paper Series

No. 643 October 2007

Overconfidence, Insurance and Paternalism

Alvaro Sandroni and Francesco Squintani

Overconfidence, Insurance and Paternalism*

Alvaro Sandroni
Northwestern University and University of Pennsylvania.†

Francesco Squintani
Universita' degli Studi di Brescia, and ELSE, University College London.‡

First Version: February 2004, This Version March 2007.

Abstract

It is well known that when agents are fully rational, compulsory public insurance may make all agents better off in the Rothschild and Stiglitz (1976) model of insurance markets. We find that when sufficiently many agents underestimate their personal risks, compulsory insurance makes low-risk agents worse off. Hence, behavioral biases may weaken some of the well-established rationales for government intervention based on asymmetric information.

The behavioral economics literature has produced broad empirical evidence that agents do not always act in their own best interest. When considering single-agent models, a possible implication of behavioral biases is paternalism: Policies designed to affect agents' choices for their own good.[1] However, this implication has not been thoroughly investigated in fully-developed market models.[2] As behavioral biases are difficult to observe, it is natural to approach this investigation in markets with asymmetric information.

This paper explores the policy implications of behavioral biases in the classic model of insurance markets with asymmetric information by Michael D. Rothschild and Joseph E.

*This paper has previously circulated under the title "Paternalism in a Behavioral Economy with Asymmetric Information" We thank the Editor, Vince Crawford, two referees, the audiences of University College London, Northwestern University, Boston University, the Stony Brook Summer Workshop 2004, Gerzensee ESSET 2006, Luca Anderlini, Mark Bils, Erik Eyster, Guillaume Frechette, Faruk Gul, Bart Lipman, Ben Lockwood, Michael Manove, Costas Meghir, Lars Nesheim and Jean Tirole for their comments.

†Kellogg School of Management, MEDS Department, 2001 Sheridan Rd., Evanston, IL 60208, USA and Department of Economics, 3708 Locust Walk, Philadelphia, PA 19104.

‡Department of Economics, Via S. Faustino, 74/B, 25122 Brescia, Italy.

Stiglitz (1976). Insurance companies are perfectly competitive and cannot observe their subscribers' risk, which may be either high or low. Some agents know their risk. We assume that some agents are overconfident: They believe that their risk is low when, in fact, it is high. This assumption is supported by robust empirical evidence that many individuals underestimate important risks, such as those associated with driving.[3] While overconfidence need not be common in all insurance markets, it is a natural first step to explore behavioral biases in the Rothschild and Stiglitz (1976) framework.

When all agents are unbiased, the Rothschild and Stiglitz (1976) model makes a strong case for government intervention. Because of asymmetric information, compulsory insurance may improve all agents' welfare.[4] A different rationale for compulsory insurance is behavioral. Individuals may underinsure because they are overconfident. Compulsory insurance does not harm unbiased agents because they want to be insured, and should be imposed on overconfident individuals for their own benefit.

Our main result shows that the asymmetric-information rationale and the behavioral rationale for compulsory insurance do not reinforce each other. When there is a significant fraction of overconfident agents, compulsory insurance ceases to improve all agents' welfare because it makes low-risk agents worse off. For instance, in the automobile insurance market, compulsory driving insurance translates into a tax on safe drivers that subsidizes unsafe drivers.[5] So, contrary to *prima facie* intuition, behavioral biases may weaken asymmetric-information rationales for government intervention because they may turn policies beneficial to all agents into wealth transfers between agents.

This unexpected result holds because overconfidence changes the equilibrium of the Rothschild and Stiglitz (1976) model qualitatively. Without overconfidence, the market equilibrium is pinned down by a binding incentive compatibility constraint. Low-risk agents' insurance is constrained to ensure separation from high-risk subscribers. High-risk agents benefit from compulsory insurance because they obtain insurance coverage at lower prices. Compulsory insurance also benefits low-risk agents because it relaxes the incentive compatibility constraint. However, when the economy has a significant fraction of overconfident agents, the incentive compatibility constraint no longer binds.[6] Compulsory insurance is equivalent to a transfer of wealth from low-risk to high-risk agents.

The incentive compatibility constraint does not bind in equilibrium because overconfident agents cannot be screened from low-risk agents. These agents share the same beliefs about their risk and so make identical decisions. In addition, we assume that insurance companies cannot directly observe agents' beliefs. Hence, the higher the fraction of overconfident agents in the economy, the higher the average risk of the pool of low-risk and overconfident agents, and the higher the price that insurance firms must offer to avoid negative profits. At high prices, these contracts become unattractive to high-risk agents. For instance, consider the extreme case with the fraction of low-risk agents (relative to the fraction of overconfident agents) is small. The insurance price for low-risk and overconfident agents is close to the insurance price for high-risk agents. Therefore, low-risk agents are better off purchasing small amounts of insurance and are hurt by compulsory insurance.

Our basic result extends beyond compulsory insurance. When the fraction of overconfident agents is significant, budget-balanced government intervention cannot weakly improve the welfare of both high-risk and low-risk agents over the *laissez-faire* equilibrium of our model, unless it changes the fraction of biased agents in the economy. This result also extends beyond overconfidence and still holds if we replace the assumption of a significant fraction of overconfident agents with the weaker assumption of a significant fraction of biased agents that can either be overconfident or underconfident. Finally, we show that policies that directly reduce overconfidence in the economy may benefit low-risk agents without harming high-risk agents. In the context of driving insurance, such policies materialize in voluntary training programs designed to help drivers improve their self-assessment skills.

The paper is organized as follows. Section I presents the model. Section II provides a graphical description of the equilibrium. Section III presents our main result informally. Section IV contains additional policy results. Section V concludes. The formal analysis is laid out in a web appendix.

Related Literature Our paper is related to two branches of behavioral economics. The first branch studies market interactions between sophisticated firms and biased consumers. Stefano DellaVigna and Ulricke Malmendier (2004), Glenn Ellison (2005) and Xavier Gabaix and David Laibson (2005) study models where consumers may have naive beliefs, overlook add-on prices, or underestimate the chance of being subject to hidden fees. They find that in competitive markets, naive consumers may be exploited to the advantage of sophisticated consumers.

Unlike these models, our naive, overconfident agents cannot be separated from low-risk agents because their beliefs are the same. This entails higher insurance prices and an efficiency loss, not only distributive effects. Ran Spiegler (2005) finds an efficiency loss in a market where consumers have a bounded ability to infer quality by sampling goods. Unlike our work, his emphasis is on equilibrium characterization, rather than policy analysis.[7]

The second related branch of behavioral economics studies the effects of behavioral biases. Roland Benabou and Jean Tirole (2002) and Botond Koszegi (2000) show that overconfident agents may strategically ignore information. Roland Benabou and Jean Tirole (2003) study incentives to manipulate self-confidence. Muhamet Yildiz (2003) studies how excessive optimism affects bargaining. Michael Manove and A. Jorge Padilla (1999) and Augustin Landier and David Thesmar (2003) study how entrepreneurs' overconfidence affects financial contracting. Kfir Eliaz and Ran Spiegler (2006) study principal-agent problems where agents may be overconfident. Olivier Compte and Andrew Postlewaite (2003) study optimal beliefs when confidence enhances task performance. Eric J. van den Steen (2004) shows that agents with different priors may overestimate their chances of success. Joel Sobel and Luis Santos-Pinto (2005) show that rational agents may develop optimistic self-assessments if they disagree on which skills determine abilities. Anil Arya and Brian Mittendorf (2004) study an example of insurance market with a monopolistic firm and underconfident agents:

the equilibrium is a pooling, full insurance outcome and the incentive compatibility contraint does not bind.

I. The Model

For each agent, there are two possible states of the world. In state 1 her wealth is W. In state 2 an accident of damage d occurs and the individual's wealth is $W - d$. An insurance contract is a pair $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$ so that the individual's wealth is $(W - \alpha_1, W - d + \alpha_2)$ when buying $\boldsymbol{\alpha}$. The amount α_1 is the premium, $\alpha_1 + \alpha_2$ is the insurance coverage, and $P = \alpha_1 / (\alpha_1 + \alpha_2)$ is the *price* of a unit of insurance. We assume that $\alpha_1 \geq 0$, $\alpha_2 \geq 0$: individuals cannot take on more risk through an insurance contract. Each agent's *risk* is the probability p that the accident occurs, which can either be high (p_H) and low (p_L), with $p_H > p_L$.

Conditional on all observable variables, there are three types of agents in the economy. *High-risk* (type H) and *Low-risk* (type L) agents know that their risks are p_H and p_L, respectively. *Overconfident* (type O) agents believe that their risk is low when in fact it is high. We let $\lambda \in (0,1)$ be the fraction of low risk agents in the economy, and $\kappa \in (0,1)$ be the fraction of overconfident agents, so that $\kappa + \lambda \leq 1$. Agents are risk averse; their expected utility is $V(W, d; p, \boldsymbol{\alpha}) = (1-p)\, U(W - \alpha_1) + pU(W - d + \alpha_2)$, where U is twice differentiable, $U' > 0$ and $U'' < 0$.[8]

The insurance market is a competitive industry of expected profit maximizing (risk neutral) companies. A contract $\boldsymbol{\alpha}$ sold to an agent with risk p yields expected profit $\pi(p, \boldsymbol{\alpha}) = (1-p)\, \alpha_1 - p\alpha_2$. We assume that the insurance firms cannot observe a subscriber's risk or beliefs, but they know κ and λ. A *perfectly competitive equilibrium* is a set of contracts A such that: (i) no contract $\boldsymbol{\alpha} \in A$ makes strictly negative expected profits, and (ii) no contract $\boldsymbol{\alpha}' \notin A$ makes strictly positive profits.

Remark. A perfectly competitive equilibrium may fail to exist in the Rothschild and Stiglitz (1976) model. A set of contracts is *locally competitive* if the insurance firms cannot make positive profits by introducing small changes in the contracts they already offer (this concept is formally defined in the appendix).[9] Any perfectly competitive equilibrium is also locally competitive, but not vice-versa. A locally competitive equilibrium always exists, and is unique, in the Rothschild and Stiglitz (1976) model and in our model as well. A perfectly competitive equilibrium exists in our model as long as the fraction of overconfident agents is above a threshold formally defined in the following section.

II. Graphical Description of Equilibrium

Equilibrium in Insurance Markets without Overconfidence For future reference, we briefly consider the model without overconfidence, i.e., $\kappa = 0$. Rothschild and Stiglitz (1976) show that the equilibrium is separating. Subscribers are screened according to the contract they choose. High-risk individuals fully insure. Their contract $\boldsymbol{\alpha}^H$ equalizes wealth across states and lies on the intersection of the 45-degree line with the zero-profit line $\pi_H = 0$. Incentive compatibility requires that high-risk subscribers (weakly) prefer contract $\boldsymbol{\alpha}^H$ to the low-risk individuals' contract $\boldsymbol{\alpha}^L$. Hence, the contract $\boldsymbol{\alpha}^L$ lies on the intersection of the zero-profit line $\pi_L = 0$ with the indifference curve I_H (through the high-risk agents' contract $\boldsymbol{\alpha}^H$). The contracts $(\boldsymbol{\alpha}^L, \boldsymbol{\alpha}^H)$ are a (unique) perfectly competitive equilibrium as long as the fraction λ of low-risk subscribers is sufficiently small. The equilibrium contracts are illustrated in Figure 1.

Equilibrium in Insurance Markets with Overconfidence We now describe equilibrium with overconfidence (i.e., $\kappa > 0$). The core of our analysis is based on two intuitive insights. The first one is that *insurance firms cannot screen between overconfident and low-risk individuals* because, at the time of purchasing insurance, both types believe that their risk is low. Given this qualification, arguments analogous to the analysis of Rothschild and Stiglitz (1976) allow us to conclude that in the unique competitive equilibrium, individuals are separated on the basis of their beliefs. High-risk individuals purchase a contract $\boldsymbol{\alpha}^H$, whereas low-risk and overconfident individuals choose a different contract $\boldsymbol{\alpha}^{LO}$. As in the case without overconfidence, high-risk individuals fully insure.

The average accident probability of overconfident and low-risk agents is

$$p_{LO} \equiv \frac{\kappa p_H + \lambda p_L}{\kappa + \lambda}.$$

Perfect competition requires that the equilibrium contract $\boldsymbol{\alpha}^{LO}$ satisfies the zero-profit condition $(1 - p_{LO})\alpha_1^{LO} - p_{LO}\alpha_2^{LO} = 0$ (in short, $\pi_{LO} = 0$). So, the price of insurance P^{LO} coincides with p_{LO}. As the fraction of overconfidence agents κ increases, the zero-profit line $\pi_{LO} = 0$ rotates counterclockwise towards the zero-profit line for high-risk types, $\pi_H = 0$.

This leads to the second insight. Unlike in the case without overconfidence, *incentive compatibility need not be binding in equilibrium.* As we argue below, it does not bind when the fraction of overconfident individuals κ is large enough relative to the fraction of low risk agents λ. In order to describe the equilibrium, we distinguish between three different cases depending on the parameters κ and λ. The three significant parameter regions are characterized by the threshold functions $\kappa_1(\lambda)$ and $\kappa_2(\lambda)$, formally defined in the appendix.

Case 1. Small Overconfidence. Assume that the fraction of overconfident agents κ is small relative to the fraction of low-risk individuals λ, i.e. $\kappa \leq \kappa_1(\lambda)$. Then, the locally competitive equilibrium contracts $(\boldsymbol{\alpha}^{LO}, \boldsymbol{\alpha}^H)$ are shown in Figure 1. The only difference from the case without overconfidence is that the contract $\boldsymbol{\alpha}^{LO}$ must lie on the zero-profit line $\pi_{LO} = 0$, since it is chosen by low-risk and overconfident agents alike. As in Rothschild and Stiglitz (1976), the contracts $(\boldsymbol{\alpha}^{LO}, \boldsymbol{\alpha}^H)$ are a (unique) perfectly competitive equilibrium if and only if the fraction λ of low-risk agents is sufficiently small.

Case 2. Intermediate Overconfidence. When the fraction of overconfident individuals is intermediate, i.e., $\kappa_1(\lambda) < \kappa < \kappa_2(\kappa)$, there is always a unique (locally and) perfectly competitive equilibrium. The equilibrium is represented in Figure 2. *The incentive compatibility constraint no longer binds.* To see this, let $\boldsymbol{\alpha}^+$ be the intersection of the zero-profit line $\pi_{LO} = 0$ with the indifference curve I_H passing through $\boldsymbol{\alpha}^H$. Note that the indifference curve of low-risk agents passing through $\boldsymbol{\alpha}^+$ is steeper than the zero-profit line $\pi_{LO} = 0$ (in contrast, in Figure 1 it was flatter). Hence, $\boldsymbol{\alpha}^+$ is no longer an equilibrium because any contract lying to the right of $\boldsymbol{\alpha}^+$ between the indifference curve I_L and the zero-profit line $\pi_{LO} = 0$ would make strictly positive profits.[10] The equilibrium contract for low risk and overconfident agents, denoted by $\boldsymbol{\alpha}^{LO}$, is determined by the tangency point of the indifference curve I_L on the zero-profit line $\pi_{LO} = 0$. Under regularity conditions, low-risk and overconfident agents' utilities decrease in κ.[11] By revealed preferences, low-risk agents' utilities are higher than high-risk agents' utilities which are higher than overconfident agents' utilities.

Case 3. Large Overconfidence. When the fraction of overconfident individuals is large, $\kappa \geq \kappa_2(\kappa)$, the incentive compatibility constraint still does not bind. The zero-profit line $\pi_{LO} = 0$ is sufficiently close to the zero-profit line $\pi_H = 0$ that it becomes flatter than the indifference curve I_L that passes through the no-insurance contract $\mathbf{0}$. Hence, a corner solution $\boldsymbol{\alpha}^{LO} = \mathbf{0}$ is obtained. In the unique locally and perfectly competitive equilibrium, low-risk and overconfident agents believe that the insurance contracts they are offered are so unfavorable that they do not insure.

III. Compulsory Insurance

Compulsory Insurance without Overconfidence A *compulsory insurance* requirement is a contract $\boldsymbol{\beta} = (\beta_1, \beta_2) > \mathbf{0}$ that makes zero profits if imposed uniformly across all agents. Each agent is required to buy contract $\boldsymbol{\beta}$ and is free to buy additional insurance $\boldsymbol{\alpha}(\boldsymbol{\beta})$ on top of $\boldsymbol{\beta}$. Formally, let $p_{LH} \equiv (1 - \lambda) p_H + \lambda p_L$ be the average probability of accident in the economy. Any compulsory insurance contract $\boldsymbol{\beta}$ that keeps the budget balanced must lie on the zero-profit line $\pi_{LH} = 0$, i.e., $(1 - p_{LH})\beta_1 - p_{LH}\beta_2 = 0$.

In the Rothschild and Stiglitz (1976) model, the introduction of compulsory insurance yields a Pareto improvement, as long as the fraction of low-risk individuals is above a threshold. To see this, note that the adoption of $\boldsymbol{\beta}$ is equivalent to a change of endowment from $(W, W - d)$ to $(W - \beta_2, W - d + \beta_1)$. Given this, the remainder of the analysis is qualitatively unchanged. High-risk agents' contracts $\boldsymbol{\alpha}^H(\boldsymbol{\beta})$ fully insure. Low-risk agents' contracts $\boldsymbol{\alpha}^L(\boldsymbol{\beta})$ lies in the intersection of the zero-profit line $\pi_L(\boldsymbol{\beta}) = \mathbf{0}$ and the indifference curve I_H passing through $\boldsymbol{\alpha}^H(\boldsymbol{\beta})$ (see Figure 3).

Compulsory insurance makes high-risk individuals better off because the terms of the compulsory contract $\boldsymbol{\beta}$ are more favorable than the terms of the equilibrium contract $\boldsymbol{\alpha}^H$. Low-risk agents pay the cost of being pooled together with high-risk individuals on the contract $\boldsymbol{\beta}$. However, compulsory insurance relaxes the incentive compatibility constraint imposed by the high-risk subscribers. This can be seen in Figure 3, as the compulsory insurance contract $\boldsymbol{\beta}$ shifts the indifference curve I_H up. When the fraction of high-risk subscribers is sufficiently small, the relaxation of incentive compatibility is large enough to make low-risk agents better off.[12]

Compulsory Insurance with Overconfidence Now consider the case in which the fraction of overconfident agents in the economy is intermediate or large, i.e. $\kappa > \kappa_1(\lambda)$.[13] Because the incentive compatibility constraint does not bind in equilibrium, result 1 below shows that the introduction of compulsory insurance cannot improve all agents' welfare over the *laissez-faire* equilibrium. Specifically, it makes low-risk individuals worse off. Unlike the case that abstracts from overconfidence, compulsory insurance now induces a transfer of wealth from low-risk agents to high-risk agents without any beneficial effect on incentive compatibility constraints.

Result 1 *Suppose that the fraction of overconfident agents in the economy is either intermediate or large (i.e., $\kappa > \kappa_1(\lambda)$). Then, any compulsory insurance contract $\boldsymbol{\beta} > \mathbf{0}$ makes low risk agents strictly worse off.*

This result may be appreciated by inspecting Figure 4. The low-risk and overconfident agents' zero-profit line $\pi_{LO} = 0$ lies below the low-risk agents' indifference curve I_L, passing through the equilibrium contract $\boldsymbol{\alpha}^{LO}$. Any budget-balanced compulsory insurance contract $\boldsymbol{\beta}$ lies on the zero-profit line $\pi_{LH} = 0$, which is strictly below the zero-profit line $\pi_{LO} = 0$. So, any contract $\boldsymbol{\alpha}^{LO}(\boldsymbol{\beta})$ purchased on top of a compulsory insurance contract $\boldsymbol{\beta}$ also lies below the zero-profit line $\pi_{LO} = 0$ and, hence, below the indifference curve I_L. Thus, low-risk agents prefer the *laissez-faire* contract $\boldsymbol{\alpha}^{LO}$ over any allocation resulting from the introduction of compulsory insurance.

IV. Further Policy Results

General Policies We now show that the logic of Result 1 extends to any incentive-compatible budget-balanced policy (paternalistic or not). We define these policies formally in the appendix. In contrast to the case without overconfidence, government intervention cannot improve all agents' welfare over the equilibrium outcome of this model.

Result 2 *Suppose that the fraction of overconfident agents in the economy is either intermediate or large (i.e.,* $\kappa > \kappa_1(\lambda)$*). Then, no incentive-compatible budget-balanced policy can weakly improve the welfare of both low- and high-risk agents over the competitive equilibrium.*

The intuition for Result 2 is as follows.[14] The equilibrium contract $\boldsymbol{\alpha}^{H}$ strictly maximizes high-risk agents' utility among contracts on the zero profit line π_H. Because the incentive compatibility constraint is not binding, the equilibrium contract $\boldsymbol{\alpha}^{LO}$ strictly maximizes low-risk agents utility among contracts on the zero-profit line $\pi_{LO} = 0$ (see Figure 4). Low-risk and overconfident agents cannot be separated by any incentive-compatible policy because they have the same beliefs. Budget-balanced government intervention cannot simultaneously assign an allocation to high-risk agents above the zero-profit line $\pi_H = 0$ and an allocation to low-risk agents above the line $\pi_{LO} = 0$. So, it cannot strictly increase the welfare of either high-risk or low risk agents without making one of the two types strictly worse off.

Underconfidence We now enrich our basic model by introducing underconfident agents who perceive that their risk is high, when, in fact, it is low. We let their fraction in the economy be $\upsilon \geq 0$, and we denote the fraction of unbiased high-risk agents by $\eta = 1 - \lambda - \kappa - \upsilon$.

The average risk of high-risk and underconfident agents is:

$$p_{HU} = \frac{\upsilon p_L + \eta p_H}{\upsilon + \eta}.$$

We assume that p_{HU} is larger than the average risk of low-risk and overconfident agents p_{LO}.

In the unique (locally) competitive equilibrium, the contract $\boldsymbol{\alpha}^{HU}$ is purchased by high-risk and underconfident agents, and the contract $\boldsymbol{\alpha}^{LO}$ by low-risk and overconfident agents. Incentive compatibility ensures that high-risk and underconfident agents do not prefer $\boldsymbol{\alpha}^{LO}$ to $\boldsymbol{\alpha}^{HU}$. The main difference with respect to the equilibrium in Section II is that high-risk and underconfident agents overinsure: $\alpha_1^{HU} + \alpha_2^{HU} > d$. These agents are less risky, on average, then they perceive to be: $p_{HU} < p_H$. Hence, they are willing to overinsure at the competitive price $P^{HU} = p_{HU}$ of contract $\boldsymbol{\alpha}^{HU}$.

Result 3, below, shows that our analysis extends beyond overconfidence. Specifically, compulsory insurance fails to make all agents in our model better off, provided that there are sufficiently many biased agents that can either be overconfident or underconfident.[15] Formally, result 3 holds when the fraction of overconfident agents κ is larger than a threshold $\bar{\kappa}(\upsilon, \lambda)$ defined in the appendix. Because the function $\bar{\kappa}(\upsilon, \lambda)$ decreases in υ, the fraction κ is larger than $\bar{\kappa}(\upsilon, \lambda)$ ($\bar{\kappa}(\upsilon, \lambda)$ may be zero) whenever the fraction of underconfident agents υ is larger than a threshold $\bar{\upsilon}(\kappa, \lambda)$.

Result 3 *Unless both fractions of overconfident and underconfident agents κ and υ are small (i.e. $\kappa \leq \bar{\kappa}(\upsilon, \lambda)$), the government cannot weakly improve the welfare of both low- and high-risk agents upon the perfectly competitive equilibrium $(\boldsymbol{\alpha}^{HU}, \boldsymbol{\alpha}^{LO})$ by means of any incentive-compatible budget-balanced policy (including compulsory insurance).*

Result 3 holds because when υ increases, the average risk p^{HU} of the pool of high-risk and underconfident agents decreases. So, when κ increases, the low-risk and overconfident agents average risk p^{LO} also increases. As either υ or κ (or both) increase, p^{HU} becomes closer to p^{LO}. In a competitive equilibrium, the prices P^{HU} and P^{LO} of the equilibrium contracts $\boldsymbol{\alpha}^{HU}$ and $\boldsymbol{\alpha}^{LO}$ coincide with p^{HU} and p^{LO}, respectively. Hence, as either υ or κ (or both) increase, the price difference between the contracts $\boldsymbol{\alpha}^{HU}$ and $\boldsymbol{\alpha}^{LO}$ decreases, and thus contract $\boldsymbol{\alpha}^{LO}$ becomes less attractive to high-risk and underconfident agents. As a result, incentive compatibility does not bind. Therefore, as in result 2, government intervention cannot improve the welfare of all agents in our model.

Training Programs We now consider policies that reduce overconfidence in the context of driving insurance. A *self-assessment training* program may change overconfident agents' beliefs. At cost $c > 0$, each overconfident agent becomes aware of her high risk with probability $q > 0$. The other agents' beliefs are not changed by the program. This leads to a reduction of the fraction of overconfident individuals in the economy. We assume that participation to the training program is voluntary.

If the training cost c is sufficiently small, the equilibrium is as follows. The terms of insurance contracts depend on attendance at the training program. Agents who do not attend the program are offered the contracts $\boldsymbol{\alpha}^{LO}$ and $\boldsymbol{\alpha}^{H}$ derived in section II. Agents who attend the program are offered $\boldsymbol{\alpha}^{H}$ and a contract $\hat{\boldsymbol{\alpha}}^{LO}$ with a lower price than $\boldsymbol{\alpha}^{LO}$, after they complete the program. The contract $\hat{\boldsymbol{\alpha}}^{LO}$ is purchased by low-risk agents and by those agents who remain overconfident despite participating in the training program. Overconfident agents who correct their beliefs, due to having attended the training program, buy contract $\boldsymbol{\alpha}^{H}$. Low-risk and overconfident agents join the training program, high-risk agents do not. To see that this is the (unique) equilibrium, note that if the training cost c is sufficiently small, the low-risk and overconfident agents are attracted to lower insurance prices and join the training program.[16] As a result, the fraction of overconfident individuals κ decreases and this results in lower insurance prices.

Low-risk agents' beliefs are not changed by the training program, but they benefit indirectly through the reduction of the insurance price. High-risk agents do not join the program and are not affected by it. So, low-risk agents are strictly better off with the voluntary training program, whereas high-risk agents are not harmed by it.[17]

When describing welfare of overconfident agents, we focus on *actual* welfare, defined as the average ex-post utility $V(W, d; p_H, \boldsymbol{\alpha})$ as a function of the equilibrium contract $\boldsymbol{\alpha}$ and the *actual* risk p_H, and where the wealth W is net of training costs. It is conceptually difficult to describe the effect of training programs on the perceived welfare of the overconfident agents who change their beliefs because of the program. However, their actual welfare increases when c is sufficiently small, because they correct their beliefs and make a better insurance choice. Agents who remain overconfident despite participating in the training program improve their actual (and perceived) welfare indirectly through the reduction of the insurance price. The above discussion is formalized in the following result.

Result 4 *Assume that the fraction of overconfidence agents in the economy is either intermediate or large (i.e., $\kappa > \kappa_1(\lambda)$). As long as benefits q are sufficiently large and costs c are sufficiently low, the introduction of a voluntary training program strictly increases the welfare of low-risk agents and the actual welfare of overconfident agents. It does not change the welfare of high-risk agents.*

V. Conclusion

In the Rothschild and Stiglitz (1976) model of insurance markets with asymmetric information, compulsory insurance may make all agents better off, provided that agents are fully rational. We build on this basic model of insurance, but we assume that a significant fraction of agents in the economy do not accurately access actual risks. In addition, we assume that insurance companies cannot directly observe agents' beliefs. Under these assumptions, compulsory insurance fails to make all agents better off because it is detrimental to low-risk agents. Our results do not deliver unqualified support for *laissez-faire* policies. Rather they show that while behavioral biases may support paternalistic policies in simple decision-theoretic models, they may also weaken asymmetric information rationales for government intervention in fully-developed market models.

We hope that these results will motivate additional studies on the interactions between different reasons for government intervention in the economy and also on the functioning of markets when agents are less than fully rational.

Footnotes

[1] In fact, behavioral economists advocate only mild forms of intervention which guarantee the possibility of opting out. See, among others, Edward D. O'Donoghue and Matthew Rabin (2003), Richard H. Thaler and Cass Sunstein (2003) and Colin F. Camerer, Samuel Issacharoff, George Loewenstein, Edward D. O'Donoghue and Matthew Rabin (2003).

[2] An exception is O'Donoghue and Rabin (2003).

[3] According to Werner F.M. De Bondt and Richard H. Thaler (1995, p. 389), "perhaps the most robust finding in the psychology of judgment is that people are overconfident." Among many papers finding evidence of overconfidence, see Howard Kunreuther et al. (1978), Linda Babcock and George Loewenstein (1997), Colin F. Camerer and Dan Lovallo (1999), Shlomo Benartzi (2001), Jay Bhattacharya, Dana P. Goldmanz and Navin Sood (2004). A brief survey of this literature is presented in our companion paper.

[4] This argument, demonstrated by Charles A. Wilson (1977) and Bev G. Dahlby (1983), is highlighted both in textbooks (e.g. Alan J. Auerbach and Martin Feldstein, 2002), and in institutional debates (e.g. Mark V. Pauly, 1994).

[5] In the context of motorist insurance, our analysis applies only to personal loss insurance, in the forms of the Personal Injury Protection and Uninsured Motorist insurance, which is mandatory in most US States (see the Summary of Selected State Laws published by American Insurance Association, 1976-2003). PIP insurance covers loss when the driver is at fault, and UM insurance covers loss caused by another driver who is at fault and not insured. Our analysis does not apply to liability insurance, which covers the losses that a driver can cause to others.

[6] This finding does not depend on the assumption of perfect competition, as demonstrated

by C. Mark Armstrong (2005) in versions of our model with either a monopolistic firm or with imperfect competition.

[7] More distently related, Paul Heidhues and Botond Koszegi (2004) provide a rationale for price stickiness in a model with loss-averse consumers.

[8] To simplify the exposition, we focus on the case in which the difference between low risk and high risk is not too small relative to the damage d. That is, we assume that

$$\frac{(1-p_L)/p_L}{(1-p_H)/p_H} > \frac{U'(W-d)}{U'(W)}.$$

[9] In a general equilibrium model, Pradeep K. Dubey and John G. Geanakoplos (2002) show the existence of an equilibrium that approximates the locally-competitive equilibrium. John G. Riley (1979) shows that the locally-competitive equilibrium coincides with a "reactive" equilibrium where firms, before introducing new contracts, anticipate that competitors will react by offering new contracts, if they generate positive profits. Charles A. Wilson (1977) proposes an alternative reactive equilibrium where loss-making contracts are removed as a reaction to newly-introduced contracts.

[10] Any such contract α makes strictly positive profits because it is purchased only by low-risk and overconfident agents and its price is larger than P^{LO}, as α lies below the zero-profit line $\pi_{LO} = 0$. Low-risk and overconfident agents prefer this contract α to α^+, because α lies above the indifference curve I_L. High-risk agents still prefer α^H to the contract α, because α lies below the indifference curve I_H.

[11] Specifically, this result holds if the coefficient of Relative Risk Aversion $-wU''(w)/U'(w)$ is smaller than the bound $(W-d)/W$ for any wealth amount $w \in [W-d, W]$.

[12] Unlike the Rothschild and Stiglitz equilibrium and the Wilson (1977) equilibrium, the Miyazaki-Wilson-Spence equilibrium cannot be improved by compulsory insurance (see Crocker and Snow (1985)). In this equilibrium, insurers are not profit maximizers: They sell loss-making contracts to high-risk agents, subsidized with profit-making contracts sold to low-risk agents.

[13] If $\kappa < \kappa_1(\lambda)$, the analysis is analogous to the case without overconfidence.

[14] Result 2 subsumes result 1 because compulsory insurance is a special case of incentive-compatible budget-balanced government policy. Thus, result 1 is demonstrated as a corollary of result 2.

[15] In our companion paper, we further explore the robustness of our results and show that they still hold (with proper qualifications) when there are more than two levels of risk in the economy.

[16] At the time they choose to join the training program, none of these agents believe that they will improve their self-assessment skill. They join only because $\hat{\alpha}^{LO}$ is cheaper than the contract α^{LO} that they would be offered if they did not attend the program.

[17] In our companion paper, we show that if participation in self-assessment training programs were compulsory, it would reduce the utility of high-risk agents.

References

[1] "Overconfidence, Insurance and Paternalism: Companion Paper", (2006) mimeo.

[2] American Insurance Association. 1976-2003. *Summary of Selected State Laws and Regulations Relating to Automobile Insurance,* New York: American Insurance Association.

[3] **Arya, Anil, and Brian Mittendorf.** 2004. "Benefits of a Slanted View: a Discussion of Disclosure Bias," *Journal of Accounting and Economics,* 38: 251-262.

[4] **Armstrong, Mark C.** 2005. "Price Discrimination on the Basis of Mistaken Priors: an Example." mimeo, University College London.

[5] **Auerbach, Alan J., and Martin Feldstein, ed.** 2002. *Handbook of Public Economics*, Vol. 4, Amsterdam: North Holland.

[6] **Babcock, Linda and George Loewenstein.** 1997. "Explaining Bargaining Impasse: The Role of Self-Serving Biases." *Journal of Economic Perspectives,* 11: 109-126.

[7] **Benabou, Roland and Jean Tirole.** 2002. "Self-Confidence and Personal Motivation." *Quarterly Journal of Economics,* 117: 871-915.

[8] **Benabou, Roland and Jean Tirole.** 2003. "Intrinsic and Extrinsic Motivation." *Review of Economic Studies,* 70: 489-520.

[9] **Benartzi, Shlomo.** 2001. "Excessive extrapolation and the allocation of 401(k) accounts to company stock." *Journal of Finance,* 56: 1747-1764.

[10] **Bhattacharya Jay, Dana P. Goldmanz and Navin Sood.** 2004. "Market Evidence of Misperceived Prices and Mistaken Mortality Risks." NBER Working Paper 9863.

[11] **Camerer, Colin F., and Dan Lovallo.** 1999. "Overconfidence and Excess Entry: and Experimental Approach." *American Economic Review,* 89: 306-318.

[12] **Camerer, Colin F., Samuel Issacharoff, George Loewenstein, Edward D. O'Donoghue and Matthew Rabin.** 2003. "Regulation for Conservatives: Behavioral Economics and the case for 'Asymmetric Paternalism'." *University of Pennsylvania Law Review* 151: 101-144.

[13] **Compte, Olivier and Andrew Postlewaite.** 2003. "Confidence-Enhanced Performance." *American Economic Review,* forthcoming.

[14] **Crocker, Keith J. and Arthur Snow.** 1985. "The Efficiency of Competitive Equilibria in Insurance Markets with Adverse Selection." *Journal of Public Economics* 26: 207-219.

[15] **Dahlby, Bev G**. 1981. "Adverse Selection and Pareto Improvements through Compulsory Insurance." *Public Choice* 37: 547-58.

[16] **De Bondt, Werner F.M., and Thaler, Richard H.** 1995. "Financial Decision-Making in Markets and Firms: A Behavioral Perspective." in R. Jarrow et al., eds., *Handbooks in Operations Research and Management, Vol. 9.* Amsterdam: Elsevier Science.

[17] **DellaVigna, Stefano and Ulricke Malmendier.** 2004. "Contract Design and Self-Control: Theory and Evidence." *Quarterly Journal of Economics,* 119: 353-402.

[18] **O'Donoghue, Edward D. and Matthew Rabin.** 2003. "Studying Optimal Paternalism, Illustrated by a Model of Sin Taxes." *American Economic Review* 93: 186-191

[19] **Dubey, Pradeep K., and John G. Geanakoplos.** 2002. "Competitive Pooling: Rothschild-Stiglitz Reconsidered." *Quarterly Journal of Economics,* 117: 1529-1570.

[20] **Ellison, Glenn.** 2005. "A Model of Add-On Pricing." mimeo, MIT.

[21] **Gabaix, Xavier and David Laibson.** 2005. "Consumer Myopia, Shrouded Attributes, and Information Suppression in Competitive Markets." mimeo, Harvard.

[22] **Eliaz, Kfir and Ran Spiegler.** 2006. "Speculative Contracts." mimeo, Tel Aviv.

[23] **Heidhues, Paul and Botond Koszegi.** 2004. "The Impact of Consumer Loss Aversion on Pricing." mimeo, Berkeley.

[24] Insurance Research Council. 2000 - 2003. *Uninsured Motorists and Public Attitude Monitoring.*

[25] **Koszegi, Botond.** 2000. "Ego Utility, Overconfidence and Task Choice." mimeo.

[26] **Kunreuther, Howard, R. Ginsberg, L. Miller, P. Sagi, P. Slovic, B. Borkan, and N. Katz.** 1978. *Disaster Insurance Protection: Public Policy Lessons,* Wiley Interscience.

[27] **Landier, Augustin and David Thesmar.** 2003. "Financial Contracting with Optimistic Entrepreneurs: Theory and Evidence." mimeo.

[28] **Manove, Michael and Atilano Jorge Padilla.** 1999. "Banking (Conservatively) with Optimists." *RAND Journal of Economics,* 30: 324–350.

[29] **Pauly, Mark V.** 1994. "Universal Health Insurance in the Clinton Plan: Coverage as a Tax-Financed Public Good." *Journal of Economic Perspectives,* 8: 45-53.

[30] **Riley, John G.** 1979. "Informational Equilibrium." *Econometrica* 47: 331-360.

[31] **Rothschild, Michael D. and Joseph E. Stiglitz.** 1976. "Equilibrium in Competitive Insurance Markets: an Essay on the Economics of Imperfect Information." *Quarterly Journal of Economics* 90: 629–649.

[32] **Sobel, Joel and Luis Santos-Pinto**. 2005. "A Model of Positive Self-Image in Subjective Assessments." *American Economic Review,* forthcoming.

[33] **Spiegler, Ran.** 2005. "Competition over Agents with Boundedly Rational Expectations." mimeo.

[34] **Thaler, Richard H. and Cass Sunstein.** 2003. "Libertarian Paternalism." *American Economic Review* 93: 175-179

[35] **Van den Steen, Eric J.** 2004. "Rational Overoptimism." *American Economic Review*, 94: 1141-51.

[36] **Wilson, Charles A**. 1977. "A Model of Insurance Markets with Incomplete Information." *Journal of Economic Theory* 16: 167-207.

[37] **Yildiz, Muhamet.** 2003. "Bargaining without a common prior: an immediate agreement theorem." *Econometrica* 71, 793-811.

Figures

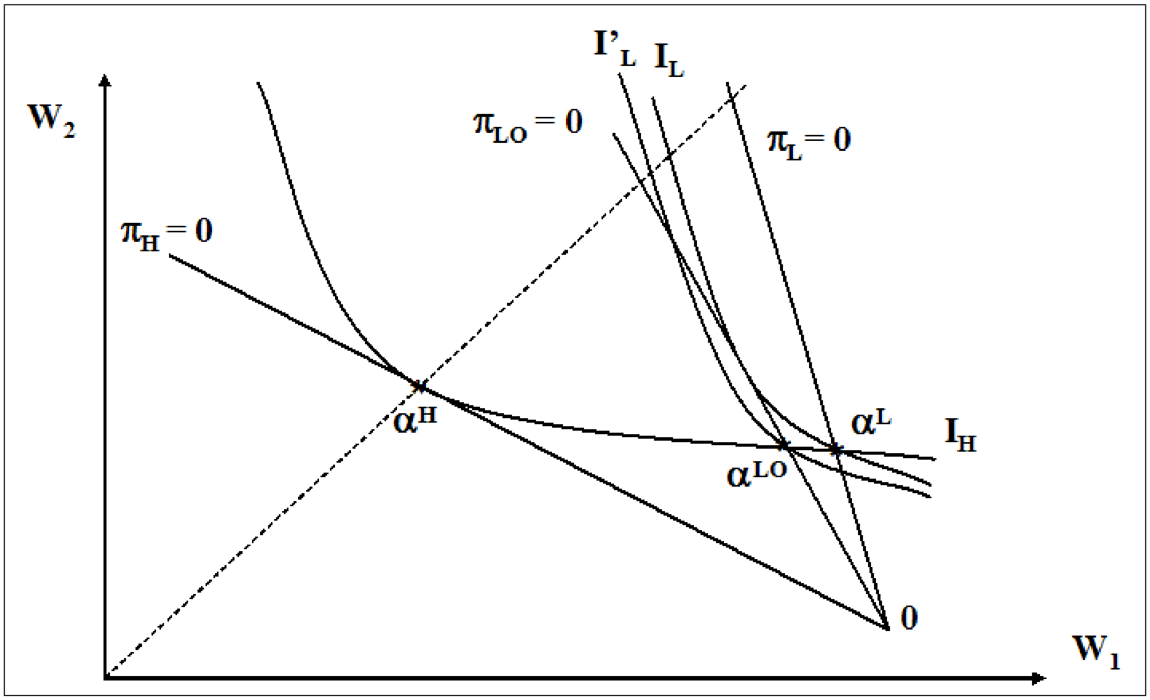

Figure 1: Equilibrium without overconfidence, and with small overconfidence.

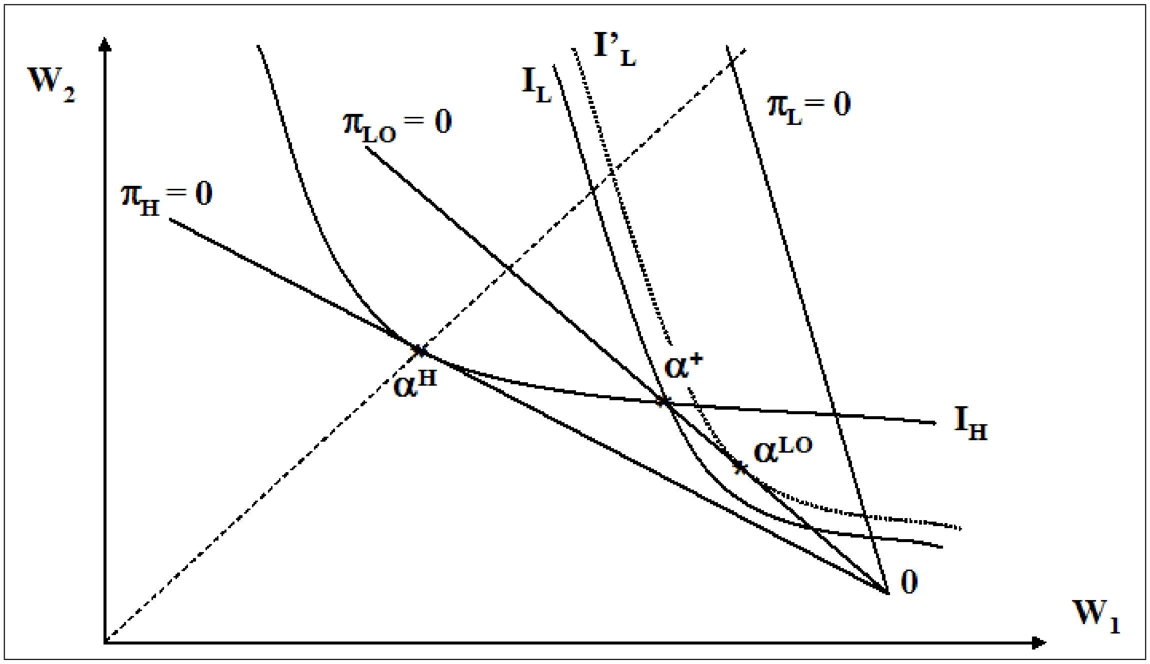

Figure 2: Equilibrium with Intermediate Overconfidence.

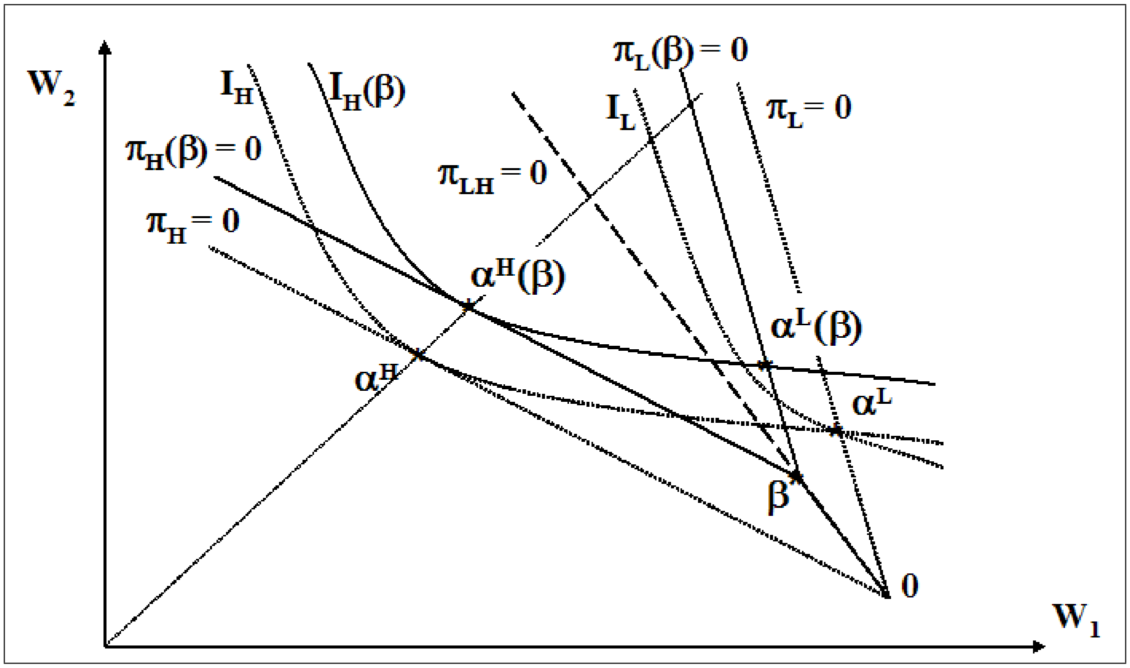

Figure 3: Compulsory Insurance without Overconfidence

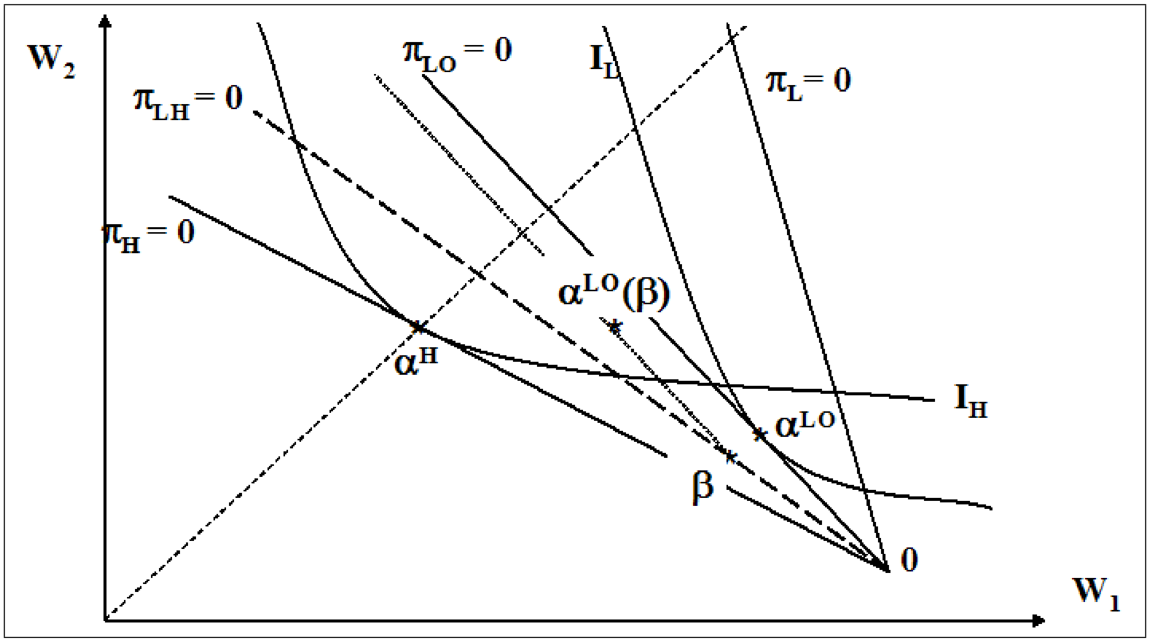

Figure 4: Compulsory Insurance with Overconfidence

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* $\boldsymbol{\alpha}^H$.

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:

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

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

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

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 $M\left(\boldsymbol{\alpha}^{LO}, p_H\right)$ is larger than $M\left(\boldsymbol{\alpha}^{LO}, p_L\right)$. 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

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

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 2.

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 $(\boldsymbol{\alpha}^H, \boldsymbol{\alpha}^{LO})$.*

Proof. Any candidate allocation $\boldsymbol{\alpha}^*$ must satisfy $V(W, d; p_H, \boldsymbol{\alpha}^*(p_H, p_H)) \geq V(W, d; p_H, \boldsymbol{\alpha}^H)$. 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. $V(W, d; p_L, \boldsymbol{\alpha}^*(p_L, p_L)) < V(W, d; p_L, \boldsymbol{\alpha}^{LO})$. ■

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, $V(W, d; p_H, \boldsymbol{\alpha}^*(p_H, p_H)) > V(W, d; p_H, \boldsymbol{\alpha}^H)$ 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 $\kappa > \kappa_1(\lambda)$, $V(W, d; p_L, \boldsymbol{\alpha}^*(p_L, p_L)) < V(W, d; p_L, \boldsymbol{\alpha}^{LO})$. ■

In order to prove Result 3, 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

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

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 $p_\alpha \leq p_{LO}$ for any equilibrium $\boldsymbol{\alpha}$ such that $\sigma_L^\alpha + \sigma_O^\alpha > 0$. Also, $p_\alpha \leq p_{HU}$ 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 3.

Proof of Result 3. The proof of Proposition .2, with obvious modifications, concludes that (i) the incentive compatibility constraint $V(W,d;p_H,\boldsymbol{\alpha}^{HU}) \geq V(W,d;p_H,\boldsymbol{\alpha}^{LO})$ does not bind in equilibrium if and only if $\kappa > \bar{\kappa}(\upsilon,\lambda)$ where $\bar{\kappa}$ solves

$$M\left(\bar{\boldsymbol{\alpha}}, p_L\right) = (1-p_{LO})/p_{LO}, \tag{.14}$$

$$p_{LO}\bar{\alpha}_2 = (1-p_{LO})\,\bar{\alpha}_1,\ \ V(W,d;p_H,\boldsymbol{\alpha}^{HU}) = V(W,d;p_H,\bar{\boldsymbol{\alpha}}), \tag{.15}$$

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

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

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 $V(W-c, d; p_H, \boldsymbol{\alpha}^H) > V(W, d; p_H, \boldsymbol{\alpha}^{LO}(\kappa))$. ■