

Discussion Papers
Department of Economics
University of Copenhagen

No. 13-12

Facilitating Consumer Learning in Insurance
Markets-What Are The Welfare Effects?

Johan N. M. Lagerlöf and Christoph Schottmüller

Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K., Denmark

Tel.: +45 35 32 30 01 – Fax: +45 35 32 30 00

<http://www.econ.ku.dk>

ISSN: 1601-2461 (E)

Facilitating Consumer Learning in Insurance Markets—What Are the Welfare Effects?

Johan N. M. Lagerlöf

Christoph Schottmüller*

University of Copenhagen and CEPR

University of Copenhagen and TILEC

November 11, 2013

Abstract

What are the welfare effects of a policy that facilitates for insurance customers to privately and covertly learn about their accident risks? We endogenize the information structure in Stiglitz's classic monopoly insurance model. We first show that his results are robust: For a small information acquisition cost c , the consumer gathers information and the optimal contracts are close to the ones in the Stiglitz model. If c is so low that the consumer already gathers information ($c < c^*$), both insurer and consumer benefit from a policy that reduces c further. For $c > c^*$, marginally reducing c hurts the insurer and weakly benefits the consumer. Paradoxically, a reduction in c that is “successful,” meaning that the consumer gathers information after the reduction but not before it, can hurt both parties. The reasons for this are that, after the reduction, (i) the cost is actually incurred and (ii) the contracts can be more distorted.

Keywords: asymmetric information, information acquisition, insurance, screening, adverse selection

JEL codes: D82, I13

*Address to both authors: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen K, Denmark; email: johan.lagerlof@econ.ku.dk and christoph.schottmuller@econ.ku.dk. We are grateful for helpful comments from seminar participants at the University of Copenhagen.

1. Introduction

What are the welfare effects of public policies that facilitate for insurance customers to learn, privately and covertly, about their accident risks? It may be tempting to presume that such policies are unambiguously welfare enhancing, or at least that they beneficial for the consumers. However, the answer to the above question is not straightforward, certainly not in a situation where insurance companies have some market power. While it is true that more accurate information can help consumers choose the most appropriate insurance policy, a reduction in the information gathering cost is also likely to affect the insurer's optimal design of the policy, possibly in a way that hurts consumers and overall welfare. Moreover, in contrast to many other adverse selection problems, like problems of procurement and price discrimination, information about the agent's type (which in our insurance application is her accident risk) is not required to achieve allocative efficiency: With a risk averse individual, full insurance is always desirable, regardless of the magnitude of the risk.

In this paper, we study formally the welfare effects of an exogenous change in insurance customers' cost of learning, privately and covertly, about their accident risks. We do this by first endogenizing the information structure in Stiglitz's (1977) classic model of a monopoly insurance market with adverse selection, and we then carry out comparative statics exercises with respect to the information gathering cost.

Although our analysis is formally consistent with many kinds of insurance problems, our preferred interpretation of the model is in terms of health insurance. In recent years, genetic tests have become available to consumers. Private companies offer the possibility to test for certain risk factors (or even to provide a full DNA sequencing).¹ Such tests allow customers to obtain, at a cost, private information concerning their health risks. In this context, we are interested in the welfare effects of subsidizing or promoting such tests. Another interpretation of our model could be that an individual has private information about her extended family's health history, but she needs to learn about how that information translates into a change in her own health risk. Alternatively, an individual may privately know the extent to which she is physically active and that such activities are likely to reduce certain health risks; but to find out which health risks that are affected and how much, she must exert some effort to find additional information. In both these examples, if the individual indeed gathers the additional information she will have access to private information about her own health risk. (This is true also if the additional information is known by the insurance company, since the company is not

¹See www.23andme.com for an example. It should be noted that the purpose of these tests is not necessarily treatment but mainly information. For example, the marker rs12425791 (which is analyzed by 23andme.com) increases the risk of having a stroke by 1.3 if it is "A"; see Ikram et al. (2009).

privity to the health history of the individual's family or the level of her physical activity.) The public policy that facilitates gathering the additional information can be thought of as the funding and launch of a website with relevant information, of a telephone line with free expert advice, or of an informational campaign.

In our model, we assume an insurance company that has a monopoly in its market.² This company and the consumer initially both know only the prior distribution of the latter's accident risk. The company first offers a menu of insurance policies. The consumer then chooses whether or not to, at a cost c , acquire private information about her accident risk. Thereafter she chooses an insurance policy from the menu or decides to remain uninsured.

The main insights that come out of our analysis are as follows. First, if the cost of information acquisition is sufficiently small ($c < c^*$, for a cutoff value $c^* > 0$), the consumer indeed gathers information and thus, at the optimum, is privately informed. This result is not obvious; Crémer and Khalil (1992) study a related problem of optimal procurement in which the agent never is induced to gather information.³ Second, again if the cost is sufficiently small ($c < c^*$), the qualitative features of the optimal menu of insurance policies are the same as in the Stiglitz (1977) model. Indeed, in the limit as the cost goes to zero, the optimal menu approaches the one in that model. This robustness result is useful, as it can serve as a justification of Stiglitz's assumption that information is, exogenously, asymmetrically distributed.

Third, yet again for sufficiently low values of the cost ($c < c^*$), our naive intuition is confirmed: A reduction in the cost of information gathering benefits *both* the insurer and the consumer. In other words, a policy that facilitates consumer learning is a Pareto improvement if, already at the outset, the consumer finds it worthwhile to learn about her risk. In practice, this may be precisely the kind of situation in which such a policy would not be undertaken, since policymakers would believe it is not needed.

If the cost of information gathering is very high ($c > c''$, for some $c'' > c^*$), it is common knowledge between the insurer and the consumer that the latter will never have an incentive to gather information and, therefore, the cost does not matter for the design of the optimal insurance policy. More interestingly, however, there are intermediate values of the cost ($c^* < c < c''$) for which the consumer does not incur the cost but it does affect the insurance policy. The fourth main insight that comes out of our analysis

²The important feature of the insurance market that we want to capture with this assumption is that the insurer has at least some market power. For empirical evidence suggesting that health insurance markets are not perfectly competitive, see Dafny (2010). To assume a monopoly is an extreme but also very tractable way of incorporating this feature into the model. It would be interesting to study the questions that we look at in an oligopoly setting, but that must await future research.

³This paper is reviewed below.

is that, for this range of cost levels, the insurance company is always worse off from a local reduction in the cost and the consumer is, depending on parameter values, either indifferent to a local cost reduction or better off from it. That is, in this range of cost levels, there is much weaker support for the naive intuition discussed above: Either there is a conflict of interest concerning the desirability of a cost reduction or such a reduction would lead to a Pareto inferior outcome (with one party being unaffected and the other being strictly worse off). Again, however, in practice this may be precisely the kind of situation in which a policy that facilitates consumer learning is indeed undertaken, since policymakers would believe it is needed.

The comparisons discussed above concern changes in c within the range $0 < c < c^*$ and within the range $c^* < c < c''$. In both cases, we concluded that the consumer is never strictly worse off from a reduction in the cost parameter (although the insurer may be). What about a cost reduction that makes c drop from some level above c^* to a level below it? The fifth main result that we are able to show, with the help of numerical examples, is that such a cost reduction can make the consumer strictly worse off.⁴ Indeed, the level of c that maximizes the consumer's expected utility, in these examples, is greater than zero and it is high enough to ensure that information is not acquired. There are two reasons why the consumer's expected utility is lower after the cost reduction: For cost levels below c^* (i) the cost is actually incurred as consumers acquire information and (ii) the contract is sometimes more distorted compared to cost levels above that cutoff value. In our examples, both (i) and (ii) contribute to the expected utility loss that the cost reduction gives rise to. Our analysis thus suggests that a public policy that "successfully" facilitates consumer learning (in the sense that it makes the consumers start to gather information) can be counterproductive in that it lowers consumer surplus.

Our paper is related to two broad strands of literature. The first one is a relatively small but growing literature that endogenizes the information structure in models of adverse selection (or screening). Instead of assuming an asymmetry between the principal and the agent in terms of access to the information itself, this literature takes as its primitive an asymmetry in the *ability to gather* information. The contract that the principal offers must therefore serve the double role of appropriately incentivizing the agent's information gathering and ensuring truthful revelation. Whether or not the principal chooses to induce information gathering, this double role of the contract often means that the contracts look different relative to what they do in the standard models of asymmetric information.

Three early contributions to this literature are Barzel (1977) and Craswell (1988) in the field of law and economics and Demski and Sappington (1987) on delegated expertise.

⁴The insurer's profits may either increase or decrease, depending on parameter values.

There are also some early papers on auction theory and information acquisition; see, e.g., Lee (1982). Somewhat more recent contributions include a series of papers by Crémer and Khalil and by Crémer, Khalil and Rochet, two of which are particularly close to our study.

The papers by Crémer and coauthors are based on a procurement setting à la Baron and Myerson (1982), where the information that can be gathered concerns the supplier's production cost. Our insurance application differs from that setting in that, in our framework, the information that the agent can acquire is not relevant for the first best allocation (full insurance is always desirable, as explained in the introductory paragraph). In Crémer and Khalil (1992) the production cost is, regardless of whether information gathering takes place, assumed to become common knowledge after the contract is signed but before production takes place, which makes the setting a bit more similar to ours (since in neither setting information gathering is required for allocative efficiency). Still, information gathering is never induced in that model, which is in contrast to our results. In Crémer et al. (1998a) the production cost does not automatically become known before production, but can only be learned by the agent if incurring the cost before signing the contract.⁵ Apart from the differences in setting, our analysis also puts more emphasis on the comparative statics exercises compared to the papers by Crémer and coauthors (and compared to other papers in the literature).⁶

The second broad strand of literature that our paper is related to is the one on insurance and adverse selection.⁷ A significant part of that literature, initiated by Rothschild and Stiglitz (1976) and Wilson (1977), is concerned with competitive markets. However, in an early and classic contribution Stiglitz (1977) modeled a monopolist insurer facing privately informed consumers; for a recent treatment of this problem, see Chade and Schlee (2012).

The papers in the insurance literature that are perhaps the closest to ours are Doherty and Thistle (1996) and Hoy and Polborn (2000). Like we do, Doherty and Thistle study the incentives of consumers in an insurance market to gather information about their accident risks. However, they assume a competitive market, which leads to the result that either the equilibrium insurance policy does not depend on the magnitude of the information gathering cost or an equilibrium of the model does not exist. Hence, in their setting, a comparative statics exercise with respect to the information gathering

⁵The authors refer to this case as *productive*, as opposed to *strategic*, information gathering.

⁶Crémer and Khalil (1994) and Crémer et al. (1998b) study settings where the agent must gather information before the contract is offered, which make those papers less related to our analysis. The analysis in Crémer et al. (1998a) has recently been extended by Szalay (2009) and Iossa and Martimort (2013).

⁷For a recent survey on adverse selection in insurance contracting, see Dionne et al. (2012).

cost is either not interesting or not possible. The main point made by Doherty and Thistle (1996) is that for consumers to have an incentive to gather costly information, insurers must not be able to observe the consumers' information. Hoy and Polborn (2000) investigate similar questions to Doherty and Thistle but assume a setting with life instead of health insurance (still in a competitive market).⁸

In a companion paper (Lagerlöf and Schottmüller, 2013), we investigate a problem that is similar to the one in the present paper, but in an alternative economic environment where the consumer can choose the precision of a noisy signal from a continuum. We discuss that paper and how it relates to the present one in our concluding discussion (Section 5).

The remainder of the paper is organized as follows. The next section describes our model. Section 3 studies some benchmarks, which will be useful for the analysis of our main model. Section 4.1 solves for the optimal menu of insurance policies in our main model, conditional on the insurer inducing information gathering; Section 4.2 does the same for the case when information gathering is not induced. Section 4.3 compares those two cases and solves for the overall optimum. In that subsection, we also present and discuss our main results. Section 5 concludes. Some of the formal proofs are relegated to three appendices.

2. The model

We model a monopoly insurance market as a principal-agent problem with adverse selection. Our setup builds on Stiglitz's (1977) classic framework, but we endogenize the information structure.

The principal of the model (P) is a risk neutral and profit maximizing monopoly insurer, and the agent (A) is a risk averse insuree who has initial wealth $w > 0$. With exogenous probability $\theta \in (0, 1)$, A has an accident, which means that she incurs (the equivalent of) a monetary loss $d \in (0, w)$. A has the opportunity to purchase an insurance policy from P . An insurance policy consists of a premium p and a net indemnity a that is paid in the event of an accident. A 's Bernoulli utility function is twice continuously differentiable and denoted by u (with $u' > 0$ and $u'' < 0$). Hence, provided that A purchases the insurance policy (p, a) , her utility is $u(w - d + a)$ if having an accident and $u(w - p)$ if not having an accident.

The accident probability is either low ($\theta = \underline{\theta}$) or high ($\theta = \bar{\theta}$), where $\underline{\theta} < \bar{\theta}$. We

⁸Ligon and Thistle (1996) model a monopoly insurance problem with both adverse selection and moral hazard. The authors compare a situation where the consumer knows her risk with one where she does not know it. However, Ligon and Thistle do not, as we do, model any information acquisition.

will refer to an agent with $\theta = \underline{\theta}$ as the *low-demand type*, as this agent has a relatively low willingness to pay for insurance. Similarly, an agent with $\theta = \bar{\theta}$ will be called the *high-demand type*. At the outset of the game, neither P nor A knows θ . Their common prior belief is that θ is low with probability $v \in (0, 1)$.

The sequence of events is as follows.

- (i) P commits to a menu of insurance policies, $\{(\underline{p}, \underline{a}), (\bar{p}, \bar{a})\}$, where the policy $(\underline{p}, \underline{a})$ is aimed at the low-demand type and the policy (\bar{p}, \bar{a}) is aimed at the high-demand type.
- (ii) A observes the menu and then makes a choice whether or not to gather information, $x \in \{0, 1\}$. If $x = 1$, A must incur a cost $c > 0$ but receives a signal that perfectly reveals the true value of θ . If $x = 0$, A incurs no cost but does not obtain any new information about θ . We assume that the cost c enters A 's payoff as an additive term.⁹ A 's choice of x is not observed by P . Nor can P observe the signal that A receives if $x = 1$.
- (iii) A decides whether to accept any insurance policy in the menu and, if so, which one.

Two remarks are in order. First, the assumption in (ii) that P cannot observe A 's choice whether to gather information adds an element of moral hazard to the model, on top of the adverse selection problem that arises if A indeed gathers private information. Second, the notation introduced in (i) *allows* for a menu that contains two distinct insurance policies. However, the notation is also consistent with the possibility that P bundles the two policies ($\underline{p} = \bar{p}$ and $\underline{a} = \bar{a}$) or shuts, say, the low-demand type down ($\underline{p} = \underline{a} = 0$). Moreover, if A chooses $x = 0$, then effectively there is only one type, which we can capture by again imposing $\underline{p} = \bar{p}$ and $\underline{a} = \bar{a}$.

P 's ex ante expected profit from the menu of insurance policies equals

$$\pi = v [(1 - \underline{\theta}) \underline{p} - \underline{\theta} \underline{a}] + (1 - v) [(1 - \bar{\theta}) \bar{p} - \bar{\theta} \bar{a}]. \quad (1)$$

The choice variables that P has available when maximizing profits are the premium and the net indemnity of the two types. When solving the model, however, it will be more

⁹This assumption, which is also made by Doherty and Thistle (1996), can be justified in two ways. First, we can think of c as a disutility of effort rather than a monetary expenditure, and suppose that the disutility of effort is independent of the wealth level. Second, we can think of the information gathering choice to be made earlier in time relative to the consumption of the wealth, and suppose again that there are no cross effects. In the concluding section, we discuss the effect on the results of instead letting c enter the argument of the utility function u .

convenient to think of P as (equivalently) choosing A 's ex post utility levels directly. Thus, let \bar{u}_A and \bar{u}_N denote A 's utility if she has an accident and if she does not have an accident, respectively, given that she has purchased the policy (\bar{p}, \bar{a}) . Similarly, let \underline{u}_A and \underline{u}_N denote A 's utility if she has an accident and if she does not have an accident, respectively, when having purchased the policy $(\underline{p}, \underline{a})$.¹⁰ Also, let h be the inverse of the function u (hence $h' > 0$ and $h'' > 0$). We can now, equivalently to (1) above, write P 's ex ante expected profit from the menu of insurance policies as follows:

$$\pi = \widehat{w} - v [(1 - \underline{\theta}) h(\underline{u}_N) + \underline{\theta} h(\underline{u}_A)] - (1 - v) [(1 - \bar{\theta}) h(\bar{u}_N) + \bar{\theta} h(\bar{u}_A)], \quad (2)$$

where $\widehat{w} \equiv w - [v\underline{\theta} + (1 - v)\bar{\theta}]d$ is A 's wealth net of the ex ante expected monetary loss associated with an accident. The net wealth level \widehat{w} thus equals, in expected terms, the total amount of resources available in this economy,¹¹ and it therefore also constitutes an upper bound on the amount of profits that P can earn. In practice, the profits will fall short of \widehat{w} , since P must leave an informational rent to A .

We will solve this game by backward induction.

3. Benchmarks

Before solving the model with an endogenous information structure, it is useful to study some benchmarks. Our first benchmark is the original Stiglitz (1977) model, in which A is exogenously and privately informed about the accident probability θ . Our second benchmark is a model where neither P nor A knows θ (and they cannot learn about it). In our third and final benchmark we assume that, from the outset of the game, both P and A know θ .

3.1. BENCHMARK 1: THE STIGLITZ MODEL

We here assume the same setup as in Section 2, except that A now privately knows the value of θ from the outset of the game (and this is common knowledge). This model was originally studied by Stiglitz (1977). Although the results and insights coming out of the model are well-known, it will be useful for our subsequent analysis to briefly review them here.

P 's problem is to choose the utility levels \underline{u}_N , \underline{u}_A , \bar{u}_N , and \bar{u}_A so as to maximize the profits in (2), subject to four constraints. First, for each agent type there is one individual rationality (IR) constraint. This ensures that the agent type prefers the insurance policy aimed at her rather than no insurance at all (i.e., $(p, a) = (0, 0)$):

$$(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A \geq \bar{U}^*, \quad (\text{IR-high})$$

¹⁰Formally, $\bar{u}_N \equiv u(w - \bar{p})$, $\bar{u}_A \equiv u(w - d + \bar{a})$, $\underline{u}_N \equiv u(w - \underline{p})$, and $\underline{u}_A \equiv u(w - d + \underline{a})$.

¹¹The resources may, however, be further reduced by A 's information gathering cost, c .

$$(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A \geq \underline{U}^*, \quad (\text{IR-low})$$

where

$$\bar{U}^* \equiv (1 - \bar{\theta}) u(w) + \bar{\theta} u(w - d) \quad \text{and} \quad \underline{U}^* \equiv (1 - \underline{\theta}) u(w) + \underline{\theta} u(w - d)$$

are the utility levels associated with the two types' outside option of not purchasing any insurance policy at all (clearly, $\bar{U}^* < \underline{U}^*$). In addition, P 's optimal menu must satisfy two incentive compatibility (IC) constraints. These ensure that each agent type prefers the policy aimed at her to the other policy:

$$(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A \geq (1 - \bar{\theta}) \underline{u}_N + \bar{\theta} \underline{u}_A, \quad (\text{IC-high})$$

$$(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A \geq (1 - \underline{\theta}) \bar{u}_N + \underline{\theta} \bar{u}_A. \quad (\text{IC-low})$$

Depending on parameter values, there are two qualitatively different ways in which P 's optimal menu of insurance policies may look like. One possibility is that P prefers to shut down the low-demand type, meaning $(\underline{p}, \underline{a}) = (0, 0)$. If so, P effectively interacts with only one type of agent, with a known accident probability $\bar{\theta}$, and of the four constraints only IR-high is relevant. In order to maximize the profits from this single agent type, P should first provide her with full insurance and then extract all of her rents. Thus, at the optimum, the high-demand type receives the same utility whether she has an accident or not, and this utility level equals her outside option utility, \bar{U}^* . It follows (cf. (2)) that P 's optimized profit in this case equals

$$\pi_{SD}^{SZ} = (1 - v) \left[w - \bar{\theta}d - h(\bar{U}^*) \right] \quad (3)$$

(where the superscript SZ is short for Stiglitz and the subscript SD stands for shutdown). A 's ex ante expected utility equals her expected outside option utility:

$$EU_{SD}^{SZ} = v \underline{U}^* + (1 - v) \bar{U}^*. \quad (4)$$

The other possibility is that P optimally interacts with both types, meaning that $(\underline{p}, \underline{a}) \neq (0, 0)$. As is well-known, here the optimal menu of insurance policies is such that the constraints IR-low and IC-high both bind, whereas IR-high and IC-low are lax. Proposition 1 characterizes the menu of optimal insurance policies.

Proposition 1 (The Stiglitz model). *Consider the Stiglitz model and suppose that P optimally interacts with both agent types. Then, at the optimum, the high-demand type is fully insured ($\bar{u}_A^{SZ} = \bar{u}_N^{SZ} \equiv \bar{u}^{SZ}$) and the low-demand type is underinsured ($\underline{u}_A^{SZ} < \underline{u}_N^{SZ}$). The ex post utility levels at the optimum (\underline{u}_A^{SZ} , \underline{u}_N^{SZ} , and \bar{u}^{SZ}) are*

implicitly defined by the two binding constraints (IR-low and IC-high) and by the equality

$$v\underline{\theta}(1-\underline{\theta})[h'(\underline{u}_N^{SZ}) - h'(\underline{u}_A^{SZ})] = (1-v)(\bar{\theta}-\underline{\theta})h'(\bar{u}^{SZ}). \quad (5)$$

Moreover, we have the following relationships between the utility levels:

$$\underline{u}_A^{SZ} < \bar{u}^{SZ} < \underline{U}^* < \underline{u}_N^{SZ}, \quad (6)$$

$$(1-\underline{\theta})\underline{u}_N^{SZ} + \underline{\theta}u_A^{SZ} = \underline{U}^* \quad \text{and} \quad \bar{U}^* < \bar{u}^{SZ}. \quad (7)$$

Proof. See Stiglitz (1977) and the supplementary material to the present paper (available on our websites).

Given the optimal utility levels \underline{u}_N^{SZ} , \underline{u}_A^{SZ} , and \bar{u}^{SZ} that are characterized in Proposition 1, P 's ex ante expected profit can be written as (cf. (2))

$$\pi^{SZ} = \hat{w} - v[(1-\underline{\theta})h(\underline{u}_N^{SZ}) + \underline{\theta}h(\underline{u}_A^{SZ})] - (1-v)h(\bar{u}^{SZ}). \quad (8)$$

Similarly, A 's ex ante expected utility equals

$$\begin{aligned} EU^{SZ} &= v[(1-\underline{\theta})\underline{u}_N^{SB} + \underline{\theta}u_A^{SB}] + (1-v)\bar{u}^{SZ} \\ &= v\underline{U}^* + (1-v)\bar{u}^{SZ}. \end{aligned} \quad (9)$$

(where the second equality sign uses the binding IR-low constraint).

3.2. BENCHMARK 2: NEITHER PARTY INFORMED

Our second benchmark is identical to the model in Section 2 except that here there is no possibility for A to learn the true state. Thus, throughout the game, both parties know only the prior distribution of θ (and this fact is common knowledge).

In this model there is effectively only one type of agent. The relevant outside option payoff for this single type of agent equals

$$EU^{NI} = v\underline{U}^* + (1-v)\bar{U}^* \quad (10)$$

(where NI is short for no information). Clearly, the optimal insurance policy in this benchmark provides full insurance for the single agent type and also leaves her with no rents; thus, $u_N = u_A = EU^{NI}$. It follows (cf. (2)) that P 's maximized ex ante expected profit in this case equals

$$\pi^{NI} = \hat{w} - h(EU^{NI}). \quad (11)$$

A 's ex ante expected utility is given by (10).

3.3. BENCHMARK 3: BOTH PARTIES INFORMED

In our final benchmark we assume that, from the start of the game, both P and A know A 's type perfectly and P can perfectly discriminate between the types. It is clear that the optimal menu of insurance policies provides full insurance for each agent type and there is also full rent extraction for each type. The latter means that the agent's payoff from the menu equals her (interim) expected utility from the outside option, which is either \bar{U}^* or \underline{U}^* . It follows (cf. (2)) that P 's maximized ex ante expected profit in this case equals

$$\pi^{FI} = \hat{w} - (1 - v)h(\bar{U}^*) - vh(\underline{U}^*) \quad (12)$$

(where FI is short for full information). Again, A 's ex ante expected utility in this benchmark equals her expected outside option utility:

$$EU^{FI} = v\underline{U}^* + (1 - v)\bar{U}^*. \quad (13)$$

3.4. COMPARING THE BENCHMARKS

The following proposition compares the ex ante expected payoffs of P and A , respectively, across the three benchmarks.

Proposition 2.

- a) P is strictly better off in Benchmark 2 (where neither party is informed) than in Benchmark 3 (where both parties are informed), and P is strictly better off in Benchmark 3 than in Benchmark 1 (i.e., the Stiglitz model). That is,

$$\pi^{NI} > \pi^{FI} > \pi_{max}^{SZ} \equiv \max\{\pi_{SD}^{SZ}, \pi^{SZ}\}.$$

- b) Suppose $\pi^{SZ} > \pi_{SD}^{SZ}$ (so that P interacts with both types in Benchmark 1). Then A is strictly better off in Benchmark 1 (i.e., the Stiglitz model) than in the two other benchmarks. And she is indifferent between those two other benchmarks. That is,

$$EU^{SZ} > EU_{SD}^{SZ} = EU^{NI} = EU^{FI}.$$

If $\pi^{SZ} \leq \pi_{SD}^{SZ}$, then A is indifferent between all three benchmarks.

Proof. See supplementary material (available on our websites).

To understand the logic behind these results, first consider the inequality $\pi^{NI} > \pi^{FI}$, found in part a) of the proposition. Recall that both with no party being informed and with full information, P 's optimally chosen policy involves full insurance. The difference

in terms of the outcome between the two benchmarks is that in the full information case the insurance premium will vary across the two types. This means that, from an ex ante point of view, also A 's realized interim utility level in the full information benchmark will vary, which hurts her because she is risk averse. Therefore, A is willing to pay less for the insurance policy under full information, which lowers P 's profit. The last inequality in part a) of the proposition says that the profit is even lower in the Stiglitz model. The reason for this is that then A has private information about her type and therefore can earn some rents at P 's expense.

The inequalities in part b) of the proposition can be understood in a similar way. A is best off when she has private information and thus is able to earn a rent. She is indifferent between the full information and no information benchmarks, as in each case her expected payoff equals her expected outside option utility.

4. Endogenous information

This section solves the model described in Section 2. We first consider the possibility that P wants to induce information gathering (Section 4.1). Thereafter, we study the possibility that P does not want to induce information gathering (Section 4.2). Finally, in Section 4.3, we compare these cases and study under what circumstances P optimally induces, and does not induce, information gathering.

4.1. INDUCING INFORMATION GATHERING

Suppose P wants to provide incentives for A to gather information ($x = 1$). As in the Stiglitz model, there are two possibilities: Either P offers one insurance policy to each type of agent or P interacts only with the high-demand type (meaning that the low-demand type remains uninsured). We can study the two cases in a single analytical framework: In order to obtain the case where the low-demand type is shut down, we simply think of that type's insurance policy as being equal to her outside option (i.e., $\underline{u}_N = u(w)$ and $\underline{u}_A = u(w - d)$). The analysis presented below is thus valid for both cases.

P will choose $(\underline{u}_N, \underline{u}_A, \bar{u}_N, \bar{u}_A)$ so as to maximize π as stated in (2), subject to altogether seven constraints. The first four constraints are identical to the ones in the Stiglitz model, namely IR-high, IR-low, IC-high, and IC-low, as stated in Section 3.1. These constraints ensure that each agent type, given that she has learned her accident probability, prefers the policy aimed at her to the other policy and to the outside option.

In addition, three further constraints must be satisfied. These concern A 's incentives at the ex ante stage prior to her having learned her type. First, A must prefer to gather

information about her type to not doing that and then (regardless of her type) choose her outside option:

$$\begin{aligned} EU_{x=1} &\equiv v[(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A] + (1 - v) [(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A] - c \\ &\geq v \underline{U}^* + (1 - v) \bar{U}^*. \end{aligned} \quad (\text{IR-ante})$$

The left-hand side of IR-ante, $EU_{x=1}$, is A 's ex ante expected payoff if gathering information, the last term being the information acquisition cost c (the interim constraints ensure that after having learned her type, A indeed chooses the policy aimed at that particular type); the right-hand side is A 's ex ante expected outside option utility. The second constraint, which we will refer to as an *information gathering constraint*, requires that A must prefer to acquire information about her type to not doing so and instead choose the policy aimed at the low-demand type (without gathering information):

$$EU_{x=1} \geq v[(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A] + (1 - v) [(1 - \bar{\theta}) \underline{u}_N + \bar{\theta} \underline{u}_A]. \quad (\text{IG-low})$$

Notice that the first term of $EU_{x=1}$ and the first term on the right-hand side of IG-low are identical and thus cancel each other out. The third constraint says that A must prefer to gather information about her type to not doing so and instead choose the policy aimed at the high-demand type:

$$EU_{x=1} \geq v[(1 - \underline{\theta}) \bar{u}_N + \underline{\theta} \bar{u}_A] + (1 - v) [(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A]. \quad (\text{IG-high})$$

Similarly to the case above, the second term of $EU_{x=1}$ and the second term on the right-hand side of IG-high cancel each other out.

There is a strong relationship between the two IG constraints, on the one hand, and the IC constraints, on the other. In particular, by inspecting the constraints, it is clear that IG-low implies IC-high. Similarly, IG-high implies IC-low. We thus have:

Lemma 1. *Suppose that IG-low and IG-high are satisfied. Then also IC-low and IC-high are satisfied.*

It turns out that also IR-ante is implied by other constraints:

Lemma 2. *IR-ante is implied by IR-low and IG-low.*

Proof. See Appendix A.

By Lemmas 1 and 2 stated above and by Lemma A1 in Appendix A, we can ignore IC-low, IC-high, IR-ante, and IR-high when solving P 's profit maximization problem.

The problem is then to maximize π with respect to $(\underline{u}_N, \underline{u}_A, \bar{u}_N, \bar{u}_A)$, subject to IR-low, IG-low, and IG-high. The Lagrangian of this problem can be written as¹²

$$\begin{aligned} \mathcal{L} = & \hat{w} - v [(1 - \underline{\theta}) h(\underline{u}_N) + \underline{\theta} h(\underline{u}_A)] - (1 - v) [(1 - \bar{\theta}) h(\bar{u}_N) + \bar{\theta} h(\bar{u}_A)] \\ & + \lambda [(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A - \underline{U}^*] - \bar{\mu} \{v [(1 - \underline{\theta}) (\bar{u}_N - \underline{u}_N) + \underline{\theta} (\bar{u}_A - \underline{u}_A)] + c\} \\ & + \underline{\mu} \{(1 - v) [(1 - \bar{\theta}) (\bar{u}_N - \underline{u}_N) + \bar{\theta} (\bar{u}_A - \underline{u}_A)] - c\}, \end{aligned}$$

where $\lambda \geq 0$ is the shadow price associated with IR-low, $\bar{\mu} \geq 0$ is the shadow price associated with IG-high, and $\underline{\mu} \geq 0$ is the shadow price associated with IG-low.

The first-order conditions of the maximization problem immediately imply that two of the constraints must bind:

Lemma 3. *IG-low and IR-low are both binding at the optimum (i.e., $\underline{\mu} > 0$ and $\lambda > 0$).*

Proof. First consider the first-order condition with respect to \bar{u}_N :

$$\frac{\partial \mathcal{L}}{\partial \bar{u}_N} = 0 \Leftrightarrow (1 - v) (1 - \bar{\theta}) h'(\bar{u}_N) = \underline{\mu} (1 - v) (1 - \bar{\theta}) - \bar{\mu} v (1 - \underline{\theta}). \quad (14)$$

Given $h' > 0$ and $\bar{\mu} \geq 0$, the first claim follows immediately from (14). The second claim is obvious if P interacts only with the high-demand type (while the low-demand type remains uninsured). Thus suppose P interacts with both types. In this case the first-order condition with respect to \underline{u}_N is

$$\frac{\partial \mathcal{L}}{\partial \underline{u}_N} = 0 \Leftrightarrow v (1 - \underline{\theta}) h'(\underline{u}_N) = \lambda (1 - \underline{\theta}) + \bar{\mu} v (1 - \underline{\theta}) - \underline{\mu} (1 - v) (1 - \bar{\theta}). \quad (15)$$

Now, add (14) and (15):

$$(1 - v) (1 - \bar{\theta}) h'(\bar{u}_N) + v (1 - \underline{\theta}) h'(\underline{u}_N) = \lambda (1 - \underline{\theta}). \quad (16)$$

This equality (using $h' > 0$) implies that $\lambda > 0$. \square

One implication of Lemma 3 is that A will not get an informational rent if P finds it optimal not to interact with the low-demand type. This follows because IG-low (which according to the lemma is binding) coincides with IR-ante if $\underline{u}_N = u(w)$ and $\underline{u}_A = u(w - d)$.

Combining the first-order conditions of P 's maximization problem also yield the following result concerning the distortion of the optimal insurance policies:

Lemma 4. *The low-demand type is underinsured ($\underline{u}_N > \underline{u}_A$) at the optimum.*

¹²If P prefers not to interact with the low-demand type, then the Lagrangian is a function of only \bar{u}_N and \bar{u}_A , as the other two utility levels are given by the outside option.

Proof. The claim is obvious if P chooses to interact only with the high-demand type. Hence, let us concentrate on the case where P interacts with both types. The first-order condition with respect to \underline{u}_A is

$$\frac{\partial \mathcal{L}}{\partial \underline{u}_A} = 0 \Leftrightarrow v\underline{\theta}h'(\underline{u}_A) = \lambda\underline{\theta} + \bar{\mu}v\underline{\theta} - \underline{\mu}(1-v)\bar{\theta}. \quad (17)$$

To prove the claim in the lemma, multiply (15) by $\underline{\theta}$ and multiply (17) by $(1-\underline{\theta})$:

$$v\underline{\theta}(1-\underline{\theta})h'(\underline{u}_N) = \lambda\underline{\theta}(1-\underline{\theta}) + \bar{\mu}v\underline{\theta}(1-\underline{\theta}) - \underline{\mu}(1-v)\underline{\theta}(1-\bar{\theta}),$$

$$v\underline{\theta}(1-\underline{\theta})h'(\underline{u}_A) = \lambda\underline{\theta}(1-\underline{\theta}) + \bar{\mu}v\underline{\theta}(1-\underline{\theta}) - \underline{\mu}(1-v)\bar{\theta}(1-\underline{\theta}).$$

Next subtract the second one of those equalities from the first one and simplify:

$$v\underline{\theta}(1-\underline{\theta})[h'(\underline{u}_N) - h'(\underline{u}_A)] = \underline{\mu}(1-v)(\bar{\theta} - \underline{\theta}). \quad (18)$$

The claim now follows from the strict convexity of h and the result in Lemma 3 that $\underline{\mu} > 0$. \square

Lemma 5. *Suppose that it is optimal for P to induce information gathering. Then, at the optimum, IG-high is lax (i.e., then $\bar{\mu} = 0$) and the high-demand type is fully insured (i.e. $\bar{u}_N = \bar{u}_A$).*

Proof. See Appendix A.

The logic behind the result that $\bar{\mu} = 0$ is that if IG-high were binding, then A , from an ex ante point of view (so without knowing her type), would be indifferent between incurring the cost c and not doing that and instead choose the insurance policy aimed at the high-demand type. If so, A would be willing to refrain from information acquisition and instead purchase a pooled alternative policy that yields the same ex ante expected utility to A as the original menu of insurance policies offered by P . By letting the alternative policy involve full insurance, and thanks to the fact that this policy does not (since it is pooled) require the information acquisition cost to be incurred, P can earn higher profits from the alternative policy. It follows that inducing information acquisition cannot be optimal if IG-high is binding. Once we have the result that $\bar{\mu} = 0$, it is quite straightforward to show, using the first-order conditions, that the high-demand type is fully insured.

We can now characterize the optimal menu of insurance policies.

Proposition 3. *Suppose that it is optimal for P to induce information gathering and interact with both types. Then the binding constraints are IR-low and IG-low. The high-demand type is fully insured ($\bar{u}_A^{SB} = \bar{u}_N^{SB} \equiv \bar{u}^{SB}$) and the low-demand type*

is underinsured ($\underline{u}_A^{SB} < \underline{u}_N^{SB}$). The utility levels at the optimum (\underline{u}_A^{SB} , \underline{u}_N^{SB} , and \bar{u}^{SB}) are implicitly defined by the two binding constraints and by the equality

$$v\underline{\theta}(1 - \underline{\theta}) [h'(\underline{u}_N^{SB}) - h'(\underline{u}_A^{SB})] = (1 - v)(\bar{\theta} - \underline{\theta}) h'(\bar{u}^{SB}) \quad (19)$$

and ex ante expected profits are given by

$$\pi^{SB} = \widehat{w} - v [(1 - \underline{\theta}) h(\underline{u}_N^{SB}) + \underline{\theta} h(\underline{u}_A^{SB})] - (1 - v) h(\bar{u}^{SB}). \quad (20)$$

If it is optimal for the principal to induce information gathering and to interact only with the high-demand type, then this type is fully insured with

$$u_N^{SD} = u_A^{SD} = u^{SD} = \bar{U}^* + \frac{c}{1 - v}$$

and ex ante expected profits are

$$\pi_{SD}^{SB} = (1 - v) \left[w - \bar{\theta}d - h \left(\bar{U}^* + \frac{c}{1 - v} \right) \right]. \quad (21)$$

Proof. See Appendix A.

Depending on which option is more profitable, P chooses to interact only with the high-demand type or with both types. P 's optimized profits (conditional on inducing information gathering) will thus equal

$$\pi_{x=1}^* \equiv \max \{ \pi^{SB}, \pi_{SD}^{SB} \}. \quad (22)$$

How do these profits relate to the profits P earns in the Stiglitz model? The following lemma, which will be useful when identifying the overall optimum in Section 4.3, provides us with some information about this.

Lemma 6. *Suppose that it is optimal for P to induce information gathering. Then P 's ex ante expected profits are decreasing in c : $\partial \pi_{x=1}^* / \partial c < 0$. Moreover, as the information gathering cost goes to zero, the profit level approaches the one in the Stiglitz model: $\lim_{c \rightarrow 0} \pi_{x=1}^* = \pi_{max}^{SZ} \equiv \max \{ \pi_{SD}^{SZ}, \pi^{SZ} \}$.*

Proof. See Appendix A.

The result that the profits are decreasing in the cost follows from the envelope theorem. The cost c does not show up directly in P 's profit function (since the cost is incurred by A); it appears only in the two IG constraints. These constraints become more stringent, the higher is the cost (because A must be compensated for the cost she incurs). Therefore, the profits must be decreasing in c , as IG-low binds.

4.2. NOT INDUCING INFORMATION GATHERING

Now suppose that P does not want to provide incentives for A to gather information ($x = 0$). A must then let the two types choose the same insurance policy, (u_N, u_A) . P 's profit function can be written as

$$\pi = \widehat{w} - [v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})] h(u_N) - [v\underline{\theta} + (1 - v)\bar{\theta}] h(u_A).$$

The problem P faces is to maximize π with respect to u_N and u_A , subject to three constraints.

First, without having access to information about her type, A must prefer to purchase the insurance policy to the alternative of remaining uninsured, which would yield the expected outside option utility:

$$\begin{aligned} EU_{x=0} &\equiv [v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})] u_N + [v\underline{\theta} + (1 - v)\bar{\theta}] u_A \\ &\geq v\underline{U}^* + (1 - v)\bar{U}^*. \end{aligned} \quad (\text{IR-ante})$$

Second, A must be willing to purchase the insurance policy without access to information about her type instead of gathering information and then buy the policy only if being the high-demand type:

$$EU_{x=0} \geq v\underline{U}^* + (1 - v)[(1 - \bar{\theta})u_N + \bar{\theta}u_A] - c. \quad (\text{IG-high})$$

The third constraint requires that A must be willing to purchase the insurance policy without access to information about her type instead of gathering information and then buy the policy only if being the low-demand type. Unsurprisingly, however, this constraint will never bind and we can thus safely disregard it.¹³

To solve P 's problem we will make use of a graphical analysis. To this end, note that the two constraints IR-ante and IG-high can be rewritten as $u_N \geq \varphi_{ante}(u_A)$ and $u_N \geq \varphi_{high}(u_A)$, respectively, where the two φ functions are linear and decreasing in u_A .¹⁴ Lemma 7 below relates the slopes of these functions to each other; it also indicates

¹³To see this, write the constraint as

$$EU_{x=0} \geq v[(1 - \underline{\theta})u_N + \underline{\theta}u_A] + (1 - v)\bar{U}^* - c. \quad (\text{IG-low})$$

and note that this constraint is implied by IR-ante if $(1 - \underline{\theta})u_N + \underline{\theta}u_A \leq \bar{U}^*$. If $(1 - \underline{\theta})u_N + \underline{\theta}u_A > \bar{U}^*$, then $(1 - \bar{\theta})u_N + \bar{\theta}u_A > \bar{U}^*$ has to hold as $u(w) - u(w - d) \geq u_N - u_A$ by $p \geq 0$. This implies

$$v[(1 - \underline{\theta})u_N + \underline{\theta}u_A] + (1 - v)\bar{U}^* - c < v[(1 - \underline{\theta})u_N + \underline{\theta}u_A] + (1 - v)((1 - \bar{\theta})u_N + \bar{\theta}u_A) = EU_{x=0}$$

which again implies IG-low.

¹⁴We have,

$$\varphi_{ante}(u_A) \equiv \frac{v\underline{U}^* + (1 - v)\bar{U}^*}{v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})} - \frac{v\underline{\theta} + (1 - v)\bar{\theta}}{v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})} u_A,$$

the crossing point between φ_{ante} and φ_{high} . First define the following shorthand notation: $\theta^e \equiv v\underline{\theta} + (1-v)\bar{\theta}$ and $k \equiv c/[v(1-v)(\bar{\theta} - \underline{\theta})]$.

Lemma 7. *We have $\varphi'_{ante}(u_A) < \varphi'_{high}(u_A) < 0$. Moreover, φ_{ante} and φ_{high} cross at the point*

$$(u_A, u_N) = (u(w-d) + (1-\theta^e)k, u(w) - \theta^e k). \quad (23)$$

Proof. See supplementary material (available on our websites).

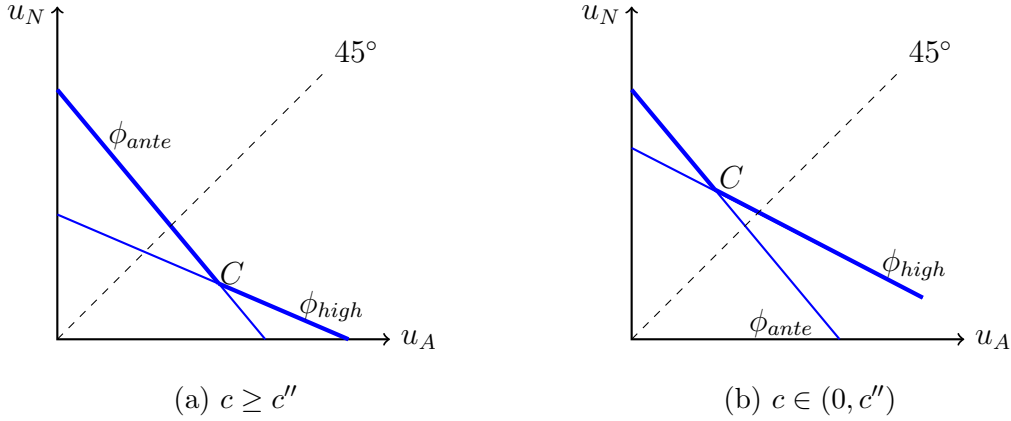


Figure 1: Not inducing information gathering

Panels (a) and (b) of Figure 1 depict the two constraints in the (u_A, u_N) -space. The two panels of the figure also show the 45-degree line along which A is fully insured. Using the information in Lemma 7, it is easy to verify that the crossing of φ_{ante} and φ_{high} (point C) is located at or below the 45-degree line if and only if

$$c \geq v(1-v)(\bar{\theta} - \underline{\theta})[u(w) - u(w-d)] \equiv c''.$$

Panel (a) depicts the case where C is located at or below the 45-degree line, and Panel (b) shows the opposite case.

An isoprofit curve has the following slope:

$$\varphi'_{ante}(u_A) \frac{h'(u_A)}{h'(u_N)}. \quad (24)$$

That is, any isoprofit curve is downward-sloping and, whenever A is fully insured ($u_A = u_N$), its slope equals the slope of IR-ante; above (below) the 45-degree line the slope of the isoprofit curve is flatter (steeper) than the slope of IR-ante. Now consider Panel

$$\varphi_{high}(u_A) \equiv \frac{vU^* - c}{v(1-\underline{\theta})} - \frac{\underline{\theta}}{1-\underline{\theta}}u_A.$$

(a) (i.e., the case where the information gathering cost c is relatively high). P 's profits are higher, the closer to the origin A 's utility levels are located. P 's optimal choice of insurance policy must of course satisfy both constraints; that is, it must lie north-east of, or exactly at, the relatively thick segments of the two straight lines in the figure. It is clear that in Panel (a), the optimum must be located on the 45-degree line at a point where the isoprofit curve is tangent to IR-ante and only that constraint is binding. At any other point on the edge of the feasible set, the isoprofit curve's slope would differ from the slope of the binding constraint and P would thus be able to increase the profits by moving closer to full insurance.

Next consider Panel (b) (i.e., the case where the information gathering cost c is relatively low). Again the edge of the feasible set is indicated by the thicker segments of the two constraints. By using the same argument as above, we can conclude that the optimum cannot be at a point where $u_A \geq u_N$: Then only IG-high would bind and this constraint is flatter than the isoprofit curve. By a similar argument, the optimum cannot be located anywhere strictly north-west of C . The only remaining possibilities are that the optimum is exactly at C or at a point between C and the 45-degree line.

To start with, let us investigate under what circumstances the optimum is exactly at C . This will be the case if and only if, at C , the slope of the isoprofit curve is weakly steeper than the slope of IG-high:

$$|\varphi'_{ante}| \frac{h' [u(w-d) + \theta^e k]}{h' [u(w) - (1 - \theta^e) k]} \geq |\varphi'_{high}|. \quad (25)$$

At $c = c''$, this inequality is satisfied strictly (because at that value of c , $u_A = u_N$ and thus the slopes of IR-ante and the isoprofit curve are identical). Moreover, it is satisfied for all $c > 0$ if and only if it is satisfied for $c = 0$. Evaluating (25) at $c = 0$ and rearranging, we have

$$\frac{u'(w)}{u'(w-d)} \geq \frac{v\underline{\theta}(1-\underline{\theta}) + (1-v)\underline{\theta}(1-\bar{\theta})}{v\underline{\theta}(1-\underline{\theta}) + (1-v)\bar{\theta}(1-\underline{\theta})}. \quad (26)$$

This inequality may or may not hold—for values of d sufficiently close to zero it does, whereas for d 's sufficiently close to w it does not. We can conclude that if the parameters are such that (26) holds, then the optimum is exactly at C for all $c \in (0, c'')$. Suppose that (26) is violated. Then—since (25) becomes more stringent, the higher c is—there must be a critical value of c , call it c' , such that for $c \in (0, c')$ the optimum is at a point along IG-high strictly between C and the 45-degree line;¹⁵ moreover, we have $0 < c' < c''$.

¹⁵The cutoff value c' is implicitly defined by

$$\varphi'_{ante} h' \left[u(w-d) + \frac{\theta^e c'}{v(1-v)(\bar{\theta}-\underline{\theta})} \right] = \varphi'_{high} h' \left[u(w) - \frac{(1-\theta^e) c'}{v(1-v)(\bar{\theta}-\underline{\theta})} \right].$$

The following proposition summarizes the results.

Proposition 4. *Assume that P chooses the optimal menu, conditional on A being induced not to gather information.*

- a) *Suppose $c \geq c''$. Then only IR-ante binds and there is full insurance: $u_A^{NI} = u_N^{NI} \equiv u^{NI} = v\underline{U}^* + (1-v)\overline{U}^*$.*
- b) *Suppose that (26) holds and $c < c''$; or that (26) is violated and $c \in [c', c'']$. Then IR-ante and IG-high bind and A is underinsured, with (u_A^{NI}, u_N^{NI}) given by (23).*
- c) *Suppose that (26) is violated and $c \in (0, c')$. Then only IG-high binds and A is underinsured, with (u_A^{NI}, u_N^{NI}) implicitly being defined by the binding constraint and by*

$$\frac{h'(u_A^{NI})}{h'(u_N^{NI})} = \frac{v\underline{\theta}(1-\underline{\theta}) + (1-v)\underline{\theta}(1-\overline{\theta})}{v\underline{\theta}(1-\underline{\theta}) + (1-v)\overline{\theta}(1-\underline{\theta})}. \quad (27)$$

Proof. See Appendix B.

We let $\pi_{x=0}^*$ denote P 's profit given the optimum in Proposition 4:

$$\pi_{x=0}^* = \widehat{w} - [v(1-\underline{\theta}) + (1-v)(1-\overline{\theta})] h(u_N^{NI}) - [v\underline{\theta} + (1-v)\overline{\theta}] h(u_A^{NI}). \quad (28)$$

The results stated in the following lemma will, in the next subsection, help us characterize the overall optimum.

Lemma 8. *Assume that P chooses the optimal menu, conditional on A being induced not to gather information.*

- a) *Suppose that $c \in (0, c'')$. Then P 's ex ante expected profits are increasing in c : $\partial\pi_{x=0}^*/\partial c > 0$.*
- b) *Suppose that $c > c''$. Then P 's ex ante expected profits are constant with respect to c : $\partial\pi_{x=0}^*/\partial c = 0$.*
- c) *As the information gathering cost goes to zero, the profit level approaches a value strictly below the ex ante expected profits in the Stiglitz model: $\lim_{c \rightarrow 0} \pi_{x=0}^* < \lim_{c \rightarrow 0} \pi_{x=1}^* = \pi_{max}^{SZ} \equiv \max\{\pi^{SZ}, \pi_{SD}^{SZ}\}$.*

Proof. See Appendix B.

Parts a) and b) of Lemma 8 say that P 's profits when not inducing information gathering are increasing in c if IG-high is binding; and if IG-high is lax then they are

constant with respect to c . The reason for this result is similar to the reason for why P 's profits in Section 4.1 are decreasing in c . There, a larger c made the IG constraints more stringent, which had a negative impact on the profits. Here, in contrast, an increase in c *relaxes* the IG constraints (P wants to ensure that A does not incur the cost, which is easier to achieve if the cost is high). Therefore, whenever IG-high is binding, P 's profits benefit from an increase in the information gathering cost.

Part c) of Lemma 8 says that, for small values of c , the profits that P can earn when not inducing information acquisition are lower than those in the Stiglitz model. The reason for this is that if A does not know her type, then P can offer only a pooled insurance policy to A . It is well known that, in the Stiglitz model with two types, offering the same contract to the two types is always dominated by separating the contracts.

The result in Lemma 8, part c), can actually be taken a step further. If we assume that inequality (26) is satisfied, then we know from above that P 's optimum is, for all $c < c''$, exactly at point C in Panel (b) of Figure 1. It is straightforward to verify that as c goes to zero, C approaches the point $(u(w), u(w - d))$. That is, in this limit, A is not insured at all, which means that P earns zero profits: $\lim_{c \rightarrow 0} \pi_{x=0}^* = 0$. In other words, when the cost of information acquisition is sufficiently close to zero, IG-high is so stringent that there exists no insurance policy that yields a positive profit for P and at the same time (i) satisfies A 's individual rationality constraint and (ii) incentivizes A not to gather information. If we were in the limit and P tried to offer a policy yielding a positive profit, A would want to learn her type (given that $c \rightarrow 0$) and then accept the offer only if being the high-demand type.

The other case, where inequality (26) is violated, is a bit less extreme. Here the optimal contract involves some (although only partial) insurance also in the limit when c goes to zero. However, in this case A receives some rents, which adds to the costs for P of trying to induce no information acquisition.

4.3. THE OVERALL OPTIMUM

P will induce information gathering if and only if the profits from doing this are higher than the profits from not inducing information gathering. In order to find the overall optimum to P 's problem, we therefore must compare the profits in (22) with those in (28).

The comparison is facilitated by Figure 2, showing the $(c, \text{profits})$ -space. First consider Panel (a) of that figure. The downward-sloping curve is the graph of P 's optimized profit, as a function of c , when information acquisition is induced. We know from Lemma 6 that this graph is continuous and downward-sloping; moreover, as c goes to zero, the graph approaches π_{max}^{SZ} , the profits in the Stiglitz model. The other curve in the figure

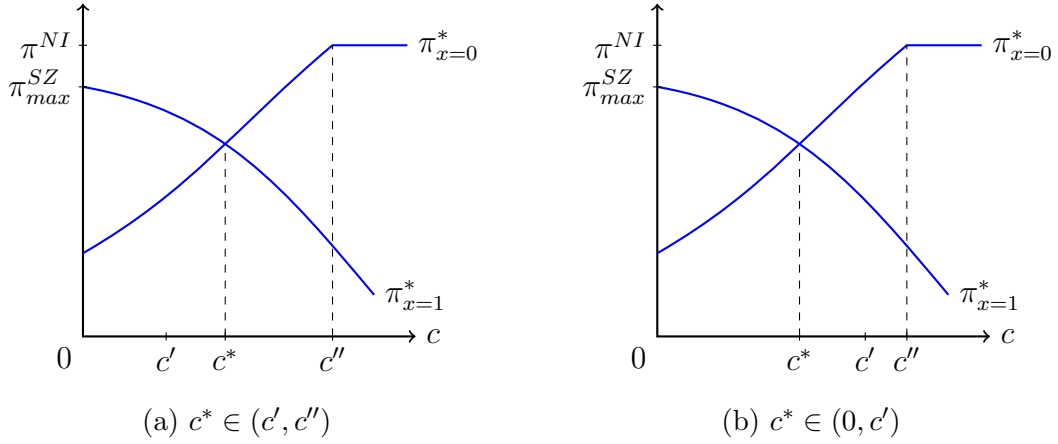


Figure 2: Overall optimum

is the graph of P 's optimized profit when information acquisition is not induced. This graph is also continuous and it starts at a point strictly below π_{max}^{SZ} ; it is then initially upward-sloping and after this, for $c \geq c''$, flat. When drawing this graph we could make use of the results in Lemma 8. Notice that, for all $c \geq c''$, P 's profits when information acquisition is not induced equal π^{NI} , the profit level in the benchmark where neither party is informed about A 's type. Moreover, from our analysis in Section 3.4, we know that $\pi^{NI} > \pi_{max}^{SZ}$.

All in all, it is clear from the figure that there exists a strictly positive cutoff value of c , which we denote by c^* , such that for all $c \in (0, c^*)$ it is optimal for P to induce information gathering and for all $c > c^*$ it is optimal for P not to do that. The cutoff value c^* is implicitly defined by the equality $\pi_{x=0}^* = \pi_{x=1}^*$.

Panel (a) illustrates the case where $c^* > c'$; that is, here the crossing of the two profit curves occurs at a value of c at which both IG-high and IR-ante are binding when P does not induce information gathering (this case corresponds to part b of Proposition 4). Panel (b) of Figure 2 shows the case where $c^* < c'$; that is, here the crossing occurs at a value of c at which only IG-high binds when P does not induce information gathering (part c of Proposition 4).

Proposition 5 below summarizes the results. We ignore the knife-edge case $c = c^*$ (where P is indifferent between inducing information gathering and not doing so).

Proposition 5. *Consider the overall optimum of the model with endogenous information acquisition.*

- a) *Suppose that $c \in (0, c^*)$. Then P induces A to gather information ($x = 1$). The menu of insurance policies is as stated in Proposition 3. Moreover, as c goes to zero, the menu of insurance policies approaches the one in the Stiglitz model.*

- b) *Suppose that $c > c^*$. Then P does not induce A to gather information ($x = 0$). The menu of insurance policies is as stated in Proposition 4.*

Proof. See Appendix C.

The result that, for low enough values of c , A gathers information stands in contrast to the result in Cremer and Khalil (1992). They study a procurement problem with endogenous information structure and show that the principal never, for any strictly positive information gathering cost, induces information gathering. The reason is that, in their model, the agent learns her type before choosing a contract from the menu even if she does not gather information. In contrast to our model, gathered information is relevant only for the participation decision but not for the contract choice.

We will use a superscript asterisk to denote the ex post utility levels at the overall optimum—similarly with P 's ex ante expected profits (π^*) and A 's ex ante expected utility (EU^*) at the overall optimum. The following proposition reports a number of comparative statics results concerning these variables.

Proposition 6. *Consider the overall optimum of the model with endogenous information acquisition.*

- a) *Suppose $c \in (0, c^*)$. Then the high-demand type is fully insured and the low-demand type is underinsured ($\bar{u}_A^* = \bar{u}_N^* \equiv \bar{u}^*$ and $\underline{u}_A^* < \underline{u}_N^*$). A lower c makes the extent of underinsurance smaller: $\partial \underline{u}_N^* / \partial c > 0$, $\partial \underline{u}_A^* / \partial c < 0$, and $\partial \bar{u}^* / \partial c > 0$. Moreover, from an ex ante perspective, a lower c (weakly) benefits both P and A : $\partial \pi^* / \partial c < 0$ and $\partial EU^* / \partial c \leq 0$.*
- b) *Suppose $c \in (c^*, c')$. Then the single type is underinsured ($u_A^* < u_N^*$). Both ex post utility levels are decreasing in c : $\partial u_N^* / \partial c < 0$ and $\partial u_A^* / \partial c < 0$; if $-\frac{u'''}{u''} \geq -2\frac{u''}{u'}$,¹⁶ then a lower c makes the extent of underinsurance larger ($\partial u_N^* / \partial c < \partial u_A^* / \partial c$). Moreover, from an ex ante perspective, a lower c hurts P and benefits A : $\partial \pi^* / \partial c > 0$ and $\partial EU^* / \partial c < 0$.*
- c) *Suppose $c > c^*$ and $c \in (c', c'')$. Then the single type is underinsured ($u_A^* < u_N^*$). A lower c makes the extent of underinsurance larger: $\partial u_N^* / \partial c < 0$ and $\partial u_A^* / \partial c > 0$. Moreover, from an ex ante perspective, a lower c hurts P and has no effect on A 's utility: $\partial \pi^* / \partial c > 0$ and $\partial EU^* / \partial c = 0$.*
- d) *Suppose $c > c''$. Then the single type is fully insured (and $\partial u^* / \partial c = 0$). Moreover, from an ex ante perspective, a lower c has no effect on P 's profits or on A 's utility: $\partial \pi^* / \partial c = 0$ and $\partial EU^* / \partial c = 0$.*

¹⁶This is a sufficient condition. With a utility function characterized by constant relative risk aversion, the condition says that the degree of relative risk aversion must not exceed one.

Proof. See Appendix C.

Propositions 5 and 6 tell us that, for sufficiently low values of the information gathering cost, the qualitative features of the optimal menu of insurance policies are the same as in the Stiglitz (1977) model. In particular, there is no distortion at the top and there is full rent extraction at the bottom. Moreover, the values of the ex post utility levels in the limit as c goes to zero approach the corresponding utility levels in the Stiglitz model. We can thus conclude that Stiglitz's classic framework appears to be robust with respect to the endogenization of asymmetric information. This is intuitive but not obvious, given Crémer and Khalil's (1992) results.

Propositions 5 and 6 also tell us that, for sufficiently large values of c , P will not induce information gathering and, thus, A will not have access to private information about her type. At first glance, this may suggest that here we are effectively in Benchmark 2 (see Section 3.2), where none of the parties is informed and there is full insurance. This is not the case, however, as long as the cost is not too large (namely, as long as $c \in (c^*, c'')$). Although A chooses not to incur the information acquisition cost, the fact that she has the opportunity to do so leads to a distortion of the insurance policy (in particular, to underinsurance). The reason for this is that the IG-high constraint binds and prevents P from providing full insurance. One can think of the situation as one of moral hazard, because the source of the problem is that A can take a non-observable action. To appropriately incentivize (not taking) this action, P chooses to underinsure A .

Proposition 6 further provides us with results concerning the welfare effects of an exogenous reduction in the cost of information gathering. These comparative statics results are interesting and potentially useful, as they can shed light on the desirability of public policies that facilitate for consumers in insurance markets to learn about their accident risks. Examples of policy interventions that fit our framework include free or subsidized tests for genetic disorders and the provision of information (through websites, free telephone lines with expert advice, or campaigns) about how health risks are affected by lifestyle choices.

We should remember that Proposition 6 deals only with cost reductions where c is, both before and after the change, located either below or above c^* (the welfare effects of moving c across the threshold c^* will be discussed below). With that in mind, however, we can conclude that Proposition 6 lends mixed support to the naive intuition that a lower cost of information gathering should be welfare enhancing. While the results are consistent with that intuition for sufficiently low costs ($c < c^*$), for somewhat higher costs the results are either ambiguous or they point in the opposite direction. Indeed, it is striking that the only situation described in Proposition 6 for which a cost reduction is beneficial for both parties is *when the cost level at the outset is so low that A already*

gathers information. It seems plausible to believe that this is precisely the kind of situation where a policymaker would not undertake the policy, as it would not be thought of as being needed. Similarly, for cost values that are so high that A does not acquire information (say $c \in (c^*, c'')$), meaning that a policymaker is likely to think of a cost-reducing policy as indeed being required and desirable, the results in Proposition 6 suggest that the policy in fact would *not* be desirable (or at least it would make one of the parties worse off). In particular, if $c > c^*$ and $c \in (c', c'')$ expected utility is constant in c while profits are increasing. Hence, a reduction in c would lead to a Pareto inferior situation.

From a normative point of view, these observations also lead us to the following policy conclusion: If empirical data suggest that consumers indeed tend to gather information about their health risks, then facilitating (further) for consumers to gather such information is welfare enhancing;¹⁷ however, if the data suggest that consumers tend not to gather information, then facilitating learning has mixed welfare effects and may indeed be undesirable (using only the Pareto criterion).

One way in which we can reconcile the results in Proposition 6 with our naive intuition (which says that lower costs should always be good) is to modify the welfare criterion, assuming a consumer surplus standard: The proposition says that A is never hurt by a local reduction in the information gathering cost.¹⁸ However, as pointed out above, Proposition 6 concerns cost reductions only within the range $c < c^*$ and within the range $c > c^*$. If a public policy lowers c from some level above c^* to a level below it, then it turns out that A can indeed be worse off. We now turn to an investigation of cost reductions that move c over the threshold c^* .

We first make the straightforward observation that if the cost is lowered from some cost level satisfying $c > c^*$ and $c \in (c', c'')$ to some cost level $c < c^*$, then the consumer

¹⁷Obviously, this is true only if the act of reducing the information acquisition cost is itself not prohibitively costly for the policymaker.

¹⁸Another way in which we can modify our welfare criterion is to disregard the profit and utility levels and instead focus on the extent to which the insurance policies in P 's menu are distorted away from full coverage. Proposition 6 shows that if the cost parameter is such that A acquires information, then a reduction in c leads to less underinsurance and, therefore, to less distortion. However, if the cost parameter is such that A does not acquire information (but still not so high that the cost does not matter for the optimal insurance policy), a reduction in the cost leads to *more severe* underinsurance and, therefore, to more distortion. (A caveat to this statement is that in the case described under b it is guaranteed to hold only if the condition $-\frac{u'''}{u''} \geq -2\frac{u''}{u'}$ is satisfied.) This result is broadly in line with what we concluded above, namely that a public policy that facilitates consumer learning is welfare enhancing for low cost levels (so low that the consumer at the outset gathers information), but not necessarily otherwise.

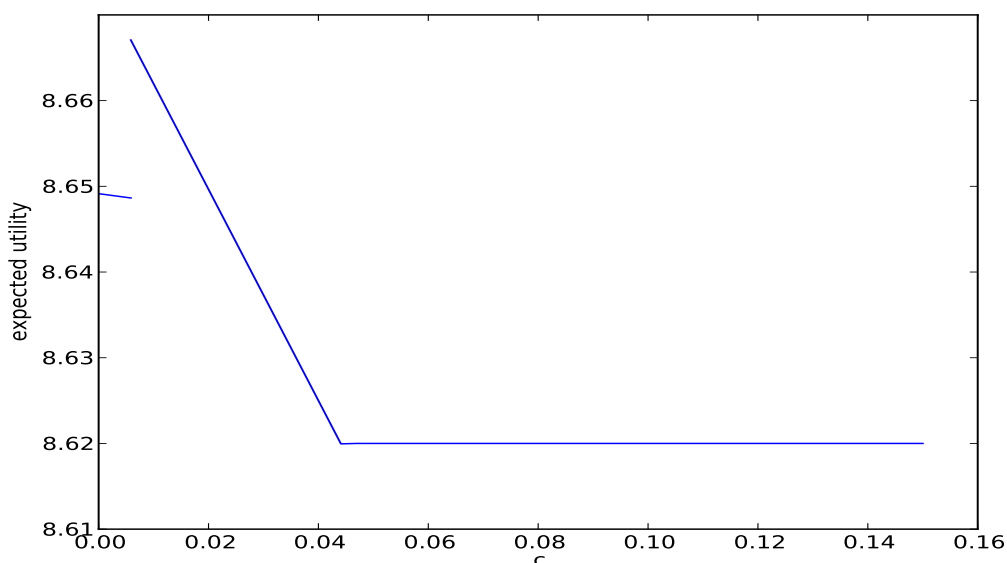


Figure 3: The consumer's expected utility (numerical example)

gains, as she here receives a positive rent only after the cost reduction.¹⁹ (Whether the insurer is better or worse off depends on the particular parameter values—see Figure 2.) Now consider the more subtle case of a change of c from some $c \in (c^*, c')$ to a $c < c^*$ (meaning that the consumer receives a rent also before the cost reduction). Here the analysis is harder, but we have explored the effects of this kind of cost reduction with the help of numerical methods.²⁰ In particular, we have verified that there exist parameter configurations for which the consumer is indeed worse off from the cost reduction. This is shown in Figure 3, which graphs the consumer's expected utility as a function of c . The calculations behind the figure assume a square-root utility function and the following parameter values: $w = 100$, $d = 51$, $\bar{\theta} = 0.7$, $\underline{\theta} = 0.4$, $\nu = 0.8$. Under these assumptions, the three endogenous cutoff values of the model are given by $c^* = 0.00585$, $c' = 0.044926$, and $c'' = 0.144$. As shown in the figure, and consistent with our analytical results, the expected utility is constant in the range $c' < c < c''$. Moreover, in each one of the ranges $0 < c < c^*$ and $c^* < c < c'$ the expected utility is (again consistent with the analytical results) decreasing in c . Interestingly, at the point $c = c^*$ the expected utility makes a jump, and it is clear from the figure that there are many possible values of the cost, before and after the change, that would make a cost reduction bad for the consumer. Indeed, the value of c that yields the highest expected utility for the consumer is strictly

¹⁹If P interacts only with the high-demand type when information is acquired (so if ν is sufficiently low), then A does not get any rent left of c^* either.

²⁰The code that we have used for these simulations (written in Python) is available on our websites.

positive and sufficiently large to ensure that the insurer does not induce information gathering.

Our numerical analysis thus shows that *a public policy that “successfully” facilitates consumer learning* (in the sense that the consumers gather information after, but not prior to, the cost reduction) *can be counterproductive in that it lowers the consumer surplus*. What is the logic behind this result? There are, for the consumer, two potential advantages with a c that is high enough not to induce information gathering. One of them is that then the cost does not have to be incurred. The other potential advantage is that, loosely speaking, the distortions in the insurance contracts can be smaller when information acquisition is not induced. In our example in Figure 3, both these reasons play a role. The consumer’s expected utility immediately right and left of the cutoff c^* equals 8.66707 and 8.64864, respectively. This means that the vertical length of the jump at c^* (i.e., the consumer’s loss in expected utility when the cost is lowered) equals $8.66707 - 8.64864 = 0.01843$. However, the cost $c^* = 0.00585$ is only about one-third (31.7 percent) of this utility difference and it can therefore not account fully for the consumer’s utility loss. The remaining two-thirds can be attributed to the fact that, overall, the contracts are more severely distorted left of c^* .²¹ We should note that in this example the insurer interacts with both consumer types when $c < c^*$. In an optimum in which the low-demand type was shut down, the consumer would not receive *any* rents left of the cutoff c^* . This reinforces the reasoning outlined above and we should thus expect, again, that the consumer’s ideal cost level is strictly positive.

5. Conclusions

This paper has investigated an insurance market in which consumers can learn about their risk. In particular, we analyzed the welfare effects of a public policy that facilitates such learning. We did this by endogenizing the information structure in the monopoly insurance model due to Stiglitz (1977). The previous literature has not carried out this kind of exercise in an insurance setting and in the presence of market power, nor has it focused on our comparative statics exercises.

We find that the Stiglitz model is robust and that, for a sufficiently small information acquisition cost, the consumer of the model is, endogenously, privately informed about her health risks. We also show that a public policy that facilitates consumer learning is welfare enhancing if the information acquisition cost is low enough to ensure that the consumer indeed incurs it. For higher cost levels, however, a cost-reducing policy can

²¹Of course, the high-demand consumer’s contract is not distorted at all left of c^* , so this is indeed “overall” or on average.

hurt welfare (using only the Pareto criterion) or otherwise it can benefit one party and hurt the other. Paradoxically, a cost reduction that is “successful,” in the sense that it ensures that the consumer starts to gather information, can in fact be a failure in that it makes both the consumer and the insurer worse off.

A couple of caveats are in order. First, in our analysis we did not explicitly incorporate a direct cost of facilitating consumer learning (i.e., the costs of providing public websites with health information or subsidizing testing for genetic disorders). Any such costs must be added to the indirect costs that we identified in our analysis. Similarly, we have not allowed for any direct benefits of learning one’s health risk, such as the possibility of taking some kind of preventive action (e.g., a woman having a mastectomy if learning that she has a high risk of developing breast cancer). For applications where such possibilities exist, the benefits of facilitating consumer learning would obviously be larger than what our analysis here suggests (still, the costs of facilitating learning that our analysis has identified would still be there and they would have to be traded off against the benefits). However, our modeling assumption is the appropriate one for disorders such as Huntington’s disease, where there is no known treatment.

In our analysis, we made some further modeling choices that are not obvious. One of these is the assumption that the information acquisition cost enters the consumer’s payoff additively. This assumption has been made previously by Doherty and Thistle (1996); we also offer a couple of possible justifications for our assumption in footnote 9. Still, it is natural to wonder how our results would be affected if one relaxed the assumption. In the supplementary material to this paper (available on our websites), we investigate the implications of an alternative setting where the cost c is deducted from A ’s wealth (and thus shows up in the argument of the utility function u). We can show that with CARA utility, which ensures that wealth effects do not affect the degree of risk aversion, most of our qualitative results reported in the present paper go through.²² When wealth effects affect risk aversion (i.e., if the utility function does not satisfy CARA), the algebra becomes intractable. However, at least for sufficiently small wealth effects, we expect most of our results to be robust with respect to the assumption that c enters as an additive term instead of as a deduction from A ’s wealth.

Another specific assumption that we made was the one saying that acquiring information is a binary decision. We do not think of this assumption as merely simplifying. We believe that, for many real world situations, the information gathering decision *is*

²²The main exception to this concerns the welfare results for the consumer when $c < c^*$. But that is the case where we, in Proposition 6, could report results that are in line with naive intuition. Hence, if the alternative assumption were to alter those results, then the new results would reinforce our point that facilitating consumer learning is not always welfare enhancing.

essentially binary—a natural example of this is a consumer’s decision whether to take a test that can detect a genetic disorder. However, one could also think of situations where consumers decide how much effort to exert when gathering information. More effort implies better, i.e. less noisy, information. In a companion paper (Lagerlöf and Schottmüller, 2013), we analyze such a model with a continuous effort choice. Several results differ in that setting.²³ First, with a continuous effort choice all insurance policies, for all parameter values, involve less than full coverage, which in particular means that there is always distortion at the top. Second, a marginal reduction in the information gathering costs can, simultaneously, strictly reduce profits and consumer surplus, also if information is acquired at the outset. The latter result is established in numerical examples only.

A broad lesson coming out of both this paper and the one with a continuum of effort levels is that a public policy that facilitates consumer learning in insurance markets can affect welfare not only directly through the consumer’s cost savings, but also via the market’s optimal response. More surprisingly, both papers show that a lower information gathering cost can lead to smaller consumer surplus and lower welfare. Hence, we offer a word of caution concerning policies that provide information to consumers or reduce the cost for them of taking, for example, genetic tests.

²³The reason why the results differ is that the logic of the continuous-effort model is quite different from the logic of the model here. In the former model there are, because of the many effort levels, a larger number of information gathering constraints, several of which are binding at the optimum. In fact, in that model it is in the interest of the principal that the agent is badly informed. This is due to something we call the *sorting effect*, which does not show up when the information gathering decision is binary.

Appendix A: Inducing information gathering

Lemma A1. *Suppose that IR-low and IC-high are satisfied. Then IR-high is satisfied with a strict inequality.*

Proof. First note that by $\underline{p} \geq 0$,²⁴ IR-low implies the following inequality:

$$(1 - \bar{\theta}) \underline{u}_N + \bar{\theta} \underline{u}_A \geq \bar{U}^*. \quad (29)$$

To see this, rewrite IR-low as follows:

$$\begin{aligned} (1 - \underline{\theta}) u(w - \underline{p}) + \underline{\theta} u(w - d + \underline{a}) &\geq (1 - \underline{\theta}) u(w) + \underline{\theta} u(w - d) \Leftrightarrow \\ \underline{\theta} [u(w) - u(w - \underline{p}) + u(w - d + \underline{a}) - u(w - d)] &\geq u(w) - u(w - \underline{p}). \end{aligned}$$

Given that $\underline{p} \geq 0$, the right-hand side of the last inequality is (weakly) positive; hence, so is the expression in square brackets on the left-hand side. This means that the inequality must still hold if we substitute $\bar{\theta}$ for $\underline{\theta}$:

$$\bar{\theta} [u(w) - u(w - \underline{p}) + u(w - d + \underline{a}) - u(w - d)] \geq u(w) - u(w - \underline{p}),$$

which is equivalent to (29). The claim in the lemma now follows immediately from IC-high and the inequality in (29). \square

Proof of Lemma 2. Rewrite the left-hand side of IR-ante as follows:

$$\begin{aligned} &v [(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A] + (1 - v) [(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A] - c \\ &\geq v \underline{U}^* + (1 - v) [(1 - \bar{\theta}) \underline{u}_N + \bar{\theta} \underline{u}_A] \\ &\geq v \underline{U}^* + (1 - v) \bar{U}^*. \end{aligned}$$

Here the first inequality uses IR-low and IG-low, and the second inequality uses (29). Since the last line equals the right-hand side of IR-ante, the claim follows. \square

Proof of Lemma 5. We first state the remaining one of the four first-order conditions:

$$\frac{\partial \mathcal{L}}{\partial \bar{u}_A} = 0 \Leftrightarrow (1 - v) \bar{\theta} h'(\bar{u}_A) = \underline{\mu} (1 - v) \bar{\theta} - \bar{\mu} v \underline{\theta}, \quad (30)$$

We now obtain the following result from the first-order conditions (14) and (30): *If IG-high is lax (binding), the high-demand type is fully insured (overinsured).* To prove this claim, multiply (14) by $\bar{\theta}$ and multiply (30) by $(1 - \bar{\theta})$:

$$\begin{aligned} (1 - v) \bar{\theta} (1 - \bar{\theta}) h'(\bar{u}_N) &= \underline{\mu} (1 - v) \bar{\theta} (1 - \bar{\theta}) - \bar{\mu} v \bar{\theta} (1 - \underline{\theta}), \\ (1 - v) \bar{\theta} (1 - \bar{\theta}) h'(\bar{u}_A) &= \underline{\mu} (1 - v) \bar{\theta} (1 - \bar{\theta}) - \bar{\mu} v \underline{\theta} (1 - \bar{\theta}). \end{aligned}$$

²⁴If $\underline{p} < 0$, P would be better off not offering the contract.

Subtracting the second one of those equalities from the first one and then simplifying yield

$$(1 - v) \bar{\theta} (1 - \bar{\theta}) [h'(\bar{u}_N) - h'(\bar{u}_A)] = -\bar{\mu}v (\bar{\theta} - \underline{\theta}).$$

The claim now follows from the strict convexity of h .

Now suppose, per contra, that IG-high is binding in the kind of situation described in the lemma. Then, by the result of the previous paragraph, $\bar{u}_A > \bar{u}_N$. We will now show that P can earn higher profits by *not* inducing information gathering ($x = 0$) and offering a pooled full-coverage insurance policy (instead of inducing $x = 1$ and offering the menu $\{(\underline{u}_N, \underline{u}_A), (\bar{u}_N, \bar{u}_A)\}$)—thus contradicting the assumption that it is optimal to induce information gathering. Let the ex ante expected utility level of this pooled full-coverage policy be denoted by u_p , where by construction $u_p = EU_{x=1}$. We can write:

$$\begin{aligned} \pi &= \hat{w} - v [(1 - \underline{\theta}) h(\underline{u}_N) + \underline{\theta} h(\underline{u}_A)] - (1 - v) [(1 - \bar{\theta}) h(\bar{u}_N) + \bar{\theta} h(\bar{u}_A)] \\ &< \hat{w} - h[v(1 - \underline{\theta}) \underline{u}_N + v \underline{\theta} \underline{u}_A + (1 - v)(1 - \bar{\theta}) \bar{u}_N + (1 - v) \bar{\theta} \bar{u}_A] \\ &= \hat{w} - h(u_p + c) < \hat{w} - h(u_p) = \pi_p, \end{aligned}$$

where the first inequality follows from the strict convexity of h and the last one from $c > 0$ and $h' > 0$. It remains to check that A indeed has an incentive to refrain from gathering information and to purchase the pooled full-coverage policy, thus receiving the payoff u_p . To this end, first note that (by construction) A 's ex ante expected utility remains unchanged. Hence, IR-ante is satisfied. Second, we must check that it is not optimal for A to gather information and then purchase the insurance policy only if learning that $\theta = \bar{\theta}$ (and remain uninsured if $\theta = \underline{\theta}$). Because IG-high is binding, we know that $u_p = [v \underline{\theta} + (1 - v) \bar{\theta}] \bar{u}_A + [v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})] \bar{u}_N$. Since $\bar{u}_A > \bar{u}_N$ and $\bar{\theta} > \underline{\theta}$, this implies $u_p < \bar{\theta} \bar{u}_A + (1 - \bar{\theta}) \bar{u}_N$. Therefore,

$$\begin{aligned} u_p &= v[(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A] + (1 - v) [(1 - \bar{\theta}) \bar{u}_N + \bar{\theta} \bar{u}_A] - c \\ &> v[(1 - \underline{\theta}) \underline{u}_N + \underline{\theta} \underline{u}_A] + (1 - v) u_p - c \\ &\geq v \underline{U}^* + (1 - v) u_p - c, \end{aligned}$$

where the last step uses IR-low. But this inequality implies that buying the pooling insurance policy without gathering information leads to a higher utility than gathering information and buying insurance only if learning that the accident risk is high. \square

Proof of Proposition 3. For the case where P interacts with both types, all the claims except for equation (19) follow from Lemmas 1-5 and from the analysis in the text. To derive equation (19), add (17) and (30), using $\bar{\mu} = 0$:

$$(1 - v) \bar{\theta} h'(\bar{u}_A) + v \underline{\theta} h'(\underline{u}_A) = \lambda \underline{\theta}. \quad (31)$$

Notice that the right-hand side of (16) multiplied by $\underline{\theta}$ equals the right-hand side of (31) multiplied by $(1 - \underline{\theta})$. We thus have:

$$\underline{\theta} [(1 - v) (1 - \bar{\theta}) h'(\bar{u}_N) + v (1 - \underline{\theta}) h'(u_N)] = (1 - \underline{\theta}) [(1 - v) \bar{\theta} h'(\bar{u}_A) + v \underline{\theta} h'(u_A)],$$

which rewritten (using $\bar{u}_A = \bar{u}_N = \bar{u}$) yields equation (19).

For the case where P interacts only with the high-demand type, note that the binding IG-low constraint becomes IR-ante. Using the fact that the offered contract has full coverage, IR-ante implies

$$u = \bar{U}^* + \frac{c}{1 - v}. \quad (32)$$

□

Proof of Lemma 6. For the case where P interacts only with the high-demand type, the claim about $\partial\pi^{SB}/\partial c$ is obvious from (21). For the other case, the (general) envelope theorem,²⁵ implies $\partial\pi^{SB}/\partial c = -\underline{\mu}^{SB} < 0$. Next turn to the proof of the limit result. By Proposition 3, \underline{u}_A^{SB} , \underline{u}_N^{SB} , and \bar{u}^{SB} are defined by the binding IR-low and IG-low and by (19). By Proposition 1, the utility levels at the optimum of the Stiglitz model (\underline{u}_A^{SZ} , \underline{u}_N^{SZ} , and \bar{u}^{SZ}) are defined by the binding IR-low and IC-high and by (5). But (5) and (19) are identical, and in the limit where $c \rightarrow 0$, IG-low and IC-high also are identical. □

Appendix B: Not inducing information gathering

Proof of Proposition 4. Almost all of the claims in the proposition are proven in the preceding text. The only thing remaining is to verify that the optimal insurance policy in part c) of the proposition must satisfy equation (27). But this equality is simply the condition that, at the optimum (u_A^{NI}, u_N^{NI}) , the slope of the iso-profit curve equals the slope of IG-high. By using (24) and the equations in footnote 14, we obtain

$$\varphi'_{ante}(u_A^{NI}) \frac{h'(u_A^{NI})}{h'(u_N^{NI})} = \varphi'_{high}(u_A^{NI}) \Leftrightarrow \frac{v\underline{\theta} + (1 - v)\bar{\theta}}{v(1 - \underline{\theta}) + (1 - v)(1 - \bar{\theta})} \frac{h'(u_A)}{h'(u_N)} = \frac{\underline{\theta}}{1 - \underline{\theta}},$$

which simplifies to the equation in the proposition. □

Proof of Lemma 8. The Lagrangian of P 's problem in Section 4.3 can be written as

$$\begin{aligned} \mathcal{L}_{x=0} &= \hat{w} - v [(1 - \underline{\theta}) h(u_N) + \underline{\theta} h(u_A)] - (1 - v) [(1 - \bar{\theta}) h(u_N) + \bar{\theta} h(u_A)] \\ &\quad + \lambda [EU_{x=0} - v\underline{U}^* - (1 - v)\bar{U}^*] \\ &\quad + \bar{\mu} \{EU_{x=0} - v\underline{U}^* - (1 - v) [(1 - \bar{\theta}) u_N + \bar{\theta} u_A] + c\} \\ &\quad + \underline{\mu} \{EU_{x=0} - v [(1 - \underline{\theta}) u_N + \underline{\theta} u_A] - (1 - v)\bar{U}^* + c\}, \end{aligned}$$

²⁵See, for example, Sydsæter and Hammond (1995, p. 678 ff.).

where $\lambda \geq 0$ is the shadow price associated with IR-ante, $\bar{\mu} \geq 0$ is the shadow price associated with IG-high, and $\underline{\mu} \geq 0$ is the shadow price associated with IG-low. By the (general) envelope theorem, we have

$$\frac{\partial \pi_{x=0}^*}{\partial c} = \frac{\partial \mathcal{L}_{x=0}}{\partial c} = \bar{\mu} + \underline{\mu}.$$

We know from Proposition 4 that, for $c < c''$, IG-high is always binding ($\bar{\mu} > 0$) and IG-low is not ($\underline{\mu} = 0$), which means that then $\partial \pi_{x=0}^*/\partial c > 0$. We also know that for $c > c''$, both IG-high and IG-low are lax ($\bar{\mu} = \underline{\mu} = 0$), which means that then $\partial \pi_{x=0}^*/\partial c = 0$. It now only remains to prove part c) of the proposition. But this is to a large extent done in the text after the proposition. The standard result referred to there, that a pooled contract is always dominated by a separating one, is proven, for example, in Stiglitz (1977). We too implicitly showed the result when we solved Benchmark 1 (i.e., the Stiglitz model) and found that P never at the optimum offers the same contract to the two types, although the formulation of P 's problem allows for this possibility. \square

Appendix C: The overall optimum

Proof of Proposition 6, part a. First consider part a) of the proposition and the comparative statics results concerning the ex post utility levels. We know that these utility levels are defined by the two binding constraints (IG-low and IR-low) and by equation (19). Multiplying IR-low by $\bar{\theta}$ and IG-low by $\underline{\theta}$, and then combining, yield:

$$(\bar{\theta} - \underline{\theta}) \underline{u}_N^* = \underline{U}^* \bar{\theta} - \left(\bar{u}^* - \frac{c}{1-v} \right) \underline{\theta}.$$

Multiplying IR-low by $1 - \bar{\theta}$ and IG-low by $1 - \underline{\theta}$, and then combining, yield:

$$(\bar{\theta} - \underline{\theta}) \underline{u}_A^* = (1 - \underline{\theta}) \left(\bar{u}^* - \frac{c}{1-v} \right) - (1 - \bar{\theta}) \underline{U}^*.$$

Equation (19) is:

$$v \underline{\theta} (1 - \underline{\theta}) [h'(\underline{u}_N^*) - h'(\underline{u}_A^*)] = (1 - v) (\bar{\theta} - \underline{\theta}) h'(\bar{u}^*).$$

Now differentiate these three equations with respect to c :

$$\begin{aligned} (\bar{\theta} - \underline{\theta}) \frac{\partial \underline{u}_N^*}{\partial c} &= - \left(\frac{\partial \bar{u}^*}{\partial c} - \frac{1}{1-v} \right) \underline{\theta}, \\ (\bar{\theta} - \underline{\theta}) \frac{\partial \underline{u}_A^*}{\partial c} &= (1 - \underline{\theta}) \left(\frac{\partial \bar{u}^*}{\partial c} - \frac{1}{1-v} \right), \\ v \underline{\theta} (1 - \underline{\theta}) \left[h''(\underline{u}_N^*) \frac{\partial \underline{u}_N^*}{\partial c} - h''(\underline{u}_A^*) \frac{\partial \underline{u}_A^*}{\partial c} \right] &= (1 - v) (\bar{\theta} - \underline{\theta}) h''(\bar{u}^*) \frac{\partial \bar{u}^*}{\partial c}. \end{aligned}$$

Write the three equations immediately above on matrix form:

$$M \begin{bmatrix} \frac{\partial \underline{u}_N^*}{\partial c} \\ \frac{\partial \underline{u}_A^*}{\partial c} \\ \frac{\partial \bar{u}^*}{\partial c} \end{bmatrix} = \begin{bmatrix} \frac{\underline{\theta}}{1-v} \\ -\frac{1-\underline{\theta}}{1-v} \\ 0 \end{bmatrix},$$

where

$$M \equiv \begin{bmatrix} \bar{\theta} - \underline{\theta} & 0 & \underline{\theta} \\ 0 & \bar{\theta} - \underline{\theta} & -(1 - \underline{\theta}) \\ v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_N^*) & -v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_A^*) & -(1 - v)(\bar{\theta} - \underline{\theta}) h''(\bar{u}^*) \end{bmatrix}.$$

The determinant of M is:

$$\begin{aligned} \det(M) &= -(\bar{\theta} - \underline{\theta}) [(\bar{\theta} - \underline{\theta})(1 - v)(\bar{\theta} - \underline{\theta}) h''(\bar{u}^*) + v\underline{\theta}(1 - \underline{\theta})^2 h''(\underline{u}_A^*)] \\ &\quad - v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_N^*) (\bar{\theta} - \underline{\theta}) \underline{\theta} \\ &= -(\bar{\theta} - \underline{\theta}) [(1 - v)(\bar{\theta} - \underline{\theta})^2 h''(\bar{u}^*) + v\underline{\theta}(1 - \underline{\theta})^2 h''(\underline{u}_A^*) + v\underline{\theta}^2(1 - \underline{\theta}) h''(\underline{u}_N^*)] \\ &\equiv D < 0. \end{aligned}$$

Cramer's rule thus gives us

$$\begin{aligned} D \frac{\partial \underline{u}_N^*}{\partial c} &= -\frac{\underline{\theta}}{1-v} [(\bar{\theta} - \underline{\theta})^2(1 - v) h''(\bar{u}^*) + v\underline{\theta}(1 - \underline{\theta})^2 h''(\underline{u}_A^*)] \\ &\quad + \frac{1 - \underline{\theta}}{1 - v} v\underline{\theta}^2(1 - \underline{\theta}) h''(\underline{u}_A^*) \\ &= -\frac{1}{1 - v} [(1 - v)\underline{\theta}(\bar{\theta} - \underline{\theta})^2 h''(\bar{u}^*) + v\underline{\theta}^2(1 - \underline{\theta})^2 h''(\underline{u}_A^*) - v\underline{\theta}^2(1 - \underline{\theta})^2 h''(\underline{u}_A^*)] \\ &= -\frac{1}{1 - v} [(1 - v)\underline{\theta}(\bar{\theta} - \underline{\theta})^2 h''(\bar{u}^*)] \Leftrightarrow \end{aligned}$$

$$\frac{\partial \underline{u}_N^*}{\partial c} = \frac{\underline{\theta}(\bar{\theta} - \underline{\theta}) h''(\bar{u}^*)}{(1 - v)(\bar{\theta} - \underline{\theta})^2 h''(\bar{u}^*) + v\underline{\theta}(1 - \underline{\theta})^2 h''(\underline{u}_A^*) + v\underline{\theta}^2(1 - \underline{\theta}) h''(\underline{u}_N^*)} > 0.$$

And

$$\begin{aligned} D \frac{\partial \underline{u}_A^*}{\partial c} &= -v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_N^*) \left[\frac{(1 - \underline{\theta})\underline{\theta}}{1 - v} - \frac{\underline{\theta}(1 - \underline{\theta})}{1 - v} \right] - (1 - v)(\bar{\theta} - \underline{\theta}) h''(\bar{u}^*) \times \\ &\quad \times \left[-\frac{(\bar{\theta} - \underline{\theta})(1 - \underline{\theta})}{1 - v} \right] = (\bar{\theta} - \underline{\theta})^2(1 - \underline{\theta}) h''(\bar{u}^*) \Leftrightarrow \end{aligned}$$

$$\frac{\partial \underline{u}_A^*}{\partial c} = -\frac{(\bar{\theta} - \underline{\theta})(1 - \underline{\theta}) h''(\bar{u}^*)}{(1 - v)(\bar{\theta} - \underline{\theta})^2 h''(\bar{u}^*) + v\underline{\theta}(1 - \underline{\theta})^2 h''(\underline{u}_A^*) + v\underline{\theta}^2(1 - \underline{\theta}) h''(\underline{u}_N^*)} < 0.$$

And

$$\begin{aligned} D \frac{\partial \bar{u}^*}{\partial c} &= v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_N^*) \left[-(\bar{\theta} - \underline{\theta}) \frac{\underline{\theta}}{1 - v} \right] + v\underline{\theta}(1 - \underline{\theta}) h''(\underline{u}_A^*) \left[-(\bar{\theta} - \underline{\theta}) \frac{1 - \underline{\theta}}{1 - v} \right] \\ &= -\frac{v\underline{\theta}(1 - \underline{\theta})(\bar{\theta} - \underline{\theta})}{1 - v} [\underline{\theta} h''(\underline{u}_N^*) + (1 - \underline{\theta}) h''(\underline{u}_A^*)] \Leftrightarrow \end{aligned}$$

$$\frac{\partial \bar{u}^*}{\partial c} = \frac{v\underline{\theta}(1-\underline{\theta})[\underline{\theta}h''(\underline{u}_N^*) + (1-\underline{\theta})h''(\underline{u}_A^*)]}{(1-v)\left[(1-v)(\bar{\theta}-\underline{\theta})^2 h''(\bar{u}^*) + v\underline{\theta}(1-\underline{\theta})^2 h''(\underline{u}_A^*) + v\underline{\theta}^2(1-\underline{\theta})h''(\underline{u}_N^*)\right]} > 0.$$

The claim about P 's ex ante expected profit follows from Lemma 6 and Proposition 5.

Let us finally prove the claim about A 's ex ante expected utility. First, if P optimally interacts with only the high-demand type, then A does not get any rent and, thus, her ex ante expected utility is independent of c . Next suppose P optimally interacts with both types. We then know that IR-low binds: $(1-\underline{\theta})\underline{u}_N^* + \underline{\theta}\underline{u}_A^* = \underline{U}^*$. Differentiating this identity yields

$$(1-\underline{\theta})\frac{\partial \underline{u}_N^*}{\partial c} + \underline{\theta}\frac{\partial \underline{u}_A^*}{\partial c} = 0 \Leftrightarrow \frac{\partial \underline{u}_N^*}{\partial c} = -\frac{\underline{\theta}}{1-\underline{\theta}}\frac{\partial \underline{u}_A^*}{\partial c}. \quad (33)$$

A 's ex ante expected utility can be written as

$$\begin{aligned} EU^* &= v[(1-\underline{\theta})\underline{u}_N^* + \underline{\theta}\underline{u}_A^*] + (1-v)\bar{u}^* - c \\ &= v\underline{U}^* + (1-v)[(1-\bar{\theta})\underline{u}_N^* + \bar{\theta}\underline{u}_A^*], \end{aligned} \quad (34)$$

where the second line makes use of the binding IR-low (see above) and the binding IG-low ($(1-v)\bar{u}^* - c = (1-v)[(1-\bar{\theta})\underline{u}_N^* + \bar{\theta}\underline{u}_A^*]$). Now differentiate the expression for A 's ex ante expected utility in (34) with respect to c :

$$\begin{aligned} \frac{\partial EU^*}{\partial c} &= (1-v)\left[(1-\bar{\theta})\frac{\partial \underline{u}_N^*}{\partial c} + \bar{\theta}\frac{\partial \underline{u}_A^*}{\partial c}\right] \\ &= (1-v)\left[(1-\bar{\theta})\left(-\frac{\underline{\theta}}{1-\underline{\theta}}\frac{\partial \underline{u}_A^*}{\partial c}\right) + \bar{\theta}\frac{\partial \underline{u}_A^*}{\partial c}\right] \\ &= (1-v)\frac{\partial \underline{u}_A^*}{\partial c}\left[\frac{\bar{\theta}-\underline{\theta}}{1-\underline{\theta}}\right], \end{aligned}$$

where the second equality makes use of (33). The last expression is strictly negative as $\partial \underline{u}_A^*/\partial c < 0$. \square

Proof of Proposition 6, part b. Here the ex post utility levels are defined by the binding IG-high and equality (27):

$$v(1-\underline{\theta})u_N^* + v\underline{\theta}u_A^* = v\underline{U}^* - c, \quad (35)$$

$$\Phi h'(u_A^*) = h'(u_N^*),$$

where

$$\Phi \equiv \frac{v\underline{\theta}(1-\underline{\theta}) + (1-v)\bar{\theta}(1-\underline{\theta})}{v\underline{\theta}(1-\underline{\theta}) + (1-v)\underline{\theta}(1-\bar{\theta})} > 1.$$

Differentiate the two identities with respect to c :

$$v(1-\underline{\theta})\frac{\partial u_N^*}{\partial c} + v\underline{\theta}\frac{\partial u_A^*}{\partial c} = -1,$$

$$\Phi h''(u_A^*) \frac{\partial u_A^*}{\partial c} = h''(u_N^*) \frac{\partial u_N^*}{\partial c}.$$

Write these two equations on matrix form:

$$\begin{bmatrix} v(1-\underline{\theta}) & v\underline{\theta} \\ h''(u_N^*) & -\Phi h''(u_A^*) \end{bmatrix} \begin{bmatrix} \frac{\partial u_N^*}{\partial c} \\ \frac{\partial u_A^*}{\partial c} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

Using Cramer's rule we have

$$\frac{\partial u_N^*}{\partial c} = -\frac{\Phi h''(u_A^*)}{v(1-\underline{\theta})\Phi h''(u_A^*) + v\underline{\theta}h''(u_N^*)} < 0, \quad (36)$$

$$\frac{\partial u_A^*}{\partial c} = -\frac{h''(u_N^*)}{v(1-\underline{\theta})\Phi h''(u_A^*) + v\underline{\theta}h''(u_N^*)} < 0. \quad (37)$$

Thus, the extent of insurance coverage is strictly decreasing in c if and only if

$$\begin{aligned} \frac{\partial u_N^*}{\partial c} - \frac{\partial u_A^*}{\partial c} &= \frac{-\Phi h''(u_A^*) + h''(u_N^*)}{v(1-\underline{\theta})\Phi h''(u_A^*) + v\underline{\theta}h''(u_N^*)} < 0 \Leftrightarrow \Phi > \frac{h''(u_N^*)}{h''(u_A^*)} \Leftrightarrow \\ &\frac{h'(u_N^*)}{h'(u_A^*)} > \frac{h''(u_N^*)}{h''(u_A^*)} \Leftrightarrow 1 > \frac{h''(u_N^*)/h'(u_N^*)}{h''(u_A^*)/h'(u_A^*)}. \end{aligned} \quad (38)$$

We know that $u_N^* > u_A^*$. Therefore, a sufficient condition for (38) to hold is that $h''(z)/h'(z)$ is weakly decreasing in z . We have

$$\frac{\partial (h''(z)/h'(z))}{\partial z} \leq 0 \Leftrightarrow h'''(z)h'(z) \leq [h''(z)]^2.$$

It is straightforward (but a bit cumbersome) to verify that $h' = 1/u'$,

$$h'' = -\frac{u''}{(u')^3},$$

$$h''' = \left(\frac{3u''}{u'} - \frac{u'''}{u''} \right) \frac{u''}{(u')^4}.$$

Therefore, a sufficient condition for (38) to hold is that

$$\begin{aligned} h'''(z)h'(z) \leq [h''(z)]^2 &\Leftrightarrow \left(\frac{3u''}{u'} - \frac{u'''}{u''} \right) \frac{u''}{(u')^5} \leq \frac{(u'')^2}{(u')^6} \\ &\Leftrightarrow \frac{3u''}{u'} - \frac{u'''}{u''} \geq \frac{u''}{u'} \Leftrightarrow -\frac{u'''}{u''} \geq -2\frac{u''}{u'}. \end{aligned}$$

Finally, the claim about A 's ex ante expected utility follows immediately from the result above in (36) and (37), and the claim about P 's ex ante expected profits follows from Lemma 8a and Proposition 5. \square

Proof of Proposition 6, part c. Here the ex post utility levels are defined by the binding IG-high in (35) and the binding IR-ante, or by equation (23) stated in Lemma 7. By inspection, we have $\partial u_N^*/\partial c < 0$ and $\partial u_A^*/\partial c > 0$. The claim about P 's ex ante

expected profits follows from Lemma 8a and Proposition 5. Finally, the claim about A 's ex ante expected utility follows from the fact that IR-ante is binding. \square

Proof of Proposition 6, part d. Here the single ex post utility level, u^* , is defined by the binding IR-ante and involves full insurance. Therefore it is independent of c and we thus have $\partial u^*/\partial c = 0$. The claim about P 's ex ante expected profits follows from Lemma 8b and Proposition 5. Finally, the claim about A 's ex ante expected utility follows from the fact that IR-ante is binding. \square

References

- Baron, D. P. and R. B. Myerson (1982). Regulating a Monopolist with Unknown Costs. *Econometrica* 50(4), 911–930.
- Barzel, Y. (1977). Some Fallacies in the Interpretation of Information Costs. *Journal of Law and Economics* 20(2), 291–307.
- Chade, H. and E. Schlee (2012). Optimal Insurance with Adverse Selection. *Theoretical Economics* 7(3), 571–607.
- Craswell, R. (1988). Precontractual Investigation as an Optimal Precaution Problem. *The Journal of Legal Studies* 17(2), 401–436.
- Crémer, J. and F. Khalil (1992). Gathering Information Before Signing a Contract. *The American Economic Review* 82(3), 566–578.
- Crémer, J. and F. Khalil (1994). Gathering Information Before the Contract is Offered: The Case with Two States of Nature. *European Economic Review* 38(3–4), 675–682.
- Crémer, J., F. Khalil, and J.-C. Rochet (1998a). Contracts and Productive Information Gathering. *Games and Economic Behavior* 25(2), 174–193.
- Crémer, J., F. Khalil, and J.-C. Rochet (1998b). Strategic Information Gathering Before a Contract is Offered. *Journal of Economic Theory* 81(1), 163–200.
- Dafny, L. S. (2010). Are Health Insurance Markets Competitive? *American Economic Review* 100(4), 1399–1431.
- Demski, J. S. and D. E. M. Sappington (1987). Delegated Expertise. *Journal of Accounting Research* 25(1), pp. 68–89.
- Dionne, G., N. Fombaron, and N. Doherty (2012, July). Adverse Selection in Insurance Contracting. Working Paper CIRRELT-2012-47. Forthcoming in the Handbook of Insurance, 2nd Edition.
- Doherty, N. A. and P. D. Thistle (1996). Adverse Selection with Endogenous Information in Insurance Markets. *Journal of Public Economics* 63(1), 83–102.
- Hoy, M. and M. Polborn (2000). The Value of Genetic Information in the Life Insurance Market. *Journal of Public Economics* 78(3), 235–252.
- Ikram, M. A., S. Seshadri, J. C. Bis, M. Fornage, A. L. DeStefano, Y. S. Aulchenko, S. Debette, T. Lumley, A. R. Folsom, E. G. van den Herik, et al. (2009). Genomewide Association Studies of Stroke. *New England Journal of Medicine* 360(17), 1718–1728.

- Iossa, E. and D. Martimort (2013, July). Hidden Action or Hidden Information? How Information Gathering Shapes Contract Design. Working Paper No. 9552, Centre for Economic Policy Analysis.
- Lagerlöf, J. N. M. and C. Schottmüller (2013, November). Monopoly Insurance with Endogenous Information. Mimeo.
- Lee, T. K. (1982). Resource Information Policy and Federal Resource Leasing. *The Bell Journal of Economics* 13(2), 561–568.
- Ligon, J. A. and P. D. Thistle (1996). Information Asymmetries and Informational Incentives in Monopolistic Insurance Markets. *Journal of Risk and Insurance* 63(3), 434–459.
- Rothschild, M. and J. Stiglitz (1976). Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information. *The Quarterly Journal of Economics* 90(4), 629–649.
- Stiglitz, J. E. (1977). Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market. *The Review of Economic Studies* 44(3), 407–430.
- Sydsæter, K. and P. Hammond (1995). *Mathematics for Economic Analysis*. Prentice Hall International.
- Szalay, D. (2009). Contracts with Endogenous Information. *Games and Economic Behavior* 65(2), 586–625.
- Wilson, C. (1977). A Model of Insurance Markets with Incomplete Information. *Journal of Economic Theory* 16(2), 167–207.