COLLUSION IN A ONE-PERIOD INSURANCE MARKET WITH ADVERSE SELECTION

Alegria, Alexander
Willington, Manuel
Collusion in a One-Period Insurance Market
with Adverse Selection

Alexander Alegría* and Manuel Willington†

July, 2011

Abstract
We show how collusive outcomes may occur in equilibrium in a one-period competitive insurance market characterized by adverse selection. We build on the Inderst and Wambach (2001) model and assume that insurance is compulsory and it involves a minimum premium and minimum coverage; these are common features in many health systems. In this setup we show that there is a range of equilibria, from the zero profit one in which low-risks implicitly subsidize high risks, to one where firms obtain profits with both types of consumers. Moreover, we show that rents only partially dissipate if we assume free entry. Along these equilibria, high risks always obtain full insurance while the low risks coverage decreases as the firms’ profits increase.

KEYWORDS: adverse selection, collusion, insurance, capacity constraints.

JEL Numbers: I11, I18, L41

CORRESPONDENT: Manuel Willington, Universidad Alberto Hurtado, Erasmo Escala 1835, Santiago, Chile.
Phone: 56-2-6920265. Fax: 56-2-6920303. Email: mwilling@uahurtado.cl

*Pontificia Universidad Javeriana de Cali. Most of the work was done while studying at Universidad Alberto Hurtado - Georgetown University.

†Universidad Alberto Hurtado - Georgetown University.
1. Introduction

The equilibrium that may exist in a one-period adverse selection insurance model was characterized in the Rothschild and Stiglitz (1976) paper (henceforth R&S). In that work, it was shown that only a separating equilibrium is possible. This occurs when low risks are only partially insured and high risks receive full coverage. This equilibrium may not exist when the proportion of low risks is sufficiently large. After this publication, a myriad of papers followed, with some of them addressing the issues of the (potential) non-existence of separating equilibria (e.g., Dasgupta and Maskin, 1986; Hellwig, 1987, 1988; Asheim and Nilssen, 1996; Inderst and Wambach, 2001) and also the non-existence of pooling equilibria (e.g., Allard et al., 1997; Newhouse, 1996).

Our starting point in this work is the Inderst and Wambach model (I&W from now on). This introduced capacity constraints to the R&S model and showed that the separating equilibrium always exists (under some mild additional assumptions). The intuition leading to this result is straightforward. The separating equilibrium failed to exist in R&S when it could be destabilized by a Pareto superior pooling contract that would be preferred—and bought—by all consumers. However, under capacity constraints, the pooling contract would be bought by only a fraction of consumers and these would be high risk consumers as they have “more to gain” and are therefore ready to face, with some probability, positive search costs if they are rationed.

We add to this model a feature which is common to many health systems: Consumers are obliged to get health insurance and pay some minimum premium and insurers must offer coverage above a minimum level. We show that in this adverse selection setup collusive outcomes may be particularly simple to
sustain: In a one period model homogeneous firms make positive profits and, even as the number of firms is arbitrarily large, aggregate profits are strictly positive.

The results obtained in the paper are somewhat consistent with an accusation brought by the Chilean national antitrust prosecutor against the five largest private health insurers in 2004. They were accused of concertedly decreasing coverage (without decreasing prices), replacing within a period of a year old plans known as “100/80” (the first number corresponds to the coverage of hospitalization costs and the second for ambulatory care expenses) by new plans known as “90/70”. In the trial, despite the fact that the coverage reduction was established, the antitrust tribunal (Tribunal de Defensa de la Libre Competencia) first and then the Supreme Court did not consider the evidence as supportive of anti-competitive practices and therefore pronounced a not guilty verdict.1

Although our contribution is mainly theoretical, we consider the issues studied here relevant for those healthcare systems that transit from traditional social security schemes –with a national single insurance funded with a fixed percentage of workers salaries– to managed competition schemes with competing insurers where the minimum premium, probably as an unintended consequence of mandating a minimum standard of insurance or to finance some scheme of solidarity in other cases, survives the reforms. Compulsory insurance with a minimum premium and (some degree of) freedom to choose the insurer is a common feature of health systems in several countries like Argentina, Chile, Colombia, Netherlands, Slovakia, Switzerland and, to some

1See Agostini et al. (2011) for a detailed analysis of the empirical evidence presented against the insurance companies and Agostini et al. (2010) for a critique to the antitrust tribunal sentence.
extent, Germany (see Londoño, 1996; Larrañaga, 1997; Tapay and Colombo, 2004; and Augurzky, 2006). Unlike these countries, the proposed reform in the US points to compulsory insurance with a minimum level of coverage, but no minimum premium constraints.

We fully characterize the set of equilibria, assuming that the minimum premium constraint is binding for low-risk consumers (i.e., the R&S separating equilibria is ruled out by this constraint). In this setup we characterize the set of subgame-perfect Nash equilibria that range from a competitive one—with zero profits for all firms—to more collusive outcomes with positive profits: As firms’ profits increase, the coverage of low risks decreases (their premiums remain at the lowest possible) and the premium for high risks increases (and they remain fully covered). Moreover, free entry may reduce the scope of equilibria, but rents never dissipate completely.

Three elements combine to obtain our result: The minimum premium constraint, firms’ capacity constraints (that limits insurees’ mobility), and adverse selection.² Given the premium constraint, if a firm were to deviate from a collusive equilibrium, it should offer a contract that attracts both low and high risks. In this context, high risks will have more to gain than low risks by “switching” to the deviating firm (because of the single crossing property of utility functions that characterizes adverse selection models), and, given the cost imposed by the probability of being rationed, only high risks will be willing to try to get insurance from the deviating firm. Naturally, if this is the case no firm will be willing to deviate. Our model clearly illustrates the dangers of

²Our base model considers two additional features that bring more realism to the model: compulsory insurance and a minimum coverage regulation. However, none of our qualitative results depends on these two additional features.
imposing minimum premium constraints in markets characterized by adverse selection and rigidities that may limit insurees’ mobility.

It is not hard to imagine how capacity constraints could help collusion in a repeated game setup, as it would limit the share of the market that the deviating firm is able to capture and, therefore, reduce the incentive to do so. But our result is not in this vein. We show how, in a one-period game, the imposition of a minimum premium constraint may facilitate collusion in the sense that it leads to multiple equilibria and many of them with positive profits for all firms. The relevant benchmark to compare our result is Inderst and Wambach’s setup whose unique equilibrium is the standard R&S’s zero-profit separating equilibrium. To this model we simply add a minimum premium constraint.

In our model there are multiple equilibria and most give positive profits. In this context, referring to those equilibria with positive profits (or to the one where they are maximized) as collusive equilibria is controversial. To be clear, in our model, firms do not act against a short-term interest while hoping to obtain larger profits in the future as is standard in models where collusion is studied in repeated games. In fact, the vast majority of literature studying how certain conditions may facilitate collusion develops in repeated game setups.

In this context, a factor identified as facilitating collusion would let the players sustain an outcome that yields larger profits (typically the monopoly outcome) with a discount factor lower than the one required in the relevant
benchmark.\textsuperscript{3} Related to this line of argument is the idea that collusion necessarily requires that firms act against short-term interests.

There are, however, several important exceptions for which it is standard to talk about collusion despite the fact that players do not act against short-term interests. Moreover, in these setups, the collusive outcome is one out of many Nash equilibria in a one period setup. Two well-known examples are the collusive results that may arise as Nash equilibria in certain auction designs (e.g., the multiple-unit auction with uniform price, Wilson, 1979, or a second-price sealed bid auction, Robinson, 1985) and also the potential role of price-matching guarantees as a collusive device (Hay, 1982, Salop, 1982).\textsuperscript{4} Another example can be found in Dastidar (2001), who shows how convex costs may allow Bertrand duopolists to act as a monopolist in a one-period game.

In recent years several papers have extended the R&S model to incorporate an additional dimension of private information (e.g., Smart, 2000; and Wambach, 2000) and found that profit making contracts can be part of the equilibrium. As pointed out by Snow (2009), for this result to hold it is required that each insurer offers a single contract. In our setup, however, profit making contracts are part of the equilibrium despite the facts that there is a

\textsuperscript{3}A standard example is how cost heterogeneity raises the critical discount factor required to sustain collusion using trigger strategies when oligopolists compete a la Cournot (a very simple presentation of this result can be found in Ivaldi et al., 2003).

\textsuperscript{4}Moorthy and Winter (2006) present a one period game where firms first announce if they offer the guarantee and then they announce their list prices. Beyond illustrating the standard argument on how monopoly prices can be sustained in equilibrium, they show that price matching guarantees can also play the role of signalling that a firm is low priced.
single dimension of private information (the risk of accident) and insurers can offer a menu of contracts.

The rest of the paper is organized as follows. In Section 2 we present our model and briefly illustrate R&S and I&W results. In Section 3 we present our results and conclusions are given in Section 4. All proofs are relegated to the Appendix.

2. The model

Our model is very similar to the one of I&W, with an additional feature: We assume all consumers are forced to buy insurance and to pay a minimum premium that has an associated minimum level of coverage. We now describe the model and introduce our notation.

Consumers and insurance contracts

There are \( N \) consumers in the economy whose expected utility function when they buy no insurance is

\[
p_x u (W - D) + (1 - p_x) u (W),
\]

where \( W \) is the individuals' initial wealth, \( u (\cdot) \) is an strictly increasing and strictly concave function, and the subindex \( x \) denotes the individual’s type \( \{H, L\} \), which determines the probability \( p_x \) that the individual suffers the loss \( D \). We assume \( 0 < p_L < p_H < 1 \). We further assume that each individual has a probability \( \gamma \in (0, 1) \) of being type \( H \).

An insurance contract in this setup is a pair \((\alpha, \beta)\), where \( \alpha \) is the premium insurees pay (in all events) and \( \beta \) is the gross indemnity. Therefore, the expected utility of a type \( x \) insured individual is

\[
U_x (\alpha, \beta) \equiv p_x u (W - D - \alpha + \beta) + (1 - p_x) u (W - \alpha).
\]
We assume that the regulation imposes a minimum premium \( \alpha \) and a minimum coverage \( \beta \) that all contracts must satisfy. In the case of health insurance, this minimum level of coverage \( \beta \) can be thought as a minimum standardized set of treatments that must be covered by the private insurers or as the level of coverage that a public insurer/health care provider gives to its affiliates.

**Firms**

We assume there are \( F \geq 3 \) risk neutral firms that by offering a contract \((\alpha, \beta)\) to a type \( x \) consumer obtain an expected profit of \( \alpha - \beta p_x \). With no loss of generality, we assume that each firm offers a menu of two incentive-compatible contracts denoted by \( \{(\alpha_L, \beta_L), (\alpha_H, \beta_H)\} \). Therefore, these satisfy

\[
p_L u (W - D - \alpha_L + \beta_L) + (1 - p_L) u (W - \alpha_L) \geq p_L u (W - D - \alpha_H + \beta_H) + (1 - p_L) u (W - \alpha_H) \quad (IC)
\]

and

\[
p_H u (W - D - \alpha_H + \beta_H) + (1 - p_H) u (W - \alpha_H) \geq p_H u (W - D - \alpha_L + \beta_L) + (1 - p_H) u (W - \alpha_L) \quad (IC)
\]

For simplicity, we assume all firms are identical, that they have a maximum capacity constraint \( k \) (logically, \( k < N \)), and that no firm is indispensable to serve all the market, therefore \(( F - 1 ) k \geq N \). Moreover, we assume that \( k \geq \left \lfloor \frac{N}{F - 1} \right \rfloor + F - 1 \).

\[ ^5 \] \([z] \) is the largest integer smaller than or equal to \( z \). The last assumption is made just to simplify the description of the symmetric equilibrium we present in Proposition 1. Obviously, for \( F \geq 2 \), \( k \geq \left \lfloor \frac{N}{F - 1} \right \rfloor + F - 1 \) implies \(( F - 1 ) k \geq N \).
Timing of the game

The timing of the game is as follows: At 0, nature reveals to each individual his type and at 1 each firm offers a menu of contracts (without knowing the type of each particular individual). Then at 2, consumers choose the firm and the contract that they will sign. If no firm faces a larger demand than its own capacity, the game ends.

Consumers could be rationed if many of them choose the same firm. We assume that in such a case all individuals face the same risk of being rationed: Let $n^j$ be the number of consumers who go to firm $j$, then the probability of being rationed is $\max\left\{0, \frac{n^j - k}{k}\right\}$. Since insurance is mandatory, all rationed consumers must seek insurance in a different firm at a cost $c > 0$. Potentially, a consumer could be rationed many times.

The precise timing of the game is as follows:

<table>
<thead>
<tr>
<th>Nature chooses ${h, l}$ for each person</th>
<th>Each firm offers a menu of contracts</th>
<th>Individuals choose firm and contract</th>
<th>Rationed consumers seek new contract</th>
<th>↓</th>
<th>↓</th>
<th>↓</th>
<th>↓</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>2'...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Rothschild and Stiglitz separating equilibrium

Figure 2.1 illustrates the separating equilibrium when the proportion of low-risk consumers is not “too large”: The origin illustrates the situation when consumers buy no insurance ($\alpha = \beta = 0$), the straight line $\pi_L$ represents all actuarially fair insurance contracts for low risk individuals (i.e., along this line, firms selling insurance only to type $l$ consumers obtain zero profits; $\alpha_L = (\beta_L - \alpha_L) p_L / (1 - p_L)$), and the line $\pi_H$ is the analogous for type $h$ consumers.
The line $\pi'$ is the relevant one when all consumers (type $h$ and $l$) are pooled in the same contract.

The pair $RS_H$ and $RS_L$ represents the equilibrium contracts for each type: Consumers self-select their respective contracts—that we will denote by $(\alpha^RS_L, \beta^RS_L)$ and $(\alpha^RS_H, \beta^RS_H)$—, firms obtain zero profits, and no firm can deviate and offer an alternative contract and make strictly positive profits.\(^6\)

The non-existence problem is illustrated in Figure 2.2: Given the large

\(^6\)It is important to highlight that this model and the extensions considered below are pure adverse selection models and do not consider any moral hazard problems. An individual could potentially behave differently if he has full or partial coverage, being less precautious the more generous is his insurance coverage. In such a setup partial coverage could be an optimal mechanism to balance the trade-off between providing incentives to reduce the risk of accident and providing insurance. In R&S’s setup and its extensions, partial coverage is a screening device that helps separate different risks.
proportion of low risks (note that \( \pi' \) is now closer to \( \pi_L \)), if all firms offer the contracts represented by \( RS_H \) and \( RS_L \), then one firm could deviate offering a contract such as \( m \). This contract would be attractive to all consumers and produce strictly positive profits (as it is above \( \pi' \)).

Is it then possible to have a pooling equilibrium? No, as if there is an equilibrium contract on the \( \pi' \) line (e.g., \( m' \)), then an alternative contract such as \( m'' \) could be offered, and that contract would be chosen only by low-risks and would therefore produce strictly positive profits.\(^7\)

**Inderst and Wambach solution**

I&W solve the non-existence problem discussed above by assuming that firms have capacity constraints and that rationed consumers must face a search cost to obtain insurance. We shall not discuss their result formally since the proof we will present for our results closely follows I&W’s logic. However, the intuition (in the problematic case depicted in Figure 2.2) is the following: If a firm now offers a contract such as \( m \), then all consumers would certainly prefer such a contract. Therefore, given that firms are capacity constrained, there will be a positive probability of being rationed and facing the search cost. Since high-risk consumers have more to gain by getting the new contract (this is a consequence of the well-known single crossing property that characterizes these models), in equilibrium of the subgame only they would go to the deviating

\(^7\)If we consider a different equilibrium concept (Wilson, 1977) or a different underlying game (Hellwig, 1987), a pooling equilibrium can exist: If in a third stage of the game we allow firms to reject applicants, then if all other firms offer a pooling contract on the \( \pi' \) line, no firm would deviate and offer \( m' \) hoping to receive low risk individuals. This is because all individuals would apply to the deviating firm, anticipating that non-deviating firms that offered the pooling contract would withdraw it as they would obtain negative profits.
firm, which as a result would then make losses since the contract $m$ is below the $\pi_H$ line.\textsuperscript{8}

3. Results

For our minimum premium assumption to have any relevance it must be binding in equilibrium. We will assume that this is the case, therefore

$$\alpha > \alpha_{RS}^{L}. \quad \text{(A1)}$$

\textsuperscript{8}I&W require two additional assumptions to construct this equilibrium: the search cost can not be “too large” (consumers must prefer to pay it rather than remaining uninsured) and it can not be “too low” (this critical value is related to the maximum capacity of the firms). We will adapt this second assumptions to our model and then state it formally in the next section. The first assumption is not required in our model as we assume that insurance is compulsory; individuals can not remain uninsured.
Moreover, we will assume that the minimum premium constraint is never binding for high risks.

In Figure 3.1 we illustrate the assumption and show the continuum of equilibria that exists when we impose this constraint. To characterize the set of equilibria, we define $\alpha_{H}^{TC}(\beta_{L})$ as the premium that makes $H$-types indifferent between the contracts $(\alpha, \beta_{L})$ and $(\alpha_{H}^{TC}(\beta_{L}), D)$, that is

$$\alpha_{H}^{TC}(\beta_{L}) = \{\alpha : U_{H}(\alpha, D) = U_{H}(\alpha, \beta_{L})\}.$$  

The set of equilibria goes from the pair $(\alpha, \beta_{L}^{Z})$ and $(\alpha_{H}^{TC}(\beta_{L}^{Z}), D)$—represented in Figure 3.1 by the points $(Z_{L}, Z_{H})$—to the pair of contracts $(\alpha, \beta_{L}^{A})$ and $(\alpha_{H}^{TC}(\beta_{L}^{A}), D)$—represented by the points $(A_{L}, A_{H})$. The first pair $(Z_{L}, Z_{H})$ is such that firms make zero profits when they get a fraction $\gamma$ of low risks and $1 - \gamma$ of high risks. By definition of $\alpha_{H}^{TC}(\beta_{L})$, the incentive compatibility constraint $(TC)$ is satisfied as an equality, therefore $\beta_{L}^{Z}$ is formally defined as

$$\beta_{L}^{Z} \equiv \{\beta : (1 - \gamma) \left[ \alpha_{H}^{TC}(\beta_{L}) - p_{H} D \right] + \gamma \left[ \alpha - p_{L} \beta_{L} \right] = 0 \}.$$  

At the other extreme, we must distinguish between two scenarios depending on the generosity of the minimum coverage $\beta$. Assume first that $\beta$ is zero or very low (this is the case depicted in Figure 3.1). For the pair of contracts represented by $(A_{L}, A_{H})$, $(TC)$ is satisfied as an equality—by definition of

$9$The existence of $\beta_{L}^{Z}$ is guaranteed: if $\beta_{L} = \alpha/p_{L}$ then profits—associated to the pair of contracts $(\alpha_{H}^{TC}(\beta_{L}), D), (\alpha, \beta_{L})$—are strictly negative (the firm gets zero profits from low risks and negative profits from high risks) and if $\beta_{L} = \alpha/p_{H}$ then profits are strictly positive (the insurer makes positive profits with both types). Therefore, since the expected profit function $(1 - \gamma) \left[ \alpha_{H}^{TC}(\beta_{L}^{Z}) - p_{H} D \right] + \gamma \left[ \alpha - p_{L} \beta_{L}^{Z} \right]$ is continuous and strictly decreasing in $\beta_{L}$, there is a unique $\beta_{L}$ such that expected profits are zero.
\[ \alpha_{\text{IC}}^T (\beta_L^A) - \text{and a firm’s profits in equilibrium are identical to what it could obtain by filling its capacity with high risks that buy the contract represented by } A_H; \text{i.e. } \left( \alpha_{\text{IC}}^T (\beta_L^A), D \right). \]

Formally, \( \beta_L^A \) is defined as

\[
\beta_L^A \equiv \left\{ \beta : \frac{N}{F} \left[ (1 - \gamma) \left( \alpha_{\text{IC}}^T (\beta_L) - p_H D \right) + \gamma \left( \alpha - p_L \beta_L \right) \right] = k \left( \alpha_{\text{IC}}^T (\beta_L) - p_H D \right) \right\}.
\]

The LHS of the above equation represents the firm’s expected profits in a symmetric equilibrium where all firms offer \( (A_L, A_H) \), while the RHS are the profits a firm could get if it fills its capacity \( k \) with \( H \)-types buying the contract \( A_H \).\(^{10}\)

Between these extrema, any pair such that \( (\text{IC}) \) holds as an equality, the high risks receive full insurance, and the low risks pay the minimum premium can be sustained as a subgame-perfect Nash equilibrium of the original game; i.e., any pair \( (\alpha, \beta_L) \) and \( \left( \alpha_{\text{IC}}^T (\beta_L), D \right) \) such that \( \beta_L \in [\beta_L^A, \beta_L^Z] \).

If, on the other hand, the minimum coverage \( \underline{\beta} \) is larger than the \( \beta_L^A \) defined above, then the range of equilibrium contracts is restricted accordingly:

From the pair of contracts that yields zero profit to insurers \(- (\alpha, \beta_L^Z) \) and \( \left( \alpha_{\text{IC}}^T (\beta_L^Z), D \right) - \) to the pair \( (\alpha, \beta) \) and \( \left( \alpha_{\text{IC}}^T (\beta), D \right) \).

To formalize this result we need to make an additional assumption that sets a lower limit to \( c \) given \( k \):

\[
(1 - \rho^M) U_H (\alpha^*, \beta^*) + \rho^M \left[ U_H \left( \alpha_{\text{IC}}^T (\beta_L^A), D \right) - c \right] < U_H \left( \alpha_{\text{IC}}^T (\beta_L^A), D \right).
\]

\( \rho^M \) is defined as the expected rationing probability a consumer would face if all type \( H \) individuals go to same firm and he also chooses to go to that particular

\(^{10}\)Note that \( c \) plays no role on the above definition; \( \beta_L^Z \) is independent of the search cost.

In the Appendix we provide a formal proof of the existence of a \( \beta_L^A \).
Figure 3.1: Continuum of Equilibrium with Minimum Premium Constraint

firm, and $U_H(\alpha^*, \beta^*)$ is the utility level a type $H$ individual could get if he is offered his most preferred contract in the set $\{(\alpha, \beta) : (\alpha = p_h \beta) \land (\alpha \geq \bar{\alpha})\}$; i.e., the most preferred contract with a low-type fair premium that also satisfies the minimum premium constraint. Informally, (A2) requires that no matter how attractive a contract offered by a deviating firm is (as long as it is “reasonable” in the sense that the insurer would not lose money if the contract is signed only by $L$-types), in a continuation equilibrium it will not be the case that all $H$-types want to go to the deviating firm.

The assumption can be rewritten as

$$\left[ U_H(\alpha^*, \beta^*) - U_H\left(\alpha_{HC}^{\alpha_L}(\beta_L^A), D\right) \right] \left(\frac{1}{\rho^M} - 1\right) < c. $$

$^1$Formally, $\rho^M = \sum_{m=0}^{N-1} \Pr(N_H = m) \max\{0, \frac{m-k}{k}\}$, where $\Pr(N_H = m) = \binom{N-1}{m} \gamma^m (1-\gamma)^{N-1-m}$ is the probability that there are exactly $m$ high-risk individuals in the population.
From the above expression it is obvious that the assumption is never satisfied if \( c = 0 \) and, on the contrary, it is always satisfied when, given the maximum capacity \( k \), the number of individuals in the economy \( N \) is large enough (\( \rho^M \) tends to one in such a case).\(^{12}\) We discuss the role of the assumption below.

We can now formally state our result.

**Proposition 1.** Assume that \((A1)\) and \((A2)\) hold. Then, there is a continuum of symmetric subgame-perfect Nash equilibria where no individual is rationed and all firms offer a menu \( \left\{ (\alpha, \beta_L), \left( \alpha^*_{IH}(\beta_L), D \right) \right\} \), where \( \beta_L \in [\max\{\beta^A_L, \beta\}, \beta^Z_L] \).

The formal proof is relegated to the Appendix.

What is the intuition for the proof? Take any of the proposed equilibrium in Proposition 1. A firm could deviate from the prescribed equilibria in three qualitatively different ways (recall \( \alpha \) is a minimum premium so no deviation can lower \( \alpha_L \)). First, it could choose a menu such that both types are worse off, but that would never be optimal because the firm would be left with no clients (recall we assumed that no firm is indispensable).

Second, it could choose a menu such that only type \( H \) consumers are better-off, but then it will have only type \( H \) customers and \( \beta_L \geq \beta^A_L \) guarantees precisely that this is not a profitable deviation.\(^{13}\)

\(^{12}\)Notice also that if \( k = N \) (i.e., the capacity constraint is irrelevant), then \( \rho^M \) is zero and \((A2)\) is never satisfied.

\(^{13}\)This is not obvious if the proposed deviation menu is such that type \( L \) consumers are as well as with the contract offered by other firms. But, since the deviating firm is offering a better contract for \( H \)-type consumers, in any continuation equilibria there must be congestion in the deviating firm, so the \( L \)-types will strictly prefer to choose a different firm.
Finally, it could deviate with a menu such that both types of consumers are better off.\footnote{Note that the minimum premium constraint being binding and the incentive compatibility for $H$-types being satisfied as an equality prevent firms from deviating to a menu that is strictly preferred only by $L$-types.} The intuition why this deviation will not pay is more subtle: By the single crossing property (that our expected utility functions satisfy), any incentive compatible deviation menu will necessarily do more for type $H$ consumers than for the $L$-types. Then, type $H$ customers will be more willing to risk being rationed, and in any continuation equilibrium the number of high risk individuals willing to go in the first place to the deviating firm is such that $L$-type customers prefer to get the original contract with probability one. Therefore, as the deviating firm will attract only type $H$ customers, the deviation will not pay.

Obviously, the requirement that $\beta_L \leq \beta_L^Z$ guarantees that firms prefer to offer the prescribed contract rather than not offering any contract at all.

The assumption $(A2)$ is sufficient to prevent the third type of deviation discussed: If, given a maximum capacity $k$, $c$ is small enough, then both types of insurees would be ready to try to get the contract offered by the deviating firm and, therefore, the deviation would be profitable.\footnote{More formally, assumption $(A2)$ guarantees that for any “reasonable” deviation a firm may take (any contract that makes negative profits even when subscribed only by $L$-types would be unreasonable) not all $H$-types will prefer to go to the deviating firm; only a fraction. Therefore, in equilibrium of the subgame, $H$-types are indifferent between the deviating firm and the equilibrium contract. Since they are also indifferent between the $H$-type and the $L$-type equilibrium contracts, the single crossing property guarantees that $L$-types strictly prefer their equilibrium contract to the one offered by the deviating firm.} In such a case, the only equilibrium would be with zero profits for all firms.\footnote{Assumption $(A2)$ implicitly defines a strictly positive relationship between our parame-}
It is important to highlight that positive profits are not directly related to search costs as long as \((A2)\) is satisfied. It is not the case that higher search costs give firms larger margin to increase prices or reduce coverage (and increase profits) since \(\beta_L^A\) does not depend on \(c\).

Among all the possible equilibriums, profits are maximized with the pair of contracts that offer the lowest possible coverage to \(L\)-types (i.e., \(\beta_L^A\) or \(\beta\)) and the largest possible premium for the \(H\)-types. This equilibrium is therefore payoff dominant. Additionally, if coverage for \(L\)-types is equal to the minimum coverage \(\beta\), it is reasonable to think that this equilibrium could be a focal-point not only because of its larger payoff but also because of the salience of its label.\(^{17}\)

How do these profits compare to a monopolist’s profits? The relevant
ters \(c\) and \(k\). In terms of \(k\), our analysis is restricted to the domain \([N/(F-1), N]\) as we assumed \(N - 1\) firms can serve the whole market and no firm is able to serve it by itself. In this domain, if \((A2)\) is satisfied then our result is valid while, if \(c\) were too low, then the only equilibrium would be with zero profits. The same is true if if there were no relevant capacity constraints \((k \geq N)\).

\(^{17}\)Alternative criterias to select among many Nash equilibriums are the notions of Strong Equilibrium (Aumann, 1959) and risk-dominance (Harsanyi and Selten, 1988). It is trivial to show that none of the Nash equilibriums is a strong one: the unique candidate to be a Strong Equilibrium is the one with largest profits, but a coalition of \(F - 1\) firms could increase its profits by offering a “slightly” more generous menu of contracts that would leave aggregate profits (almost) unaltered and the non-deviating firm with zero profits.

The notion of risk dominance proposed by Harsanyi and Selten has been extended by Peski (2010) to multi-player games with many actions (potentially a continuum of actions). If we analyze the stage game at which firms offer the insurance contracts (taken as given continuation payoffs such that the firms that offer more generous contracts get \(H\)-type individuals), it is straightforward to show that none of the equilibriums is Cardinal Generalized Risk Dominant.
benchmark here is a monopolist who is able to serve the whole market (i.e., it is not capacity constrained) and faces a minimum premium and minimum coverage regulation. Trivially, the monopolist would offer the pair of contracts \( \{(\alpha, \beta), (\alpha_H^T(\beta), D)\} \), so his profits would be equal to aggregate profits in our setup if \( \beta_L^A \leq \beta \) or larger otherwise. Logically, this profit level may well be larger than what a monopolist facing no minimum premium and minimum coverage constraint could obtain, since the participation constraints for both types are modified (potentially weakened).

**Entry analysis**

Our previous proposition characterized the set of equilibria assuming that the number of firms was given. Since in (almost) all the equilibria firms obtain a strictly positive profit, it is worth analyzing the potential effect of entry on the equilibrium set. It turns out that, as the number of firms is larger, the set of equilibrium allocations shrinks, but it never reduces to the zero profit equilibrium contracts \((Z_L, Z_H)\). Corollary 1 characterizes the set of equilibrium contracts when there are infinitely many firms.

**Corollary 1.** Assume \((A1)\) and \((A2)\) hold. As the number of firms tends to infinity, there is a continuum of symmetric subgame-perfect Nash equilibria where no individual is rationed and all firms offer a menu \(\{(\alpha, \beta_L), (\alpha_H^T(\beta_L), D)\}\) such that \(\beta_L \in \max\{\beta_L^{A,\infty}, \beta_L^Z, \beta_L^T\}\), where \(\beta_L^{A,\infty}\) is such that the contract \(\alpha_H^T(\beta_L^{A,\infty}), D\) makes zero profits when subscribed only by high risk individuals; i.e., \(\alpha_H^T(\beta_L^{A,\infty}) = p_H D\).

The formal proof is presented in the Appendix and it is straightforward from Proposition 1. The only difference between Proposition 1 and its corollary is that \(\beta_L^A\) is substituted by \(\beta_L^{A,\infty}\). Recall that \(\beta_L^A\) in Proposition 1, assuming the
minimum coverage is very low, restricts the level of profits firms can make by reducing further and further the coverage level for $L$-types (and increasing the premium for $H$-types), because at some point if $\beta_L$ is too low and $\alpha_H$ too high a firm would prefer to specialize in type $h$ consumers. As the number of firms tends to infinity, the profits for each of them at any symmetric equilibrium obviously approach zero. Therefore, if all firms were offering contracts that yielded positive profits for both high and low risks, then a firm would choose to specialize in high-risk customers, fill its capacity and make larger profits!

The set of equilibria is therefore restricted to menus such that the contract for $H$-types makes non-positive profits. Figure 3.2 illustrates this result: The contract $\left(\alpha, \beta_L^{A,\infty}\right)$ is represented by the point $A_L^\infty$ and $\left(\alpha_H^{TC}(\beta_L), D\right)$ is, naturally, point $RS_H$.

The intuition why aggregate profits are not zero when the number of
firms grows larger is straightforward: If all incumbent firms offer the menu 
\[ \left\{ \left( a_L, \beta_L^{A, \infty} \right), \left( a_H, \beta_L^{A, \infty}, D \right) \right\} \] -assuming \( \beta_L^{A, \infty} \geq \beta \)- new entrants can not undercut incumbents, as any menu that is preferred by \( L \)-types would also be preferred by \( H \)-types and, as discussed in the proof of Proposition 1, would in fact attract only \( H \)-types and, therefore, make negative profits.

Two remarks about this result are in order. First, in a long run analysis the entry of new firms is just one of the factors (arguably the most relevant one) that may change, but there may be others. In particular, some factors that restrict capacity in the short-run (e.g., capital requirements) may not do it in the long run, but other factors may be relevant both in the long and short run (e.g., the network of health care providers). For our result to hold we do not require that the long term capacity of firms remain exactly the same, all we need is that the new capacity limit (say \( k' \)) is such that assumption A2 is satisfied. Of course, if capacity constraints become irrelevant in the long term then our result (and I&W’s existence result as well) will not hold.

In the long run, capacity constraints can be more or less relevant depending on the relationship between the insurer and health care providers. Traditional or indemnity health insurance plans that simply pay physicians or hospitals on a fee-for-service scheme may face lower capacity constraints than, for example, managed care organizations (e.g., PPOs and HMOs) that take a more active role on the provision of health care. Given (1996) and Wholey et al. (1996) show that scale economies for HMO’s are exhausted relatively soon. In a more recent study, for the Israeli health insurance market, Schmuelli and Messika (2010) show that diseconomies of scale are relevant for the largest health insurer. According to Robinson (1999), diseconomies of scale for managed care
plans arise from increasing bureaucratization (e.g., increasing agency problems with managers or among different groups within the firm).\textsuperscript{18}

Second, we are not formally modeling an entry decision process; ours is a comparative static analysis. In a natural extension of our one-period model, we could have every period a set of potential entrants that must decide whether to enter or not before all firms offer their contracts. A subgame-perfect equilibrium for the repeated game would be that all potential entrants do enter every period, all firms offer every period the same menu \( \left\{ (\alpha, \beta_L), (\alpha_{H}^{TC}(\beta_L), D) \right\} \) where \( \beta_L \in \left[ \max \left\{ \beta_{L_{\text{max}}}, \beta \right\}, \beta_{L_{\text{max}}} \right] \), and consumers choose the contract intended for their type and a firm in such a way there is no congestion in equilibrium.\textsuperscript{19}

\textsuperscript{18}Diseconomies of scale may be relevant for “much simpler” insurance businesses as well. Cummins and Zi (1998) find that most of the largest life insurance firms in the US face decreasing returns to scale. Cummins and Rubio-Misas (2006) show that the wave of mergers of life insurance companies in Spain that followed a deregulation of the industry allowed many firms to benefit from economies of scale, but it shows also that the largest companies ended up operating under decreasing returns to scale. For non-life insurance in Malaysia, Yin (2009) finds that increasing returns to scale are followed by constant returns and then by decreasing returns, and there is no single efficient scale as it depends on the mix of products. For the property/casualty insurance industry, Hanweck and Hogan (1996) find evidence of scale economies for small firms but diseconomies for those firms in the upper quartile size class.

\textsuperscript{19}If there are many potential entrants we could end up with more firms than customers... Such an extreme result could be easily corrected if we assume, as in Smart (2000), there is a positive entry cost firms must afford and/or that firms are not risk neutral when the number of customers per firm is too low. In such a case the total number of firms would be finite and, no matter how small is the entry cost (or the degree of risk aversion), the set of equilibrium contracts is characterized by Corollary 1.
Non-mandatory insurance

The model was developed assuming insurance is mandatory, which is a common feature for health insurance in many countries, but may not be relevant for other insurance markets (or countries). We discuss here how our results change if we assume that insurance is not mandatory and, therefore, individuals may choose between no insurance at all and the offered contracts that, we assume, must satisfy the minimum premium and coverage constraints defined by $\alpha$ and $\beta$.

In terms of our model, this assumption can be easily fitted including a participation constraint for $L$-types. Logically, this constraint may reduce the set of contracts that insurers can offer in equilibrium. Let $\beta^{LR}$ be the coverage level such that $L$-types are indifferent between the contract $(\alpha, \beta^{LR})$ and not buying insurance at all, then all equilibriums in which $\beta_L < \beta^{LR}$ must be ruled out.\(^{20}\)

In terms of Proposition 1 we could simply impose $\beta_L \in \left[ \max \left\{ \beta^A_L, \beta, \beta^{LR} \right\}, \beta^Z_L \right]$, but there are two caveats to be considered. First, it could be the case that $\beta^{LR} > \beta^Z_L$ in which case there would be no equilibrium with both types buying insurance.

Second, since we are assuming that insurance is not mandatory, we must assume the search cost $c$ is not “too large” to rule out the following case: If $c$ is such that a rationed consumer prefers to remain uninsured rather than paying $c$, then it could be possible for a firm to deviate offering a contract that, despite being attractive for both types, only $L$-types are willing to visit\(^{20}\)

\(^{20}\)Note that $(\alpha, \beta^{LR})$ being preferred by $L$-types to not having insurance guarantees that all equilibrium contracts for $H$-types, since they must satisfy (TC), will also satisfy their individual participation constraint.
the deviating firm and risk being rationed, since they face a lower cost of remaining uninsured if rationed.

Formally, we need to assume that

\[ U_H \left( a_H^{IC} \left( \max \{ \beta_L^A, \beta_L^{IR} \} \right), D \right) - c > U_H (0, 0); \quad (A3) \]

i.e., a rationed \( H \)-type consumer prefers to visit another firm and pay \( c \) to get his equilibrium contract rather than remaining uninsured.

**Proposition 2.** Assuming that \((A1), (A2), \) and \((A3)\) hold, then:

1. If \( \beta^{LR} \leq \beta_L^Z \), there is a continuum of symmetric subgame-perfect Nash equilibria where no individual is rationed and all firms offer a menu \( \left\{ (\alpha, \beta_L) \left| \alpha_H^{IC} (\beta_L), D \right\} \right\} \), where \( \beta_L \in \left[ \max \{ \beta_L^A, \beta_L^{IR} \}, \beta_L^Z \right] \).

2. If \( \beta^{LR} > \beta_L^Z \), then in equilibrium all firms offer the single insurance contract \((p_H D, D)\) which is chosen only by \( H \)-types. \( L \)-types remain uninsured.

The formal proof is omitted. The first part is almost identical to the proof of Proposition 1.\(^{21}\) As to the second part, it is immediate that \( \beta^{LR} > \beta_L^Z \) implies there is no menu of contracts that is incentive compatible, individually rational for \( L \)-types, and gives non-negative expected profits. Therefore, in equilibrium competing firms will offer the contract that maximizes \( H \)-types expected utility subject to a non-negative expected profits constraint; i.e., a contract that is actuarially fair and gives complete coverage.

\(^{21}\)In that proof we used the fact that a rationed consumer would, in the following round, get the equilibrium contract since he had no choice but to get insurance. In the case of Proposition 2, assumption \((A3)\) guarantees this is also the case.
A public insurer

The role of a public insurer can be easily fitted in our model. In many countries private insurers coexist with a public insurer, and the latter can be considered as the relevant level of minimum coverage that private insurers must provide. This is the case, for example, of health insurance in Chile where there exist a minimum level of coverage that all insurers must provide (defined as a set of diagnoses and related treatments that must be covered), but the coverage provided by the public insurer is broader than this minimum.\textsuperscript{22}

In terms of our model, we only need to reinterpret $\underline{\beta}$ as the coverage offered by the public insurer that, we assume, offers a single insurance contract $(\underline{\alpha}, \underline{\beta})$. If we assume the public insurer has no capacity constraints, the contract $(\underline{\alpha}, \underline{\beta})$ would become the default option that consumers have, and this may also affect the equilibrium set.

Naturally, if $\underline{\beta}$ is “too low”, then nothing relevant changes and the equilibrium set is the one characterized in Proposition 1. If $\underline{\beta}$ is not too low, then the equilibrium contracts are those pairs $\left\{(\underline{\alpha}, \beta_L), \left(\alpha_H, (\beta_L), D\right)\right\}$ that satisfy $\beta_L \in \left[\max \{\beta_L^A, \underline{\beta}\}, \beta_L^Z\right]$ and the following condition:

$$U_H (\alpha_H, D) - c > U_H (\underline{\alpha}, \underline{\beta}). \quad (A3')$$

This condition is analogous to assumption $(A3)$, but in this case the default option is $(\underline{\alpha}, \underline{\beta})$ rather than $(0, 0)$.

\textsuperscript{22}See Fischer and Serra (1996) for a description of the Chilean health insurance system and Beteta and Willington (2009) for an analysis of the role of the recently introduced minimum mandatory health insurance plan (known as Plan Auge in Chile).
4. Conclusion and discussion

We have presented a model that considers several characteristics of many health insurance markets: The adverse selection problem that firms face, mandatory insurance, minimum coverage, minimum premium, and, as a straightforward extension, the presence of a public insurer. We have built on the model of Inderst and Wambach (2001), which by assuming capacity constraints for insurers and search costs for insurees, solves the (potential) problem of non-existence of equilibrium in the R&S model.

We have added to this model a minimum premium constraint and a minimum coverage (with mandatory insurance) and have shown that for most equilibria firms obtain positive profits. Moreover, aggregate profits do not fully dissipate even if the number of firms is arbitrarily large.

The intuition why such outcomes can be sustained (even in a one-period game) is as follows: First of all, because of the minimum premium constraint, firms cannot attempt to attract low risks by reducing their coverage and premium. Price competition is therefore limited. Instead, in order to attract low risks they must increase coverage (and possibly the premium as well), being aware that high risks will be also attracted. This strategy would pay (just as in the R&S model the separating equilibria could be destabilized by a contract that attracts both types) if they could attract a sufficiently large fraction of low risks, but the combination of the capacity constraint and the adverse selection problem determines that only high risks would show up to the deviating firm.

What determines the maximum level of profits that can be sustained in equilibrium? If the minimum coverage required by law is very low, then profits in equilibrium are restricted by the profits a firm could earn by deviating from
the equilibrium offering a contract only for high risks and filling its capacity. Logically, the larger the spare capacity that firms have in equilibrium, the larger the profits from deviating will be, and the smaller the aggregated profits that can be sustained in equilibrium. The presence of a public insurer that charges the minimum premium and/or the imposition of a relevant minimum coverage associated to the minimum premium could restrict further the set of equilibriums by eliminating those with lower coverage.

In our setup –where firms do have a capacity constraint–, it is the imposition of the minimum constraint that makes the collusive outcomes sustainable (and very robust, they are achieved in a one-period model). Of course the imposition of a minimum premium may have other merits that are not addressed in our model (e.g. to finance a solidarity fund), so our result can not be interpreted as a recommendation to eliminate this kind of constraints; but these other merits should be weighted against the negative result we obtained. The case of Chile is worth noting, as in the private health insurance system –that covers mainly high-income and low-risk individuals– each person or family group is supposed to finance his/her own coverage (there is no solidarity involved), but is obliged to pay a minimum premium.
A. Appendix

Proof of the existence and uniqueness of $\beta_L^A$ (see footnote 10). Recall the definition of $\alpha_{H}^{TC}(\beta_L)$ as the $\alpha$ such that the contracts $(\alpha, D)$ and $(\alpha, \beta_L)$ satisfy $TC$ as an equality. Note that if $\beta_L$ is such that $\alpha_{H}^{TC}(\beta_L) = p_H D$, then no firm would want to specialize in $H$-types. That is,

$$\frac{N}{F} \left[ \gamma \left( \alpha_{H}^{TC}(\beta_L) - p_H D \right) + (1 - \gamma) (\alpha - p_L \beta_L) \right] > k \left( \alpha_{H}(\beta_L) - p_H D \right) = 0.$$  

Then, all we need to show is that there is a $\beta_L$ such that

$$\frac{N}{F} \left[ \gamma \left( \alpha_{H}^{TC}(\beta_L) - p_H D \right) + (1 - \gamma) (\alpha - p_L \beta_L) \right] < k \left( \alpha_{H}^{TC}(\beta_L) - p_H D \right), \quad (*)$$

and, by continuity, we will conclude there is a $\beta_L$, say $\beta_L^A$, such that the two sides of $(*)$ are equal.

To show that a $\beta_L$ such that $(*)$ holds exists, we rewrite the inequality as

$$\frac{N}{F} \left[ (1 - \gamma) (\alpha - p_L \beta_L) \right] < \left( k - \frac{N}{F} \gamma \right) \left( \alpha_{H}^{TC}(\beta_L) - p_H D \right)$$

and since

$$\alpha_{H}^{TC}(\beta_L) - p_H D > \alpha - p_H \beta_L$$

it is sufficient to show that there is a $\beta_L$ such that

$$\frac{N}{F} \left( 1 - \gamma \right) (\alpha - p_L \beta_L) < \left( k - \frac{N}{F} \gamma \right) (\alpha - p_H \beta_L);$$

which is satisfied for $\beta_L = 0$ (recall $k > \frac{N}{F}$ was assumed).

To show that there is a unique value of $\beta_L$ such that the two sides of $(*)$ are equal note that the derivative of the difference between the LHS and the RHS of $(*)$ can be written as

$$\frac{N}{F} \left[ \gamma \frac{d\alpha_{H}^{TC}(\beta_L)}{d\beta_L} - (1 - \gamma) p_L \right] - k \frac{d\alpha_{H}^{TC}(\beta_L)}{d\beta_L}, \quad (***)$$
where
\[
\frac{\alpha_H^U (\beta_L)}{d \beta_L} = -p_H u' (W - \alpha - D + \beta_L) = -p_H \lambda (\beta_L)
\]

where \( \lambda (\beta_L) \equiv \frac{u'(W - \alpha - D + \beta_L)}{u'(W - \alpha_H^U (\beta_L))} \) is greater than 1 by definition of \( \alpha_H^U (\beta_L) \) and the strict concavity of \( u(\cdot) \). Since we have assumed that \( k > \frac{N}{F} \), the expression (***) is greater than
\[
\frac{N}{F} [-\gamma p_H \lambda (\beta_L) - (1 - \gamma) p_L] + \frac{N}{F} p_H \lambda (\beta_L) = \frac{N}{F} (1 - \gamma) [p_H \lambda (\beta_L) - p_L] > 0.
\]

**Proof of Proposition 1.** By defining an ordered set of firms \( F = \{1, 2, \ldots, F\} \) and individuals \( N = \{1, 2, \ldots, N\} \), a symmetric equilibrium can be constructed which is characterized by:

1) In equilibrium, all firms offer the same menu of contracts such that the incentive compatibility for high risks is satisfied as an equality and \( \beta_L \in \left[ \max \{\beta_L^A, \beta\}, \beta_L^Z \right] \). We denote the contracts for \( L \)-types and \( H \)-types \((\alpha, \beta_L)\) and \((\alpha_H^U (\beta_L), D)\) respectively.

2) In equilibrium, the first \( \left[ \frac{N}{F} \right] \) individuals visit firm 1, the second \( \left[ \frac{N}{F} \right] \) individuals visit firm 2, etc. Each of the last \( N - F \left[ \frac{N}{F} \right] \) individuals follows a mixed strategy in which each firm is chosen with a probability of \( \frac{1}{F} \).

Proof:

I) Naturally, individuals have no incentives to deviate since all firms offer the same menu of contracts and in the proposed equilibrium there is no rationing. Moreover, high risk individuals are indifferent between their contract and the one designed for low types and therefore can not gain by choosing a contract intended for \( L \)-types with positive probability. The same is true for \( L \)-types as they strictly prefer their contract.

II) Suppose now that a firm deviates. With no loss of generality, the deviating
menu must be incentive compatible (if one of the options is preferred by both types, that would be equivalent to the deviating firm offering a single contract, which is equivalent to offer a menu with two identical options which, of course, is incentive compatible). For simplicity we assume firm 1 deviates:

1. Offering a menu such that high risks are better off and low-risks are worse off could be profitable if the prospect of filling its capacity with high risks were better than the equilibrium payoff. However, the requirement that \( \beta_L \geq \beta_L^A \) guarantees that this is not the case. By definition of \( \beta_L^A \) for \( \beta_L = \beta_L^A \) it holds that

\[
\frac{N}{F} \left[ \gamma \left( \alpha_H^{\overline{T}} (\beta_L) - p_H D \right) + (1 - \gamma) (\alpha - p_L \beta_L) \right] - k \left( \alpha_H^{\overline{T}} (\beta_L) - p_H D \right) = 0;
\]

and, since the LHS of the above equation is increasing in \( \beta_L \) (see the argument made for the proof of the uniqueness of \( \beta_L^A \)), the deviation is not profitable for any \( \beta_L \geq \beta_L^A \); i.e.,

\[
\frac{N}{F} \left[ \gamma \left( \alpha_H^{\overline{T}} (\beta_L) - p_H D \right) + (1 - \gamma) (\alpha - p_L \beta_L) \right] \geq k \left( \alpha_H^{\overline{T}} (\beta_L) - p_H D \right); \quad \forall \beta_L \geq \beta_L^A.
\]

2. Offering a menu such that both types are worse off would give the deviating firm a payoff of zero, since \((F - 1) k \geq N\). (We assume that for this history the first \( \lfloor \frac{N}{F-1} \rfloor \) individuals visit firm 2, the second \( \lfloor \frac{N}{F-1} \rfloor \) individuals visit firm 3, etc.)

3. Consider now a deviation such that both types are better off (and the contract designed for low risks is such that \( \alpha \geq p_L \beta \)) and denote by \((\alpha_L^*, \beta_L^*)\) and \((\alpha_H^*, \beta_H^*)\) the contracts offered by the deviating firm for low
and high types respectively. These contracts therefore satisfy

\[
U_L (\alpha_L', \beta_L') \geq U_L (\alpha, \beta_L) \\
U_H (\alpha_H', \beta_H') \geq U_H \left( \alpha_H' (\beta_L), D \right)
\]

and the incentive compatibility constraints:

\[
U_L (\alpha_L', \beta_L') \geq U_L (\alpha_H', \beta_H') \\
U_H (\alpha_H', \beta_H') \geq U_H (\alpha_L', \beta_L').
\]

We construct the following continuation equilibrium in which only high risks choose to visit the deviating firm with positive probability:

a. All low risks among the first \( \left\lfloor \frac{N}{F-1} \right\rfloor \) visit firm 2, all low risks among the second \( \left\lfloor \frac{N}{F-1} \right\rfloor \) visit firm 3, etc. They all get \( U_L (\alpha, \beta_L) \).

b. All high risks among the first \( \left\lfloor \frac{N}{F-1} \right\rfloor \) visit firm 2 with probability \( \phi \) and firm 1 with probability \( 1 - \phi \), all high risks among the second \( \left\lfloor \frac{N}{F-1} \right\rfloor \) visit firm 3 with probability \( \phi \) and firm 1 with probability \( 1 - \phi \), etc.

c. Note that if individuals behave according to 1. and 2. rationing can occur only in the deviating firm. For those rationed, we specify that they go to their corresponding firms in the next period and get their equilibrium contract.

d. For types \( H \) to play the specified mixed strategy, the probability \( \phi \) must be such that \( H \)-types are indifferent between visiting firm 1 (LHS of the equation below) or getting their equilibrium expected
utility $U_H \left( \alpha_H^{tc} (\beta_L), D \right)$; that is:
\[
\rho(\phi) \left( U_H \left( \alpha_H^{tc} (\beta_L), D \right) - c \right) + (1 - \rho(\phi)) U_H (\alpha'_H, \beta'_H) = U_H \left( \alpha_H^{tc} (\beta_L), D \right); \quad (#)
\]
where $\rho(\phi)$ is the expected rationing probability when all $H$-types mix with probability $\phi$.

The existence of such a $\phi$ (potentially different for different deviations) is guaranteed by assumption A.2: Note that $\rho(1) = \rho^M$ and $U_H (\alpha^*, \beta^*) \geq U_H (\alpha'_H, \beta'_H)$, therefore
\[
\rho(1) (U_H (\alpha_H, D) - c) + (1 - \rho(1)) U_H (\alpha'_H, \beta'_H) < U_H (\alpha_H, D)
\]
and, since $\rho(0) = 0$,\[
\rho(0)(U_H (\alpha_H, D) - c) + (1 - \rho(0)) U_H (\alpha'_H, \beta'_H) > U_H (\alpha_H, D)
\]
by continuity there must exist a $\phi$ such that the equality holds.

e. We need to show that a low risk individual has no incentive to visit firm 1. He would get\[
\rho(\phi) (U_L (\alpha_L, \beta_L) - c) + (1 - \rho(\phi)) U_L (\alpha'_L, \beta'_L),
\]
$(U_L (\alpha'_L, \beta'_L)$ is the utility he gets if accepted in Firm 1, $U_L (\alpha, \beta_L) - c$ is what he gets if rationed, and $\rho(\phi)$ is the probability of being rationed). We need to show that the above expression is not greater than $U_L (\alpha, \beta_L)$, that is
\[
(1 - \rho(\phi)) (U_L (\alpha'_L, \beta'_L) - U_L (\alpha, \beta_L)) \leq \rho(\phi) c.
\]
Using $(#)$ and considering that $U_H \left( \alpha_H^{tc} (\beta_L), D \right) = U_H (\alpha, \beta_L)$ and $U_H (\alpha'_H, \beta'_H) \geq U_H (\alpha'_L, \beta'_L)$, it is sufficient to show that
\[
U_L (\alpha'_L, \beta'_L) - U_L (\alpha, \beta_L) \leq U_H (\alpha'_L, \beta'_L) - U_H (\alpha, \beta_L).
\]
After some straightforward manipulations – and replacing $U_i (\cdot)$ – the above inequality can be written as

$$(p_H - p_L) [u(W - \alpha) - u(W - \alpha'_L)] \geq$$

$$(p_L - p_H) [u(W - D - \alpha'_L + \beta'_L) - u(W - D - \alpha + \beta_L)].$$

Note that the LHS is non-negative as $\alpha'_L \geq \alpha$. On the other hand, given that $U_L (\alpha'_L, \beta'_L) \geq U_L (\alpha, \beta_L)$ and $\alpha'_L \geq \alpha$, it must be true that $-\alpha'_L + \beta'_L \geq -\alpha + \beta_L$ and therefore the RHS is not positive. Therefore, since only high types would be willing to visit the deviating … rm – and $\beta_L \geq \beta^A_L$ holds – the deviating firm would be worse off.

\begin{itemize}
\item \textbf{Proof of Corollary 1.}  By definition $\beta^A_L$ satisfies

$$\alpha^T_H (\beta^A_L) = \frac{(1 - \gamma) (\alpha - p_L \beta^A_L)}{Fk - \gamma} + p_H D.$$ 

Note that by assumption $Fk > N > N\gamma$, therefore $\alpha^T_H (\beta^A_L)$ is greater than $p_H D$ for any $F$, $\alpha^T_H (\beta^A_L)$ is decreasing in $F$, and as $F$ goes to infinity, $\alpha^T_H (\beta^A_L)$ – the largest premium high-risks may pay in equilibrium – goes to $p_H D$. Therefore, a firm that fills its capacity with high risks would make zero profits. \end{itemize}
References


