barcelona gse

graduate school of economics

Multidimensional Screening in a Monopolistic Insurance Market

Pau Olivella

Fred Schroyen

March 2012

Barcelona GSE Working Paper Series

Working Paper nº 619

Multi-dimensional screening in a monopolistic insurance market*

Pau Olivella[†] and Fred Schroyen[‡]

March 31st, 2012

Abstract: In this paper, we consider a population of individuals who differ in two dimensions: their risk type (expected loss) and their risk aversion. We solve for the profit maximizing menu of contracts that a monopolistic insurer wants to put out on the market. We find that it is never optimal to fully separate all the types. Secondly, a sufficiently high risk aversion heterogeneity means that some high risk people (the risk tolerant ones) will get lower coverage than some low risk people (the risk averse ones). Thirdly, we show that when the average man and woman differ only in risk aversion, gender discrimination may lead to a Pareto improvement in the insurance market.

JEL code: D82, G22

Keywords: insurance markets, asymmetric information, screening, gender discrimination, positive correlation test.

Acknowledgements. The paper has benefited from presentations at the 6th European Health Economics Workshop (Liège, 2005), the Public Economic Theory Meeting (Marseille, 2005), the IHEA World Congress (Barcelona, 2005) the 33th EGRIE meeting (Barcelona, 2006), the HEB-HERO Health Economics Workshop (Oslo, 2009) and from seminar presentations at CORE (Louvain-la-Neuve, 2007), HECER (Helsinki, 2008) and Boston University (2009). We are grateful to Catarina Goulão, Eric Nævdal and INSERT EGRIE DISCUSSANT HERE, Eirik Kristiansen and Gaute Torsvik for discussions and comments. Olivella acknowledges support from the Government of Catalonia project 2005SGR00836 and the Ministerio de Educación y Ciencia, project ECO2009-07616 and CONSOLIDER-INGENIO 2010 (CSD2006-0016). Schroyen acknowledges the hospitality of CODE (Universitat Autónoma de Barcelona) and CORE (Université catholique de Louvain), and financial support from Health Economic Bergen through an SNF grant.

[†]Departament d'Economia i d'Historia Economica and CODE, Universitat Autónoma de Barcelona, Edifici B, E-08193 Bellaterra (Spain) and Barcelona GSE. E-mail: pau.olivella@uab.es.

[‡]Department of Economics, Norwegian School of Economics & Business Administration, Helleveien 30, N-5045 Bergen (Norway) and Health Economics Bergen (HEB. E-mail: fred.schroyen@nhh.no.

1 Introduction

Individuals who seek insurance differ from each other in many respects. At least two of these differences are of central importance for insurance companies and for insurance market outcomes: the distribution of losses that insurance takers face, and their willingness to bear the risk of those losses.[1] Empirically, heterogeneity in the second characteristic is not negligible. Aarbu and Schroyen (2011), for example, find that the degree of relative risk aversion among Norwegians averages above four with a standard deviation of about three.

Insurance market theory has primarily focussed on the consequences of private information on the loss distribution, and to a lesser extent on the case in which information on risk aversion is private, but has rarely studied situations in which private information applies to both characteristics.[2] Moreover, analysis of the two-dimensional private-information problem has been restricted to competitive markets; i.e., a setting in which several insurers compete for clients. In this paper, we study the opposite setting by asking how a monopolist would design a contract menu intended to attract agents who hold not only private information on their loss distribution, but also on their risk preferences.

Adding risk aversion heterogeneity to the analysis of insurance markets calls for a multidimensional hidden information model. Such an analysis is technically not straightforward, because the existence of private information in two or more dimensions implies that the ordering of agents according to their willingness to pay for extra coverage becomes endogenous. In other words, the ordering depends on the contract. To see this, consider two contracts: one with very partial coverage and one with almost full coverage. When offered the former contract, a highly risk-averse agent facing a low risk may be more willing to pay for additional coverage than a risk-tolerant agent facing a high risk, while the situation could be the other way around for the latter contract. Technically, the indifference curves of these "intermediate"

[1] A third factor, that will not be discussed here, is the moral stance of insurees, determining the amount of false claims that insurers have to deal with each year.

[2] Rothschild and Stiglitz (1976) analyse a perfectly competitive insurance market with private information on the distribution of losses. Stiglitz (1977, Sections 3 and 4), and Landsberger and Meilijson (1996) analyse a monopolist insurer. Stiglitz (1977, Section 5) and Landsberger and Meilijson (1994) analyse the outcomes under monopoly when private information is held on risk attitude.

insurance takers cross twice, and this invalidates standard solution methods.[3]

There is a scant literature on solutions to multidimensional screening problems. One branch of this literature is methodological and deals with a principal–agent setting, as we do—see, e.g., the "user's guide" by Armstrong and Rochet (1999). It turns out, however, that our insurance problem does not lend itself to being solved by the techniques proposed therein, the main reason being that our problem has two hidden characteristics, but only one instrument—the degree of coverage.[4] A second branch of literature deals with multidimensional screening in insurance markets, but restricts itself to competitive markets. In this literature, it is usually assumed that each insurance company offers a single contract.[5] In a monopolistic setting such as ours, such a restriction would render the analysis trivial and unrealistic. By assuming that the monopolist offers a menu of contracts, the relative

[3] Jullien *et al.* (2007) analyse whether the single crossing property holds in the general monopolistic screening model with moral hazard and in which agents differ in their risk preferences. For more information on the role of this property in a competitive insurance market with the same informational assumptions and moral hazard, see De Donder and Hindriks (2009).

[4] Armstrong and Rochet (1999) study a problem in which the agent has quasilinear and separable preferences on two action levels and a transfer. The principal has similar preferences but she is unsure whether the agent has a high or low valuation for either of the two activities. A contract specifies a transfer and two activity levels. In our problem, there is only one "activity", i.e., insurance coverage. An agent's willingness to pay for coverage depends on both her risk level and risk aversion. On the other hand, the insurer's willingness to offer coverage depends on the level of risk, but not on the agent's risk aversion. Risk aversion only indirectly determines contract profitability through the rents that must be left for incentive compatibility reasons. To sum up, we have a screening problem with two hidden characteristics, one of which is a common value, and with one instrument. This also makes our problem different from those of Armstrong (1999) (who incorporates one instrument and two common value characteristics) and Dana (1993) (two instruments, two common value characteristics).

[5] This *de facto* means that the main results are driven by the lack of order between what we refer to as "intermediate types"; namely, by those whose indifference curves cross twice. This explains why some authors only consider these intermediate types (e.g., Wambach, 2000). Although Smart (2000) and Villeneuve (2003) consider the full set of types, they maintain the assumption that each company offers a single contract per risk class. In some models with multidimensional private information, it is possible to reduce the dimensionality by a so-called type aggregator–see, e.g., De Donder and Hindricks (2007) for a political economy social insurance context. In a context closer to ours, Crocker and Snow (2011) assume that an insurance taker may face different kinds of loss (fire, theft, etc.). Since the probability for each kind of loss is private information, the insurer must engage in multidimensional screening. However, because risk classification based on observables is assumed sufficiently fine, the problem can be reduced to single dimensional private information.

proportion of the non-intermediate types play a role that is as crucial as the non-single crossing of intermediate types' indifference curves. Hence, the problem of the failure of the single crossing condition—brought about by the intermediate types—is compounded in the monopolistic setting by the necessity of dealing with non-intermediate types in the design of the optimal menu of contracts.

Our main objective is to characterize this optimal menu. We establish three results: (i) it is always optimal to pool some of the types (i.e., full separation of types is never optimal); (ii) unlike in the one-dimensional case, exclusion of some high-risk individuals from insurance may be optimal; and (iii) some low-risk individuals may end up with more coverage than some high-risk individuals.

Next, we address two issues that have received much recent attention. The first one is methodological. In testing for the presence of asymmetric information in insurance markets, the question is whether the absence of significant positive correlation between risk and coverage (i.e., the absence of adverse selection) should be taken as indicative of the absence of asymmetric information. Chiappori *et al.* (2006) derive the testable prediction that in a *sufficiently competitive* insurance market with asymmetric information, the observable risk should be related to coverage in a positively *monotonic* way. Notice that this is stronger than requiring a positive *correlation* between coverage and risk. We show when this result goes through in our monopolistic setting, and when it does not. In the latter case, we also show when risk and coverage can be statistically positively correlated, and when they cannot. In this sense, our results corroborate the role of the sufficient competition assumption for the Chiappori *et al.* (2006) result. Our analysis also adds to the list of possible explanations for the lack of evidence supporting the existence of asymmetric information the combination of market power and preference heterogeneity.[6] Other explanations in the (growing) list are: (i) endogenous heterogeneity in risks because of moral hazard (see, e.g., Cutler *et al.*, 2008); (ii) endogenous wealth heterogeneity (Netzer and Scheuer, 2007); and (iii) the insurer having privileged information on risks (Villeneuve, 2000).

The second issue concerns the possible welfare consequences of the ban on the use of gender discrimination in insurance that will take effect from December 2012 in the European Union. This ban extends the principle of

[6] Chiappori *et al.* (2006) propose a local argument for a negative correlation between risk and coverage to arise in the case of monopoly. Our analysis provides instead a full characterization.

equal treatment of women and men in the access to and the supply of goods and services to the insurance industry.[7,8] This will surely affect the insurance sector, because of the common practice of differentiating premia according to gender when underwriting life, health and car accident risks. Regarding life insurance, it has been argued that if one controls for lifestyle, environmental factors, and social class, "the difference in average life expectancy between men and women lies between zero and two years" and therefore that "the practice of insurers to use sex as a determining factor in the evaluation of risk is based on ease of use rather than on real value as a guide to life expectancy." (Commission of the European Communities, 2003: 6) Not surprisingly, European insurer carriers have reacted fiercely to the proposed ban, arguing that removing gender would weaken their ability to assess risk and that gender-neutral calculation would increase the premia for many of their products, especially for women (Financial Times, November 3, 2003, p. 2). We show that even if—as the Commission claims—gender does not provide any information on the underlying risk, if it does provide (imperfect) information on an individual's risk aversion (as empirical research suggests), then allowing the monopolist to condition the terms of the insurance contract on gender may be Pareto improving. We provide sufficient conditions for such an improvement to arise.

From a technical point of view, we have taken a new approach to the analysis of screening insurance takers that simplifies the problem and is appealing from a modelling point of view. Rather than following the standard set-up in which the individual faces the possibility of a single monetary loss, we assume that the loss is normally distributed and that agents differ in their expected losses, which can be high or low. If the insurance indemnity is linear in the loss, as is the case under a reimbursement insurance scheme with a constant co-insurance rate, the final income will also be normally distributed. Endowing agents with a utility function that displays constant absolute risk

[7] Council Directive 2004/113/EC of 13 December 2004 implementing the principle of equal treatment between men and women in the access to and supply of goods and services (Official Journal of the European Union 2004 L 373, p. 37)

[8] The gender directive of 2004 did provide for a derogation that allowed member states to permit gender-specific differences in insurance premiums and benefits in so far as gender is a determining risk factor that can be substantiated by relevant and accurate actuarial and statistical data. In March 2011, however, the European Court of Justice declared this derogating provision in the Directive to be invalid on the grounds that the use of risk factors based on sex in connection with insurance premiums and benefits is incompatible with the principle of equal treatment for men and women under European Union Law (European Court of Justice, 2011).

aversion, which also can be high or low, means that their preferences over uncertain income prospects can be represented as mean–variance preferences. An important consequence of this approach is that preferences over insurance contracts become quasilinear in the insurance premium and therefore in the information rent. Readers familiar with contract theory will acknowledge the usefulness of linearity in the information rent in specifying the incentive compatibility constraints. An additional advantage of mean–variance preferences is that they allow for an explicit characterization of the optimal menu of contracts.

The limitations of our approach follow immediately from these assumptions. We do not consider insurance contracts with either a deductible or a cap because such features would destroy the normality of net income. Second, the normality assumption implies a positive likelihood of negative losses, although this problem may be rendered of secondary importance by considering sufficiently high means and/or low variances for the losses. Perhaps the most important objection is that we have no skewness in the loss distribution, and in particular no strictly positive probability mass for a zero loss. Nevertheless, these are minor limitations when compared with the considerable advantages the approach offers for characterizing the solution to a two-dimensional screening problem. To economize on space, our general characterization is restricted to a non-positive correlation between risk size and risk aversion.

The remainder of the paper is organized as follows. In Section 2, we model the preferences of insurance takers and specify reimbursement contracts. In Section 3, we set up the problem faced by a monopolistic insurer. In Section 4, we characterize the optimal menu of contracts when insurees only differ in risk levels or risk aversion, as well as considering the case of perfect positive correlation. In Section 5, we assume that insurees differ in both respects simultaneously and discuss the five regimes (for contract menus) that may be optimal. For each regime, we characterize the optimal set of co-insurance rates. In Section 6, we determine which regime is dominating for which part of the parameter space. In Section 7, we interpret the testable prediction of Chiappori *et al.* (2006) in the light of our results. In Section 8, we trace out the consequences of allowing the monopolistic insurer to gender discriminate. Section 9 concludes the paper.

Except when otherwise stated, we have relegated all proofs of lemmas, theorems and propositions to our companion paper (Olivella and Schroyen, 2011).

2 Insurance takers and reimbursement contracts

Insurance takers

We assume that individuals are endowed with initial wealth e and a negative exponential von Neumann–Morgenstern utility function defined on final wealth y: $u(y) = -\exp(-ry)$, where $r > 0$ is the (constant) degree of absolute risk aversion. Initial wealth is subject to a random loss z that follows a normal distribution with mean μ and variance σ^2.

Agents have access to reimbursement insurance. A typical reimbursement contract pays out a compensation of $1 - c$ per Euro loss, in return for a premium P. *Ex post*, final wealth is then given by

$$y = e - P - cz, \tag{1}$$

which *ex ante* is also normally distributed. We will express a contract C as a pair of a co-insurance rate c and a premium P: $C = (c, P)$.

It is well known that under the assumptions made, the expected utility of the agent is representable by the certainty equivalent (CE) wealth function $U = \mathrm{E}(y) - \frac{r}{2}\mathrm{var}(y)$. By replacing the mean and variance of final wealth, CE wealth is given by

$$U = e - P - c\mu - \frac{r}{2}c^2\sigma^2. \tag{2}$$

From now on, we write $\nu \overset{\text{def}}{=} r\sigma^2$, and assume that this product can be either high or low, and likewise for the expected loss: $\mu \in \{\mu_L, \mu_H\}$ and $\nu \in \{\nu_L, \nu_H\}$,where $\mu_L < \mu_H$ and $\nu_L < \nu_H$. The model can thus be interpreted in two ways: either individuals are equally risk averse but their losses have different variances, or the loss variance is identical but individuals have different degrees of risk aversion. Throughout, we adhere to the second interpretation and will refer to ν as risk aversion.

A person with characteristics (μ_i, ν_j) is said to be of type ij. The share of ij individuals in the population is given by α_{ij} $(i, j = H, L, \sum_{i,j}\alpha_{ij} = 1)$. We denote by $\alpha_{k\cdot}$ the fraction of individuals with expected loss μ_k $(\alpha_{k\cdot} = \alpha_{kL} + \alpha_{kH})$; likewise, $\alpha_{\cdot k}$ is the fraction of individuals with risk aversion ν_k $(\alpha_{\cdot k} = \alpha_{Lk} + \alpha_{Hk})$.

Incentive compatible contracts

When a person of type ij ($i, j \in \{H, L\}$) signs the contract $C = (c, P)$, her CE wealth, is

$$U^{ij}(c, P) \stackrel{\text{def}}{=} e - P - c\mu_i - \frac{1}{2}c^2\nu_j. \tag{3}$$

If instead she decides to remain uninsured, her CE wealth becomes $e-\mu_i-\frac{1}{2}\nu_j$, which is of course equivalent to accepting the contract $(c, P) = (1, 0)$, under which the agent bears the full loss but pays no premium. The CE rent that the agent enjoys from contract (c, P) is then

$$R_{ij}(c, P) \stackrel{\text{def}}{=} U^{ij}(c, P) - U^{ij}(1, 0) = -P + (1 - c)\mu_i + \frac{1}{2}[1 - c^2]\nu_j. \tag{4}$$

Hence, the rent decreases with the co-insurance rate both via the expected loss and via risk aversion (if $c > 0$).

The marginal willingness to pay for a slightly lower co-insurance rate c is

$$MWP^{ij}(c) \stackrel{\text{def}}{=} -\frac{\mathrm{d}P}{\mathrm{d}c}|_{\mathrm{d}U^{ij}=0} = \mu_i + c\nu_j, \tag{5}$$

which increases linearly in c.

Indifference curves in the contract space (c, P) are thus concave in c, and downward sloping for non-negative co-insurance rates. In addition, individuals with higher expected losses and/or greater risk aversion have a higher marginal willingness to pay. Figure 1 illustrates the indifference curve that passes through the no-insurance point $N = (1, 0)$. Given that the slope of the indifference curve when it passes the P-axis is μ, it is easy to decompose the total willingness to pay for full insurance into the expected loss and the risk premium $\nu/2$.

–Figure 1 here–

When agent ij signs a contract intended for agent kl, the rent that the former receives is given by

$$R_{ij}(c_{kl}, P_{kl}) = -P_{kl} + (1 - c_{kl})\mu_i + \frac{1}{2}(1 - c_{kl}^2)\nu_j. \tag{6}$$

It is useful to define the following function:

$$\delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l) \stackrel{\text{def}}{=} (1 - c_{kl})(\mu_i - \mu_k) + \frac{1}{2}(1 - c_{kl}^2)(\nu_j - \nu_l). \tag{7}$$

Suppose now that type kl is truthful and receives rent $R_{kl}(c_{kl}, P_{kl})$. Which rent does ij obtain when choosing the contract for kl? Using (4) and (7), the answer is given by

$$R_{ij}(c_{kl}, P_{kl}) \stackrel{\text{def}}{=} R_{kl}(c_{kl}, P_{kl}) + \delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l). \tag{8}$$

Thus, by pretending to be type kl, type ij can obtain type kl's rent plus δ.

To see the usefulness of contract distortion, let us fix the rent that a truthful type kl receives under the contract (c_{kl}, P_{kl}). A marginal increase in the co-insurance rate for kl, $\mathrm{d}c_{kl} > 0$, would have to be compensated by a marginal decrease in the premium P_{kl}. This has the following effect on the rent for the mimicker ij:

$$\frac{\partial R_{ij}(c_{kl}, P_{kl})}{\partial c_{kl}}|_{\mathrm{d}R_{kl}=0} = \frac{\partial}{\partial c_{kl}}\delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l) = -(\mu_i - \mu_k) - c_{kl}(\nu_j - \nu_l).$$

Thus, the rent for ij goes down to the extent that: (i) ij is mimicking a type with a lower risk; and (ii) ij is mimicking a type with lower risk aversion. The intuition is the following. When raising the co-payment of a low risk (or risk-tolerant) individual, the decrease in the premium needed to compensate him is not too large, because of the small likelihood of needing that co-payment (or because of the low valuation of the increase in the variance of final wealth). However, a person with a higher risk level or greater risk aversion who is tempted by this contract will dislike this change. This explains why increasing a co-insurance rate for some types will lower the rents of all those mimicking (and the mimickers of these mimickers) who have a higher risk, and will increase the rent of all those mimicking (and the mimickers of these mimickers) who have lower risk aversion.

From now on, we simply write R_{ij} for $R_{ij}(c_{ij}, P_{ij})$ $(i, j = L, H)$. Self-selection between contracts (c_{ij}, P_{ij}) and (c_{kl}, P_{kl}) then requires that

$$R_{ij} \geq R_{kl} + \delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l),$$
$$R_{kl} \geq R_{ij} + \delta(c_{ij}, \mu_k - \mu_i, \nu_l - \nu_j),$$

which, taken together, imply $0 \geq \delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l) + \delta(c_{ij}, \mu_k - \mu_i, \nu_l - \nu_j)$, or, using (7),

$$\int_{c_{kl}}^{c_{ij}} [(\mu_i - \mu_k) + c(\nu_j - \nu_l)]\,\mathrm{d}c \leq 0.$$

A necessary condition for incentive compatibility between contracts Hj and Lj $(j = H, L)$ is that

$$\int_{c_{Lj}}^{c_{Hj}} \Delta\mu \mathrm{d}c \leq 0 \Longleftrightarrow c_{Hj} \leq c_{Lj}, \tag{9}$$

with $\Delta\mu \overset{\text{def}}{=} \mu_H - \mu_L > 0$. Similarly, incentive compatibility between contracts iH and iL $(i = H, L)$ requires that

$$\int_{c_{iL}}^{c_{iH}} c\Delta\nu \mathrm{d}c \leq 0 \Longleftrightarrow c_{iH} \leq c_{iL}, \tag{10}$$

with $\Delta\nu \overset{\text{def}}{=} \nu_H - \nu_L$ and where it is assumed that $c \geq 0$ (on which more below).

The double dimensionality leads in general to double crossing of the indifference curves of types HL and LH. Solving $MWP^{HL}(c) = MWP^{LH}(c)$ for c yields $c = \frac{\Delta\mu}{\Delta\nu}$. That is, in the (c, P) space, the locus of tangency points between HL's and LH's indifference curves is a vertical line at $\frac{\Delta\mu}{\Delta\nu}$. For lower co-insurance rates, HL's indifference curve crosses that of LH downwards from above, while for higher rates, this happens from below. The quadratic expressions for CE wealth ensure that if a crossing occurs at a rate c^- to the left of $\frac{\Delta\mu}{\Delta\nu}$, then the second crossing occurs at c^+, at the same distance to the right of $\frac{\Delta\mu}{\Delta\nu}$—see Figure 2. Hence, if we say that the indifference curves of HL and LH form a lens, then $\frac{c^+ + c^-}{2} = \frac{\Delta\mu}{\Delta\nu}$ is the position of this lens, while $\ell \overset{\text{def}}{=} c^+ - c^-$ is its size.[9]

–Figure 2 here–

Next, we introduce two crucial variables for characterizing the profit maximizing set of contracts, as follows:

$$D \overset{\text{def}}{=} \frac{\Delta\mu}{\nu_L} \in (0, \infty) \text{ and } x \overset{\text{def}}{=} \frac{\nu_L}{\nu_H} \in (0, 1].$$

The ratio D measures, in a unit-free fashion, the difference in risk between two types.[10] The ratio x measures the degree of similarity along the risk-aversion dimension. Using this notation, the locus of tangency points is therefore located at $D\frac{x}{1-x}$, so that for sufficiently small x, the tangency of the intermediate types' indifference occurs at a co-insurance rate below unity. This makes it possible that both crossings become relevant for the analysis.

[9] The right- (left-) crossing co-insurance rate is given by $c^+(c^-) = \frac{\Delta\mu}{\Delta\nu} + (-)\sqrt{\left(\frac{\Delta\mu}{\Delta\nu}\right)^2 + 2\frac{U^{HL} - U^{LH}}{\Delta\nu}}$, where $U^{HL}(U^{LH})$ is the CE wealth for $HL(LH)$. Hence, the size of the lens, defined as $c^+ - c^-$, is $\ell = 2\sqrt{\left(\frac{\Delta\mu}{\Delta\nu}\right)^2 + 2\frac{U^{HL} - U^{LH}}{\Delta\nu}}$, a dimensionless number.

[10] Because the coefficient of absolute risk aversion (r) measures twice the risk premium per unit of variance, we can conclude that the risk premium of a low-risk-averse type (RP_L, say) equals $\frac{1}{2}\nu_L$. Therefore, $D = \frac{\Delta\mu}{2RP_L} = \frac{1}{2}\frac{\Delta\mu/\mu_L}{RP_L/\mu_L}$.

3 The insurance company

We consider a single, risk-neutral insurer with monopoly power on the market for reimbursement contracts. Her expected profits when an agent of type ij has accepted a reimbursement contract (c, P) is given by

$$\pi^{ij}(c, P) = P - \beta\mu_i = P - (1-c)\mu_i. \tag{11}$$

Therefore, the iso-profit associated with type ij has slope $-\mu_i$ in the contract space (c, P).

With *full* information, the monopolist will provide ij with full insurance ($c_{ij} = 0$) at a premium that sets her rent equal to zero. Hence, using (4), $P_{ij} = \mu_i + \frac{1}{2}\nu_j$. This yields a per capita payoff equal to $\pi = \frac{1}{2}\nu_j$. The tangency line in Figure 1 thus corresponds to the highest feasible iso-profit line, and the profit that the insurer makes can be read off from the dashed vertical axis on the right- hand side. Under full information, the insurer can extract the entire risk premium $\nu/2$. In what follows, we will characterize the optimal co-insurance rates and the optimal rents. The corresponding premia can then be found with the help of (4).

Given (11), the insurer's total profit is equal to $\sum_{i,j} \alpha_{ij}\pi^{ij}(c_{ij}, P_{ij})$. From (4) and (11)—both evaluated at (c_{ij}, P_{ij})–and recalling that we can write R_{ij} for $R_{ij}(c_{ij}, P_{ij})$ (i.e., type ij's rent when truthful), we can express the insurer's total profit as

$$\sum_{i,j} \alpha_{ij} \left[\frac{1}{2}[1 - c_{ij}^2]\nu_j - R_{ij}\right]. \tag{12}$$

This objective function is to be maximized with respect to (c_{ij}, R_{ij}) ($ij = H, L$), subject to the usual voluntary participation and incentive compatibility constraints.

As in most of the literature, to these constraints we add two additional sets of constraints that are needed to avoid false claims (see, e.g., Picard, 2000). If a co-insurance rate is negative, the insurer refunds more than 100% of the losses, and the insuree will obviously have a strong incentive to overstate the size of the loss. On the other hand, if a co-insurance rate exceeds unity, the agent will have to be paid to accept such a contract (i.e., a negative premium). Once the agent has accepted the insurance, he would have to pay the insurer as well as bearing the loss once it occurs. It is clear that he would have strong incentives to understate the size of the loss (or even hide the loss altogether). Hence, we constrain co-insurance rates to lie in the interval $[0, 1]$.

The monopolist thus solves the following problem:

$$\max_{\{c_{ij}, R_{ij}\}} \sum_{i,j} \alpha_{ij} \left[\frac{1}{2}[1 - c_{ij}^2]\nu_j - R_{ij} \right], \text{ s.t.} \tag{13}$$

$$R_{ij} \geq 0 \qquad (i, j, = H, L) \tag{14}$$

$$R_{ij} \geq R_{kl} + \delta(c_{kl}, \mu_i - \mu_k, \nu_j - \nu_l) \quad (i, j, k, l, = H, L) \tag{15}$$

$$0 \leq c_{ij} \leq 1 \qquad (i, j, = H, L) \tag{16}$$

The first set of constraints ensures voluntary participation, while the second ensures that all types self-select. The third set comprises the (reduced form) *ex ante* and *ex post* moral hazard constraints.

The following theorem provides the usual result of no-distortion-at-the-top (full insurance for the HH type) and no-rents-at-the-bottom. Except when otherwise stated, all proofs are relegated to our companion paper (Olivella and Schroyen, 2011).

Theorem 1 *At the optimum solution,(i)* $c_{HH} = 0$ *and (ii)* $R_{LL} = 0$.

Before characterizing the rest of the solution to the two-dimensional screening problem, it is useful to first consider the one-dimensional case.

4 One-dimensional screening

There are three instances in which screening becomes unidimensional. In the first instance, all agents have the same risk aversion; i.e., $\nu_H = \nu_L = \nu$. This is the standard monopoly problem with just two types when insurees either bear a low or a high expected loss. The type distribution can be described by a single parameter $\alpha_{H\cdot}$, the proportion of high risks in the population. We have the following theorem.

Theorem 2 *When all agents have the same risk aversion, the optimal menu has* $c_{H\cdot} = 0$ *and* $c_{L\cdot} = \min\{D\frac{\alpha_{H\cdot}}{1-\alpha_{H\cdot}}, 1\}$.

The full insurance contract giving L zero rent would be selected by H as well. At a zero co-insurance rate, the slope of H's indifference curve is steeper than that of L. If the insurer increases $c_{L\cdot}$ above zero, this will create a second-order reduction in profit from L, but a first-order gain in profit from H because the latter can be charged a strictly higher premium (for full insurance). Hence, it pays to start distorting Ls contract. The optimal co-insurance rate balances the gain in profit from H ($\alpha_{H\cdot}\Delta\mu$) with the loss in profits from L ($(1-\alpha_{H\cdot})\nu$). Notice that it may pay to exclude type L whenever $\alpha_{H\cdot} \geq 1/(1+D)$; i.e., whenever the proportion of low loss agents is sufficiently small—as expected.

The second instance in which the screening problem becomes unidimensional is when individuals differ in risk aversion only. Let $\alpha_{\cdot H}$ instead be the proportion of highly risk-averse types; i.e., those with $\nu = \nu_H (> \nu_L)$. We have the following theorem.

Theorem 3 *When all agents face the same expected loss, the optimal menu has* $c_{\cdot H} = 0$, *and* $c_{\cdot L} = \begin{cases} 0 & \text{if } x > \alpha_{\cdot H} \\ 1 & \text{otherwise.} \end{cases}$

This result is less standard. With only differences in risk aversion, the optimal solution is always at the corner. Either the low type is excluded or he receives full insurance. The reason for this "bang-bang" solution is that, unlike in the different risk scenario, at a zero co-insurance rate, both H's and L's indifference curves are *tangential* to one another. Hence, distorting L's contract by raising the co-insurance rate now results in a *second-order* gain in profit from H, and it is the second-order condition that determines whether $c_{\cdot L} = 0$ is a local maximum or minimum.

The final instance of unidimensional screening arises when risk levels and risk aversion are perfectly positively correlated. As it transpires from (5), we have $MWP^{HH}(c) > MWP^{LL}(c)$ for any c. The two types are therefore once again unambiguously ordered.

Theorem 4 *When the two characteristics are perfectly positively correlated, the optimal menu has* $c_{HH} = 0$ *and* $c_{LL} = \begin{cases} \min\{D\frac{\alpha_{HH}x}{x-\alpha_{HH}}, 1\} & \text{if } x > \alpha_{HH} \\ 1 & \text{otherwise.} \end{cases}$.

We now turn to the two-dimensional screening problem.

5 Two-dimensional screening

From now on, we let individuals not only differ in their risk levels, but also in their risk aversion. The insurance company then faces the following bivariate probability distribution of types:

	ν_L	ν_H	
μ_L	α_{LL}	α_{LH}	$\alpha_{L\cdot}$
μ_H	α_{HL}	α_{HH}	$\alpha_{H\cdot}$
	$\alpha_{\cdot L}$	$\alpha_{\cdot H}$	1

The correlation between risk (μ) and risk aversion (ν) plays an important role in the analysis. This is given by

$$\mathrm{corr}(\mu,\nu) = \frac{E(\mu - E\mu)(\nu - E\nu)}{\sigma_\mu \sigma_\nu} = \frac{\alpha_{HH}\alpha_{LL} - \alpha_{LH}\alpha_{HL}}{\sqrt{\alpha_{L\cdot}\alpha_{H\cdot}}\sqrt{\alpha_{\cdot L}\alpha_{\cdot H}}}.$$

In what follows, we let ρ represent the numerator of the correlation expression: viz., $\rho \stackrel{\text{def}}{=} \alpha_{HH}\alpha_{LL} - \alpha_{LH}\alpha_{HL}$.

To parameterize the distribution of types, we use the triplet $(\alpha_{H\cdot}, \alpha_{HH}, \rho)$, and have the remaining fractions determined by

$$\alpha_{HL} = \alpha_{H\cdot} - \alpha_{HH}, \tag{17}$$

$$\alpha_{LH} = \alpha_{HH}\frac{1-\alpha_{H\cdot}}{\alpha_{H\cdot}} - \frac{\rho}{\alpha_{H\cdot}}, \text{ and} \tag{18}$$

$$\alpha_{LL} = (\alpha_{H\cdot} - \alpha_{HH})\frac{1-\alpha_{H\cdot}}{\alpha_{H\cdot}} + \frac{\rho}{\alpha_{H\cdot}}. \tag{19}$$

Non-negativity of α_{LH} and α_{LL} requires that $-\alpha_{HL}(1-\alpha_{H\cdot}) \le \rho \le \alpha_{HH}(1-\alpha_{H\cdot})$. The feasible set of distribution parameters is then

$$\begin{aligned}\mathcal{A}_0 = \big\{&(\alpha_{H\cdot}, \alpha_{HH}, \rho) \in [0,1]^2 \times R \mid \alpha_{HH} \le \alpha_{H\cdot} \\ &\text{and } -(\alpha_{H\cdot} - \alpha_{HH})(1-\alpha_{H\cdot}) \le \rho \le \alpha_{HH}(1-\alpha_{H\cdot})\big\}.\end{aligned}$$

The other parameters of the model, D and x, pertain to the characteristics of the insurance takers. This part of the parameter space is denoted as the types set $\mathcal{T}_0$:

$$\mathcal{T}_0 = \{(D,x) \in R_+ \times (0,1)\}.$$

It turns out that D and x are sufficient to describe the problem—we can discard the original parameters μ_i and ν_j $(i,j = H,L)$.[11]

In our analysis, we focus on the case in which the correlation of characteristics is non-positive ($\rho \le 0$). Arguably, this is the most empirically relevant situation: highly risk-averse individuals tend to take more precautions and

[11] The fact that four parameters can be reduced to two follows from the fact that we can normalize ν_L to unity, and because, in the monopolist problem only, $\Delta\mu$ matters—see (15).

are thereby less likely to experience losses. Our model could be seen as a reduced form of a more general model in which individuals have initially taken such precautions before going to the insurance market. Second, there is a pragmatic reason for this restriction: under negative correlation, the typology of the equilibrium set of contracts is already complex, but mostly invariant to the *degree* of negative correlation. By contrast, with positive correlation, the degree of correlation starts to matter for characterizing the optimal contract menus in the parameter space. Thus, we restrict the set of distribution parameters to

$$\mathcal{A}_1 = \{(\alpha_{H\cdot}, \alpha_{HH}, \rho) \in \mathcal{A}_0 \text{ and } \rho \leq 0\}.$$

The monotonicity conditions (9) and (10) imply that there are only two possible orderings of co-insurance rates, as follows:

$$\text{Order 1: } 0 = c_{HH} \leq c_{HL} \leq c_{LH} \leq c_{LL} \leq 1, \tag{20}$$

$$\text{Order 2: } 0 = c_{HH} \leq c_{LH} \leq c_{HL} \leq c_{LL} \leq 1. \tag{21}$$

Lemma 1 *If Order 1 applies with $c_{HH} < c_{LH}$, it is optimal to pool HL with HH if and only if $x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}$.*

This result is intuitive. With Order 1, the only type that may envy the contract for HL is HH. Thus, the choice of c_{HL} is only governed by weighing the profits from these two types. Because they have the same risk levels, we can apply Theorem 3 to this subgroup. Given that the fraction of highly risk-averse individuals in this group is $\frac{\alpha_{HH}}{\alpha_{H\cdot}}$, the result follows.

In our technical companion paper, we show that no more than five regimes solve the monopolist's problem. By a regime, we mean a menu of contracts satisfying certain pooling or separation properties and coverage rankings. The five regimes are listed in Table 1, and distinguished as to whether the degree of separation of the low-risk types, measured as $c_{LL} - c_{LH}$, is larger or smaller than the size of the lens formed by the indifference curves of HL and LH (ℓ).

Table 1. The five equilibrium regimes.

Regime	Order	separation degree of low-risk types: $c_{LL}-c_{LH}$	pooling of HL	Range[a] for x	Comments
A	1	$0 = c_{LL}-c_{LH}<\ell$	with HH	$1-\alpha_{LL}< x \leq 1$	HL pooled with HH since $\frac{\alpha_{HH}}{\alpha_{H\cdot}}<1-\alpha_{LL}$
M	1	$0 < c_{LL}-c_{LH}<\ell$	with $\begin{cases} HH \text{ if } x\geq\frac{\alpha_{HH}}{\alpha_{H\cdot}} \\ LH \text{ if } x<\frac{\alpha_{HH}}{\alpha_{H\cdot}} \end{cases}$	$x_{BM}(D)<x$ $<1-\alpha_{LL}$	Only for high D; $c_{LL}=1$
B	1	$0 < c_{LL}-c_{LH}=\ell$	with $\begin{cases} HH \text{ if } x\geq\frac{\alpha_{HH}}{\alpha_{H\cdot}} \\ LH \text{ if } x<\frac{\alpha_{HH}}{\alpha_{H\cdot}} \end{cases}$	$\min\{\frac{1-\alpha_{LL}}{1+\alpha_{LL}},\frac{1}{1+2D}\}<x$ $<\min\{1-\alpha_{LL},x_{BM}(D)\}$	$c_{LL}=1$ for high D
C	1	$0 <\ell< c_{LL}-c_{LH}$	with HH, HL and LH at $c=0$	$x_{CE}(D)<x<$ $\min\{\frac{1-\alpha_{LL}}{1+\alpha_{LL}},\frac{1}{1+2D}\}$	$c_{LL}=1$ for high D
E	2	$0 < c_{LL}-c_{LH}<\ell$	with LL	$0 < x < x_{CE}(D)$	$c_{LL}=1$ for high D or low x

[a]The functions $x_{BM}(\cdot)$ and $x_{CE}(\cdot)$ are defined below in the discussion of Fig. 8.

Note that **Regime E** distinguishes itself from the others in that Order 2 applies. Note also that full separation is *never* optimal. In the case of Order 1, the first part of Lemma (1) indicates that HL should be pooled with either HH or LH. In the case of Order 2, the suboptimality of full separation follows from the following Lemma (proven in the Appendix).

Lemma 2 *(suboptimality of full separation under Order 2) Suppose that HH is indifferent between her own contract and that for LH, but strictly dislikes that for HL, and suppose that LH is indifferent between her own contract and that for HL, but strictly dislikes that for LL, and suppose that HL is indifferent between her own contract and that for LL. Then, profit can be increased by pooling HL with either LL or LH.*

In the companion paper, we prove the first main result, stated below.

Proposition 1 *The five menu structures listed in Table 1 are potential solutions to the monopolist problem. If $\rho \leq 0$, no other menu structures can be optimal. In particular, full separation is never optimal.*

We now give a characterization of each regime. In the next section, we explain when it pays for the insurer to move from one regime to another.

- **Regime A**

This regime pools the high-risk types at full insurance, and the low-risk types at high, but partial, insurance. Figure 3 illustrates. (In this figure and those that follow, solid/dashed indifference curves refer to high/low risk aversion, while bold/thin indifference curves refer to high/low risks).

–Figure 3 here–

Denoting the co-insurance rate for the low-risk types as $c^A_{L\cdot}$, **Regime A** is described by

$$c^A_{L\cdot} = \min\{D\frac{\alpha_{H\cdot}}{1-\alpha_{H\cdot}}, 1\}, \text{ and } c^A_{HH} = c^A_{HL} = 0.$$

This policy corresponds to one under which individuals differ only in their risk dimensions (Theorem 2). Below, we argue that **Regime A** is optimal if x is sufficiently large (i.e., when heterogeneity in risk aversion is weak), more specifically when $x \geq 1 - \alpha_{LL}$. Because a non-positive correlation ensures that $1 - \alpha_{LL} > \frac{\alpha_{HH}}{\alpha_{H\cdot}}$, it follows from Lemma 1 that it is always optimal to pool HL with HH in **Regime A**.

From now on, we restrict the type space $\mathcal{T}_0$ further by imposing an upper bound $\overline{D}_A$ on D,

$$\overline{D}_A \stackrel{\text{def}}{=} \frac{1-\alpha_{H\cdot}}{\alpha_{H\cdot}};$$

that is,

$$\mathcal{T}_1 = \{(D, x) \in \mathcal{T}_0 \mid D \leq \overline{D}_A\}.$$

This restriction ensures that $c^A_{L\cdot} < 1$. In other words, it *rules out exclusion of the low-risk types when individuals are almost equally risk averse.* This condition ensures that our model encompasses the market situation described by Stiglitz (1977).

Given that $x \geq 1 - \alpha_{LL} > \alpha_{H\cdot}$, it follows that when **Regime A** applies, the pooling of the low-risk types happens at a "low" co-insurance rate, viz., $c^A_{L\cdot} < D\frac{x}{1-x}(= \frac{\Delta\mu}{\Delta\nu})$.

- **Regime M**

This regime gives full insurance to HH, insures LH at a small but positive co-insurance rate, but excludes LL. Type HL is pooled with HH if $x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}$; otherwise, this type is pooled with LH (cf Lemma 1). Figure 4 (drawn for $x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}$) illustrates this regime.

–Figure 4 here–

Below, we show that $x < 1 - \alpha_{LL}$ is a necessary condition for **Regime M** to be optimal.

The optimal values for the co-insurance rates are given by

$$c_{LL}^M = 1, c_{LH}^M = \begin{cases} D\frac{\alpha_{H\cdot}x}{\alpha_{H\cdot}(1-x)+\alpha_{LH}x} & \text{if } x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}, \\ D\frac{\alpha_{H\cdot}x}{\alpha_{HL}+\alpha_{LH}} & \text{if } x \leq \frac{\alpha_{HH}}{\alpha_{H\cdot}}, \end{cases}$$
$$c_{HH}^M = 0, \text{ and } c_{HL}^M = \begin{cases} 0 & \text{if } x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}, \\ D\frac{\alpha_{H\cdot}x}{\alpha_{HL}+\alpha_{LH}} & \text{if } x \leq \frac{\alpha_{HH}}{\alpha_{H\cdot}}. \end{cases}$$

Thus, the difference between **A** and **M** is that the low-risk types (LH and LL), are now separated from one another, but the degree of separation is "small", in the sense that $c_{LL} - c_{LH} < \ell$, the size of the lens.

- **Regime B**

In this regime, the two low-risk types are separated by positioning them on each side of the lens. That is, they satisfy $c_{LH}+c_{LL} \equiv 2\frac{Dx}{1-x}$. We may distinguish between **Regime Bf** and **Regime Bp**, depending on whether LH obtains full insurance ($c_{LH} = 0$) or partial insurance ($c_{LH} > 0$), respectively. For the latter regime, we can also make a distinction based on whether LL individuals are included (**BpI**: $c_{LL} < 1$) or excluded (**BpX**: $c_{LL} = 1$) from insurance. Lemma 1 can be applied to determine whether HL should be pooled with HH ($x \geq \frac{\alpha_{HH}}{\alpha_{H\cdot}}$) or LH ($x < \frac{\alpha_{HH}}{\alpha_{H\cdot}}$). The three panels of Figure 5 (drawn for $x \geq \frac{\alpha_{HH}}{\alpha_{H\cdot}}$) illustrate.

–Figure 5a, 5b, 5c here–

Regime B may be summarized as follows:

$$c_{HH}^B = 0,\ c_{LH}^B = \begin{cases} 2D\frac{x}{1-x} - 1 & \textbf{(BpX)}, \\ D\frac{(1+\alpha_{LH}+\alpha_{LL})x-(1+\alpha_{LH}-\alpha_{LL})}{(1-\alpha_{H\cdot})(1-x)} & \textbf{(BpI)}, \\ 0 & (\textbf{Bf}), \end{cases} \tag{22}$$

$$c_{HL}^B = \begin{cases} 0 & \text{if } x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}, \\ c_{LH}^B & \text{if } x \leq \frac{\alpha_{HH}}{\alpha_{H\cdot}}. \end{cases} \text{ and } c_{LL}^B = \begin{cases} 1 & \textbf{(BpX)}, \\ D\frac{2\alpha_{LH}+\alpha_{H\cdot}(1-x)}{(1-\alpha_{H\cdot})(1-x)} & \textbf{(BpI)}, \\ 2D\frac{x}{1-x} & (\textbf{Bf}). \end{cases}$$

- **Regime C**

Regime C is one under which everybody is fully insured, except for the LL individuals who face a very high co-insurance rate (**CI**) or are even excluded (**CX**). Moreover, the screening between LH and LL is now very thorough in the sense that $c_{LL} - c_{LH} > \ell$. Consequently, $c_{LL} \geq 2\frac{\Delta\mu}{\Delta\nu}$. This regime is illustrated in Figure 6.

–Figure 6 here–

Regime C thus balances a high premium income from the "upper" types with the loss in profit from distorting LL's contract. Intuitively, with few LL individuals around, such distortion is attractive, and with hardly any of them around, it is even optimal to exclude them altogether.

We can summarize **Regime C** as follows:

$$c_{HH}^C = c_{HL}^C = c_{LL}^C = 0, c_{LL}^C = \begin{cases} D\frac{1-\alpha_{LL}}{\alpha_{LL}} & \textbf{(CI)}, \\ 1 & \textbf{(CX).} \end{cases}$$

- **Regime E**

A common feature of all previous regimes is that Order 1 applies ($c_{HL} \leq c_{LH}$). In **Regime E**, the opposite is true: HL's contract is now severely distorted by being pooled with LL. This makes room for increasing the distortion on LH, which, in turn, allows the insurer to extract more rent from HH individuals. Again, if there are few low-risk-averse individuals around, it may pay to exclude these individuals from the market (**EX**), otherwise they are included but receive limited insurance (**EI**). Figure 7 illustrates.

–Figure 7 here–

Separation of LH from LL is once more minimal: $c_{LL} - c_{LH} < \ell$. Under Order 1, separation of LL from LH is carried out to increase the profits from HH, HL and LH at the cost of a lower profit from LL. Under Order 2, HL is pooled with LL so as to extract more rent from the highly risk-averse types, HH and LH. Across these two types, rent extraction is optimized in the standard way (cf Theorem 2).

Denoting the common co-insurance rate for HL and LL as $c_{\cdot L}$, the optimal co-insurance rates for **Regime E** are

$$c_{HH}^E = 0, c_{LH}^E = D\frac{\alpha_{HH}x}{\alpha_{LH}}, \text{ and } c_{\cdot L}^E = \begin{cases} D\frac{x\alpha_{HL}}{x-\alpha_{\cdot H}} & \textbf{(EI)}, \\ 1 & \textbf{(EX)}. \end{cases} \tag{23}$$

This concludes the presentation of the five regimes, or contract menus. Loosely speaking, one can say that the degree of separation between LH and LL, viz., $c_{LL} - c_{LH}$, increases as one moves from **A** $(c_{LL} - c_{LH} = 0)$ into **M** $(0 < c_{LL} - c_{LH} < \ell)$, into **B** $(0 < c_{LL} - c_{LH} = \ell)$, and further into **C** $(0 < \ell < c_{LL} - c_{LH})$. In **Regime E**, the degree of separation becomes minor again, but **E** is qualitatively different because it makes use of a different order.

6 Comparison of regimes

Having established the optimal co-insurance structure for each regime, we now investigate for which (D, x) combinations each of the regimes becomes optimal. The precise comparisons are relegated to the technical companion paper. Here, we limit ourselves mainly to a graphical presentation by partitioning the (D, x) space into subspaces according to which regime secures the monopolist the highest profit. We first establish the optimal menu in the neighbourhoods of the upper and lower boundaries for x. Thereafter, we sketch the optimal menus for intermediate values of x.

When there is no heterogeneity in risk aversion, we know from Theorem 2 that **Regime A** is optimal. By a continuity argument, this is also true for small differences between ν_H and ν_L. Low-risk types will be partially insured while high-risk types obtain full insurance.

Theorem 5 *As $x \to 1$, the optimal contract menu is defined by **Regime A**.*

Inspection of (23) shows that for small enough x, it is optimal to exclude the two types with low-risk aversion (HL and LL) in **Regime E**. For the other regimes, one of the risk-tolerant types (i.e., HL) continues to buy insurance. However, if x approaches zero, the willingness to pay for insurance among highly risk-averse types (HH and LH) becomes infinitely larger than that among low-risk-averse types. Therefore, it cannot be optimal to keep providing the latter with insurance, as this constrains the premia that can be charged to the former.

Theorem 6 *As $x \to 0$, the optimal contract menu is defined by **Regime EX**.*

Before we consider when the other regimes become optimal, note that because **Regimes A**, **M**, **B**, and **C** all share the same order (Order 1), when moving from one regime into the adjacent one, at least one of the co-insurance rates changes continuously. **Regime E**, on the other hand, makes use of Order 2. The move from this regime into the adjacent one makes all co-insurance rates jump (except for c_{HH}, which is always zero). Identification of the borderline of regime E is then only possible by comparing the maximal profit functions. This is explained in more detail in our technical companion paper.

We now explain in a heuristic way the optimal regimes for intermediate x. For this purpose, suppose that x is close to unity to begin with; i.e., there is initially hardly any difference in risk aversion, but then x steadily falls in value, which signifies increased heterogeneity in risk aversion. With x close to unity, the optimal menu is given by **Regime A**. The size of the lens is $\ell = 2\left(D\frac{x}{1-x} - c^A_{L\cdot}\right) = 2\left(D\frac{x}{1-x} - D\frac{1-\alpha_{H\cdot}}{\alpha_{H\cdot}}\right)$. As x falls, ν_H starts to exceed ν_L. This makes it optimal to start screening the LL from the LH types: by providing LL with less coverage (at a lower premium), LH (and therefore also the high-risk types HH and HL) can be charged a higher premium. However, because LL was initially pooled with LH at the left-hand crossing, a marginal increase in c_{LH} is impossible for incentive compatibility reasons. What is possible is to move LL from the left-hand crossing to the co-insurance rate corresponding to the right-hand crossing, and adjusting her premium to keep her rent at zero. This becomes optimal when $x < 1 - \alpha_{LL}$; then, **Regime BpI** takes over. This is possible as long as the lens is not too big, i.e., if $c^A_{L\cdot} + \ell = 2D\frac{x}{1-x} - D\frac{1-\alpha_{H\cdot}}{\alpha_{H\cdot}} \leq 1$. However, when D is large, the previous reshuffling would involve a co-insurance rate for LL that exceeds unity (and a negative premium). Because this is ruled out, the best the insurer can do is to exclude LL and to extract all the rent from LH. This is what happens in **Regime M**. It can be shown that in this regime, the right-hand crossing, i.e., $c^M_{LH} + \ell$, is increasing in x. Hence, as x falls further, then at some stage, this right-hand-side crossing will coincide with the no-insurance point $(1, 0)$. This happens when x falls short of $x_{BM}(D)$.[12] At that point, **Regime BpX** takes over from **Regime M**. See Figure 8.

–Figure 8 here–

When x falls far enough, it pays to increase the wedge between c_{LL} and c_{LH}, even in excess of ℓ. That is, the point at which **Regime C** takes over

[12]That is, $x_{BM}(D)$ solves $2\frac{Dx}{1-x} - c^M_{LH} = 1$.

from **Regime B**. In the companion paper, we show that

$$\pi^C > \pi^B \Longleftrightarrow x < \min\{\frac{1-\alpha_{LL}}{1+\alpha_{LL}}, \frac{1}{1+2D}\},$$

where α_{LL} is given by (19).

When there is substantial heterogeneity in risk aversion (when x is very small), it becomes profitable to screen the highly risk averse as a group from the low-risk-averse group. The only way to implement this is by switching to Order 2. Then, **Regime E** is optimal. In the companion paper, we also show that there exists a function, $x_{CE}(\alpha_{H\cdot}, \alpha_{HH}, \rho, D)$, that is non-increasing in D, with $x_{CE}(\alpha_{H\cdot}, \alpha_{HH}, \rho, D) < \frac{1}{1+2D}$ for any $D \in [0, \overline{D}_A]$, such that

$$\pi^E > \pi^C \Longleftrightarrow x < x_{CE}(\alpha_{H\cdot}, \alpha_{HH}, \rho, D).$$

This function is found by comparing the maximal profit under **Regime C** with the maximal profit under **Regime E**. Because there is continuity when switching from **B** to **C**, whereas there is discontinuity when switching from **C** to **E**, the question arising is whether π^C can be dominated by π^E for any x that makes **C** dominate **B**. In other words, does it make more sense to compare π^E with π^B? This is illustrated in Figure 9. The profit function π^E intersects with π^C at $\widehat{x} < \min\{\frac{1-\alpha_{LL}}{1+\alpha_{LL}}, \frac{1}{1+2D}\}$, while the function $\widetilde{\pi}^E$ dominates π^C, indicating that once x falls short of $\widetilde{x}$, **Regime B** should be replaced by **Regime E**.

–Figure 9 here–

For each value of ρ (≤ 0), we can define a region $\mathcal{R}(\rho)$ in the $(\alpha_{H\cdot}, \alpha_{HH})$ space such that this is indeed what happens:

$$\mathcal{R}(\rho) = \{(\alpha_{H\cdot}, \alpha_{HH}) \in [0,1]^2 : x_{CE}(\alpha_{H\cdot}, \alpha_{HH}, \rho, D)|_{\text{small } D} \geq \frac{1-\alpha_{LL}(\alpha_{H\cdot}, \alpha_{HH}, \rho)}{1+\alpha_{LL}(\alpha_{H\cdot}, \alpha_{HH}, \rho)}\}$$

Bundles in this region can be shown to be feasible.[13] Figure 10a–c displays $\mathcal{R}(\rho)$ for $\rho = 0, -\frac{1}{30}, -\frac{2}{30}$. Thus, if $\rho = -\frac{2}{30}$, then for almost all $(\alpha_{H\cdot}, \alpha_{HH})$ that are feasible in combination with this value for ρ (the area delineated by the dashed line), it transpires that the interval for which **Regime C** is optimal, $[x_{CE}(\alpha_{H\cdot}, \alpha_{HH}, \rho, D), \min\{\frac{1-\alpha_{LL}(\alpha_{H\cdot}, \alpha_{HH}, \rho)}{1+\alpha_{LL}(\alpha_{H\cdot}, \alpha_{HH}, \rho)}, \frac{1}{1+2D}\}]$, is non-empty. For $\rho \leq -.089$, "almost all" can be replaced by "any".

[13] That is, if $(\alpha_{H\cdot}, \alpha_{HH}) \in \mathcal{R}(\rho)$ for some $\rho \leq 0$, then $(\alpha_{H\cdot}, \alpha_{HH}, \rho) \in \mathcal{A}_1$.

–Figure 10a, 10b, 10c here–

We therefore restrict the set of distribution parameters further to

$$\mathcal{A}_2 = \{(\alpha_{H\cdot}, \alpha_{HH}, \rho) \in \mathcal{A}_1 \text{ and } (\alpha_{H\cdot}, \alpha_{HH}) \notin \mathcal{R}(\rho)\}.$$

However, from the previous discussion, $\mathcal{A}_2$ is almost as large as $\mathcal{A}_1$.

We now provide the second main result of the paper.

Proposition 2 *Suppose that $(\alpha_{H\cdot}, \alpha_{HH}, \rho) \in \mathcal{A}_2$. Then, the optimal menu structure as a mapping from $\mathcal{T}_1$ into the menu set is as illustrated in Figure 8.*[14]

Recall that D measures the incentive for μ_H-type individuals to mimic μ_L-type individuals, normalized by (twice) the risk premium of the latter. A high co-insurance rate discourages the former group from applying for the contracts intended for the latter, and thus allows insurers to charge the former group more for full insurance. Regimes **M**, **BX**, **CX**, and **EX** all exclude LL; they become optimal for high levels of D.

On the other hand, x measures the extent of similarity in risk aversion. Dissimilarity warrants a contract menu that screens low-risk-averse consumers from highly risk-averse ones. The latter group is much more willing to pay for insurance coverage, and the monopolist takes advantage of this. Such screening is absent in **Regime A** and maximal in **Regime EX** , under which all risk-tolerant individuals are excluded from coverage. The result is a market with only highly risk-averse customers, who have private information on their expected losses. The standard screening problem thus applies.

We conclude this section by plotting in Figure 11 the optimal co-insurance rates for LL and LH underlying the different regimes (assuming that $D < \frac{(1-\alpha_{H\cdot})\alpha_{LL}}{\alpha_{LH}+(1-\alpha_{H\cdot})(1-\alpha_{LL})}$ such that **Regime M** can be ignored). The analysis of Section 8 is based on this figure.

–Figure 11 here–

[14] As explained in the previous section, whether HL is pooled with HH or with LH in **Regimes Bp** and **M** depends on whether x exceeds or falls short of $\frac{\alpha_{HH}}{\alpha_{H\cdot}}$. Given that $\alpha_{\cdot H} = \frac{\alpha_{HH}}{\alpha_{H\cdot}} - \frac{\rho}{\alpha_{H\cdot}}$, with non-positive correlation, $\frac{\alpha_{HH}}{\alpha_{H\cdot}}$ will never exceed $\alpha_{\cdot H}(< \frac{1+\alpha_{LH}-\alpha_{LL}}{1+\alpha_{LH}+\alpha_{LL}})$. Figure 8 is based on $\alpha_{H\cdot} = .6, \alpha_{HH} = .2$ and $\rho = 0$. Hence, $\frac{\alpha_{HH}}{\alpha_{H\cdot}} = \alpha_{\cdot H}$, and all (D, x) combinations in the regions for **BpX** and **M** have pooling of HL with HH rather than with LH.

7 The positive correlation test

Chiappori *et al.* (2006) showed that a common prediction of any model of a competitive insurance market with asymmetric information is a strictly positive relationship between the degree of coverage and the expected loss across contracts. This is quite a strong result, and we refer to it as *positive monotonicity* (PM). This property implies a positive correlation between coverage and risk, but the converse is not true.

In the empirical literature on testing for asymmetric information in insurance markets, researchers typically rely on estimating the correlation coefficient between coverage and the expected loss, and then use a one-sided test to determine whether this coefficient is statistically significantly positive (see, e.g., Cohen and Siegelman, 2010; Finkelstein and McGarry, 2006). The empirical evidence on positive correlation is somewhat weak; there is even evidence of negative correlation in some markets.[15] This is quite surprising, because the result of Chiappori *et al.* (2006) is general; conditional on the competition assumption, it holds for any combination of moral hazard and adverse selection in underlying risk.[16]

As mentioned in the introduction, there are proposed theoretical explanations for this lack of evidence. One is the so-called "cherry picking argument" (Chiappori and Salanié, 2000) or "propitious selection" (Hemenway, 1990), which combines adverse selection in risk preference (but not in the underlying risk) with moral hazard. The argument is that if individuals take precautions having purchased insurance, then highly risk-averse individuals

[15] The later phenomenon is termed "advantageous or propitious selection". In regard to life insurance, see, e.g., Cawley and Philipson (1999) and McCarthy and Mitchell (2003), and in long-term care, see, e.g., Finkelstein and McGarry (2006). On the other hand, Olivella and Vera (2011) show that in duplicate or substitutive private health insurance systems (in which the public and (competitive) private insurance sectors offer the same portfolio of services), if (but not only if) there is heterogeneity in risks only, then propitious selection into private insurance should be observed if and only if information on risks is symmetric.

[16] When the literature refers to moral hazard, this could encompass two distinct phenomena. One relates to individuals who enjoy more coverage having less of an incentive to undertake precautionary behaviour, which makes them *observationally* more risky. The other arises because one does not necessarily observe actual risk but the usage of, say, health services. Because coverage implies a lower cost of accessing these services, individuals may use more of these services because they have more coverage, not because they are more risky. Notice that both types of moral hazard reinforce the positive correlation. Of course, one of the econometric issues is that, even after observing some positive correlation, it is hard to disentangle the adverse selection and either of the two moral hazard effects.

will increase their coverage as well as take more precautions, everything else being equal. This may then result in a negative correlation between observed risk and coverage.[17]

Because the optimal menu in a monopoly market with two-dimensional screening may display Order 2, PM will cease to hold; for a subset of types (LH and HL), coverage is negatively related to risk size. This of course does not imply that the *correlation* between risk and coverage is negative, because PM does hold for other subsets of types (between HH on the one hand and LH and LL on the other). This also suggests that a sufficiently negative correlation between risk and risk aversion (i.e., a sufficiently small ρ) ensures a negative correlation between coverage and risk size. This is shown below.

In fairness, Chiappori *et al.* (2006) pointed out that in a monopoly, the PM property may be violated. They do this by starting with a model in which only preference heterogeneity exists (cf. Section 4), and then by introducing an infinitesimal amount of exogenous risk heterogeneity that is perfectly negatively correlated with risk heterogeneity (i.e., the more risk-averse agents have a slightly smaller accident probability). We show that the PM property does not hold whenever **Regime E** applies, *even if* the underlying risk and risk preference are independently distributed.

Translated into our setting, the Chiappori *et al.* (2006) proposition may be stated as follows:

Consider two contracts C_a and C_b that are offered on the market. Suppose that: (i) C_a gives more coverage then C_b, i.e., $c_a < c_b$; and (ii) the per capita profit generated by contract a does not exceed that of contract b, $\pi(C_a) \leq \pi(C_b)$. Then, (iii) the expected loss to those consumers signing up for contract a should exceed the expected loss of those consumers signing up for contract b, i.e., $\mu(C_a) \geq \mu(C_b)$.

It is easy to see that property (iii) is satisfied in all regimes except for **Regime E**. In that regime, the contract for LH has more coverage than the contract for the low-risk-averse individuals (LL and HL). The PM property would then require that $\mu(C_{LH}) = \mu_L > \mu(C_{\cdot L}) = \left(\frac{\alpha_{HL}}{\alpha_{\cdot L}}\mu_H + \frac{\alpha_{LL}}{\alpha_{\cdot L}}\mu_L\right)$, which is obviously violated. The culprit is the violation of condition (ii): C^E_{LH} generates a higher per capita profit than does $C^E_{\cdot L}$. This can be seen as

[17] See, e.g., Jullien et al. (2007), De Donder and Hindriks (2009), and Finkelstein and McGarry (2006).

follows:

$$\begin{aligned}
\pi(C^E_{\cdot L}) &= \frac{\alpha_{HL}}{\alpha_{\cdot L}}\left\{\frac{1}{2}[1-(c^E_{\cdot L})^2]\nu_L - (1-c^E_{\cdot L})\Delta\mu\right\} + \frac{\alpha_{LL}}{\alpha_{\cdot L}}\frac{1}{2}[1-(c^E_{\cdot L})^2]\nu_L \\
&= \frac{1}{2}[1-(c^E_{\cdot L})^2]\nu_L - \frac{\alpha_{HL}}{\alpha_{\cdot L}}(1-c^E_{\cdot L})\Delta\mu, \\
\pi(C^E_{LH}) &= \frac{1}{2}[1-(c^E_{LH})^2]\nu_H - \frac{1}{2}[1-(c^E_{LH})^2]\Delta\nu \\
&= \frac{1}{2}[1-(c^E_{LH})^2]\nu_L.
\end{aligned}$$

Because $c^E_{LH} < c^E_{\cdot L}$, it follows that $\pi(C_{LH}) > \pi(C_{\cdot L})$, irrespective of which optimal values the co-insurance rates take under **Regime E**.

Performing a positive correlation test on our model would amount to calculating the covariance across contracts between $1 - c(C)$ and $\mu(C)$, as follows:

$$cov(1 - c(C_{ij}), \mu(C_{ij})) = \sum_{i,j} \alpha_{ij}(1 - c_{ij})\mu_i - \sum_{i,j} \alpha_{ij}(1 - c_{ij}) \sum_{i,j} \alpha_{ij}\mu_i.$$

As the second part of the following proposition shows, when the optimal regime is **EI**, this covariance is negative only if the correlation between expected loss and risk aversion, ρ, is sufficiently negative.

Proposition 3 *(i) For a sufficient degree of heterogeneity in risk aversion, such that* ***Regime E*** *prevails, some low-risk individuals (LH) purchase more coverage than do some high-risk individuals (HL). (ii) In the case of* ***Regime EI****,* $cov(1 - c(C_{ij}), \mu(C_{ij})) < 0$ *if and only if* $\rho < -\alpha_{HH}\left(x - \frac{\alpha_{HH}}{\alpha_{H\cdot}}\right) (< 0)$.

In other words, the advantageous selection among LH and HL, described by part (i), may exactly offset the standard adverse selection, such that any correlation between risk and coverage vanishes. Finkelstein and McGarry (2006) show that the long-term care insurance market may suffer from asymmetric information, despite the absence of evidence for a positive correlation between risk and coverage. Our model helps in interpreting this evidence.

8 Gender discrimination

Crocker and Snow (1986) have shown that imperfect categorical discrimination in insurance—such as gender discrimination—always expands the efficiency frontier. Hoy (1982) showed how categorization based on a signal may lead to a Pareto improvement in a competitive insurance market if the signal conveys information about the level of risk. In this section, we ask when such efficiency gains arise in a monopolistic market structure. We show that a Pareto improvement is possible if the signal, such as gender, is informative about risk aversion.

Let us write $p(\mu, \nu, g)$ as the likelihood function that an arbitrary insuree has an expected loss of μ, risk aversion of ν and gender $g \in \{m, w\}$. A monopolist who is allowed to condition on gender will, for each gender, g, design an optimal contract menu based on the risk-aversion ratio, x, the risk difference parameter, D, and the probability matrix,[18]

$$\begin{pmatrix} p(\mu_L, \nu_L|g) & p(\mu_L, \nu_H|g) \\ p(\mu_H, \nu_L|g) & p(\mu_H, \nu_H|g) \end{pmatrix}.$$

We now assume that risk aversion is a sufficient statistic for gender with respect to the expected loss:

Condition S $\quad p(\mu|\nu, g) = p(\mu|\nu).$

Condition S means that within a given risk-aversion class, the observation of a person's gender carries no extra information about the risk class to which this person belongs.[19] In general, sufficiency is not enough to break the link between gender and expected loss. If female drivers are highly risk averse, and if this attitude leads them to careful driving, then there will still be a connection between gender and expected loss. This last connection is broken by the assumption that expected loss is independently distributed of risk aversion—i.e., risk aversion has no impact on driving. This allows us to state the following result.

Lemma 3 *If the likelihood function $p(\cdot)$ satisfies Condition S and if μ and ν are independently distributed, then μ and g are also independently distributed: $p(\mu|g) = p(\mu)$.*

Thus, these assumptions support the conclusion reached by the European Commission—that gender is insignificant in explaining risk type.

[18] We assume that the support of the distribution of types does *not* vary with the signal. Alternatively, the support could be made dependent on the signal. This, in effect, amounts to assuming that the support consists of more than four (μ, ν)-pairs, some of which have zero probability mass, depending on the observation of the signal.

[19] Two equivalent formulations of Condition S are: $p(g|\mu, \nu) = p(g|\nu)$ and $p(\mu, \nu|g) = p(\mu|\nu) \cdot p(\nu|g)$.

Because a gender-discriminating firm will use the probability functions $p(\mu, \nu|g)$ $(g = m, w)$, rather than the single function $p(\mu, \nu)$, to design menus, and because profits and consumer rents depend on the co-insurance rate c_{LL}, it is important to determine the effect of $p(\cdot)$ on c_{LL}. For this purpose, let us assume that $D < \frac{(1-\alpha_{H\cdot})\alpha_{LL}}{\alpha_{LH}+(1-\alpha_{H\cdot})(1-\alpha_{LL})}$ so that we can ignore **Regime M**. From proposition 2, it follows that without discrimination, the upper boundaries of regimes **C**, **Bf**, and **BpI** are determined by the parameters α_{LL} and α_{LH}. Fixing x, $\alpha_{L\cdot}$ (and therefore $\alpha_{H\cdot} = 1 - \alpha_{L\cdot}$) allows one to trace out the optimal value of c_{LL} as a function of α_{LL}. This yields Figure 12, in which it is assumed that $\frac{1-x}{1+x} < \frac{1}{2}(1+\alpha_{L\cdot})(1-x)$. This means that the curve for LL's optimal co-insurance rate is flat for some range of α_{LL} values; this is equivalent to assuming that

$$\frac{1-\alpha_{L\cdot}}{1+\alpha_{L\cdot}} < x. \tag{24}$$

–Figure 12 here–

Let us define ω_L (ω_H) as the likelihood that an arbitrary person with low (high) risk aversion is a female; i.e., $\omega_L \overset{\text{def}}{=} p(w|\nu_L)$ and $\omega_H \overset{\text{def}}{=} p(w|\nu_H)$. If half the population are women, then $p(w) = \omega_H \alpha_{\cdot H} + \omega_L(1 - \alpha_{\cdot H}) = \frac{1}{2}$.

There is now ample evidence that men are on average less risk averse than women.[20] For our model, this means that $p(\nu_L|w) < p(\nu_L) < p(\nu_L|m)$. An insurance company that is allowed to gender discriminate, having observed the customer's gender g, will update the probability α_{LL} in the following way:

$$\alpha_{LL|m} \overset{\text{def}}{=} p(\mu_L, \nu_L|m) = p(\mu_L|\nu_L) \cdot p(\nu_L|m) = p(\mu_L) \cdot p(\nu_L|m) \ (< \alpha_{LL}), \text{ and}$$
$$\alpha_{LL|w} \overset{\text{def}}{=} p(\mu_L, \nu_L|w) = p(\mu_L|\nu_L) \cdot p(\nu_L|w) = p(\mu_L) \cdot p(\nu_L|m) \ (> \alpha_{LL}).$$

where the first equality sign follows from Condition S and the second follows from independence. Thus, $\alpha_{L\cdot}$ and $\alpha_{H\cdot}$ do not change when gender is observed.

Suppose now that (24) holds, and suppose that the proportion of LL individuals as a whole, the proportion among men, and the proportion among

[20]See Hartog *et al.* (2002), Cohen and Einav (2007), Eckel and Grossman (2008), Kimball *et al.* (2008) and Aarbu and Schroyen (2011).

women, are $\alpha_{LL}, \alpha_{LL|m}$, and $\alpha_{LL|w}$, respectively. Suppose also that these proportions are as illustrated in Figure 13.

–Figure 13 here–

Then, we can conclude that because

$$\frac{1-x}{1+x} < \alpha_{LL|w} < \alpha_{LL} < \frac{1}{2}(1+\alpha_{L\cdot})(1-x), \tag{25}$$

the co-insurance rate for LL women will remain at its no-discrimination value, and the rents of LH women, HL women, and HH women will not change because of discrimination (and LL women continue to receive zero rent). On the other hand, because

$$\frac{1-x}{1+x} < \alpha_{LL} < \frac{1}{2}(1+\alpha_{L\cdot})(1-x) < \alpha_{LL|m}, \tag{26}$$

the optimal co-insurance rate for LL men will drop below its no-discrimination value, and therefore, all men will receive more rent when offered the optimal contract menu for men (except LL men, who continue to receive zero rent).[21] The insurance company will increase its total profits because of finding it optimal to choose a new menu for its male clientele—it could have stuck to the same menu as in the no-discrimination case. Thus, a Pareto improvement is possible by allowing gender discrimination. As can be seen from the figure, conditions (25) and (26) are not only sufficient for a Pareto improvement, but also necessary. We summarize this result in the following proposition.

Proposition 4 *Suppose that Condition S holds, and that μ and ν are independently distributed. Suppose that (24) holds. For given values of $x, \alpha_{L\cdot}$, and D, allowing gender discrimination will lead to a Pareto improvement in the insurance market if and only if conditions (25) and (26) hold.*

Condition (24) is satisfied when the proportion of low-risk individuals, $\alpha_{L\cdot}$, is not too small in relation to x.

The intuition for Proposition 4 is the following. Because men are on average less risk averse, the "male" market consists of more LL types than does the overall market. This makes the distortion of the LL contract that

[21]Because the optimal menu for men is of the type **Bp**, the rents are given as follows: $R_{HH} = R_{HL} + \frac{1}{2}\Delta\nu, R_{HL} = R_{LL} + (1 - c_{LL})\Delta\mu$, $R_{LH} = R_{LL} + \frac{1}{2}(1 - c_{LL}^2)\Delta\nu$, and $R_{LL} = 0$.

was optimal for the entire market too costly for the "male" market: offering LL men a lower co-insurance rate (in return for a higher premium) increases profits from this market segment sufficiently to compensate for the lower rents extracted from the "higher" male types. Hence, all men benefit, and so does the monopolist.

9 Conclusion

In this paper, we studied the outcome in a monopolistic insurance market when the insurer is only aware of the statistical distribution of the expected loss and the level of risk aversion of its customers. We formulated a mean–variance model that results in quasilinear preferences over contracts; we identified the five contract menus that emerge in equilibrium; and for each menu, we derived the optimal co-insurance rates. Next, we identified for each menu the subset of parameter values for which that menu is optimal. We did this under non-positive correlation between the two characteristics.

We found:

- it is never optimal to fully separate all the types. In other words, there will always be some pooling of types in equilibrium;
- the greater is the heterogeneity in the expected loss, the more it pays to screen the low-risk from the high-risk types, by imposing a high co-insurance rate on the former;
- the greater is the heterogeneity in terms of risk aversion, the more it pays to screen the low-risk-averse from the highly risk-averse by imposing a high co-insurance rate on the former; and
- the property of positive monotonicity between coverage and expected loss need no longer hold—neither does the property of positive correlation.

We also identified an open set of parameter values such that when the female distribution of risk aversion first order stochastically dominates the male distribution, allowing gender discrimination results in a Pareto improvement in this market. Hence, our analysis shows that one should be careful when abolishing gender categorization; even when gender itself does not (statistically) affect the expected level of losses or claims, it may affect the outcome in an imperfectly competitive insurance market so that nobody gains and some participants become worse off.

References

[1] Aarbu, K O and F Schroyen (2011) “Mapping risk aversion in Norway using hypothetical income gambles”, discussion paper 13/09 (revised Sept.), Norwegian School of Economics.

[2] Armstrong, M (1999) “Optimal regulation with unknown demand and cost functions”, *Journal of Economic Theory* **84**, 196–215.

[3] Armstrong, M & J-C Rochet (1999) ”Multi-dimensional screening: a user’s guide”, *European Economic Review* **43**, 959–979.

[4] Cawley, J and T Philipson (1999) “An empirical examination of information barriers to trade in insurance”, *American Economic Review* **89**, 827–846.

[5] Chiappori, P-A, B Jullien, B Salanié and F Salanié (2006) “Asymmetric information in insurance: general testable implications”, *Rand Journal of Economics* **37**, 783–798.

[6] Chiappori, P-A, and B Salanié (2000) “Testing for asymmetric information in insurance markets”, *Journal of Political Economy* **108**, 56–78.

[7] Cohen, A and L Einav (2007) “Estimating risk preferences from deductible choice”, *American Economic Review* **97**, 745–788.

[8] Cohen, A and P Siegelman (2010) “Testing for adverse selection in insurance markets”, *Journal of Risk and Insurance* **77**, 39–84.

[9] Commission of the European Communities (2003) *Proposal for a Council Directive Implementing the Principle of Equal Treatment between Women and Men in the Access to and the Supply of Goods and Services* (Council Directive 2003/657) (available at http://europa.eu.int/eur-lex/en/com/pdf/2003/com2003_0657en01.pdf).

[10] Crocker, K and A Snow (1986) ”The efficiency effects of categorical discrimination in the insurance industry”, *Journal of Political Economy* **94**, 321–344.

[11] Crocker, K and A Snow (2011) "Multidimensional screening in insurance markets with adverse selection", *Journal of Risk and Insurance* **78**, 287-307.

[12] Cutler, D, A Finkelstein and K McGarry (2008) "Preference heterogeneity and insurance markets: explaining a puzzle of insurance", *American Economic Review (Papers and Proceedings)* **98**, 157–162.

[13] Dana, J (1993) "The organization and scope of agents", *Journal of Economic Theory* **59**, 288–310.

[14] De Donder, P and J Hindriks (2007) "Equilibrium social insurance with policy-motivated parties", *mimeo*, University of Toulouse.

[15] De Donder, P and J Hindriks (2009) "Adverse selection, moral hazard and propitious selection", *Journal of Risk and Uncertainty* **38,** 73–86.

[16] Eckel, C and P J Grossman (2008) "Men, women and risk aversion", ch. 113 in C Plot and V Smith (eds) *Handbook of Experimental Economics Results vol. 1* (Amsterdam: North Holland).

[17] European Court of Justice (2011) *Press release No 12/2011-1 March 2011* (available at http://curia.europa.eu/jcms/upload/docs/application/pdf/2011-03/cp110012en.pdf).

[18] Finkelstein, A and K McGarry (2006) "Multiple dimensions of private information: evidence from the long-term care insurance market", *American Economic Review* **96**, 938–958.

[19] Hartog, J, A Ferrer-i-Carbonell and N Jonker (2000) "Linking measured risk aversion to individual characteristics", *Kyklos* **55**, 3–26.

[20] Hemenway, D (1990) "Propitious selection", *Quarterly Journal of Economics* **105**, 1063–1069.

[21] Hoy, M (1982) "Categorizing risks in the insurance industry", *Quarterly Journal of Economics* **97**, 321–336.

[22] Jullien, B, B Salanie and F Salanie (2007) "Screening risk averse agents under moral hazard", *Economic Theory* **1**, 151–169.

[23] Kimball, M, C Sahm and M Shapiro (2008) “Imputing risk tolerance from survey responses”, *Journal of the American Statistical Association* **103**, 1028–1038.

[24] Landsberger, M and I Meilijson (1994) “Monopoly insurance under adverse selection when agents differ in risk aversion”, *Journal of Economic Theory* **63**, 392–407.

[25] Landsberger, M and I Meilijson (1996) “Extraction of surplus under adverse selection: the case of insurance markets”, *Journal of Economic Theory* **69**, 234–239.

[26] McCarthy, D and O Mitchell (2003) “International adverse selection in life insurance and annuities”, NBER Working Paper 9975.

[27] Netzer, N and F Scheuer (2007) “Competitive Screening in Insurance markets with endogenous wealth heterogeneity”, mimeo, University of Konstanz, 2007.

[28] Olivella, P and F Schroyen (2011) “The principal–agent risk sharing problem with two-dimensional hidden information: a complete characterisation”, mimeo, Norwegian School of Economics.

[29] Olivella P and M Vera-Hernández (2011) “Testing for asymmetric information in private health insurance”, mimeo, Universitat Autónoma de Barcelona.

[30] Picard, P (2000) “Economic analysis of insurance fraud”, In G. Dionne (ed.) *Handbook of Insurance*. New York: North-Holland, 2000.

[31] Rothschild, M and J Stiglitz (1976) ”Equilibrium in competitive insurance markets: an essay on the economics of imperfect information”, *Quarterly Journal of Economics* **90**, 629–649.

[32] Smart, M (2000) ”Competitive insurance markets with two unobservables”, *International Economic Review* **41**, 153–169.

[33] Stiglitz, J (1977) ”Monopoly, non-linear pricing and imperfect information: the insurance market”, *Review of Economic Studies* **44**, 407–430.

[34] Villeneuve, B (2003) ”Concurrence et antisélection multidimensionelle en assurance”, *Annales d’Économie et de Statistique* **69**, 119–142.

[35] Villeneuve, B (2000). "The consequences for a monopolistic insurer of evaluating risk better than customers: the adverse selection hypothesis reversed", *The Geneva Papers on Risk and Insurance Theory* **25**, 65–79.

[36] Wambach, A (2000) "Introducing heterogeneity in the Rothschild–Stiglitz model", *Journal of Risk and Insurance* **67**, 579–591.

Appendix

Proof of Lemma 2.

Suppose that full separation under Order 2 is optimal. This situation is depicted in Figure 14.

–Figure 14 here–

First note that c_{LL} must exceed $\frac{\Delta\mu}{\Delta\nu} = D\frac{x}{1-x}$ because, otherwise, LH and HL could not have been separated.

The profits from the different types are as follows.

$$\begin{aligned}
\pi_{HH} &= \frac{1}{2}(1-c_{HH}^2)\nu_H - (1-c_{LH})\Delta\mu + (1-c_{HL})\Delta\mu - \frac{1}{2}(1-c_{HL}^2)\Delta\nu - (1-c_{LL})\Delta\mu \\
\pi_{HL} &= \frac{1}{2}(1-c_{HL}^2)\nu_L - (1-c_{LL})\Delta\mu \\
\pi_{LH} &= \frac{1}{2}(1-c_{LH}^2)\nu_H + (1-c_{HL})\Delta\mu - \frac{1}{2}(1-c_{HL}^2)\Delta\nu - (1-c_{LL})\Delta\mu \\
\pi_{LL} &= \frac{1}{2}(1-c_{LL}^2)\nu_L
\end{aligned}$$

Weighting with the respective population proportions gives the following first derivatives with respect to the co-insurance rates.

$$\begin{aligned}
\frac{\partial\pi_{tot}}{\partial c_{HH}} &= -\alpha_{HH}c_{HH}\nu_H, \frac{\partial\pi_{tot}}{\partial c_{LH}} = \alpha_{HH}\Delta\mu - \alpha_{LH}\nu_H c_{LH}, \\
\frac{\partial\pi_{tot}}{\partial c_{HL}} &= -\alpha_{.H}\Delta\mu + \alpha_{.H}c_{HL}\Delta\nu - \alpha_{HL}c_{HL}\nu_L, \\
\frac{\partial\pi_{tot}}{\partial c_{LL}} &= (1-\alpha_{LL})\Delta\mu - \alpha_{LL}c_{LL}\nu_L
\end{aligned}$$

The solution for c_{LL} is $c_{LL} = \min\{D\frac{1-\alpha_{LL}}{\alpha_{LL}}, 1\}$. A necessary condition that $c_{LL} > D\frac{x}{1-x}$ is $x < 1-\alpha_{LL}$. Only if $c_{LL} > D\frac{x}{1-x}$ is there room to separate LH from HL. Because

$$\frac{\partial\pi_{tot}}{\partial c_{HL}} = -\alpha_{.H}\Delta\mu + [\alpha_{.H}(1-x) - \alpha_{HL}x]\nu_H c_{HL},$$

total profit is strictly concave in c_{HL} if and only if $x \geq \frac{\alpha_{\cdot H}}{1-\alpha_{LL}}$. In that case, the optimal solution for c_{HL} is $c_{HL} = \min\{D\frac{\alpha_{\cdot H}x}{\alpha_{\cdot H}(1-x)-\alpha_{HL}x}, 1\}$.

By monotonicity, the only possibility for full separation arises when $c_{HL} = D\frac{\alpha_{\cdot H}x}{\alpha_{\cdot H}(1-x)-\alpha_{HL}x} < 1$. It remains then to check whether $c_{HL} < c_{LL}$. Suppose first that $c_{LL} = D\frac{1-\alpha_{LL}}{\alpha_{LL}} < 1$. Then,

$$c_{HL} < c_{LL} \Longleftrightarrow x < \frac{\alpha_{\cdot H}(1-\alpha_{LL})}{\alpha_{\cdot H}\alpha_{LL} + (1-\alpha_{LL})^2}.$$

Because $\frac{\alpha_{\cdot H}(1-\alpha_{LL})}{\alpha_{\cdot H}\alpha_{LL}+(1-\alpha_{LL})^2} < \frac{\alpha_{\cdot H}}{1-\alpha_{LL}}$, this condition contradicts the assumption that $x \geq \frac{\alpha_{\cdot H}}{1-\alpha_{LL}}$. Suppose next that $c_{LL} = 1$. Then,

$$c_{HL} < c_{LL} \Longleftrightarrow x < \frac{\alpha_{\cdot H}}{1-\alpha_{LL} + D\alpha_{\cdot H}}.$$

Again, this contradicts the assumption that $x \geq \frac{\alpha_{\cdot H}}{1-\alpha_{LL}}$. Hence, $c_{HL} = c_{LL}$, meaning that HL is pooled with LL.

On the other hand, if total profit is strictly convex in c_{HL}, it pays to move c_{HL} either down to c_{LH} or up to c_{LL}. Hence, full separation is never optimal. ■

Proof of proposition 3, part (ii).

Under **Regime EI**, $c^E_{\cdot L} = D\frac{x\alpha_{HL}}{x-\alpha_{\cdot H}}$, with $x > \alpha_{\cdot H} = \alpha_{LH} + \alpha_{HH}$. Using the definition of α_{LH}, this condition on x is equivalent to $x > \frac{\alpha_{HH}}{\alpha_{H\cdot}} - \frac{\rho}{\alpha_{H\cdot}}$. Therefore,

$$\rho > -\alpha_{H\cdot}\left(x - \frac{\alpha_{HH}}{\alpha_{H\cdot}}\right). \tag{27}$$

In addition, because $\rho \leq 0$, a necessary condition on x is that

$$x > \frac{\alpha_{HH}}{\alpha_{H\cdot}}. \tag{28}$$

Substituting the co-insurance rates into the covariance formula for the expressions given by (23), and making use of the formulae for α_{HL}, α_{LH} and α_{LL} ((17)-(19), enables one to write the covariance between coverage and risk as

$$cov = \frac{\rho + \alpha_{HH}\left(x - \frac{\alpha_{HH}}{\alpha_{H\cdot}}\right)}{\rho + \alpha_{H\cdot}\left(x - \frac{\alpha_{HH}}{\alpha_{H\cdot}}\right)}\alpha^2_{H\cdot}(\Delta\mu)^2\frac{x}{\nu_L}.$$

Consider now Figure 9 in the main text. Let D be small enough for Regime EI to prevail. Given (28), all the terms in the round brackets in the above expression are positive. From (27), the denominator is positive. Therefore,

$$cov < 0 \Longleftrightarrow \rho < -\alpha_{HH}\left(x - \frac{\alpha_{HH}}{\alpha_{H\cdot}}\right).$$

■

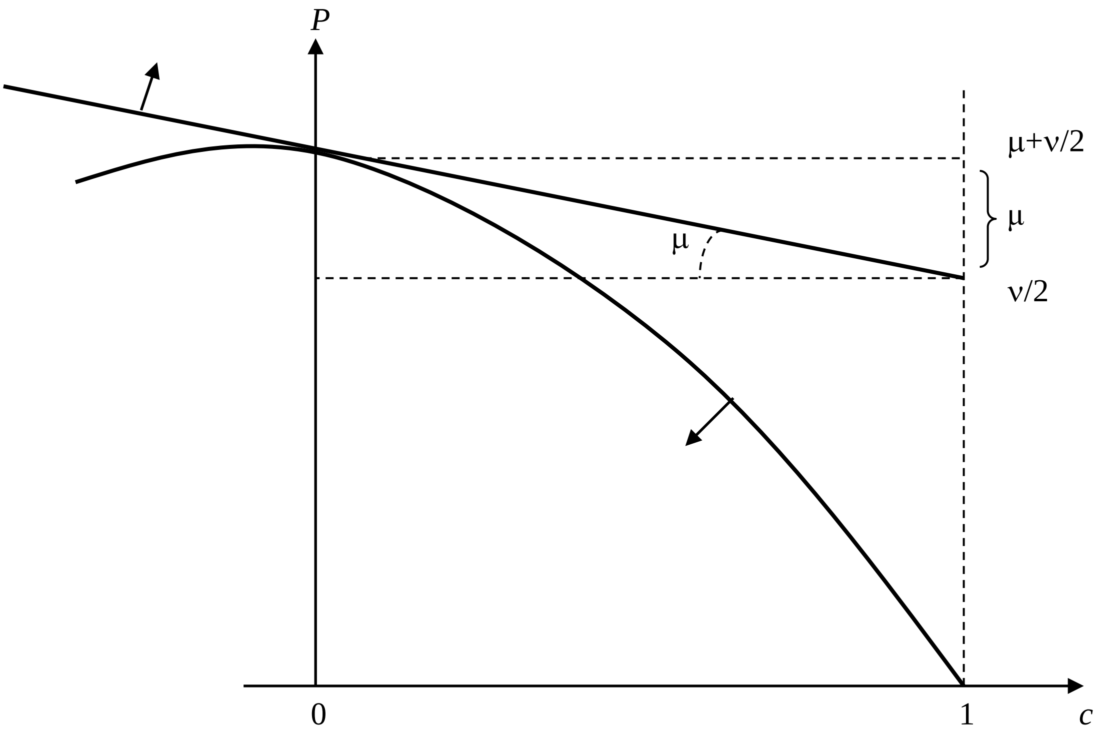

Figure 1. An indifference curve and an iso-profit line in the (*c*,*P*)-space

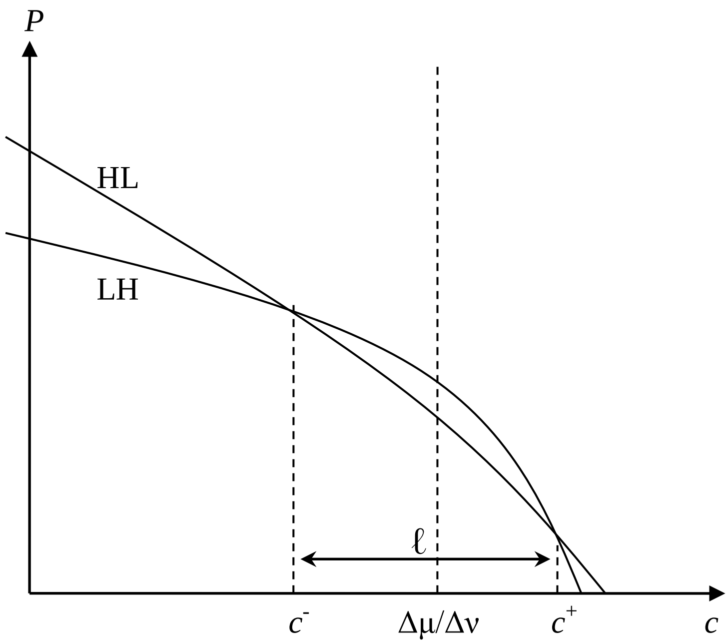

Figure 2. The indifference curves of HL and LH cross twice

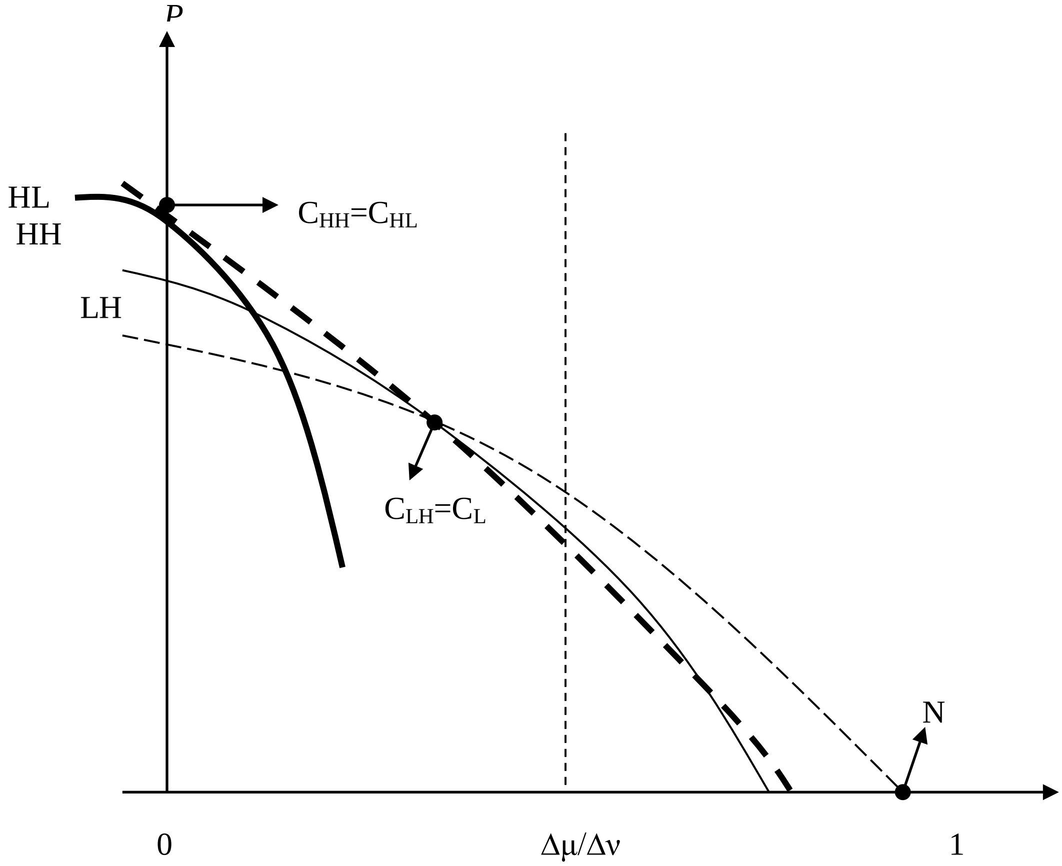

Figure 3. Regime A

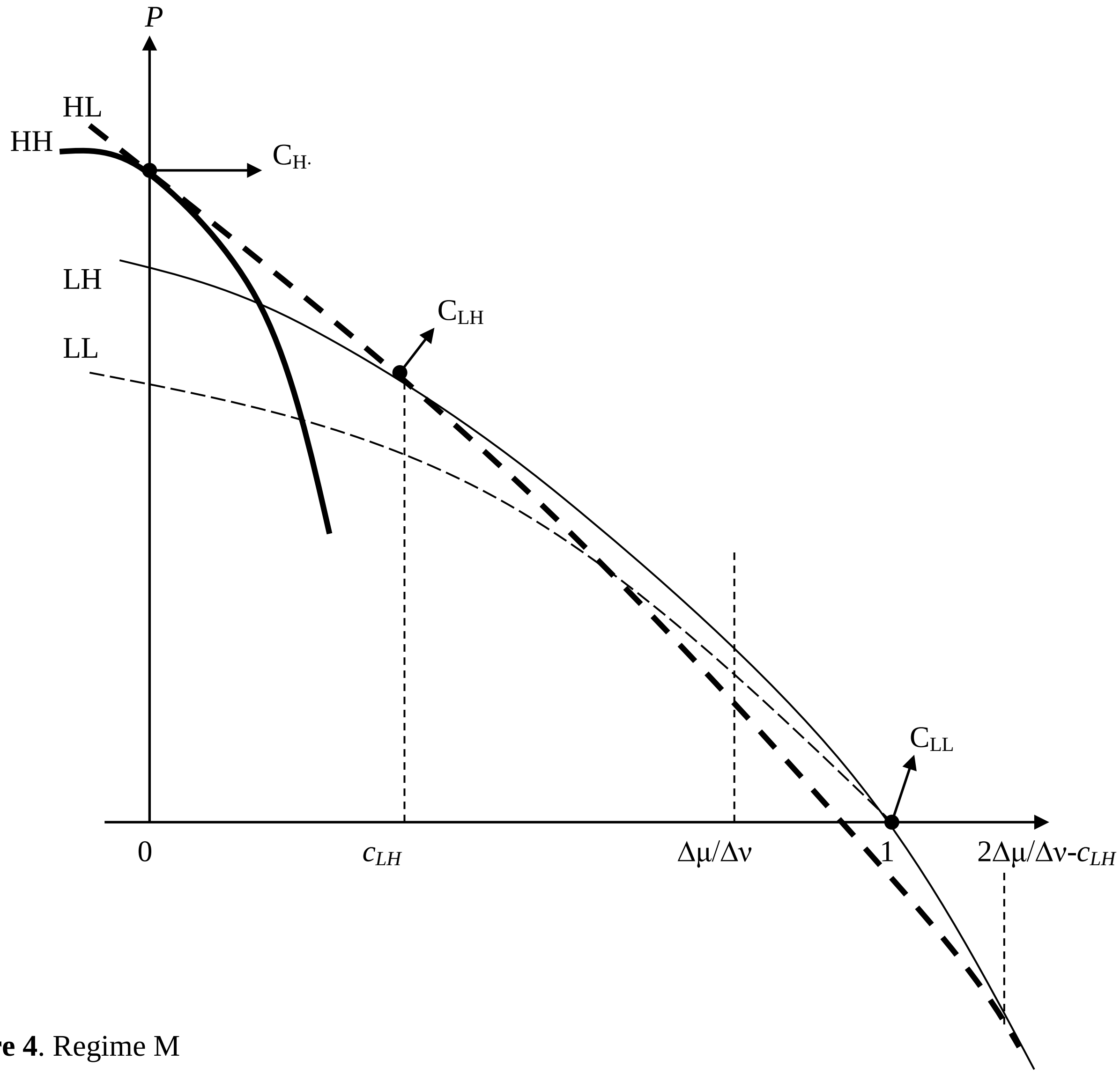

Figure 4. Regime M

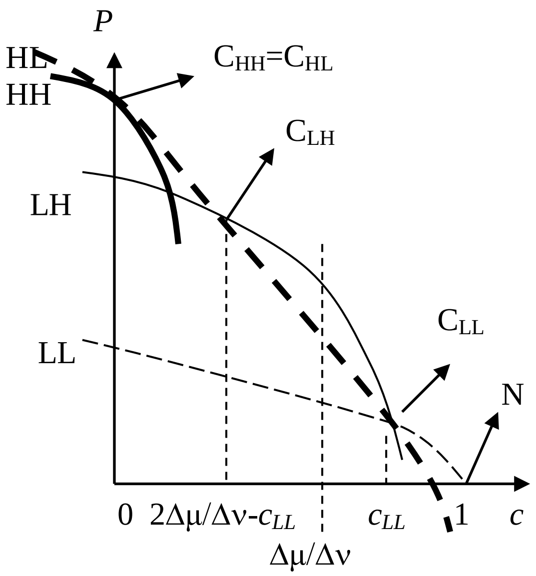

Figure 5a. Regime BpI

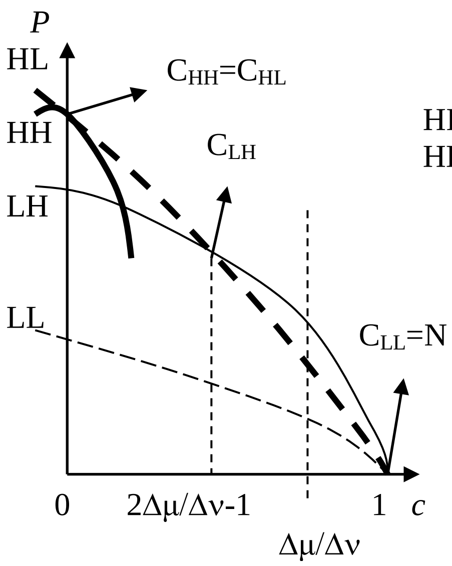

Figure 5b. RegimeBpX

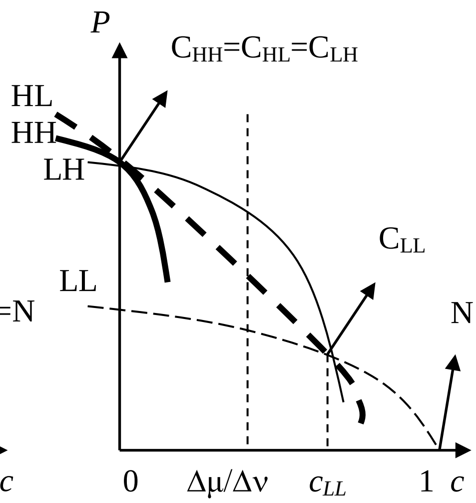

Figure 5c. Regime Bf

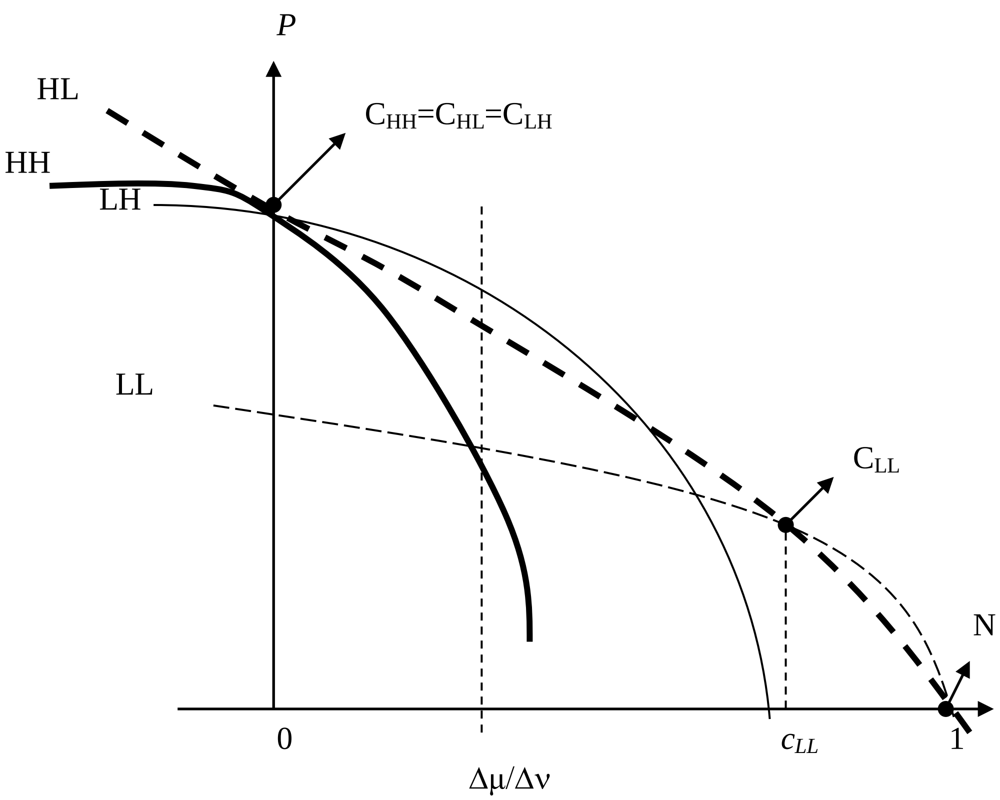

Figure 6. Regime C

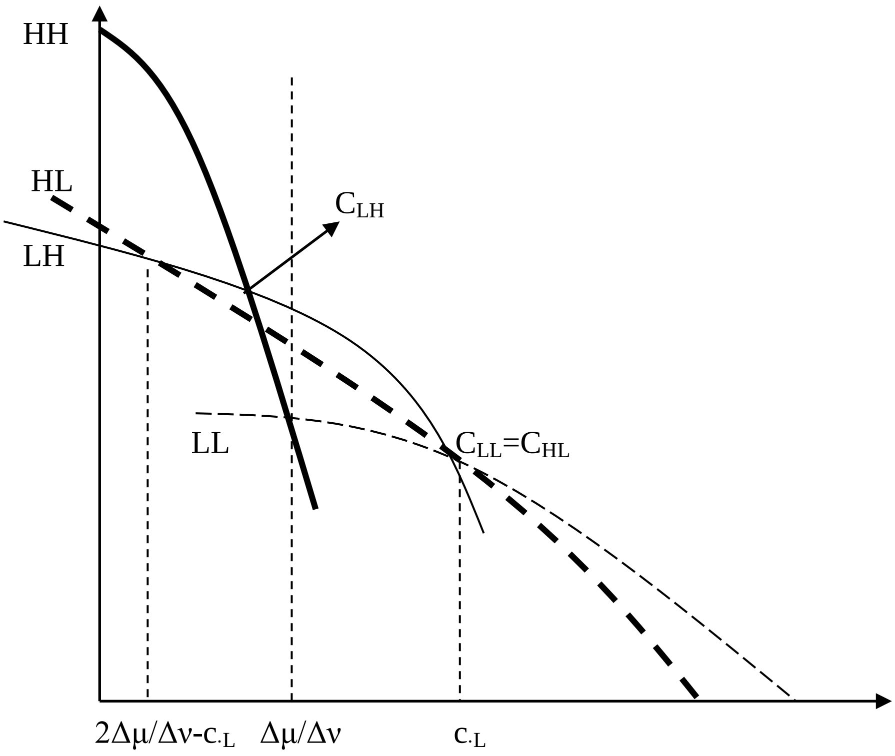

Figure 7. Regime E

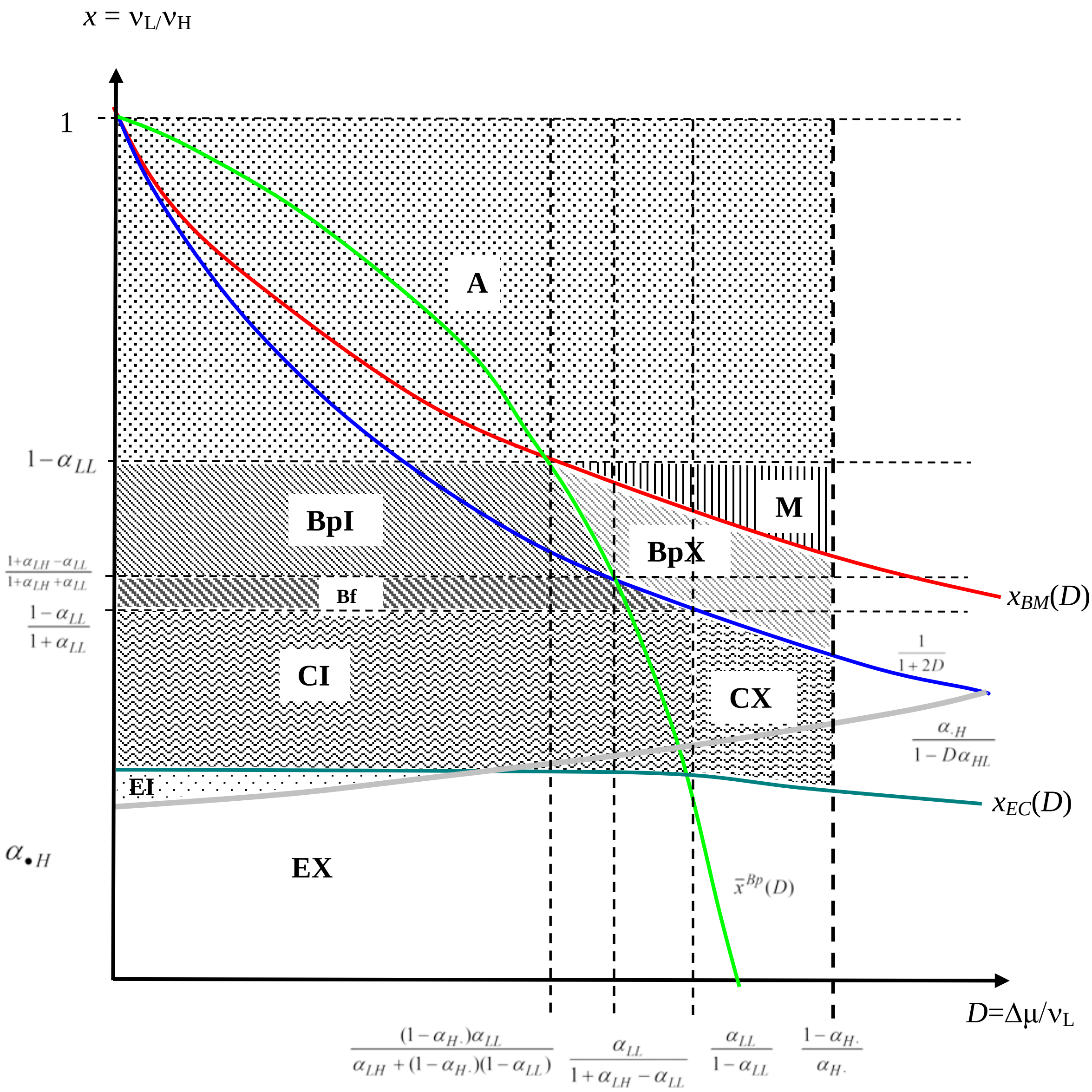

Figure 8. Optimal regimes in the (D,x)-space. x_{BM} (D) and x_{CE}(D) are explained in the text. $\bar{x}^{Bp}(D)$ is found by setting the expression for c_{LL}^{B} for Regime BpI (see eq (22)) equal to 1

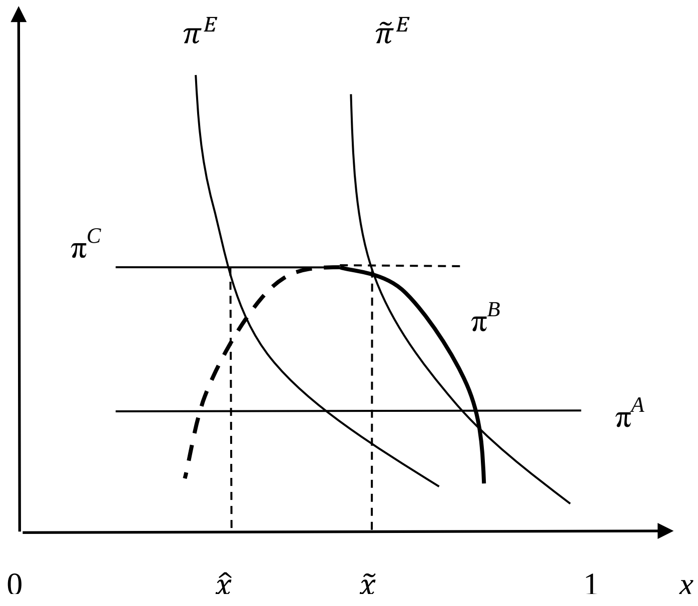

Figure 9. When profits of Regime E are given by $\tilde{\pi}^E$, Regime C is entirely dominated by Regime E.

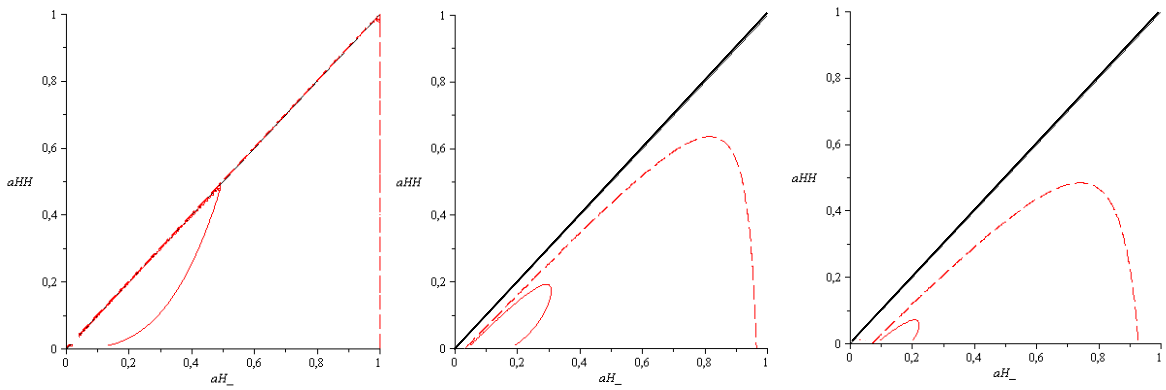

Figure 10. For ($\alpha_{H\cdot}$,α_{HH})-values inside the solid line area, Regime C ceases to occur when ρ equals 0 (a), -.0333 (b), -.0666 (c). The dashed line area gives all combinations of ($\alpha_{H\cdot}$,α_{HH}) such that for the selected ρ the triplet ($\alpha_{H\cdot}$,α_{HH},ρ) belongs to $\mathcal{A}_o$.

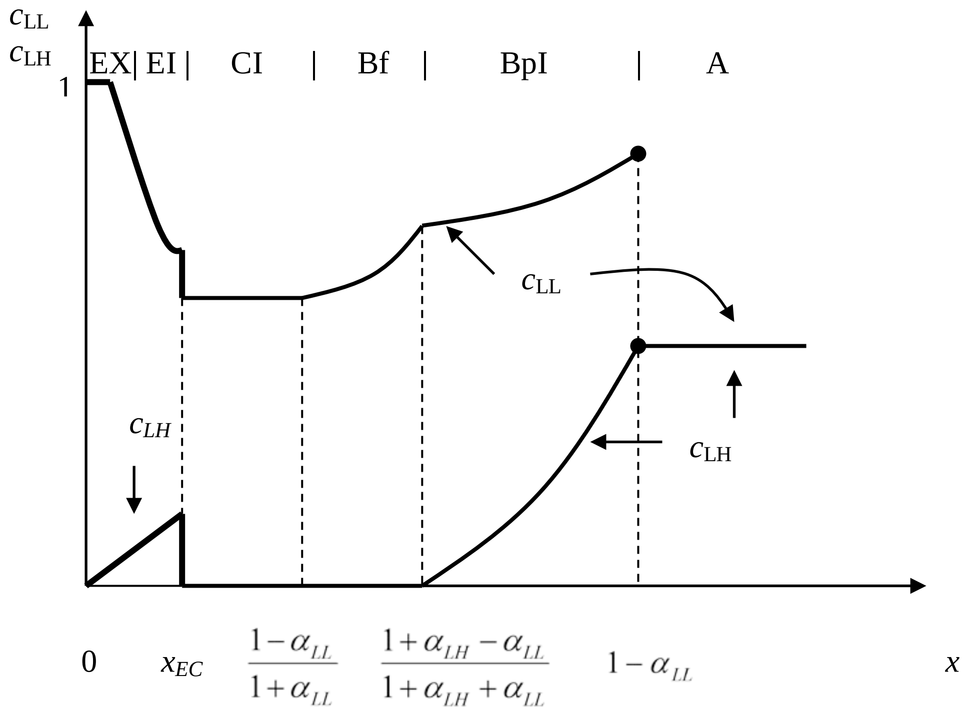

Figure 11. Optimal coinsurance rates for *LH* and *LL* as a function of *x*

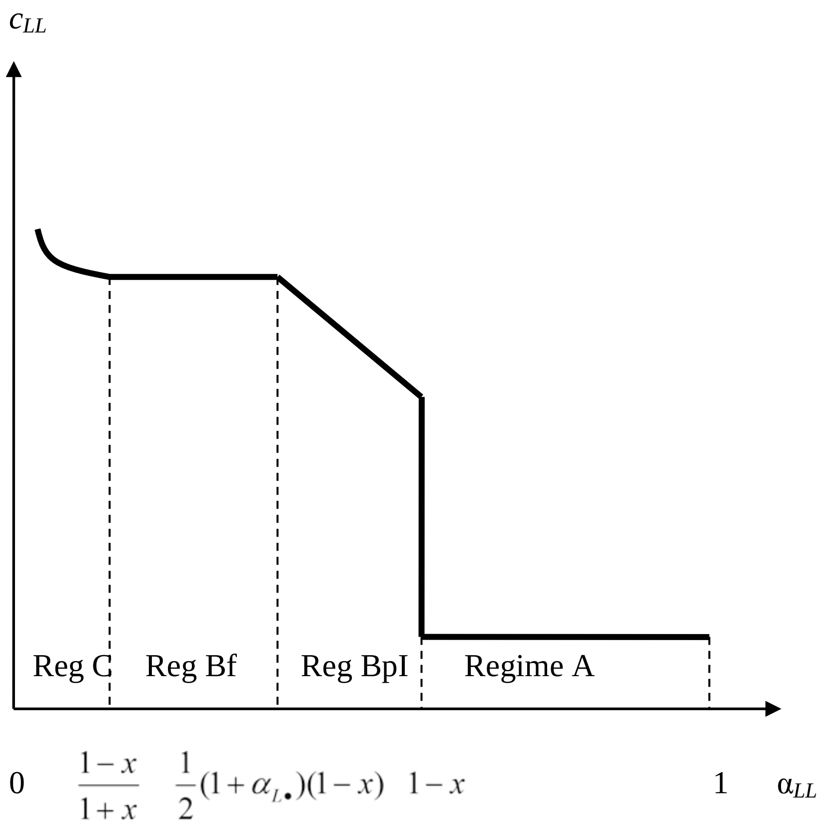

Figure 12. The optimal coinsurance rate for LL as a function of α_{LL}

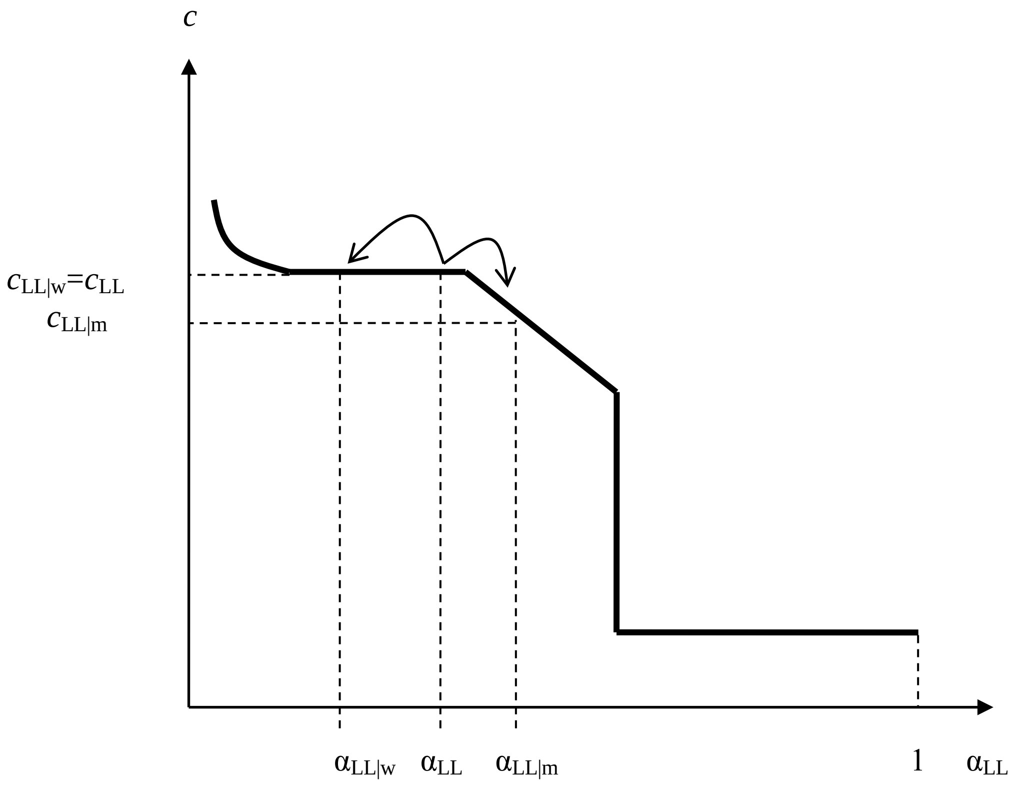

Figure 13. *A priori* and updated probability of type *LL* and corresponding coinsurance rates

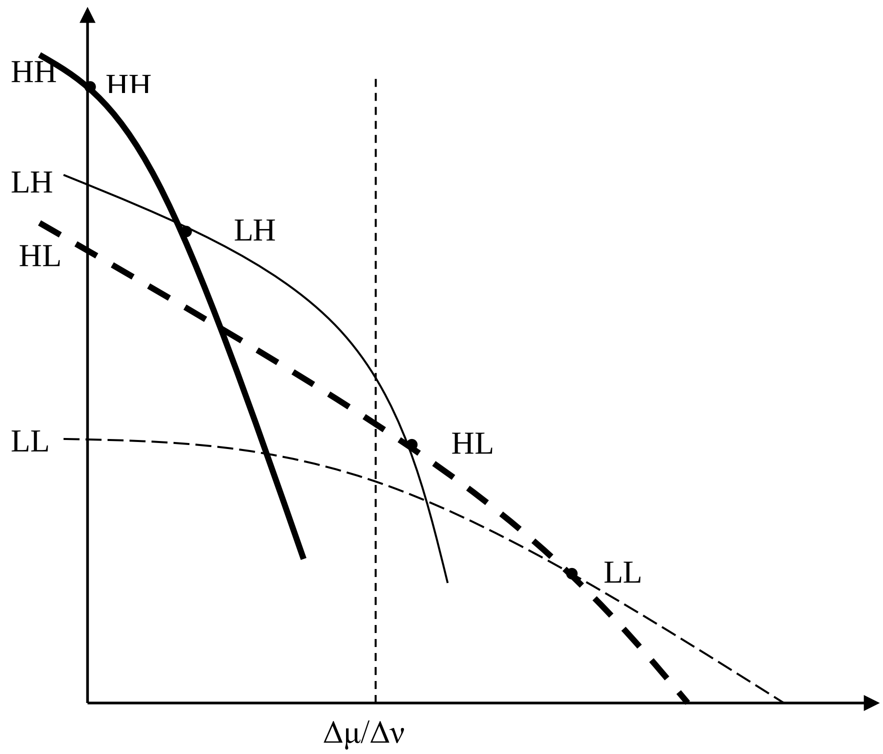

Figure 14. A full-separation menu under Order 2