Economics of insurance

Renaud Bourlès

2020-2021

This course is mainly addressed to third year students at Centrale Marseille ("Data and Economic and Financial Decisions" track) and to Master 2 students at Aix-Marseille University (AMSE and ISMA Masters).

This aim of these lectures is to provide with the basis of the theoretical analysis of insurance pricing. We will first analyze the simple case in which all the individuals are identical (Chapter 1) before studying the implications of heterogeneity in terms of risk exposure, be it observable (Chapter 2) or not (Chapter 3). This analysis will be supplemented by a study of insured's behavior in presence of insurance (Chapter 4) and some extensions through exercises (Chapter 5).

Contents

Bibliography

- Henriet, D. and Rochet, J.-C., "Microéconomie de l'assurance", Economica, 1990.
- Picard, P., "Economic Analysis of Insurance Fraud", Handbook of Insurance, 2nd edition, G. Dionne. (ed), Springer, 2013.
- Rees, R. and Wambach, A., "The Microeconomics of Insurance", Foundations and Trends in Microeconomics, 4(1-2), 2008.
- Schlessinger, H., "The Theory of Insurance Demand", Handbook of Insurance, G. Dionne. (ed), Kluwer, 2000.

Chapter 1

The single risk model

To start apprehending the pricing of insurance contract, we begin by considering that all the potential insureds are identical, and in particular that they all face the same risk. This model, called Mossin's model, allows to introduce the notions of full insurance, partial insurance and deductible; and to analyze how insurance demand varies with price and revenue.

1.1   Mossin's model

Consider an individual, with revenue R, who faces a risk of accident – that corresponds to a monetary loss or damage D – that can occur with probability p. Denote by $u(.)$ her vNM utility function (strictly increasing, concave and C^2).
Without insurance, the expected utility of such an individual would write:

$$V_0 = pu(R - D) + (1 - p)u(R)$$

In that context, an insurance contract is a couple $z = (\Pi, q)$, and corresponds to the payment of a coverage q in case of damage against a premium Π.

We assume $q \in [0, D]$. Full coverage will correspond to the case $q = D$, whereas if $q < D$ the quantity $D - q$ will be called the deductible. The expected utility of an individual buying a contract (Π, q) therefore equals:

$$V(\Pi, q) = pu(R - D + q - \Pi) + (1 - p)u(R - \Pi)$$

with $R - D + q - \Pi \equiv A$ the wealth if the damage occurred and $R - \Pi \equiv N$ the wealth if it didn't.

The assumption $q \leq D$ implies $A \leq N$ and make sure that individuals don't intentionally provoke or overstate damage (we will come back to those phenomenons – called moral hazard – in chapter 4).

To represent these contracts and the corresponding expected utility in the plan (q, Π) let us study the effects of the terms of the contact on $V(.)$:

$$\frac{\partial V(\Pi, q)}{\partial \Pi} = -[pu'(R - D + q - \Pi) + (1 - p)u'(R - \Pi)] < 0, \text{ and}$$

$$\frac{\partial V(\Pi, q)}{\partial q} = pu'(R - D + q - \Pi) > 0$$

In the (coverage, premium) plan, expected utilities are therefore increasing to the South-East. The expected utility increases with coverage and decreases with premium. Moreover, the indifference curves (i.e. the set of contracts providing with the same level of expected utility) are increasing and concave in the (q, Π) plan. Using the implicit function theorem $\left(\left. \frac{\partial \Pi}{\partial q} \right|_{V(\Pi, q) = k} = - \frac{\partial V / \partial q}{\partial V / \partial \Pi} \right)$ we indeed have :

$$\left. \frac{\partial \Pi}{\partial q} \right|_{V(\Pi, q) = k} = \frac{pu'(R - D + q - \Pi)}{pu'(R - D + q - \Pi) + (1 - p)u'(R - \Pi)} > 0 \qquad (*)$$

$$\text{and } \left. \frac{\partial^2 \Pi}{\partial q^2} \right|_{V(\Pi, q) = k} = \frac{p(1 - p)u''(R - D + q - \Pi)u'(R - \Pi)}{[pu'(R - D + q - \Pi) + (1 - p)u'(R - \Pi)]^2} < 0$$

An individual will then accept a contract $z = (\Pi, q)$ if it provides him with higher expected utility than without insurance. Plotting the indifference curve $V(\Pi, q) = V(0, 0)$ gives us the set of acceptable contracts from the insureds' point of view.
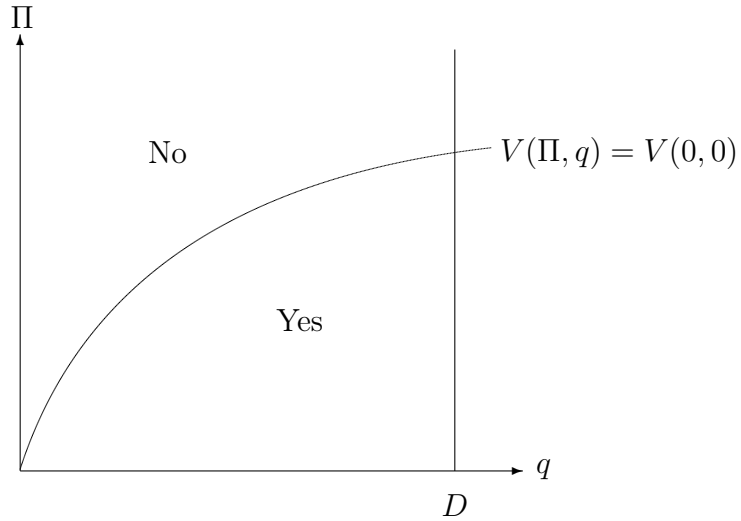
Let's now analyze the behavior of the insurance company. For each insurance policy sold, the benefit is the premium Π whatever the state (be there a damage or not), whereas the cost amounts to the coverage only in case of damage. Assuming that there may exist a transaction cost λ (called loading factor) the effective effective cost is $(1 + \lambda)q$ in case of damage. For each sold policy, the expected profit of the insurer therefore writes $\Pi - p(1 + \lambda)q$.

The area in which the insurance transaction is beneficial to both parties can then be represented in the (q, Π) plan as follows :
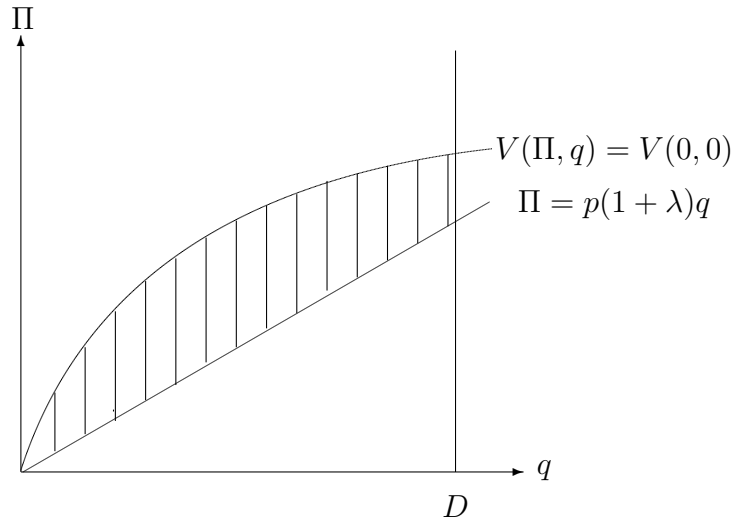
Consider first $\lambda = 0$. The slope of the isoprofit curve (i.e. the set of contracts leading to the same level of expected profit) is then equal to p. Every euros of coverage costs, in expectation, $p \in$ to the insurer. Remark now that the willingness to pay of the insured (the extra premium an individual is ready to pay for an extra euro of coverage) is everywhere greater than p except for $q = D$ (where it equals p). Indeed,

- At $q = D$,

$$u(A) = u(N) \Rightarrow \left.\frac{\partial \Pi}{\partial q}\right|_{V(\Pi,q)=k} = -\frac{\partial V/\partial q}{\partial V/\partial \Pi} = p$$

- whereas if $q < D$,

$$
\begin{aligned}
A < N \quad &\Rightarrow \quad u'(A) > u'(N) \text{ as } u \text{ is concave.} \\
&\Rightarrow \quad pu'(A) + (1-p)u'(N) < u'(A) \\
&\Rightarrow \quad \left.\frac{\partial \Pi}{\partial q}\right|_{V(\Pi,q)=k} = \frac{\partial V/\partial q}{\partial V/\partial \Pi} > p
\end{aligned}
$$

Therefore the insurer would gain by increasing the coverage up to $q = D$ as, for $q < D$, the insured is willing to pay more for a extra euro of coverage than what it costs to the insurer.

This gives the central result as Mossin's model:

Theorem. *In the absence of loading factor ($\lambda = 0$), full insurance is optimal.*

Remark. *This holds as well under monopoly (maximal profit) as under perfect competition (zero profit).*

Under monopoly, the insured has the same ex-post revenue in both states of nature (damage / no damage) as $q = D$ and the same expected utility than without insurance (no surplus). We then talk about **certainty equivalent**.

Definition. *The certainty equivalent of a risk $\widetilde{x}$ is the constant:*

$$ e(\widetilde{x}) \quad / \quad \mathbb{E}(u(\widetilde{x})) = u(e(\widetilde{x})) $$

In our case, it is therefore $e / pu(R - D) + (1 - p)u(R) = u(e)$, that is $e = R - \Pi_{\max}$.

The **risk premium** is then defined as the difference between expected wealth and the certainty equivalent (it is the amount the insured is willing to pay to be fully insured). Here, the expected revenue is $\mathbb{E} = p(R - D) + (1 - p)R = R - pD$ and the risk premium equals $R - pD - R + \Pi_{\max} = \Pi_{\max} - pD$.

Π
Π_max
p
risk premium
pD
q
D
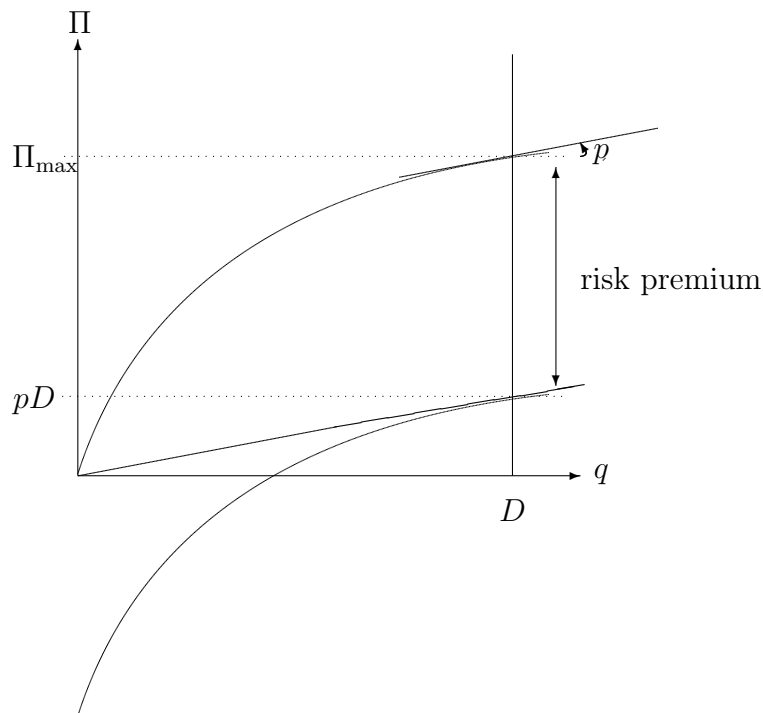
If we now consider the loading factor λ to be strictly positive (*i.e.* that it exists transaction costs) the slope of the isoprofit curve is greater than p (it equals $(1 + \lambda)p$) and the optimum is reached for $q < D$.

Theorem. *We then can complete the previous theorem*

- *In the absence of loading factor ($\lambda = 0$), full insurance is optimal.*

- *If the loading factor is strictly positive ($\lambda > 0$), partial insurance only is optimal ($q < D$).*

Remark. *The model can be simplified by assuming that the premium is a linear function of coverage: $\Pi = \pi q$, as it is for example the case under competition ($\pi = (1 + \lambda)p$). With this simplification, the choice of the insured amounts to choosing the coverage that maximizes her expected utility:*

$$V(q) = pu(R - D + (1 - \pi)q) + (1 - p)u(R - \pi q)$$

meaning that she chooses q^* such that:

$$V'(q^*) = p(1 - \pi)u'(R - D + (1 - \pi)q^*) - (1 - p)\pi u'(R - \pi q^*) = 0 \qquad \text{(CPO)}$$

(the concavity of $u(\cdot)$ ensuring the second order condition). We then obtain

$$u'(R - D + (1 - \pi)q^*) = \frac{\pi}{p}\frac{1 - p}{1 - \pi}u'(R - \pi q^*)$$

and recover previous theorem

- *if $\pi = p$ (that is $\lambda = 0$ under competition) then $q^* = D$*

- *if $\pi > p$ (that is $\lambda > 0$ under competition) then $q^* < D$*
  (as $u(\cdot)$ is concave)

This simplification also allows to study the link between the optimal demand for insurance (q^*) and exogenous variables as the revenue R, the damage probability p, the size of the loss D, or even the price of insurance (π, made exogenous under this simplification). Indeed, the optimality condition (CPO) can be written as:

$$F(q^*; p, D, R, \pi) \equiv p(1 - \pi)u'(R - D + (1 - \pi)q^*) - (1 - p)\pi u'(R - \pi q^*) = 0$$

with $\frac{\partial F}{\partial q^*} < 0$.

Therefore, using the implicit function theorem, we are able to study the impact on q^* of changes in p, D, R and π.

1.2 Wealth effect

Let's study first how the optimal coverage varies with revenue R. We have:

$$\frac{dq^*}{dR} = -\frac{\partial F/\partial R}{\partial F/\partial q^*}$$

and

$$\frac{\partial F}{\partial R} = p(1-\pi)u''(A^*) - \pi(1-p)u''(N^*)$$

with $A^* = R - D + (1-\pi)q^*$ and $N^* = R - \pi q^*$. Thus, $\frac{\partial F}{\partial R}$ and $\frac{dq^*}{dR}$ cannot be signed in general. However, using the optimality condition (CPO), one gets:

$$p(1-\pi) = \pi(1-p)\frac{u'(N^*)}{u'(A^*)}$$

and we can write:

$$\frac{\partial F}{\partial R} = \pi(1-p)u'(N^*)\left(\frac{u''(A^*)}{u'(A^*)} - \frac{u''(N^*)}{u'(N^*)}\right)$$

Therefore $\frac{\partial F}{\partial R} \geq 0$ if and only if $-\frac{u''(N^*)}{u'(N^*)} \geq -\frac{u''(A^*)}{u'(A^*)}$. As $N^* \geq A^* \ \forall q^* \leq D$, we can state the following result:

Proposition. *The optimal coverage is increasing with wealth (and insurance is a normal good) if the index of absolute risk aversion is (weakly) increasing with wealth. On the contrary, if the index of absolute risk aversion is strictly decreasing, the optimal coverage is strictly decreasing with wealth, and insurance is an inferior good.*

Given that this last assumption is the one generally chosen to model behavior under risk, this would suggest that insurance is an inferior good. The intuition is pretty simple : if the propensity to take risk increases with revenue then, everything else being equal, the demand for insurance should decrease with wealth.
However, this doesn't seem to be the case in reality and most of the statistical evidences show that the demand for insurance increases with wealth. This nevertheless doesn't necessarily contradict the previous theoretical result that has been established "everything else being equal". Indeed, in real life, the size of the loss D is likely to be increasing with wealth. As, everything else being equal, the demand for insurance is increasing with the size of the loss ($\partial F/\partial D = -p(1-\pi)u''(A^*) > 0$ and $dq^*/dD = -\frac{\partial F/\partial D}{\partial F/\partial q} > 0$), the above empirical observation can be rationalized with the present model.

1.3 Price effect

Let's now analyze how insurance demand varies with the unit price of coverage, or premium rate: π.

$$\frac{\partial F}{\partial \pi} = -pu'(A^*) - (1-p)u'(N^*) + q\left[(1-p)\pi u''(N^*) - p(1-\pi)u''(A^*)\right]$$

Again, the effect is ambiguous. One cannot be sure that insurance demand decreases when the premium per unit of coverage increases. However,

$$\frac{dq^*}{d\pi} = -\frac{\partial F/\partial \pi}{\partial F/\partial q} = \frac{pu'(A^*) + (1-p)u'(N^*)}{\partial F/\partial q} + q^* \frac{\partial F/\partial R}{\partial F/\partial q} = \frac{pu'(A^*) + (1-p)u'(N^*)}{dF/dq} - q\frac{\partial q^*}{\partial R}$$

meaning that one can isolate a (negative) "substitution effect" from a "wealth effect". If $\frac{dq^*}{dR} > 0$, then $\frac{dq^*}{d\pi} < 0$. Using the previous proposition, we end up with:

Proposition. *The demand for insurance decreases with the premium rate if the index of absolute risk aversion is increasing or constant (with wealth). However, when risk aversion is (sufficiently) decreasing, the demand can increase with price (insurance is then a Giffen good).*

Again, the intuition for this result is pretty simple: an increase in premium increases the relative price of state A ("accident") with respect to state N ("no accident"), what – for constant utility – tends to decrease the demand for coverage. However, the increase in premium also leads to a decrease in wealth (for a given level of coverage) and therefore increases the demand for insurance if risk aversion is decreasing in wealth.

One can moreover find another sufficient condition for q^* to be decreasing in π writing $\frac{\partial F}{\partial \pi}$ as:

$$\begin{aligned}\frac{\partial F}{\partial \pi} &= pu'(A^*)\left[-1 - \frac{(1-\pi)qu''(A^*)}{u'(A^*)}\right] + (1-p)u'(N^*)\left[-1 + \frac{q\pi u''(N^*)}{u'(N^*)}\right] \\ &= pu'(A^*)\left[-1 - \frac{A^*u''(A^*)}{u'(A^*)}\right] + (1-p)u'(N^*)\left[-1 - \frac{N^*u''(N^*)}{u'(N^*)}\right] \\ &\quad + pu''(A^*)(R-D) + (1-p)u''(N^*)R\end{aligned}$$

Thus, if the index of relative risk aversion $\left(-\frac{cu''(c)}{u'(c)}\right)$ is always lower than 1 (what is mostly invalidated by recent studies[1]), demand for insurance decreases with the "premium rate".

Again, these comparative statics results are computed assuming that the other exogenous variables are kept constant. In particular, variations in premium are supposed to be independent from variations in the probability of damage. Yet, these two variables are certainly not independent due to supply effects (see above). Moreover, holding the premium rate constant, an increase in the probability of damage increases the demand for insurance ($\frac{\partial F}{\partial p} > 0$). Therefore, accounting for the response of supply (via π), an increase in the probability of damage will in general have an ambiguous effect on the amount of coverage bought.

---

[1]cf. Meyer, D. and Meyer, J. (2005), "Relative Risk Aversion: What Do We Know?", *The Journal of Risk and Uncertainty*, 31(3), 243-262 or Chetty, R. (2006), "A New Method of Estimating Risk Aversion", *The American Economic Review*, 96(5), 1821-1834.

Chapter 2

Product differentiation

The reality of insurance markets is of course far more complex than the previous model, with a large number of contracts and different pricing. For now, we will try to understand the causes of such a diversity by introducing alternately heterogeneity among agents and asymmetric information. We will also study how an insurance company accounts for the differences among insureds and how it tries to spot these differences.

2.1   Introducing heterogeneity in Mossin's model

Let us consider a model in which a population of individuals faces the same risk of damage, represented by a monetary loss D. Individuals only differ regarding the probability of damage p_i. They are identical with respect to other parameters. They have the same vNM utility function $u(.)$ and the same initial revenue R. Heterogeneity in the probability of damage affects both the cost for the insurer and the preferences of the agents:

$$V(p_i, z) = p_i u(R - D + q - \Pi) + (1 - p_i)u(R - \Pi)$$

If we assume that both the insurer and the insured have full information on the probability of damage and that the loading factor is null, the results of the previous section are easily extendable:

- Under perfect competition($\max_z V(p_i, z)$ under zero profit constraint), type i agents pay a premium $\Pi = p_i q$ for a coverage q and optimally choose full insurance $q = D$

- In a monopolistic setting, the insurance company offers to type i agents the contract that maximizes its profit among those acceptable for the agents ($\max_{q,\Pi} \Pi - p_i q \ / \ V(p_i, z) \geq V_0$). We can easily show that the solution corresponds to full insurance: $q = D$ at a price such that the constraint is binding.

In these two cases, there are as many prices as there are agent types.

Previous reasoning however assumes that the insurer perfectly knows the probability of damage for all the insureds. This is of course not the case in reality. The insurers still developed methods (called scoring methods) to evaluate these probabilities based on insureds' observable characteristics (age, gender, wealth, socio-economic classification, car's power for automobile insurance...).

2.2 Measuring the probability of damage: scoring methods

Scoring methods are statistical methods that allow to assign ex-ante a probability of damage to an incoming insured. These methods therefore enable insurers to estimate the p_i's of the previous model. Using the history of previous contracts, insurance companies evaluate the impact of each of the observable variables on the probability of damage.
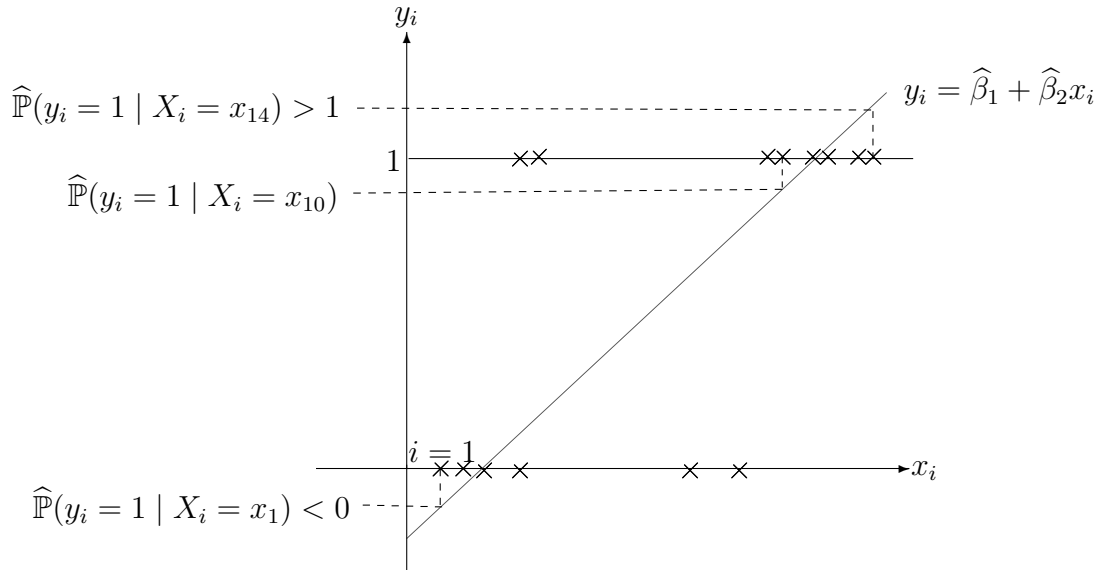
To do so, define a dummy variable Y_i for the occurrence of damage in the current year for individual i ($Y_i = 1$ if the insured i experienced a damage, $Y_i = 0$ otherwise) and X_i the vector of her characteristics. We want to guess the effect of each of the characteristics X_{it} on Y_i. The binary nature of Y_i however prevents us from applying standard linear regressions, that is to look for the vector of coefficients β, such that

$$Y_i = X_i'\beta + u_i, \qquad \text{with } u_i \sim_{iid} \mathcal{N}(0,1)$$

Indeed, even if this model can be interpreted in terms of probabilities:

$$
\left.
\begin{aligned}
\mathbb{E}(Y_i \mid X_i) &= \mathbb{E}(X_i'\beta \mid X_i) + \underbrace{\mathbb{E}(u_i \mid X_i)}_{=0} = X_i'\beta, \text{ and} \\
\mathbb{E}(Y_i \mid X_i) &= 1.\mathbb{P}(Y_i = 1 \mid X_i) + 0.\mathbb{P}(Y_i = 0 \mid X_i) \\
&= \mathbb{P}(Y_i = 1 \mid X_i)
\end{aligned}
\right\} \Rightarrow X_i'\beta = \mathbb{P}(Y_i = 1 \mid X_i)
$$

nothing constraints the estimated probabilities (called the scores): $\widehat{\mathbb{P}}_i = \widehat{\mathbb{P}(Y_i = 1} \mid X_i) = X_i'\widehat{\beta}$ to lie between 0 and 1 (see. next figure, in the case of unidimensional X_{it}).

y_i

$\widehat{\mathbb{P}}(y_i = 1 \mid X_i = x_{14}) > 1$ ----

$y_i = \widehat{\beta}_1 + \widehat{\beta}_2 x_i$

1

$\widehat{\mathbb{P}}(y_i = 1 \mid X_i = x_{10})$ ---

$i = 1$

x_i

$\widehat{\mathbb{P}}(y_i = 1 \mid X_i = x_1) < 0$ ----

That is why we don't specify a linear relationship (between Y_i and X_{it}s) in scoring models. To solve this issue (and constrain the estimated probability to lie between 0 and 1), we look for a function $F(X_i'\beta)$ such that:

- $\lim_{v \to +\infty} F(v) = 1$

- $\lim_{v \to -\infty} F(v) = 0$

- F is continuously differentiable and $\frac{dF(v)}{dv} > 0$

We can then specify $\mathbb{E}(Y_i \mid X_i) = F(X_i'\beta)$ and $\frac{\partial P(Y_i = 1 \mid X_i)}{\partial X_{it}} = \frac{\partial F(X_i'\beta)}{\partial X_{it}} = f(X_i'\beta)\beta_t$

Two solutions for F have been particularly studied:

- the Probit model, that uses the cumulative distribution function of the standard Gaussian distribution:

$$
\begin{aligned}
\mathbb{P}(Y_i = 1 \mid X_i) = F(X_i'\beta) &= \Phi(X_i'\beta) \\
&= \int_{-\infty}^{X_i'\beta} \frac{1}{\sqrt{2\pi}} e^{-t/2} dt
\end{aligned}
$$

- the Logit model, that uses the cumulative distribution function of the logistic distribution:

$$
\begin{aligned}
\mathbb{P}(Y_i = 1 \mid X_i) = F(X_i'\beta) &= \frac{\exp(X_i'\beta)}{1 + \exp(X_i'\beta)} \\
&= \frac{1}{1 + (\exp(X_i'\beta))^{-1}} = \frac{1}{1 + (\exp(-X_i'\beta))}
\end{aligned}
$$

To analyze the characteristics of these models and understand the similarities, it is useful to recall some of the properties of Gaussian and logistic distributions.

- If $Z \sim \mathcal{N}(0,1)$, its c.d.f. is denoted $F_Z(z) = \Phi(z)$ and its density writes $f_Z(z) = \phi(z) = \frac{1}{\sqrt{2\pi}}e^{-t/2}$. We then get $E(Z) = 0$ and $\text{Var}(Z) = 1$. Although the standard Gaussian distribution has the advantage of being symmetric and of having been widely studied, it has the drawback of not having an explicit expression for its c.d.f. (but only an integral expression).

- On the contrary, if $Z \sim \text{Logistic}$, its c.d.f. is perfectly determined and writes $F_Z(z) = \frac{e^z}{1+e^z}$. Then $f_Z(z) = \frac{e^z}{(1+e^z)^2}$, $f_z(z) = F_Z(z)(1 - F_Z(z))$ (what will be useful in the following), $E(Z) = 0$ and $\text{Var}(X) = \frac{\pi^2}{3}$. The logistic distribution therefore has the advantage of being very close to the standard Gaussian while having a c.d.f. with a simple analytical expression (see next figure, in which the logistic distribution has also been reduced).
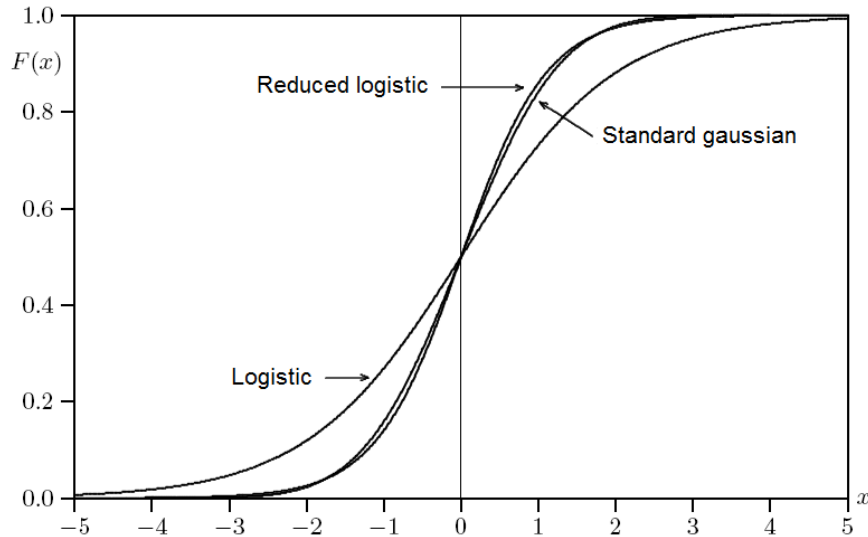
On top on being statistically simple, the Logit and Probit models can be theoretically justified.

The Probit model can indeed be reached by a model with an unobserved (or latent) variable Y_i^*. Assume that $Y_i^* = X_i\beta + u_i$ with $u_i \sim iid\mathcal{N}(0,1)$ but cannot be observed. Still one can observe the sign of Y_i^* through the binary variable Y_i defined as: $Y_i = 1$ if $Y_i^* > 0$ and $Y_i = 0$ if $Y_i^* \leq 0$. Then

$$\begin{aligned} \mathbb{P}(Y_i = 1) &= \mathbb{P}(Y_i^* > 0) = \mathbb{P}(X_i\beta + u_i > 0) = 1 - \mathbb{P}(u_i \leq -X_i'\beta) \\ &= 1 - \Phi(-X_i'\beta) = \Phi(X_i'\beta). \end{aligned}$$

The Logit model can also be reached by such a model with error distributed as a logistic. It can however be derived more easily by assuming that the observed characteristics (X_i) don't determine p_i but $\ln(p_i/(1-p_i))$ (that is the log of the odd ratio):

$$\ln\left(\frac{p_i}{1 - p_i}\right) = X_i'\beta$$

We then obtain:

$$p_i = \frac{\exp(X_i'\beta)}{1 + \exp(X_i'\beta)}$$

2.3 Estimating scoring models

Once the model chosen, the "scores" can be computed after estimating the βs by maximum likelihood.

As

$$Y_i = \left\{ \begin{array}{c} 1 \\ 0 \end{array} \right.$$

there exists two types of contributions to the likelihood:

- $i/Y_i = 1 : L_i(\beta) = \mathbb{P}(y_i = 1 \mid X_i) = F(X_i'\beta)$

- $i/Y_i = 0 : L_i(\beta) = \mathbb{P}(y_i = 0 \mid X_i) = 1 - F(X_i'\beta)$

The contribution to the likelihood then writes

$$\forall i, L_i(\beta) = \mathbb{P}(y_i = 1 \mid X_i)^{Y_i}.\left[1 - \mathbb{P}(y_i = 1 \mid X_i)\right]^{1-Y_i} = F(X_i'\beta)^{Y_i}.\left[1 - F(X_i'\beta)\right]^{1-Y_i}$$

and

$$L(\beta) = \prod_{i=1}^{n} F(X_i'\beta)^{Y_i}.\prod_{i=1}^{n}\left[1 - F(X_i'\beta)\right]^{1-Y_i}$$

or more simply (using log-likelihood)

$$l(\beta) = \sum_{i=1}^{n}\left[Y_i \ln(F(X_i'\beta)) + (1 - Y_i)\ln\left(1 - F(X_i'\beta)\right)\right]$$

14

The maximum likelihood is then obtained for β such that:

$$
\begin{aligned}
\frac{\partial l(\beta)}{\partial \beta} &= \sum_{i=1}^{n} \left[Y_i \frac{1}{F(X_i'\beta)} f(X_i'\beta) X_i - (1 - Y_i) \frac{1}{1 - F(X_i'\beta)} f(X_i'\beta) X_i \right] = 0 \\
&\Leftrightarrow \sum_{i=1}^{n} \left[\frac{Y_i - F(X_i'\beta)}{F(X_i'\beta)(1 - F(X_i'\beta))} \right] f(X_i'\beta) X_i = 0
\end{aligned}
$$

What simplifies in the case of Logit model (for which $f_z(z) = F_Z(z)(1 - F_Z(z))$) into:

$$
\sum_{i=1}^{n} \left[Y_i - F(X_i'\beta) \right] X_i = 0
$$

$$
\Leftrightarrow \sum_{i=1}^{n} \left[Y_i - \frac{\exp(X_i'\beta)}{1 + \exp(X_i'\beta)} \right] X_i = 0
$$

Remark. *Rather than estimating the probability that (at least) one damage occurs in the year, insurers can also account for the possibility that several damages occur the same year for the same policyholder. It then estimates the probability for each possible number of damage ($Y_i = 0, 1, ..., J$), using either:*

- *a multivariate Logit:*
  We then want

$$
\begin{cases}
\ln \left(\frac{\mathbb{P}(Y_i = j | X_i)}{\mathbb{P}(Y_i = l | X_i)} \right) = X_i'(\beta_j - \beta_l) \forall j \forall l \\
\sum_{j=0}^{L} \mathbb{P}(Y_i = j \mid X_i) = 1
\end{cases}
$$

  what gives

$$
\begin{cases}
\mathbb{P}(Y_i = J \mid X_i) = \frac{1}{1 + \sum_{j=0}^{J-1} \exp\left[X_i'(\beta_j - \beta_J) \right]} \\
\mathbb{P}(Y_i = j \mid X_i) = \frac{\exp\left[X_i'(\beta_j - \beta_J) \right]}{1 + \sum_{l=0}^{J-1} \exp\left[X_i'(\beta_l - \beta_J) \right]} \forall j = 0, ..., J - 1
\end{cases}
$$

- *a Poisson model: $\mathbb{P}(Y_i = k) = \frac{\lambda_i^k}{k!} e^{-\lambda_i}$ with $\lambda_i = X_i'\beta$ the expectation and variance of the number of damages over the period for insured i.*

Chapter 3

Unobservable criteria

In spite of scoring methods, some unobserved heterogeneity will remain. In other words, the insurance company cannot observe all the characteristics that impact the probability of damage. In this chapter, we analyze the consequences of this unobserved heterogeneity and how insurers can account for it. We more precisely assume that the insureds differ (exogenously) in their risk exposure – i.e. in their probability of damage – without their insurer being able to know it. We still consider that the insureds perfectly knows this probability.

3.1   The adverse selection problem

To analyze the consequences of an insurer not being able to distinguish among agents, let us consider a setting in which it only knows the average risk exposure p_M. In a competitive setting (and in the absence of loading factor), the insurer would then offer a premium $\Pi = p_M q$. This premium would be too high for the individuals less exposed than the average (those with $p_i < p_M$), in the sense that it would be higher than their willingness to pay for full insurance. Those individuals would then decide not to be insured or to buy partial insurance (depending on the available contracts). On the contrary, this premium would be lower than what the agents more exposed to risk ($p_i > p_M$) are willing to pay for full insurance. Those would even like to be "more than fully" insured.

In the simple case of only two types of risk (*i.e.* of agents) p_H and p_L (with $p_L < p_H$) in proportion λ_H and $\lambda_L = 1 - \lambda_H$ respectively (such that $p_M = \lambda_H p_H + \lambda_L p_L$), the situation can be represented as follows:
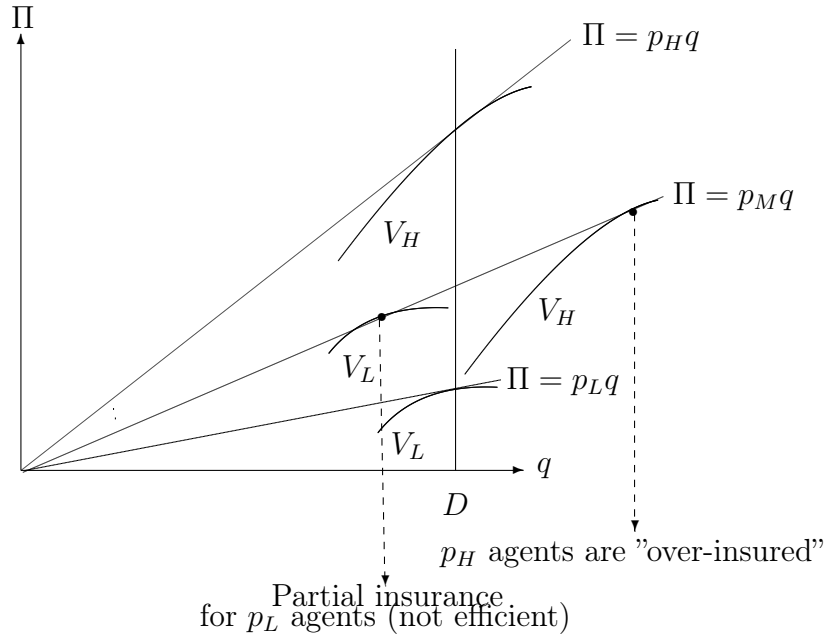
Because of unobserved heterogeneity, the expected profit of the insurer is then negative, the premium being calculated on the basis of all insureds buying the same amount of insurance. In cases where only full insurance contracts are offered, the insurer may even end up insuring high-risk agents only. We refer to this mechanism as **adverse selection** (as the insurer then "selects" the less profitable insureds).

Remark. *It should be noticed that this situation is rare in reality, where we generally observe the reverse: situations in which less exposed agents are "well" insured whereas more risky are badly insured (or ever not insured at all). This might be explained either by the fact that in reality, the asymmetry of information go the other way around (the insurer knowing the risk better than the insured)[2]; or because differences about risk exposure are correlated with differences in behavior, the more risk tolerant agents being those who both buy less insurance and take more risk.[3] We will study more deeply such behaviors that can reduce or increase risk exposure in the next chapter.*

[2]see B. Villeneuve (2000), "The Consequences for a Monopolistic Insurer of Evaluating Risk Better than Customers : The Adverse Selection Hypothesis Reversed", *Geneva Papers on Risk and Insurance Theory*, 25, 65-79

[3]see D. de Meza and D. Webb (2001), "Advantageous selection in insurance markets", *The RAND Journal of Economics*, 32(2), 249-262

3.2 Self-selection: the Rothschild-Stiglitz model

To solve this adverse selection issue, insurers need to differentiate the contracts it offers to "high-risk" (the agents with probability p_H) and "low-risk" (those with probability p_L) individuals. As it cannot distinguish between the two types, it has to build contracts such that the insureds would differentiate themselves. We refer to such mechanism as "self selection". Each type of insured / of risk will choose optimally the contact intended for her.
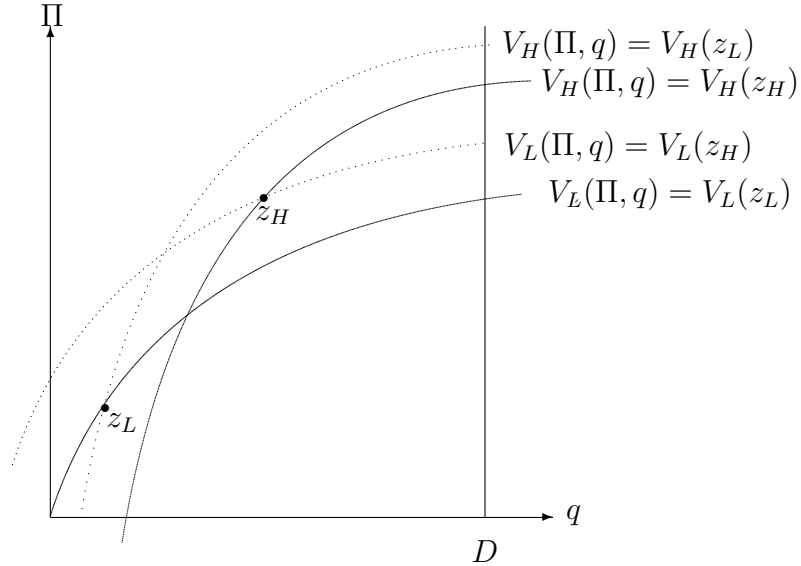
Definition. *A self-selective menu is a couple of contracts (z_L, z_H) such that*

$$\begin{cases} V_H(z_H) \geq V_H(z_L) \\ V_L(z_L) \geq V_L(z_H) \end{cases}$$

that is (Π_H, q_H, Π_L, q_L) such that

$$\begin{cases} p_H u(R - D + q_H - \Pi_H) + (1 - p_H)u(R - \Pi_H) \geq p_H u(R - D + q_L - \Pi_L) + (1 - p_H)u(R - \Pi_L) \\ p_L u(R - D + q_L - \Pi_L) + (1 - p_L)u(R - \Pi_L) \geq p_L u(R - D + q_H - \Pi_H) + (1 - p_L)u(R - \Pi_H) \end{cases}$$

Type H agents (the high-risk) will then prefer the contract z_H, whereas low-risk (type B) will prefer the contract z_L.

In a competitive setting, at equilibrium, expected profit must be zero **on each contract**. We will then talk about "neutral pricing".

Definition. *A neutral menu is a self-selective menu for which expected profit on each contract is zero: $\Pi_i = p_i q_i \ \forall i = L, H$*
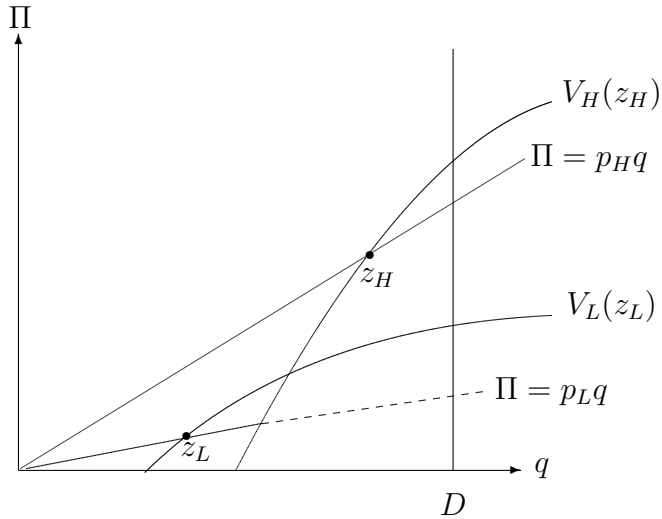
In the above figure, (z_L, z_H) is a neutral menu. However it won't be the case if z_L would lie on the doted line (high-risk would then prefer z_L to z_H) or if there would be an intersection between the $V_L(z_L)$ curve and the zero-profit line $\Pi = p_H q$ (then, an insurance company would be able to attract both high-risk and low-risk while making strictly positive profit).

It exists an infinity of such neutral menus. Still, the Rothschild-Stiglitz theorem highlights that one of them Pareto-dominates all the others (none of the other neutral menu is mutually preferred by both types).

Theorem. *(Rothschild - Stiglitz, 1976) There exists a unique neutral menu that Pareto-dominates all the other neutral menu. It is defined by:*

$$\forall i \quad \Pi_i^* = p_i q_i^*$$
$$q_H^* = D \qquad and \qquad V_H(z_H^*) = V_H(z_L^*)$$
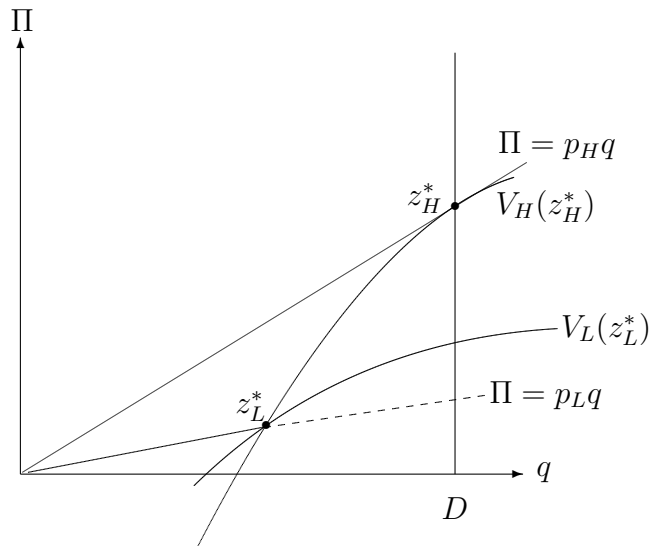
We thus can talk about a "best neutral menu" that is obtained by (i) providing full insurance to high-risk, and (ii) offering to low-risk a contract with a deductible exactly sufficient to prevent the high-risk to choose it: z_L^*. Low-risk therefore bear an external effect. Contrary to low-risk, high-risk are offered their preferred contract.

This menu seems to be a good candidate for equilibrium, as no insured has an interest in choosing the other contract, and no insurance company can offer a mutually preferred menu (according to the previous theorem). To ascertain that this "best neutral menu" is indeed an equilibrium, we however need to define an equilibrium notion, that is to specify the out-of-equilibrium evolution of the "game". The simplest equilibrium concept has been defined by Rothschild and Stiglitz in 1976 (we study alternative equilibrium concepts in the next section):

Definition. *A Rothschild-Stiglitz equilibrium is a set of no-deficit contracts $z = (z_L, z_H)$, such that it doesn't exist any contract that would make non-negative expected profit, if jointly offered.*

Theorem. *If a Rothschild-Stiglitz equilibrium exists, it must be $z^* = (z_B^*, z_H^*)$*

We have already seen that z^* was the "best" neutral menu, that is the best self-selective (or separating) menu. z^* would then attract all the insureds (and make zero profit) if offered jointly to any other separating menu. No separating menu other than z^* can therefore be a Rothschild-Stiglitz equilibrium. To prove the above theorem, we now need to show that no pooling contract (where $z_L = z_H$) can be an equilibrium, *i.e.* that for all pooling contract, there exists a (separating) contract that would make non-negative profit when jointly offered.
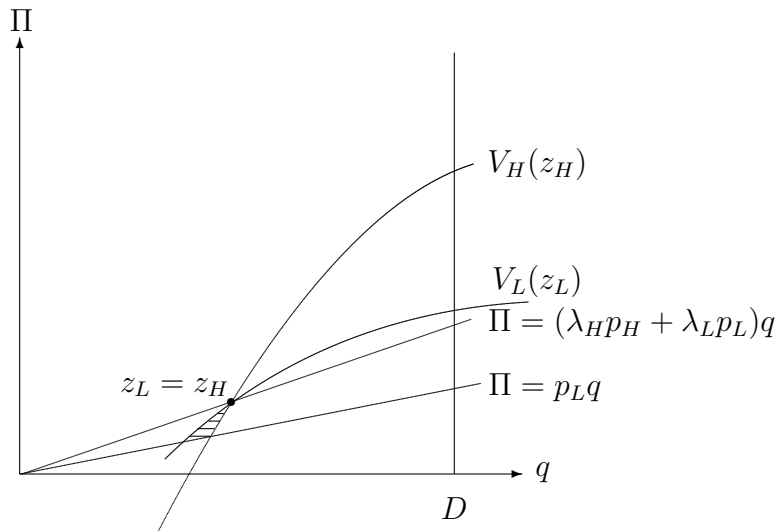
Start from any pooling contract. An insurance company can then build a profitable contract by offering lower coverage and premium (in the dashed area of the above figure) and attracting low-risk individuals only. This phenomenon is called "cream-skimming". All pooling contracts being subject to cream-skimming by a separating contract, only separating menus can be equilibrium. z^* is thus the only candidate to equilibrium.

However, it can exists settings in which no equilibrium exists, *i.e.* in which a contract can be profitable when offered in addition to z^*.

Theorem. *No Rothschild-Stiglitz equilibrium exists if the share of low risk is too high ($\lambda_L > k$).*

We have already seen that only (z_L^*, z_H^*) can be a Rothschild-Stiglitz equilibrium. For such an equilibrium to exist, no contract offered in addition to (z_L^*, z_H^*) should make non negative expected profit. As z^* is the "best" separating menu, it can only be the case of a pooling contract. One can show – as illustrated in the following figure – that when the share of low-risk is (too) high, a pooling contract can attract (all) the insureds and make a positive profit. When this is the case, graphically, the zero profit line of an insurer offering a pooling contract ($\Pi = (\lambda_H p_H + \lambda_L p_L)q = p_M q$) is close to the one of contracts offered to low-risk only ($\Pi = p_L q$).
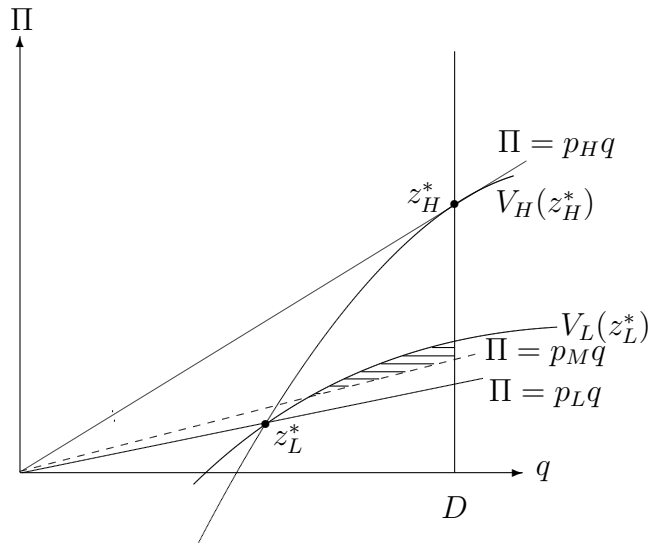
Then, offering a contract in the dashed area (between the $\Pi = p_M q$ line and the indifference curve $V_L(\Pi, q) = V_L(z_L^*)$), an insurer would make a positive expected profit (as it lies above the $\Pi = p_M q$ line) and attract both types of insureds (as it lies to the South-Est of both the $V_L(\Pi, q) = V_L(z_L^*)$ and the $V_H(\Pi, q) = V_H(z_H^*)$ curves). However, we have already seen that every pooling contract (even in this area) are subject to cream-skimming. Therefore, no pooling equilibrium will exist because of profitable cream-skimming and neither will a separating equilibrium when the share of high-risk is too low. This because, as stated before, the best separating contract (z^*) entails a welfare loss for low-risk (with a low coverage), in order to make high-risk pay the highest premium rate (obtained for full insurance). Such a strategy is then optimal only if the share of high-risk is high.

3.3 Equilibrium existence

The absence of equilibrium in the Rothschild-Stiglitz model when $\lambda_L > k$ has received a lot of interest and several solutions have been proposed to restore existence. These solutions concern either (i) the equilibrium definition (Wilson 1977)[4], (ii) the nature of insurance contracts (Picard 2014)[5] or (iii) the consideration of limited liability for the insurer, that is the possibility of failure (Mimra and Wambach 2015)[6].

- The first development, proposed by Wilson (1977), concern a modification of the equilibrium concept and the definition of the so-called "anticipative" equilibrium. In this model, a potential entrant anticipates the possibility that the incumbent withdraws contracts.

  Definition. *A Wilson equilibrium is a set of no-deficit contracts $z = (z_L, z_H)$, such that it doesn't exist any contract that would make non-negative expected profit if jointly offered, once existing contracts that lose money (if they exist) withdrawn.*

  This "adaptive" notion of equilibrium allows to restore existence whatever the proportion of low-risk. Indeed, under this definition, when (z_B^*, z_H^*) is dominated by a pooling contract, cream-skimming (of this polling contract) is no more profitable. A pooling contract can then be an equilibrium. This is because cream-skimming of a pooling contract (the fact of attracting low-risk only) leads to the "default" of the insurer offering such a contract (as it lies below the $\Pi = p_H q$ line). After the withdrawal of this (pooling) contract that loses money, the insurer that "deviated" ends up insuring both types. As its contract necessarily lies below the $\Pi = p_M q$ line (to cream-skim low-risk, see figure page 20), offering a deviating contract is no more profitable "once existing contracts that lose money are withdrawn". According to Wilson's definition, a polling equilibrium therefore exists when $\lambda_L > k$.

  Theorem. *A Wilson equilibrium always exists. The corresponding menu is unique:*

  - *if the share of low-risk is low ($\lambda_L \leq k$), it is the best neutral menu z^**
  - *if the share of low-risk is too high ($\lambda_L > k$), it is the pooling contract with zero expected profit that satisfies the most the low-risk, that is $\{\overline{z_L}, \overline{z_L}\}$, with $\overline{z_L}$ such that:*

  $$V_L(\overline{z_L}) = \max_q V_L(p_M q, q)$$

[4]C. Wilson, C. (1977), "A model of insurance markets with incomplete information", Journal of Economic Theory, 16(2), 167-207.

[5]P. Picard (2014), "Participating Insurance Contracts and the Rothschild-Stiglitz Equilibrium Puzzle", Geneva Risk and Insurance Review, 39, 153-175.

[6]W. Mimra and A. Wambach (2019), "Endogenous Insolvency in the Rothschild-Stiglitz Model", Journal of Risk and Insurance, 2019, 86, 165-181.
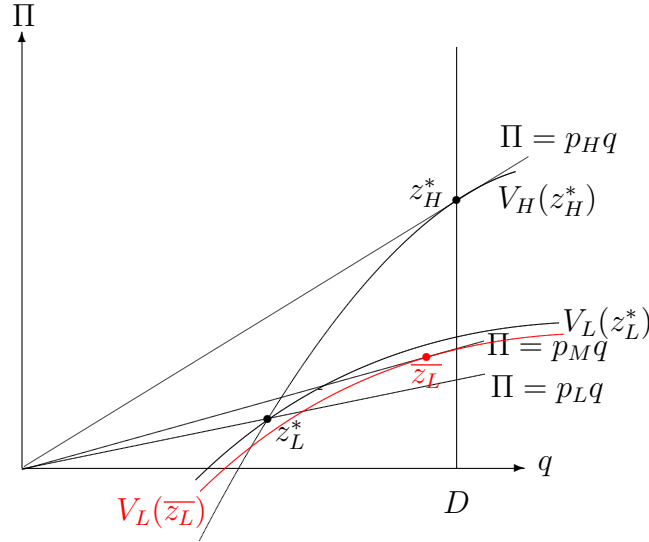
- More recently, Picard (2014) has proposed a solution based on the nature of insurance contract (instead of the equilibrium concept). He shows that the existence issue vanishes when allowing insurers to offer participative contracts. Then, the fact that the insureds receive part of the profit and pay part of the losses acts as an implicit threat against cream-skimming. Indeed, if the high-risk owe a participative contract, when the low-risk change insurer, the welfare of the high-risk decreases, what makes it difficult for the deviating (or entrant) insurer to attract low-risk only. A Rothschild-Stiglitz equilibrium exists even when the share of low-risk is high: it is then a fully participative contract (profit and losses are totally redistributed) – as offer by mutual companies – that corresponds in terms of expected welfare to $(\overline{z_L}, \overline{z_L})$. When the share of low-risk is low, the equilibrium is still (z_L^*, z_H^*) offered in a non participative framework.

- In the same vein, Mimra and Wambach (2015) have shown that a Rothschild-Stiglitz equilibrium exists whatever the share of low-risk if the insureds take into account the possibility that their insurer fails (without changing the definition of an equilibrium). Again, this modification makes cream-skimming non profitable. Indeed, the escape of low-risk from a pooling contract then worsen the expected utility of high-risk by increasing the probability that their insurance fails. By the same mechanism as before, the contract supposed to attract low-risk then becomes attractive for high-risk also, and no cream-skimming is feasible. We then have as Rothschild-Stiglitz equilibrium: (z_L^*, z_H^*) if $\lambda_L \leq k$ and $(\overline{z_L}, \overline{z_L})$ if $\lambda_L > k$.

Chapter 4

Moral hazard

The above analysis doesn't account for actions the insureds can take to individually reduce the risk they face. Such actions are yet key to insurance pricing as they impact the expected profit of the insurer, but also as the insurance contract can impact choices regarding it. We consider in this section actions that can reduce (i) the size of the loss D (as for example wearing a seat-belt or setting fire extinguishers), in which cases we will talk about self-insurance, and (ii) the probability of damage p (as for example using ABS, or driving carefully) in which cases we will talk about self-protection. Although comparable from the viewpoint of the insured, these two kind of actions have different consequences for the insurer. The cost function is affected by self-protection, whereas it is not by self-insurance, insofar as the size of the loss is observed by the insurer.

4.1   Self-insurance and its consequences

Let c be the cost of self-insurance activities, which impact on the size of the loss is described by the function $D(.)$. We assume $D'(c) < -1$ (*i.e.* $(c + D(c))' < 0$) and $D''(\cdot) > 0$, meaning that self-insurance activities have positive but decreasing return.

Without insurance, the expected utility of an individual then writes:

$$V_0(c) = pu(R - D(c) - c) + (1 - p)u(R - c)$$

and is concave under the above assumptions. The level of self-insurance c_0 chosen by uninsured individual then solves:

$$V_0'(c_0) = -pu'(R - D(c_0) - c_0)(1 + D'(c_0)) - (1 - p)u'(R - c_0) = 0$$

That is

$$-(D'(c_0) + 1) = \frac{1 - p}{p} \times \frac{u'(R - c_0)}{u'(R - D(c_0) - c_0)}$$

Consider now that this individual can buy insurance at a unit price π (by coverage unit). Denoting by c her self-insurance level and q the coverage she buys, her expected utility is:

$$V(q,c) = pu(\underbrace{R - D(c) - c + (1-\pi)q}_{\equiv A}) + (1-p)u(\underbrace{R - c - \pi q}_{\equiv N})$$

and her optimal decisions solve:

$$\begin{cases} \frac{\partial V}{\partial c}(q^*, c^*) & = pu'(A)(1 + D'(c^*)) + (1-p)u'(N) = 0 \\ \frac{\partial V}{\partial q}(q^*, c^*) & = (1-\pi)pu'(A) - \pi(1-p)u'(N) = 0 \end{cases}$$

that is

$$-(D'(c^*) + 1) = \frac{1-p}{p} \times \frac{u'(N)}{u'(A)} = \frac{1-\pi}{\pi}$$

$D(\cdot)$ being convex, this means that c^* is increasing in π. Namely, the spending on self-insurance increases (and coverage decreases) when insurance becomes more costly. **Self-insurance and insurance are therefore substitutes**. Remark moreover that when $\pi = 1$, $q^* = 0$ (an insured would even want $q = -\infty$) and we find back the non-insurance case $c^* = c_0$. As c^* is increasing in π, this implies $\forall \pi < 1$, $c^* < c_0$. **The access to insurance market decreases the incentives to self-insure**.

4.2 Self-protection and moral hazard

As it impacts the probability of damage, self-protection activities affect the expected cost for the insurer. The optimum will thus be for the insurer to include in the contracts clauses that would give the insureds an incentive to take up these actions that reduces the frequency of damage. Most of the time, however, self-protection actions (for example: driving carefully) are not observable by the insurer. It is thus difficult for it (even with scoring methods) to discern moral hazard (hidden action) from adverse selection (hidden characteristics).

We focus here on a context of "pure" moral hazard, by assuming that all the insureds are identical. The only issue for the insurer is then to give the insureds an incentive to take off self-protection actions although that have a cost (either psychological or monetary). We moreover assume that these actions are unobservable, both ex-ante and ex-post.

The only instrument the insurer can use is then the level of deductible. The level of deductible has to achieve a compromise between the insureds willingness to be insured and the insurer wish to incite to self-protection.

We will work here on the simplest model in which individuals only choose whether or not to undertake a self-protection effort e that will decrease her probability of damage from $\bar{p}$ to $\underline{p}$, at a utility cost c. The expected utility of an uninsured individual is then:

$$V_0(e) = p(e)u(R - D) + (1 - p(e))u(R) - c(e)$$

with

$$e \in \{0, 1\}, p(0) = \bar{p} > p(1) = \underline{p} \text{ and } c(0) = 0 < c(1) = c$$

Considering an insurance contract z, with a premium Π and a coverage q, expected utility becomes:

$$V(e, \Pi, q) = p(e)u(R - D - \Pi + q) + (1 - p(e))u(R - \Pi) - c(e)$$

If effort were observable, the insurer could make it play in the contact. If we restrict to non deficit contracts $(\Pi \geq p(e)q)$, the optimal contract under competition would consist in full insurance $(q = D)$ and actuarial premium $(\Pi = p(e)q)$.

In this case, an insured would provide effort if:

$$u(R - \underline{p}D) - c > u(R - \bar{p}D)$$
$$\Leftrightarrow \quad c < c^* = u(R - \underline{p}D) - u(R - \bar{p}D)$$

On the contrary, an uninsured individual would provide effort if:

$$\underline{p}u(R - D) + (1 - \underline{p})u(R) - c > \bar{p}u(R - D) + (1 - \bar{p})u(R)$$
$$\Leftrightarrow \quad c < \bar{c} = (\bar{p} - \underline{p})[u(R) - u(R - D)]$$

There is no systematic relationship between $\bar{c}$ and c^*. The access to insurance market can or cannot increase the incentive to provide self-protection effort. To see it, let's define $\Delta = \bar{c} - c^*$. Then, as $u(\cdot)$ is concave, Jensen's inequality gives:

$$\lim_{\underline{p} \to 0} \Delta = u(R - \bar{p}D) - [\bar{p}u(R - D) + (1 - \bar{p})u(R)] > 0 \text{ and}$$
$$\lim_{\bar{p} \to 1} \Delta = [\underline{p}u(R - D) + (1 - \underline{p})u(R)] - u(R - \underline{p}D) < 0$$

Assume now that $c < \min\{\bar{c}, c^*\}$, and that e is unobservable by the insurer. Let's analyze the shape of insureds' preferences in the (q, Π) plan.

To do so, we have to consider the set of contracts that provide the insured with the same level of expected utility k **whether she makes the effort or not**, that is $\{(\Pi, q)/V(1, \Pi, q) = k \text{ or } V(0, \Pi, q) = k\}$. Then, we will need to study, for each contract in this set, if she has an interest to make the effort.

Remark first that for $q = D$ (full insurance) the utility gross of the effort cost is independent from the effort level: $V(e, \Pi, D) + c(e) = u(R - \Pi) \; \forall e$. Thus, at $q = D$, $V(1, \Pi, D) = V(0, \Pi, D) - c$. Moreover, we know from chapter 1 that the slope of the curve $V(1, \Pi, q) = k$ equals $\underline{p}$ at $q = D$, when the one of $V(0, \Pi, q) = k$ equals $\bar{p}$. Lastly, one can easily show using chapter 1 (equation (*)), that the slope of the indifference curve is increasing in p $\left(\left(\frac{\partial \Pi}{\partial q}\Big|_{V(\Pi, q) = k}\right)'_p > 0\right)$. The slope of $V(0, \Pi, q) = k$ is thus higher than the one of $V(1, \Pi, q) = k$ for every level of coverage.

Therefore, the indifference curve $V(1, \Pi, q) = k$ will first be above $V(0, \Pi, q) = k$ and then below.
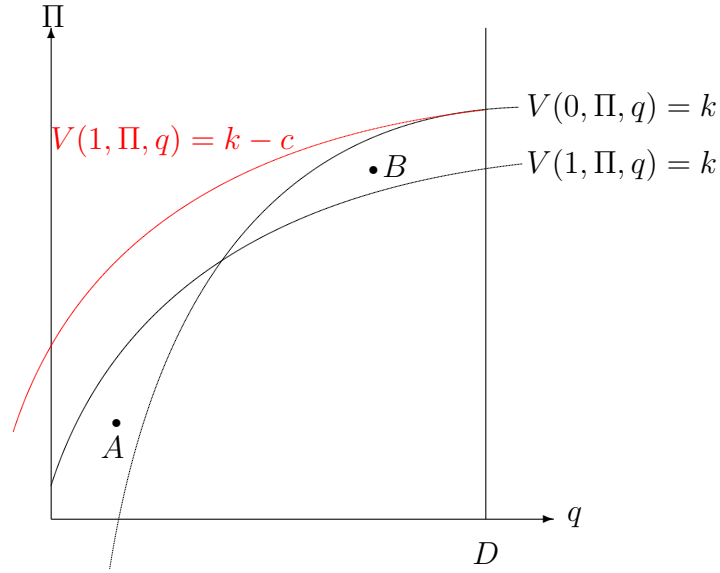
For low levels of coverage (point A), the insured would prefer to make the effort ($V(1, \Pi, q) > V(0, \Pi, q)$) whereas when coverage is high, her expected utility would be higher when not providing effort (point B). This can be formalized in the following proposition.

Proposition. *For every premium Π, there exists a threshold level of coverage ($L(\Pi)$) such that:*

- *the insured makes the effort when the coverage is lower than this threshold ($e = 1$ if $q \leq L(\Pi)$) and*

- *doesn't make it when it is higher ($e = 0$ if $q > L(\Pi)$)*

The function $L(\Pi)$ is moreover increasing: the higher the premium, the higher the level of coverage below which the insured makes the effort.

Proof. Let

$$
\begin{aligned}
W(q, \Pi) &\equiv V(0, \Pi, q) - V(1, \Pi, q) \\
&= c - (\bar{p} - \underline{p})\left[u(R - \Pi) - u(R - \Pi + q - D)\right]
\end{aligned}
$$

The insured will make the effort if and only if $W(\cdot)$ is negative. Now $W(\cdot)$ is continuous and strictly increasing in q. We can then define $L(\Pi)$ by:

$$
W(L(\Pi), \Pi) = 0
$$

The first result follows directly. $L(.)$ being increasing comes from the implicit function theorem:

$$
\frac{\partial W}{\partial q}(L(\Pi), \Pi) > 0 \text{ already stated}
$$

$$
\frac{\partial W}{\partial \Pi}(L(\Pi), \Pi) = -(\bar{p} - \underline{p})[u'(R - \Pi + L(\Pi) - D) - u'(R - \Pi)] < 0
$$

$$
\Rightarrow \quad \frac{dL}{d\Pi} = -\frac{\frac{\partial W}{\partial \Pi}(L(\Pi), \Pi)}{\frac{\partial W}{\partial q}(L(\Pi), \Pi)} > 0
$$

This gives the following indifference curves, with angular points.
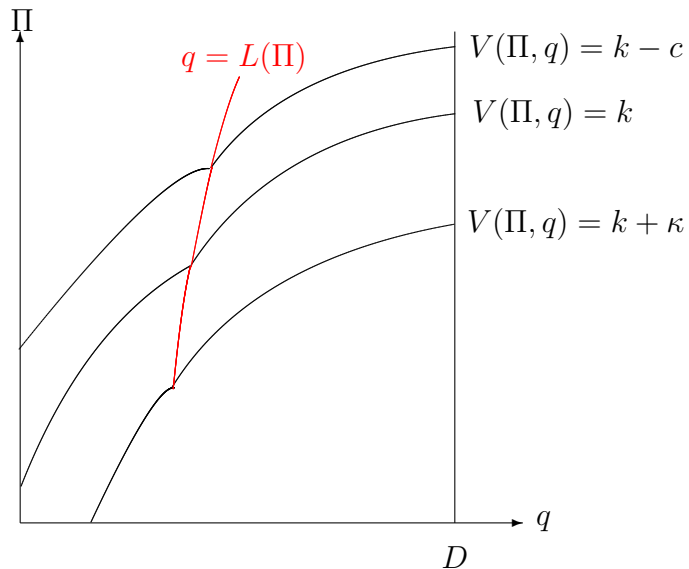
Now that we know more about the indifference curves, let's study the optimal contract under unobservable efforts.

As all the insureds are identical, it seems natural to restrict attention to anonymous allocations (that is to offer the same contract to all individuals). Given the above proposition, a contract $z = (q, \Pi)$ would be feasible only if:

$$\begin{cases} \Pi \geq \underline{p}q & \text{and} \quad q \leq L(\Pi) \quad \text{(incentive contracts), or} \\ \Pi \geq \overline{p}q & \text{and} \quad q \leq D \quad\quad \text{(no-effort contracts)} \end{cases}$$
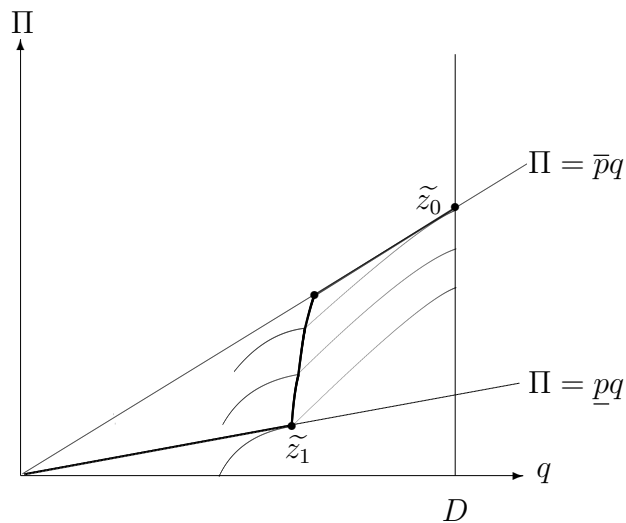
As the indifference curves, the set of feasible contracts have angular points.

Among these feasible contracts, let's now focus on the second best, optimal from the point of view of the insureds. We thus look for the contract that maximizes insureds' expected utility $(V = \max_{e \in \{0,1\}} V(e, \Pi, q))$ over the set of feasible contracts. That would be the contract offered by insurers that doesn't observe effort under perfect competition. Using previous notions, one easily gets the following proposition.

Proposition. *It exists a threshold cost of effort $\tilde{c}$ (lower than c^*), such that the second best optimal contract $\tilde{z}$ is characterized by:*

- *full insurance and no incentives if the effort cost is above the threshold: $\tilde{z} = \tilde{z}_0 = (D, \bar{p}D)$ and $e = 0$ if $c > \tilde{c}$*

- *partial insurance and incentives if the effort cost is below the threshold. In this case, the optimal contract is the preferred contract among those incentive compatible and actuarially fair. If $c \leq \tilde{c}$, $e = 1$ and $\tilde{z} = \tilde{z}_1 = (\underline{q}, \underline{\Pi})$ with $(\underline{q}, \underline{\Pi})$ such that $\underline{q} = L(\underline{\Pi})$ and $\underline{\Pi} = \underline{p}q$*

Proof. $\tilde{z}$ has to lie on the inferior frontier of the feasible set (as it maximizes expected utility). This frontier has three parts (as $c < \bar{c} \Rightarrow L(0) < 0$)

- the left part of the $\Pi = \underline{p}q$ line on which the maximum of V is reached at $\tilde{z}_1$, intersection with $q = L(\Pi)$

- a part of the $q = L(\Pi)$ curve (between $\Pi = \underline{p}q$ and $\Pi = \bar{p}q$) on which the maximum of V lies on $\tilde{z}_1$

- the right part of the $\Pi = \bar{p}q$ line on which the maximum of V is reached at $\tilde{z}_0 = (D, \bar{p}D)$

The proposition then comes from the comparison of $V(\tilde{z}_0)$ with $V(\tilde{z}_1)$. By construction, at $\tilde{z}_1$, the insured is indifferent between making the effort and not:

$$
\begin{aligned}
V(\tilde{z}_1) &= \underline{p}u[R - D + q_1(1 - \underline{p})] + (1 - \underline{p})u(R - \underline{p}q_1) - c \\
&= \bar{p}u[R - D + q_1(1 - \underline{p})] + (1 - \bar{p})u(R - \underline{p}q_1)
\end{aligned}
$$

In particular, q_1 is implicitly defined by:

$$
\frac{c}{\bar{p} - \underline{p}} = u(R - \underline{p}q_1) - u(R - D + q_1(1 - \underline{p}))
$$

The right hand side of the equality being decreasing in q_1, q_1 is itself a decreasing function in c $\left(\frac{\partial q_1}{\partial c} < 0\right)$. Now, we have:

$$
V(\tilde{z}_0) > V(\tilde{z}_1) \Leftrightarrow u(R - \bar{p}D) > \bar{p}u[R - D + q_1(1 - \underline{p})] + (1 - \bar{p})u(R - \underline{p}q_1)
$$

The right hand side of the equality being increasing in q_1, this proves the proposition. Indeed: $\frac{\partial[V(\tilde{z}_0) - V(\tilde{z}_1)]}{\partial c} = \frac{\partial[V(\tilde{z}_0) - V(\tilde{z}_1)]}{\partial q_1} \frac{\partial q_1}{\partial c} > 0$, that is $V(\tilde{z}_0) > V(\tilde{z}_1)$ for high value of c (one can easily check that it holds for $c = c^*$ as, by definition, the insureds then even prefer $\tilde{z}_0$ to $(\underline{p}D, D)$) and $V(\tilde{z}_0) < V(\tilde{z}_1)$ for low values of c (as it is the case for $c = 0$, when insureds provide the effort whatever the contract). $\square$

We can summarize these results trough as follows:

Cost	$\widetilde{c}$	c^*
First best	Full insurance Effort	Full insurance No effort
Second best	Partial insurance Effort	Full insurance No effort

When the cost of self-protection is low enough, partial insurance allows to give the insured an incentive to provide effort. However, when this cost is too high, making the insured participate to her risk is too costly. It is then better to insure her fully, even if it means her not making the effort.

We can moreover show that, when insurers can be guaranteed exclusivity (an insurer cannot insure a client who already has a contract with another company), it exists a unique equilibrium. It corresponds to a unique contract: the second best exposed above. The exclusivity assumption allows to prevent situations in which, in spite of competition, insurers can make a positive profit by offering contracts on the $q = L(\Pi)$ curve (and $z \neq \widetilde{z}_1$)

4.3   Ex-post moral hazard: the case of insurance fraud

Up to now, we have only considered actions that were undertook before the realization of the risk (ex-ante). However, one can also envision that the fact of being insured can promote ex-post behavior (after the realization of the risk) aiming at increasing the (declared) size of the loss, or even to declare not occurred damage. We then refer to insurance fraud or ex-post moral hazard.

The existence of such frauds and their size are by essence difficult to measure. However, various studies have estimated that frauds represent between 10% and 20% of reported damage. "The Insurance Information Institute" has even estimated that insurance frauds have cost to American insurers about 30 millions dollars in 2004.

From a theoretical point of view, insurance fraud becomes possible as soon as the insured can falsify a statement without the insurer being able to easily verify it. If, for example, an insured can costlessly falsify a damage and if it is not possible (or infinitely costly) for her insurer to verify it; an insured (not constrained by moral considerations) would always have an interest in declaring the level of damage leading to the highest coverage. Anticipating this, the insurer would propose a unique level of coverage, whatever the size of damage. If, at least, the occurrence of damage is verifiable, this coverage will be positive. Otherwise, the insurance market totally vanishes.

To make the problem more interesting and more realistic, we study in the following two less extreme scenarios. First, we analyze situations in which falsification is costless but verification possible (and costly); and then turn to a setting where verification is impossible but falsification costly.

Costly verification and deterministic audit

Let us consider the previous model with a random damage.

An insured with revenue R faces a risk of damage $\widetilde{D}$, distributed according to a c.d.f. $F(D)$ (with support $[0, \overline{D}]$). As the event $D = 0$ can be reach with non zero probability, $\widetilde{D}$ follows a mixed distribution with a mass at 0.

We assume here that the size of the loss (the realization) D is known by the insured only, but that the insurer can discover it at a cost c. We will then talk about audit. The insurer doesn't have the same information if it audits or not. On the one hand, if it decides not to audit, it only has the information declared by the insured, denoted $\widehat{D}$. The coverage will then be denoted $q_N(\widehat{D})$. On the other hand, in case it audits, the insurer knows both $\widehat{D}$ and D, the true size of damage. The coverage might then depend on both D and $\widehat{D}$, and include a form of punishment if $D \neq \widehat{D}$. We then denote the coverage by $q_A(D, \widehat{D})$. As previously, Π represents the premium asked by the insurer.

Let us analyze first the case of deterministic audit, that is the case in which for each level of (declared) damage, the insurer has to decide whether it audits or not. We denote by M the set of audited damages. In this case, we obtain the following result.

Lemma. *Every contract is Pareto-dominated (based on expected utility and expected profit) by an incentive compatible contract (with $\widehat{D} = D$) for which:*

- *non audited damages are all covered the same way: $q_N(D) = q_0 \ \forall D \in \overline{M}$*

- *audited damages give higher coverage than non audited ones: $q_A(D, \widehat{D}) = q_A(D) > q_0$ $\forall D, \widehat{D} \in M$*

The above Lemma can be demonstrated starting from any contract $\delta^0 = (\ M^0\ ,\ q_N^0(\widehat{D})\ ,$ $q_A^0(D, \widehat{D})\ ,\ \Pi^0\ )$. Let's consider first the set $\overline{M^0}$. In this set, the insureds will not be audited (and knows it). Let's note $q_0 = \max\limits_{D \in \overline{M^0}} q_N^0(D)$ and $D_0 = \arg\max\limits_{D \in \overline{M^0}} q_N^0(D)$. q_0 is the highest coverage an insured can get without being audited and D_0 is the size of the damage to be declared to get this coverage. Thus, it is clear that for every damage leading to a coverage lower than q_0, the insured is better of declaring D_0. The same payoff than δ^0 can then be obtained by keeping the same premium Π^0 with $q_N(D) = q_0 \forall D \in \overline{M}$ and $\overline{M} = \overline{M^0} \cup \{D/q_A(D, D) \leq q_0\}$, what would decrease the audit costs.

Moreover, we can easily end up with an incentive mechanism (that is a mechanism leading to no fraud) in the audit zone by setting $q_A(D) = \max\limits_{\widehat{D}} q_A^0(D, \widehat{D}) \ \forall D, \widehat{D} \in M$ (in this zone, the damage will be automatically audited and the insurer will always know the true value). Note however that for the mechanism to work, the insurer has to credibly commit to auditing in the set M, although it will never discover any fraud.

The power of the above Lemma comes from the fact that we can then focus on such incentive contracts. In a competitive setting, the optimal contract will therefore be the contract $(M\ ,\ q_0$ $,\ q_A(D)\ ,\ \Pi)$ that maximizes an insured expected utility:

$$V = \int_M u(R - \Pi - D + q_A(D, D))dF(D) + \int_{\overline{M}} u(R - \Pi - D - q_0)dF(D)$$

under the insurer participation constraint (non negative expected profit)

$$\Pi - \int_M [q_A(D, D) + c] \, dF(D) - \int_{\overline{M}} q_0 dF(D) \geq 0$$

and the incentive (to tell the true) constraint: $q_A(D, D) > q_0 \; \forall D \in M$.

We can then show that:

(i) because of the concavity of $u(\cdot)$, it is optimal to offer minimal and constant coverage when losses are low and to cover more widely high losses: $M =]m, \overline{D}]$ with $m \in [0, \overline{D}]$, and

(ii) in cases in which the damage is verified, the **marginal** loss is fully insured: $q_A(D) = D - k$ $\forall D \in M$. Indeed, systematic verification makes incentive useless in the set M.

Moreover, one can easily understand that if the occurrence of a damage $(D > 0)$ is not observable, it is optimal to set $q_0 = 0$. We thus have the following result:

Proposition. *In the case of deterministic audit, an optimal insurance contract has the following properties:*

- *only the losses higher than a given amount are covered and systematically audited ($M =]m, \overline{D}]$ and $q(D) = 0$ if $D \leq m$).*

- *when a damage is covered, the marginal loss if perfectly covered ($q(D) = D - k$ if $D > m$, with $0 < k < m$)*
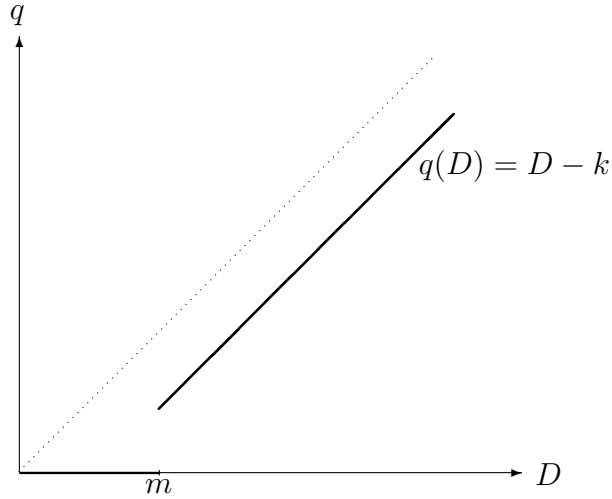

The fact that the total loss isn't fully insured $(k > 0)$ comes from the maximization program. Marginal utility when the loss is compensated has to equal marginal expected utility (an increase in the coverage in one state has to be compensated by an increase in premium in all states). As losses are not compensated for $D < m$, ex-post wealth are then lower than $R - \Pi$. Thus marginal expected utility is higher than $u'(R - \Pi)$ as $u(\cdot)$ is concave, and $R - \Pi - D + q_A(D)$ has to be lower than $R - \Pi$.

Remark. • *If it is possible to verify costlessly whether a damage occur $(D > 0)$ then $q(0) = 0$, $q(D) = q_0 > 0$ $\forall 0 < D \leq m$ and $q(D) = D$ $\forall D > m$, with $0 < q_0 < m$. As the insurer is no more ""constrained" not to cover small losses $(D < m)$, full insurance is optimal when damages are audited.*
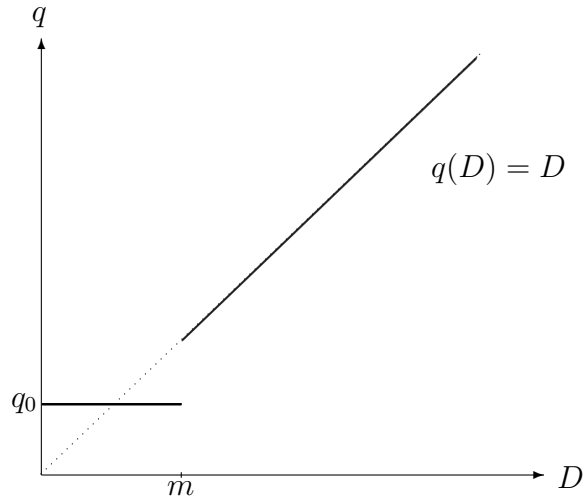
• *We get back to usual contracts with deductibles $(q(D) = \sup\{0, D - m\}$ with $m > 0$ and $M =]m, \overline{D}])$ if we add the possibility for the insured to intentionally increase the size of the loss (and if the insurer cannot distinguish the part intentionally created). Indeed, the two above contracts are the subject to intentional increase, because of their discontinuity.*
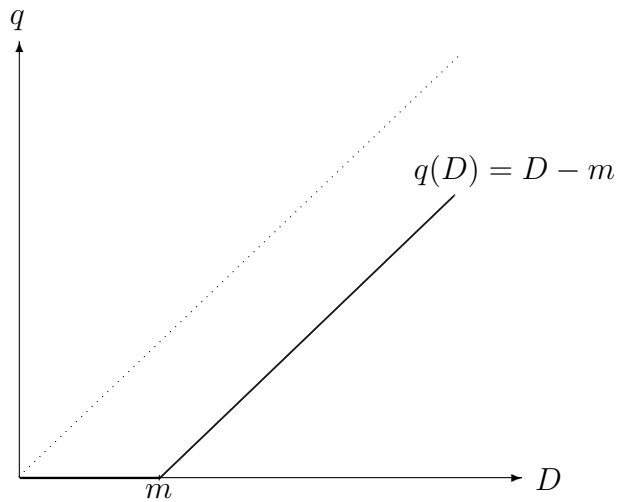
Costly verification and random audit

We have assumed up to now that the insurer had to choose the set of damage levels it commits to always audit. We can alternatively consider that it chooses, for each level of reported damage, an audit probability $\gamma(\widehat{D})$ (or equivalently a share of damage to audit). We then cannot have $q_A(D, \widehat{D}) = q_A(D) \ \forall \widehat{D}$ as then fraud would be beneficial in expectation. It is thus necessary – to give the insured an incentive to tell the truth – for fraud to be punished. Obviously, the simplest strategy for the insurer would be to infinitely punish fraudsters, what would guarantee the absence of fraud. This is however often impossible, mostly because of limited commitment (we cannot ask the insureds for more than what she has). We thus assume, to simplify, that $q_A(D, \widehat{D}) = -B \ \forall \widehat{D} \neq D$. Under this assumption (we mostly need punishment to be independent of $\widehat{D}$), one can show that – as in the deterministic audit case – we can restrict attention to incentive compatible contracts. We thus focus on contracts such that:

$$(1 - \gamma(D))u(R - \Pi - D + q_N(D)) + \gamma(D)u(R - \Pi - D + q_A(D, D))$$
$$\geq (1 - \gamma(\widehat{D}))u(R - \pi - D + q_N(\widehat{D}) + \gamma(\widehat{D})u(R - \Pi - D - B) \qquad \text{(IC)}$$
$$\forall D, \widehat{D} \in [0, \overline{D}]^2$$

Then, in a competitive setting, the optimal contract is obtained by maximizing insureds' expected utility:

$$V = \int_0^{\overline{D}} [(1 - \gamma(D))u(R - \Pi - D + q_N(D)) + \gamma(D)u(R - \Pi - D + q_A(D, D))] \, dF(D)$$

under the incentive constraints (IC) and the participation constraint of the insurer:

$$\Pi - \int_0^{\overline{D}} [(1 - \gamma(D))q_N(D) + \gamma(D)(q_A(D, D) + c)] \, dF(D) \geq 0$$

The complexity of this program comes from the fact that we cannot identify which incentive constraint(s) binds at the optimum. However, assuming a CARA utility function $\left(u(C) = -\frac{1}{\alpha}e^{-\alpha C}\right)$ greatly simplifies this issue. In this case, the factor $-\frac{1}{\alpha}e^{-\alpha(R-\Pi-D)}$ vanishes from (IC). In other terms, the use of a CARA utility function, allows to get read of wealth effects in the choice to fraud or not, that is from the choice between lotteries $[q_N(D), (1 - \gamma(D)); q_A(D, D), \gamma(D)]$ and $[q_N(\widehat{D}), (1 - \gamma(\widehat{D})); -B, \gamma(\widehat{D})]$. A direct consequence of these simplifications is that it is then enough – to give the insureds an incentive to tell the truth – that the (IC) constraint is satisfied for the lowest level of coverage. It is usually the case when $D = 0$ and $q_N(0) = q_A(0, 0) = 0$. The incentive constraint then becomes (replacing $\widehat{D}$ by D): $\forall D \in [0, \overline{D}]$,

$$u(R - \Pi) \geq (1 - \gamma(D))u(R - \pi + q_N(D)) + \gamma(D)u(R - \Pi - B) \qquad \text{(IC.CARA)}$$

What amounts to:

$$\gamma(D) = \frac{u(R - \Pi + q_N(D)) - u(R - \Pi)}{u(R - \Pi + q_N(D)) - u(R - \Pi - B)}$$

We thus optimally obtain an audit probability $\gamma(D)$ increasing in the level of coverage $q_N(D)$.

It is moreover possible to show[7] that we obtain in the CARA case a generalization of the result obtained with deterministic audit:

Proposition. *If insureds have a CARA utility function and the punishment B is high enough[8] the optimal contract is of the form:*

- $q_A(D) = q_N(D) = 0$ *if* $D \leq m$

- $q_A(D) = D - k$ *and* $q_N(D) = D - k - \eta(D)$ *if* $D > m$ *with*
  $\eta(D) > 0$, $\eta'(D) < 0$, $\eta(m) = m - k$ *and* $\eta(D) \to 0$ *when* $D \to \infty$

with $m > 0$ *and* $0 < k < m$

---
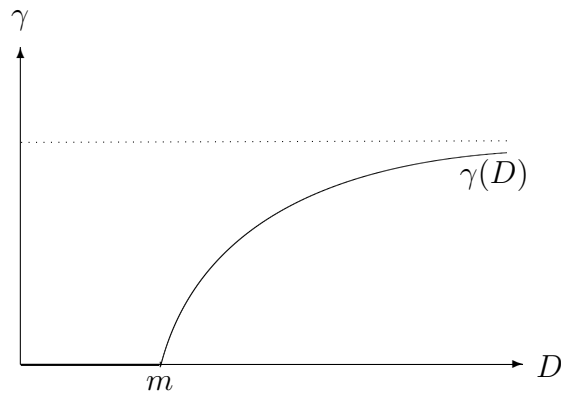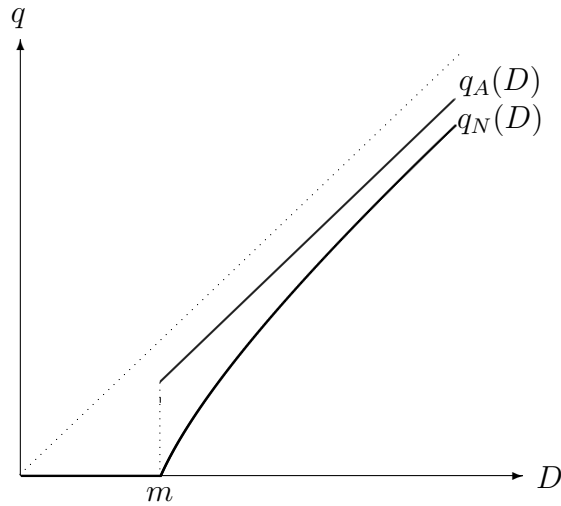
[7]cf. M.-C. Fagart and P. Picard (1999), "Optimal Insurance Under Random Auditing", The Geneva Papers on Risk and Insurance Theory, 24, 29-54.
[8]Otherwise we are back in the deterministic case.

Observe first that **the optimal contract involve a deductible**. It is optimal not to cover low level of coverage, as it would imply having to pay the audit cost with a given probability. For highest level of damage, that exists **two sorts of coverage**: one if the damage is not audited and the other, higher, if it is. In the last sort, marginal damage is fully insured. This is because $q_A(\cdot)$ has no incentive role in (IC.CARA). Coverage $q_N(D)$ is moreover optimally lower than $q_A(D)$. This comes from the fact that – through the incentive constraint – an increase in $q_N(D)$ increases the audit probability $\gamma(D)$, what ultimately increases premium, through the participation constraint of the insurer.

Remark. *The use of a CARA utility function greatly simplifies the model as risk aversion then doesn't depend on the size of the loss. If, instead, the insured risk aversion where decreasing with wealth, the incentive to fraud would be higher for small level of damage. It would then be optimal to offer a positive coverage even for small level of damage to reduce this incentive.*

Costly falsification

Another branch of the economic analysis of insurance fraud assumes that it is impossible for the insurer to verify a fraud, but that it is costly for the insured to falsify her statement. Formally, we assume that it will cost a amount $c(\widehat{D} - D)$ for an insured suffering a damage D to declare a damage $\widehat{D}$, with $c(0) = 0$, $c'(\cdot) > 0$ and $c''(\cdot) > 0$. In this case and using previous notations, the expected utility of an insured suffering a damage D and chooses to declare $\widehat{D}$ writes:

$$V = \int_0^{\overline{D}} u\left(R - \Pi - D + q(\widehat{D}) - c(\widehat{D} - D)\right) dF(D)$$

The insured would then optimally choose $\widehat{D}(D)$ such that:

$$u'\left(R - \Pi - D + q\left(\widehat{D}(D)\right) - c\left(\widehat{D} - D\right)\right) . \left[q'\left(\widehat{D}(D)\right) - c'\left(\widehat{D}(D) - D\right)\right] = 0$$

In words, she will then choose to overstate her damage as long as her marginal utility to do so exceeds the marginal cost of fraud.

Several characteristics of the optimal contract can be obtained with the above equation.

- If $c'(\cdot) < 1$ (that if overstating is damage of 1 euro cost less than one euro) optimal insurance is never full ($q'(\widehat{D}) < 1 \; \forall \widehat{D}$)

- If $c'(0) = 0$, insurance fraud can only be avoided by offering a fix coverage (*i.e.* independent of the size of the damage). If we want to increase the coverage for high level of damage, fraud will necessarily appear.

- If $c'(0) = \delta > 0$, the optimal contract is of the form $q(D) = q_0 + \delta D$ with $q_0 > 0$. Small damages are over-insured whereas, as $q'(D) = \delta < 1$, high damages are under-insured.

Extensions and empirical evidence

The models presented above obviously have several limitations and can be extended in various directions. Among those the most studied, the existence of "moral" concerns preventing the insureds to fraud, are pretty easy to implement. Things become more complicated as one assumes that the insureds have different behaviors with respect to fraud, some being "honest", the other "opportunistic". We then have a model with both moral hazard and adverse selection.

As already raised, it is by essence difficult to estimate the real weight of insurance fraud. Some studies still try to analyze to what extend such fraud can be detected in data. To do so they build on a simple assumption: if fraud has a high enough probability to succeed, one should observe higher damages when (everything else being equal) the deductible is higher. Results are then quite significant. It is for exemple the case of a study on automobile insurance in Canada[9]. It indeed turns out that, in the absence of witness, reported damage are 24.6% to 31.8% higher with a 500 $ deductible than for a 250 $ deductible (whereas the difference is not significant in the presence of witness). This results thus seems to be caused by fraud (and not to adverse selection) as it is highly linked to the presence of a witness.

---

[9]G. Dionne and R. Gagné (2001), "Deductible Contracts Against Fraudulent Claims: Evidence from Automobile Insurance", *Review of Economics and Statistics*, 83(2) 290-301.

Chapter 5

Extensions and exercises

5.1 Extensions of Mossin's model

Mossin's model with default

Let us consider the standard Mossin's model (revenue R, loss D with probability p, Von Neumann Morgenstern utility function u(.)) without loading factor ($\lambda = 0$). Let q be the coverage paid by the insurer in case of loss. We however assume that this coverage is only paid with probability σ, as with probability $1 - \sigma$, the insurer company fails.

1. Compute the actuarially fair premium taking into account the default probability.

2. Explain (intuitively) why it isn't realistic.

3. Assuming that the effective premium is the actuarially fair one computed above, find the optimal coverage chosen by an insured (NB : the premium being fixed, there is no need to analyze indifference curves).

4. Show that, at the optimum, the marginal utility in one state can be obtained as a linear combination of marginal utilities in the two others.

5. Infer if insurance is full at the optimum.

6. How can we explain this result?

7. One can show that an increase in risk aversion then doesn't necessarily increase coverage. Explain intuitively why.

Mossin's model with exceptions

Let us consider the standard Mossin's model (revenue R, loss D with probability p, Von Neumann Morgenstern utility function u(.)) without loading factor ($\lambda = 0$). Let q be the coverage paid by the insurer in case of damage. There is unluckily several exceptions in the contract. In case of loss, the coverage is only paid with probability σ. With probability $1 - \sigma$, we fell in an exception and the loss is not covered. However, in these cases, the insurer pay the premium back to the insured (what make this exercise different from the previous one).

1. Compute the actuarially fair premium.

2. Write the first and second order conditions of the maximization program of an insured when the premium is actuarially fair.

3. Infer if insurance is full at the optimum.

4. Compare with the result of the previous exercise and comment.

5.2 Insurance demand and exogenous risk

Let us consider an individual with revenue R and an increasing and concave utility function $u(.)$ who faces two independent risks:

- *loose an amount D with probability p*

- *loose an amount K with probability θ.*

An insurance company offer to insure at rate π the first risk (q units of coverage against this risk can then be bought at price $\Pi = \pi.q$). The second risk is however assumed to be uninsurable (for example because it isn't diversifiable). The aim of this exercise is to analyze the effect of this exogenous risk on insurance demand. Recall that in the absence of this second risk, the optimal coverage q_0^* is defined by:

$$\frac{u'(R - D + (1 - \pi)q_0^*)}{u'(R - \pi q_0^*)} = \frac{\pi}{p}\frac{1 - p}{1 - \pi}$$

1. Write the expected utility reach for a coverage q in the presence of the second risk.

2. Write the equation that q_1^*, the optimal coverage in the presence of the second risk, has to verify.

3. Show that, when the premium is actuarially fair ($\pi = p$), $q_1^* = q_0^*$

4. Show that it is also the case ($q_1^* = q_0^*$) with a positive loading factor ($\pi > p$), if $u(C) = -\frac{1}{\alpha}e^{-\alpha C}$ (CARA utility function)

5. Using a first order Taylor expansion in K around 0, show that for K small enough:

   - $q_1 > q_0$ if risk aversion is decreasing (DARA)
   - $q_1 < q_0$ if risk aversion is increasing (IARA)

5.3 On the value of genetic information

Consider the standard Mossin's model (revenue R, loss D with probability p, Von Neumann Morgenstern utility function $u(.)$). Here the loss D is associated with the "cost" of an illness linked to a specific gene. The gene has two forms, one favorable L: in which case the probability to develop illness is p_L, and one unfavorable H in which case the probability is p_H. A priori, the share of type H genes in the population is b (and the one of type L, $(1 - b)$). Without genetic screening, the individuals don't know their type.

1. A private insurance can insure against the at stake illness. What is the actuarial premium offered for full insurance? (The insurer has no information on types but knows p_L, p_H, b and D)

2. What is the expected utility then achieved?

Now assume that each individual can, without cost, perform a genetic test that provides her with her type. On its side, the insurer can use this test to price but cannot require its realization.

The insurer a priori offers 3 premiums. One if the test is L: π_L, one without test: π_M and one if the test is H: π_H, with $\pi_L \leq \pi_M \leq \pi_H$.

3. Has an individual having performed the test an interest to show its result to its insurer?

4. What are then the premiums effectively chosen?

5. Should an individual take the test?

6. Compute the premiums effectively chosen if pricing is actuarially fair.

7. What is then the ex-ante (before test) expected utility?

8. What can we infer from this about the value of genetic information?

5.4 Genetic information and self-insurance

[Drawn from Crainich, D., "Self-Insurance With Genetic Testing Tools", *Journal of Risk and Insurance*, 84(1), 73-94, 2017.]

We just saw, in the previous exercise, that genetic information might have negative value. We however neglected then the effect this information may have on preventive behavior.

Let us consider the previous model (revenue R, utility function $u(\cdot)$ increasing and concave, and a population shared into two categories H and L based on genetic) and include a preventive effort. In case of disease, a treatment is available at a cost D paid by patients. We only focus here on financial consequences of disease. The severity of the disease and thus the financial cost of the associated treatment can be reduced through self-insurance efforts undertaken before the potential appearance of the disease (still in the same unique theoretical period, as in the course). We denote by c the amount spent on these efforts. The treatment cost is then $D(c)$ with $D'(c) < 0$ and $D''(c) > 0$.

Insurance is available at an actuarially fair premium and individuals chooses the share of loss it insures, denoted q. Individuals' self-insurance actions are assumed to be observed by insurers. Insurance premia and indemnities are thus contingent on self-insurance actions and are respectively equal to $p_i.q.D(c)$ and $qD(c)$ (with $i = L, H, M$ according to the applicants' type and to whether this information is provided to insurers; with a priori $p_M = bp_H + (1 - b)p_L$).

Assume first that genetic tests are not available. Individuals and insurers have no knowledge of genetic types. Insurance contracts are written and individuals' decisions are made according to the average probability of disease.

1. Show that individuals then all choose to be fully insured ($q^* = 1$).

2. Derive from this the condition defining the optimal level of self-insurance.

Assume now that genetic tests are available, that all individuals take them and that insurers know the results. It then sell type-specific contracts.

3. Show that both type of individuals then all choose to be fully insured ($q^* = 1$).

4. Derive from this the condition defining the optimal level of self-insurance for each type and compare it to the one found without genetic information.

We finally assume that genetic tests are available, that all individuals take them but that insurers cannot obtain the results. We thus have asymmetric information of types.

5. In which cases can an insurance company offer a separating menu? Are both types then fully insured (without computation)?

6. Use these results to compare the optimal efforts to those found in question 4. Conclude that asymmetric information can lead to an increase in self-insurance efforts.

5.5 Health risks and bidimensional utility

In the particular case of health risks, one cannot reasonably limit the analysis to monetary losses linked to illness. We instead use in these cases utility functions with two arguments $u(w, h)$ that depends on both wealth w and health h. We assume that utility is increasing and concave in both arguments: $u_1'(\cdot) > 0$, $u_2'(\cdot) > 0$, $u_{11}''(\cdot) < 0$ and $u_{22}''(\cdot) < 0$. We are interested here on the role played by the crossed derivative $u_{12}''(\cdot)$ (or $u_{21}''(\cdot)$).

We consider the case of an individual facing health risk. If she is healthy, her revenue is R and her health status is denoted S_H. She can fall sick with probability p. In this case, she suffer both a loss of revenue (the cost of treatment) and a worsening of her health status. Without insurance her revenue is then $R - D$ and her health status is denoted $S_B < S_H$.

Health risk and insurance

We consider first a one-period model in which the individual can buy an insurance that cover all or part of the cost of treatment. More precisely, she can buy, at the actuarial rate ($\Pi = p.q$), an insurance that covers $q \in$ of the treatment cost if she fall sick.

1. Write the expected utility achieved by an individual who buys $q \in$ of coverage.

2. Write the equation that the optimal coverage q^* has to verify.

3. Show that an individual chooses full coverage $q = D$ when her utility function is separable: $u(w, h) = f(w) + g(h)$

4. In the general case $u(w, h)$, under which condition on $u_{12}''(\cdot)$ the individual would like to be more than fully insured? **Interpret**.

Health risk and prevention

*We consider now a case **without insurance** in which the individual can exert an effort of prevention that reduce the risk of falling sick. To do so we build a **two-periods** model:*

- *During the first period, the individual is healthy: $h = S_H$. She can exert a effort of prevention at a **monetary** cost e. If she chooses a level of prevention e (continuous) her revenue is $R - e$ (this model is thus different from the one developed in section 4.2). This effort will decrease the probability of illness occurrence $p(e)$ (with $p'(\cdot) < 0$ and $p''(\cdot) > 0$) at the following period.*

- *During the second period, if the individual doesn't fall sick (with probability $1 - p(e)$), her revenue is R and her health status S_H. However, if she fall sick (with probability $p(e)$), she suffer a treatment cost D (her revenue is $R - D$) and her health status becomes $S_B < S_H$.*

1. Write the expected utility achieved by the considered individual (assuming that she doesn't discount the second period).

2. Write the equation that defines her optimal prevention effort e^* under the form $F(e^*, D, S_B, S_H) = 0$

3. Using the partial derivatives of $F(\cdot)$, state how this level of prevention varies with D and S_B. Interpret.

4. We now want to know how e^* varies with S_H. Show that if $u''_{12}(\cdot) < 0$, e^* is increasing in S_H. **Interpret**.

5.6   Life insurance and savings

[Drawn from Lockwood, L. M., "Bequest motives and the annuity puzzle", *Review of Economic Dynamics*, 15(2), 226-243, 2012.]

Consider the wealth allocation decision of an individual who lives two periods with probability p and lives one period otherwise. In the first period, the individual chooses how much of his revenue, R, to consume (c_1), save (s), and annuitize (π), $c_1 + s + \Pi = R$. Saving, $s \geq 0$, earns an interest rate i regardless of whether the individual lives. Annuities, $\pi \geq 0$, earn a larger return than saving if the individual lives, $r > i$, but return nothing if the individual dies. In old age (period 2), if alive, the individual receives no income on top of his accumulated saving and annuities. She then splits her wealth between consumption (c_2) and bequest (b_2): $c_2 + b_2 = (1 + r)\pi + (1 + i)s$. Bequests if the individual dies young is $b_1 = (1 + i)s$. The choice variables of our individual are then π, s and c_2.

1. Write consumptions and bequests in each state of nature c_1, c_2, b_1, b_2 as a function of the choice variables (c_2, π, s) and the exogenous variables (i, r, R).

The individual expected utility accounts for both her consumptions and bequests, and doesn't discount the second period. More precisely we assume that she maximizes

$$u(c_1) + (1-p)v(b_1) + p[u(c_2) + v(b_2)]$$

with $u(\cdot)$ $(u'(\cdot) > 0$, $u''(\cdot) < 0)$ her own utility function, and $v(\cdot)$ $(v'(\cdot) > 0$, $v''(\cdot) < 0)$ the utility function of her legatee.

We also assume that annuities are sell at a fair price, that might include a loading factor λ (in cases the annuity is paid); and that the insurer earns the same returns than the private individuals on its investments.

2. Show that this last assumption gives $(1+r) = \frac{(1+i)}{p(1+\lambda)}$ (taking into account the timing of cash-flows).

3. What are the optimal savings (s^*) and annuitization (π^*) when $\lambda = 0$? Remark that we then have $b_2 = b_1 = (1+i)s^*$. Conclude on the respective motives for savings and annuitization.

4. How do that change when $\lambda > 0$? (Show that then $b_2 < b_1$ and comment on motives.)