On the design of optimal parametric insurance

Alexis Louaas and Pierre Picard*

19 August 2022

Abstract

We analyze parametric insurance in an asymmetric information setting. The insurance payout is a function of a publicly observable parameter vector, while the actual loss is private information of the policyholder. The parameter vector yields a loss index, which is the best estimate of the loss, the basis risk being the random difference between the actual loss and the loss index. We show that the design of optimal parametric insurance depends on whether the loss index and the basis risk are independently distributed or not, and we analyze how insurance demand is affected by the size of the basis risk and by the attitude toward risk of the policyholder.

*CREST-Ecole Polytechnique.

1 Introduction

Parametric insurance consists in conditioning the indemnity paid to the policyholder not on the financial value of the losses incurred, but on publicly observable information correlated with these losses. This information may be parametric *stricto sensu*, as for instance in crop insurance when the payment to the farmer depends on average rainfall in a specific area during a given period. It may also take the form of a modeled-loss index reflecting the specific exposure of the policyholder. This is the case in property insurance when the payout depends on an index corresponding to the potential damages of the policyholder, such as the expected loss calculated on the basis of the wind speed of a hurricane measured at various points along its path, or according to the magnitude and epicenter of an earthquake.

The main advantage of parametric insurance is to eliminate the moral hazard issue and to avoid the claim-handling costs associated with the assessment of policyholders' actual losses. The primary concern is the basis risk retained by the policyholder, i.e. the fact that the parametric insurance trigger does not exactly match his actual losses. Parametric insurance covers are offered by direct insurers and they are also widely used as triggers in alternative risk transfer mechanisms, particularly catbonds. They now play an important role in the coverage of agriculture climate-related risks (particularly in developing countries) and of property catastrophic risks, and they now tend to spread over a larger range of risk lines.[1]

It is in the area of agriculture risk management that parametric insurance has been most widely studied. Without being exhaustive, this includes the analysis of area-yield crop insurance by Miranda (1991) and Barnett and al. (2005), the interaction with

[1]For illustrative purposes, Power Protective Re Ltd, a parametric catbond sponsored by the Los Angeles Department of Water and Power and launched in 2020, highlights this spreading of parametric insurance to new areas, by being the first wildfire catastrophe bond issuance to benefit a municipal utility. In a different area, the coverage of offshore wind farms in Taiwan (see Liao et al., 2021) also illustrates the usefulness of parametric insurance for climate-related risks.

the poverty issue in lower-income countries by Barnett and Mahul (2007), Chantarat et al. (2007), Chantarat et al. (2013) and Skees (2008), the effect of insurance on the adoption of new technologies by Mobarak and Rosenzweig (2013), Carter et al. (2016), Biffis and Chavez (2017) and Biffis et al. (2022), and the statistical analysis of the basis risk by Carter et al. (2017) and Kusuma et al. (2018). The design of optimal parametric insurance in a microeconomic setting hitherto has received much less attention. As we will see below, there is a logical link between this issue and the insurance demand problem with background risk, as studied by Gollier (1996). Clarke (2016) analyses parametric insurance in an expected utility setting, with the main conclusion that the basis risk may make it unattractive for strongly risk-averse policyholders. Bryan (2019) considers the case where the policyholder may be ambiguity averse. Using data from two RCTs conducted in Malawi and Kenya, he shows that the income loss from ambiguity aversion may be substantial. Teh and Woolnough (2019) analyze how parametric triggers can be compared, and they determine a partial order ranking for any risk averse individual.

Our objective in what follows is to analyze parametric insurance by formulating it as the optimal solution to a risk-sharing problem under asymmetric information: the policyholder has private information on the loss incurred, while the insurance payout depends on a publicly observable information that imperfectly reflects the actual loss. This public information takes the form of a multi-dimensional parameter vector, and the issue then is how it should be used to define the indemnity paid to the policyholder. The basis risk is the difference between the loss incurred and the conditional expected loss based on this public information, and it affects the quality of the parametric insurance cover. While the context of imperfect information is often implicit in the approach to parametric insurance, several important issues emerge when the problem is stated in that way. Firstly, should the insurance payout be a function of the conditional expected loss (i.e., the best estimate of the loss based on the parameter vector), or is it optimal to condition this payout on the parameter vector itself. Secondly, how can we characterize the optimal parametric indemnity schedule? In particular, does it look like commonly observed policies, such as index-based deductible contracts, or is it different? Thirdly, under which conditions does a change in public information, inducing a change in the basis risk, improve the efficiency of the risk-sharing mechanism? Fourthly, does the usual relationship between insurance demand and risk aversion extend to the case of parametric insurance? As we will see, the answers to these questions depend heavily on whether the parameter vector and the basis risk are independently distributed or not, which in turn depends on the stochastic relationship between public and private information.

The rest of the paper explores these issues and it is organized as follows. Section 2 presents our general setting, with private and public information, the latter taking the form of a parameter vector that defines the set of feasible parametric insurance contracts. We highlight the specificity of index-based insurance, i.e., the case where the parametric cover depends on the best estimate of the loss (called the loss index), and we define the basis risk. Section 3 focuses attention on the information structure based on the parameter vector. We show that a finer (i.e. more precise) information structure improves the quality of the parametric cover and also that it induces a lower basis risk in the sense of Rothschild and Stiglitz (1970). Section 4 characterizes the optimal parametric insurance, depending on whether the basis risk and the parameter vector are independent or not, respectively, and we show that conclusions strongly differ in both cases. We particularly focus attention on the shape of the indemnity schedule, and on how it is affected by the policyholder's attitude toward risk, characterized by risk aversion and prudence. Section 5 concludes, and Section 6 includes the proofs.

2 Setting

2.1 Model

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, with states of the world $\omega \in \Omega$. We consider a risk-averse individual who incurs a state-contingent loss : his risk exposure is defined by random variable $X(.) : \Omega \to \mathbb{R}_+$, with loss $X(\omega)$ in state ω. The state of the world - and thus the loss $X(\omega)$ - is privately observed by the individual. However, in each state ω, a signal $Y(\omega) \in \mathcal{S}$ is publicly observed, where $\mathcal{S} \subset \mathbb{R}^n$ is a measurable space. The random variable $Y(.) : \Omega \to \mathcal{S}$ defines a multi-dimensional state-dependent public information $Y(\omega)$, that will be called the *parameter vector* in what follows. Parametric insurance consists in conditioning the insurance payout on the parameter vector $Y(\omega) \in \mathcal{S} \subset \mathbb{R}^n$ rather than on the loss $X(\omega)$. We note that $\{\mathcal{S}, Y\}$ defines an *information structure* because observing signal $y \in \mathcal{S}$ implies $\omega \in \mathcal{O}(y)$, with $\mathcal{O}(y) = Y^{-1}(y)$, and $\mathcal{P} =\{\mathcal{O}(y), y \in \mathcal{S}\}$ is a partition of Ω.

A *parametric insurance contract* is defined by an indemnity function $I(.) : \mathcal{S} \to \mathbb{R}_+$ that specifies the insurance payout $I(y)$ as a function of parameter vector $y \in \mathcal{S}$. When the individual purchases such a parametric cover at price P (the insurance premium), his random final wealth is

$$W_f = w_0 - X + I(Y) - P,$$

where w_0 is his initial wealth. The individual's attitude toward risk is characterized by a twice-differentiable von Neumann-Morgenstern utility function $u(.)$, such that $u' > 0, u'' < 0$, and his expected utility is written as

$$\mathbb{E}u(W_f) = \mathbb{E}u(w_0 - X + I(Y) - P). \tag{1}$$

The insurance premium is given by

$$P = (1 + \sigma)\mathbb{E}I(Y) \tag{2}$$

where σ is the loading factor, with $\sigma \geq 0$. An optimal parametric insurance contract maximizes $\mathbb{E}u(W_f)$ with respect to P and $I(.) : \mathcal{S} \to \mathbb{R}_+$ subject to the pricing rule (2).

Let $Y(\Omega) \subset \mathcal{S}$ be the set of possible parameter vectors. Let us define $Z(.) : Y(\Omega) \to \mathbb{R}_+$ by

$$Z(y) = \mathbb{E}[X(\omega) \mid Y(\omega) = y] \quad \text{for all } y \in Y(\Omega), \tag{3}$$

In what follows, $Z(y)$ is called the *loss index* associated with parameter vector $y \in \mathcal{S}$. In words, the loss index $Z(y)$ is the conditional expected value of the loss when parameter vector y is observed, with $\mathbb{E}[X \mid Z = z] = z$ for all $z \in \mathbb{R}_+$.

We also define $\widetilde{\varepsilon}(.) : \Omega \to \mathbb{R}$ by

$$\widetilde{\varepsilon}(\omega) = X(\omega) - Z(Y(\omega)) \text{ for all } \omega \in \Omega, \tag{4}$$

with

$$\mathbb{E}[\widetilde{\varepsilon}(\omega) \mid Y(\omega) = y] = 0 \quad \text{for all } y \in Y(\Omega).$$

Thus, $\widetilde{\varepsilon}(\omega)$ is the difference between the true loss $X(\omega)$ and the loss index $Z(Y(\omega))$ in state ω, and it is called the *basis risk*. In words, the basis risk $\widetilde{\varepsilon}$ is a zero-mean random variable corresponding to the difference between the loss X and its conditional expected value $Z(Y)$.

Frequently, under parametric insurance, the indemnity paid to the policyholder is a function of the loss index $Z(y)$ induced by parameter vector $y \in \mathcal{S}$, and not a function of the parameter vector y itself. This is at least to make interpretation easier, because in that case the insurance payout is a function of what the policyholder is expected to have lost, given the available information included in the parameter vector. We will refer to such a case as *index-based insurance*. Hence, a parametric insurance contract defined by $P, I(.), \mathcal{S}$ is index-based when there exists $J(.) : \mathbb{R}_+ \to \mathbb{R}_+$, such that $I(y) = J(Z(y))$ for all $y \in \mathcal{S}$.

We have $X \equiv Z + \widetilde{\varepsilon}$ (where, here and in what follows, "$\equiv$" means "equal in distribution"), which shows that an optimal index-based insurance contract maximizes

$$\mathbb{E}u(W_f) = \mathbb{E}u(w_0 - Z - \widetilde{\varepsilon} + J(Z) - P), \tag{5}$$

with respect to P and $J(.) : \mathbb{R}_+ \to \mathbb{R}_+$, subject to

$$P = (1+\sigma)\mathbb{E}J(Z). \tag{6}$$

Interestingly, this is formally equivalent to searching for the optimal cover of an individual with insurable risk exposure Z and non-insurable background risk $\widetilde{\varepsilon}$ as studied by Gollier (1996).

2.2 Independence between loss index and basis risk

As will be set out in detail below, the characterization of an optimal parametric insurance contract strongly depends on whether or not the loss index Z and the basis risk $\widetilde{\varepsilon}$ are stochastically independent. The assumptions underlying these two cases become particularly clear when the parameter vector $Y(\omega)$ observed in state ω is a publicly observable subvector of ω. Assume that $\Omega = \Omega_1 \times \Omega_2$, with $\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega$ and let $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ two probability spaces, with $\mathcal{F} = \mathcal{F}_1 \otimes \mathcal{F}_2$. Assume further that $\mathcal{S} = \Omega_1$ and $Y(\omega) = \omega_1$, meaning that the parameter vector coincides with component ω_1 of state vector ω.

When $X(\omega)$ depends additively on ω_1 and ω_2, i.e., $X(\omega) = X_1(\omega_1) + X_2(\omega_2)$, and $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$, i.e., $(\Omega, \mathcal{F}, \mathbb{P})$ is a product probability space combining $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ and $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$, we have

$$Z = X_1 + \mathbb{E}X_2,$$

and

$$\widetilde{\varepsilon} = X - Z = X_2 - \mathbb{E}X_2,$$

which does not depend on the observable parameter vector Y that only reveals ω_1. In that case, the loss index Z and the basis risk $\widetilde{\varepsilon}$ are independently distributed.

Conversely, if the observable and non-observable components of the state vector $\omega = (\omega_1, \omega_2)$ affect the loss $X(\omega)$ either non-additively or in a non-independent way, then generically Z and $\widetilde{\varepsilon}$ are not independent.

2.3 Illustrative examples

We may illustrate the above through the cases of crop insurance and hurricane insurance. In each case, we assume $\Omega = \Omega_1 \times \Omega_2$, with $\omega = (\omega_1, \omega_2), \omega_1 \in \Omega_1, \omega_2 \in \Omega_2$, with $\mathcal{S} = \Omega_1$ and $Y(\omega_1) = \omega_1$.

2.3.1 Crop insurance

Consider the case of a farmer facing uncertain meteorological circumstances and other hazards affecting his crop. Here, w_0 is the harvested crop value under optimal conditions, and loss X is the decrease in this value due to adverse events. The parameter vector corresponds to the intensity of daily precipitations from planting to harvesting, publicly observed through satellite data. Hence, ω_1 and ω_2 correspond to information about raining and to other farm-specific random events (e.g., pest attack, local flood or hail storm), respectively. Assuming that there are 180 days from plantation to harvest gives $\Omega_1 = \mathbb{R}_+^{180}$, and $\omega_1 = (\omega_1^1, ..., \omega_1^{180})$ where ω_1^i is the precipitation intensity on day $i = 1, ..., 180$ with $\mathcal{S} = \mathbb{R}_+^{180}$.

Crop growth simulation models show that effective rainfall (i.e., the difference between rain water and evapotranspiration) is the primary source of soil moisture under rainfed agriculture. Shortage of soil moisture creates crop water stress and reduces growth, with heterogenous effects according to the crop type and to rainfall characteristics, with non-linear effects of rainfall on crop growth. Hence, crop yield depends on raining, but other hazards may also affect crop growth. In this example, $Z(y)$ is the expected decrease in crop yield beyond w_0 when the raining trajectory $y = (y^1, ..., y^{180})$ has been observed. An index-based crop insurance contract would specify the indemnity paid to the farmer, as a function of this expected decrease in crop yield.[2] In this example, Z and $\widetilde{\varepsilon}$ are independent random variables if the rainfall trajectory and other hazards affect the crop yield independently and additively.

2.3.2 Hurricane insurance

Consider an individual located in an area subject to hurricanes, with property at risk w_0. The track of a hurricane is characterized by the longitude and latitude coordinates of its center, and by the speed and direction of wind. Assume that satellite imagery provides this four-dimension information m times along the path of the hurricane. We have $\omega_1 = (\omega_1^1, ..., \omega_1^m) \in \Omega_1 = \mathbb{R}_+^{4m}$ where $\omega_1^i \in \mathbb{R}_+^4$ is the information provided by satellite data, with $i = 1, ..., m$, and $\mathcal{S} = \mathbb{R}^{4m}$. Furthermore, ω_2 corresponds to other factors that may affect damages from a hurricane, e.g., seasonal tidal range favoring storm surge or torrential rains causing flooding and triggering landslides, when a weakening hurricane interacts with inland weather features. A vulnerability model then relates the data on the hurricane track (i.e., ω_1) and the potential damages incured in a given territory, thereby leading to the loss index $Z(y)$.[3] Here, if the hurricane track and seasonal or local factors affect damages to property independently and additively, then Z and $\widetilde{\varepsilon}$ are independently distributed, and this is not the case otherwise.

[2] As a practical illustration of the construction of a yield index using rainfall data, see for instance Omondi et al. (2021): they analyze how expected crop growth in Kenya depends on satellite weather data, including onset days, rainfall depths, dry spells, and rainfall occurrence for four crop growth stages.

[3] Typically, a hurricane model yields an equivalent local wind speed (i.e., wind speed over the territory with property at risk) for each measurement point of the hurricane track. A usual assumption consists in postulating that total damages due to wind depend on the maximum equivalent local wind along the hurricane track. The modelling of damages associated with hurricanes shows that wind produces damages when their speed exceeds a threshold that depends on the property at risk, and over this threshold property damages may increase rapidly with wind speed. See Katz (2002), Pielke (2007), Nordhaus (2010) and Emanuel (2011).

3 Comparing information structures

As defined above, basis risk $\widetilde{\varepsilon} = X - Z$ is the random difference between the true loss incurred by the policyholder and the loss index, i.e., the parameter-based best estimate of this loss. It is thus intuitive that a larger basis risk corresponds to a less efficient parametric insurance coverage. This section relates this intuition with the information structure defined by the parameter vector. To do so, let us consider two information structures $(\mathcal{S}_1, Y_1)$ and $(\mathcal{S}_2, Y_2)$ - i.e., two definitions of the parameter vector - with $\mathcal{P}_i = \{\mathcal{O}_i(y_i), y_i \in \mathcal{S}_i\}$ the partition of Ω generated by $(\mathcal{S}_i, Y_i)$, where $\mathcal{O}_i(y_i) = Y_i^{-1}(y_i)$ for $y_i \in \mathcal{S}_i$ and $i = 1, 2$. The corresponding loss index and basis risk are denoted Z_i and $\widetilde{\varepsilon}_i$ for $i = 1, 2$, respectively, with $\mathbb{E}\widetilde{\varepsilon}_i = 0$, and

$$Z_1 + \widetilde{\varepsilon}_1 \equiv Z_2 + \widetilde{\varepsilon}_2 \equiv X. \tag{7}$$

In what follows, we say that information structure $(\mathcal{S}_1, Y_1)$ weakly dominates information structure $(\mathcal{S}_2, Y_2)$ if the optimal parametric-insurance contract based on $(\mathcal{S}_1, Y_1)$ is weakly prefered to the optimal contract based based on $(\mathcal{S}_2, Y_2)$, for any increasing concave utility function, and dominance is strong if, in addition, optimal expected utility is stricly larger for at least one utility function. According to Definition 1, information structure $(\mathcal{S}_1, Y_1)$ is finer than information structure $(\mathcal{S}_2, Y_2)$ if, whatever the state of nature ω, observing parameter vector $y_1 = Y_1(\omega) \in \mathcal{S}_1$ provides a more precise information on ω than observing $y_2 = Y_2(\omega) \in \mathcal{S}_2$, an equivalent characterization being provided by Lemma 1. Proposition 1 states that, in such a case, $(\mathcal{S}_1, Y_1)$ weakly dominates $(\mathcal{S}_2, Y_2)$, the dominance being strong when using $(\mathcal{S}_1, Y_1)$ allows to increase (respect. decrease) the insurance payout in states where the loss is larger (respect. lower), such a discrimination between states being impossible under information structure $(\mathcal{S}_2, Y_2)$. Proposition 2 then shows that information structure dominance takes the form of a lower basis risk with more risky loss index, risk comparison being in the sense of Rothschild-Stiglitz (1970).[4]

Definition 1 *Information structure 1 is finer than information structure 2, when, for all $y_2 \in \mathcal{S}_2$, there exists a set $\mathcal{K}(y_2) \subset \mathcal{S}_1$ such that*

$$\mathcal{O}_2(y_2) = \cup_{y_1 \in \mathcal{K}(y_2)} \mathcal{O}_1(y_1),$$

with $\{\mathcal{K}(y_2), y_2 \in \mathcal{S}_2\}$ a partition of $\mathcal{S}_1$.

Hence, when signal y_2 is perceived under information structure $(\mathcal{S}_2, Y_2)$, then a signal y_1 is perceived in $\mathcal{K}(y_2) \subset \mathcal{S}_1$ under information structure $(\mathcal{S}_1, Y_1)$. Signal y_2 provides information that $\omega \in \mathcal{O}_2(y_2)$ and signal $y_1 \in \mathcal{K}(y_2)$ provides information that $\omega \in \mathcal{O}_1(y_1)$ with $\mathcal{O}_1(y_1) \subset \mathcal{O}_2(y_2)$, hence a better information provided by $(\mathcal{S}_1, Y_1)$ than by $(\mathcal{S}_2, Y_2)$. We may then express y_2 as a function of y_1 through a function $y_2 = \Phi(y_1)$ such that $\mathcal{K}(y_2) = \Phi^{-1}(y_2)$. Lemma 1 shows that the existence of such a function $\Phi(.) : \mathcal{S}_1 \to \mathcal{S}_2$ is equivalent to the fact that $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$.

Lemma 1 *Information structure $(\mathcal{S}_1, Y_1)$ is finer than information structure $(\mathcal{S}_2, Y_2)$ if and only if there exists a function $\Phi(.) : \mathcal{S}_1 \to \mathcal{S}_2$, such that*

$$Y_2(\omega) = \Phi(Y_1(\omega)) \quad \textit{for all} \quad \omega \in \Omega.$$

As an illustration of Lemma 1, consider the case where $\Omega = \{\omega^1, \omega^2, \omega^3\}, \mathcal{S}_1 = \{y_1^1, y_1^2, y_1^3\}$ and $\mathcal{S}_2 = \{y_2^{12}, y_2^3\}$. Assume $Y_1(\omega^i) = y_1^i$ for $i \in \{1, 2, 3\}$ and $Y_2(\omega^1) = Y_2(\omega^2) = y_2^{12}$ and $Y_2(\omega^3\} = y_2^3$. Hence, signal Y_1 perfectly reveals the state of nature ω, while Y_2 only reveals either $\omega \in \{\omega^1, \omega^2\}$ or $\omega = \omega^3$. Hence $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$, and parameter vector Y_2 can be expressed as a function of Y_1 through function $\Phi(.) : \mathcal{S}_1 \to \mathcal{S}_2$ defined by $\Phi(y_1^1) = \Phi(y_1^2) = y_2^{12}$ and $\Phi(y_1^3) = y_2^3$. Equivalently, $\mathcal{K}(.) = \Phi^{-1}(.)$, with $\mathcal{K}(y_2^{12}) = \{y_1^1, y_1^2\}$ and $\mathcal{K}(y_2^3) = \{y_1^3\}$. In words, knowing Y_1 allows us to know Y_2, but it does not work the other way.

[4] In what follows, when lotteries are compared, the increasing-risk criterion is always in the sense of Rothschild and Stiglitz (1970).

Proposition 1 *Assume that* $(\mathcal{S}_1, Y_1)$ *is finer than* $(\mathcal{S}_2, Y_2)$. *In that case,* $(\mathcal{S}_1, Y_1)$ *weakly dominates* $(\mathcal{S}_2, Y_2)$. *Furthermore, assume that there exist positive-probability sets* $\mathcal{A}_2 \subset \mathcal{S}_2$, $\mathcal{A}_1^1(y_2), \mathcal{A}_1^2(y_2) \subset \mathcal{K}(y_2) \subset \mathcal{S}_1$ *with* $\mathcal{A}_1^1(y_2) \cap \mathcal{A}_1^2(y_2) = \phi$ *for all* $y_2 \in \mathcal{A}_2$, *such that (i):* $X(\omega^1) > X(\omega^2)$ *if* $Y^1(\omega^1) \in \mathcal{A}_1^1(y_2)$ *and* $Y^1(\omega^2) \in \mathcal{A}_1^2(y_2)$ *for* $y_2 \in \mathcal{A}_2$, *and (ii): For some concave utility function* $u(.)$, *we have* $I_2^*(y_2) > 0$ *when* $y_2 \in \mathcal{A}_2$, *where* $I_2^*(.) : \mathcal{S}_2 \to \mathbb{R}_+$ *is the optimal indemnity schedule under information structure* $(\mathcal{S}_2, Y_2)$.*Then* $(\mathcal{S}_1, Y_1)$ *strongly dominates* $(\mathcal{S}_2, Y_2)$.

Proposition 1 is cumbersome, but its intuition is simple. Obviously, when $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$, then any indemnity schedule $I_2(Y_2)$ based $(\mathcal{S}_2, Y_2)$ can be replicated by another indemnity schedule $I_1(Y_1) = I_2(\Phi(Y_1))$ based on $(\mathcal{S}_1, Y_1)$, hence the weak dominance property. More specifically, in its second part, Proposition 1 postulates that there exists a positive-probability set $\mathcal{A}_2 \subset \mathcal{S}_2$ such that any parameter vector $y_2 \in \mathcal{A}_2$ is the image of subsets $\mathcal{A}_1^1(y_2)$ and $\mathcal{A}_1^2(y_2) \subset \mathcal{S}_1$ by function $\Phi(.)$. Hence, information structure $(\mathcal{S}_1, Y_1)$ separates the states ω leading to $\mathcal{A}_1^1(y_2)$ from those leading to $\mathcal{A}_1^2(y_2)$, which cannot be done through $(\mathcal{S}_2, Y_2)$. Assume that the policyholder's loss is larger in the first case than in the second one, and start from the optimal parametric insurance contract based on Y_2. Increasing the insurance payout when $y_1 \in \mathcal{A}_1^1(y_2)$ and decreasing it when $y_1 \in \mathcal{A}_1^2(y_2)$, while keeping the expected payment unchanged, increases the risk-averse policyholder's expected utility for an unchanged insurance premium. This is possible when the utility function is such that the optimal parametric insurance contract based on $(\mathcal{S}_2, Y_2)$ provides positive coverage when $y_2 \in \mathcal{A}_2$.

Proposition 2 *If* $(\mathcal{S}_1, Y_1)$ *is a finer information structure than* $(\mathcal{S}_2, Y_2)$, *then* $\widetilde{\varepsilon}_2$ *is more risky than* $\widetilde{\varepsilon}_1$, *and* Z_1 *is more risky than* Z_2.

Proposition 2 shows that finer information takes the form of smaller basis risk and more risky loss index, the increasing-risk criterion being in the sense of Rothschild-Stiglitz (1970), in both cases. A completely uniformative parameter vector would lead to a constant loss index $Z = \mathbb{E}X$ equal to the unconditionnal expected loss, while the variations of an informative loss index reproduce the change in incurred losses more or less precisely. When the information structure is finer, the parameter vector provides a more precise information on the state, and the index reproduces more closely the changes in the loss, with less residual uncertainty, hence a more variable (more risky) loss index and a lower basis risk.

In Propositions 1 and 2, we have taken as our starting point the informational background of the parametric insurance problem, i.e., a probability space with a loss level in each state combined with an information structure, and we have analyzed the consequences of a finer information structure. Such an improvement in the publicly-available information allows the policyholder to reach a higher expected utility, the parameter vector providing a more precise information on losses, with a smaller basis risk. Since having access to a finer information may be costly, the design of an optimal parametric-insurance cover may then be analyzed as the outcome of a cost-benefit analysis, in which the expected-utility gain attributable to a finer information structure with a lower basis risk has to be balanced against the corresponding cost. In other words, the information structure itself may be a matter of choice.

One is also tempted to look at the reverse question: when we compare randomly-distributed parameter vectors with associated loss indices, do a smaller basis risk and a more risky loss index guarantee a higher optimal expected utility? Proposition 3 says that this is true, but only under very restrictive conditions.

Proposition 3 *Consider loss index* Z_i *and basis risk* $\widetilde{\varepsilon}_i$ *induced by information structures* $(\mathcal{S}_1, Y_1)$ *and* $(\mathcal{S}_2, Y_2)$*, respectively. Assume that* Z_i *and* $\widetilde{\varepsilon}_i$ *are independently distributed for* $i = 1$ *and* 2*. Assume also that* $\widetilde{\varepsilon}_2 \equiv \widetilde{\varepsilon}_1 + \widetilde{\eta}$ *with* $\mathbb{E}\widetilde{\eta} = 0$ *and* $Z_2, \widetilde{\eta}, \widetilde{\varepsilon}_1$ *pairwise independent. Then* $(\mathcal{S}_1, Y_1)$ *dominates* $(\mathcal{S}_2, Y_2)$*.*

Proposition 3 considers the case of two information structures $i = 1$ and 2, each of them inducing independently distributed loss index Z_i and basis risk $\widetilde{\varepsilon}_i$. It is assumed that $\widetilde{\varepsilon}_2$ is distributed as the sum of $\widetilde{\varepsilon}_1$ and of a zero-mean additive noise $\widetilde{\eta}$ with

$Z_2, \widetilde{\eta}, \widetilde{\varepsilon}_1$ pairwise independent. The fact that $\widetilde{\eta}$ and $\widetilde{\varepsilon}_1$ are independent implies that $\widetilde{\varepsilon}_2$ is more risky than $\widetilde{\varepsilon}_1$. Using (7) gives $Z_1 - Z_2 \equiv \widetilde{\eta}$, and the fact that $\widetilde{\eta}$ and Z_2 are independent implies that Z_1 is more risky than Z_2. Finally, the independence of Z_2 and $\widetilde{\varepsilon}_1$ conveys the assumption that the less precise loss index Z_2 is not informative about the residual risk associated with the more precise loss index Z_1, with precision in the sense of lower basis risk. In other words, there is no cross-effect through which the less precise loss index would provide information on the basis risk of the more precise loss index. Proposition 3 establishes that, under these assumptions, any risk-averse policyholder will reach a higher expected utility when parametric insurance is based on the information structure that sustains Z_1 than on the one that leads to Z_2. Needless to say, these are strong assumptions. As will be illustrated below through examples, under less restrictive assumptions, a larger basis risk is not synonymous of a less efficient parametric insurance coverage.

4 Comparative statics

4.1 Case where the parameter vector and the basis risk are independently distributed

Let us define indirect utility $v(w)$ by

$$v(w) \equiv \mathbb{E}_{\tilde{\varepsilon}}[u(w - \tilde{\varepsilon})],$$

with $v' > 0$, $v'' < 0$. When parameter vector Y and basis risk $\widetilde{\varepsilon}$ are independent random variables, we may write

$$\begin{aligned} \mathbb{E}u(W_f) &= \mathbb{E}_Y[\mathbb{E}_{\tilde{\varepsilon}}u(w_0 - Z(Y) + I(Y) - P - \tilde{\varepsilon})] \\ &= \mathbb{E}_Y[v(w_0 - Z(Y) + I(Y) - P)], \end{aligned}$$

This is analogous to the standard approach to risk analysis under independent background risk: when facing a zero-mean independent background risk $\tilde{\varepsilon}$, the individual's attitude toward the risk affecting his insurable wealth is the same as if there were no background risk and his utility function were $v(.)$ instead of $u(.)$. In other words, in that case, the optimal parametric insurance contract maximises $\mathbb{E}_Y[v(w_0-Z(Y)+I(Y)-P)]$ with respect to $I(.) : Y \to \mathbb{R}_+$ and P, subject to (2). This is very similar to a standard optimal insurance problem, with random loss $Z(Y)$ and utility function $v(w)$, the only difference being that the insurance payout depends on the determinants of the loss $Y \in \mathcal{S}$ rather than on the loss itself $Z(Y) \in \mathbb{R}_+$. It is very intuitive and confirmed by the proof of the following proposition that two parameter vectors $y_1, y_2 \in Y$ such that $Z(y_1) = Z(y_2)$ should lead to the same indemnity. Thus the optimal contract is index-based and it maximizes

$$\mathbb{E}u(W_f) = \mathbb{E}v(w_0 - Z + J(Z) - P),$$

with respect to P and $J(.) : \mathbb{R}_+ \to \mathbb{R}_+$, subject to (3). We know that the optimal solution to such a problem is a straight deductible contract, unless there is no loading, in which case full insurance would be optimal,[5] and thus we have the following.

Proposition 4 *If Y and $\tilde{\varepsilon}$ are independently distributed, then the optimal parametric insurance contract is index-based. The insurance payout $J(Z)$ is equal to the conditional expected loss Z if $\sigma = 0$, and it $z_0 > 0$ if $\sigma > 0$. In other words $J(Z) = \max\{Z - z_0, 0\}$, with $z_0 = 0$ if $\sigma = 0$ and $z_0 > 0$ if $\sigma > 0$.*

The analogy with the optimal insurance problem under an independent background risk allows us to answer the simple but controversial following question: considering two individuals with the same risk exposure X and the same publicly observable parameter vector Y, does the more risk averse one purchase more parametric insurance? Put differently, does the standard result according to which more risk aversion means more insurance demand, also applies in the case of parametric insurance?

[5] See Gollier (1996).

To answer this question, consider two individuals indexed by $h = 1, 2$, with utility functions $u_1(w)$ and $u_2(w)$, indirect utility functions $v_1(w)$ and $v_2(w)$, and optimal deductibles z_{01} and z_{02}, respectively. According to standard results in the theory of insurance demand, the larger the index of absolute risk aversion, the larger the demand for insurance, and thus, under constant loading, the lower the deductible. Let us denote $A_u(w) = -u''(w)/u'(w)$ and $A_v(w) = -v''(w)/v'(w)$ the Arrow-Pratt index of absolute risk aversion, for direct and indirect utility function $u(.)$ and $v(.)$, respectively. The question we are asking is whether $A_{u_2}(w) > A_{u_1}(w)$ for all w implies $z_{02} < z_{01}$. Since the optimal deductible maximizes the policyholder's expected indirect utility, we know that $z_{02} < z_{01}$ if $A_{v_2}(w) > A_{v_1}(w)$ for all w. Consequently, the larger the degree of risk aversion (for utility function u), the larger the demand for parametric insurance if $A_{u_2}(w) > A_{u_1}(w)$ implies $A_{v_2}(w) > A_{v_1}(w)$.

When this last property holds, we say that the background risk preserves comparative risk aversion in the sense of Arrow-Pratt. It has been shown in the literature on background risks that additional assumptions are required for this to be true. This is the case, in particular, when $h = 1$ and/or $h = 2$ displays decreasing risk aversion. This is also true if one reinforces the comparison of risk aversion by following the approach of Ross (1981).[6] Hence, either by postulating decreasing absolute risk aversion, or by comparing risk aversion in the manner of Ross, we may conclude that the existence of an independent background risk preserves comparative risk aversion. When at least one of these two assumptions hold, we say that risk aversion is strongly comparable. The following Proposition states that, in such a setting, the more risk

[6] $h = 2$ is said to be more risk averse than $h = 1$ in the sense of Ross (1981), if there exists a positive scalar λ and a decreasing and concave function g such that $u_2(w) = \lambda u_1(w) + g(w)$ for all w. It can be shown that comparative risk aversion in the sense of Ross (1981) implies comparative risk aversion in the sense of Arrow-Pratt, i.e., $A_{u_2}(w) > A_{u_1}(w)$, but the reverse is not true. When $h = 2$ is more risk averse than than $h = 1$ in the sense of Ross, then $A_{v_2}(w) > A_{v_1}(w)$. See Propositions 24 and 25 in Gollier (2004).

averse the individual, the larger his demand for parametric (index-based) insurance.[7]

Proposition 5 *When risk aversion is strongly comparable, the optimal index-based insurance coverage* $J(Z) = \max\{Z - z_0, 0\}$ *is increasing in risk aversion (i.e., the larger the risk aversion, the lower the deductible* z_0*) if* $\sigma > 0$*, and it is equal to the conditional expected loss* $J(Z) = Z$ *independently from risk aversion when* $\sigma = 0$*.*

4.2 Case where the parameter vector and the basis risk are not independently distributed

4.2.1 Optimal indemnity schedule

When Y and $\widetilde{\varepsilon}$ are not independent, the optimal parametric insurance contract maximizes

$$\mathbb{E}u = \mathbb{E}_Y\{\mathbb{E}_{\widetilde{\varepsilon}}[u(w_0 - Z(y) - \widetilde{\varepsilon} + I(y) - P) \mid Y = y\},$$

with respect to $I(.) : \mathcal{S} \to \mathbb{R}_+$ and P, subject to (2). Proposition 5 characterizes the optimal solution to this problem when $u''' > 0$, i.e., when the individual is downside risk averse (or prudent).

Proposition 6 *When* $u''' > 0$*, the optimal parametric-insurance indemnity schedule is written as* $I(Y) = \max\{0, \widehat{Z}(Y) - \widehat{z}_0\}$*, where the trigger is the adjusted risk* $\widehat{Z}(Y)$ *such that* $\widehat{Z}(Y) > Z(Y)$ *and the deductible* $\widehat{z}_0$ *is such that* $\widehat{z}_0 = 0$ *if* $\sigma = 0$ *and* $\widehat{z}_0 > 0$ *if* $\sigma > 0$*. For any* $y_1, y_2 \in \mathcal{S}$ *such that* $I(y_1), I(y_2) > 0$*, if the conditional probability distribution of* $\widetilde{\varepsilon}$ *corresponds to a larger risk when* $Y = y_2$ *than when* $Y = y_1$*, then* $\widehat{Z}(y_2) - Z(y_2) > \widehat{Z}(y_1) - Z(y_1)$*.*

[7] Proposition 4 seems to contradict the results of Clarke (2016) about the non-monotonicity of parametric insurance demand as a function of risk aversion. However, Clarke (2016) restricts attention to a single-value loss, with a two-value random parameter Y, a case contemplated below, in which Y and $\widetilde{\varepsilon}$ are *not* independently distributed.

Corollary 1 *Assume that the conditional distribution of* $\widetilde{\varepsilon}$ *given* $Y = y$ *only depends on* $Z(y)$*. Then, when* $u''' > 0$ *the optimal parametric insurance is index-based and it is written as* $I(Y) = \max\{\xi(Z(Y)) - \widehat{z}_0, 0\}$*, with* $\xi' > 1$ *(respect.* $\xi' > 1$*) if an increase in* $Z(Y)$ *makes the conditional distribution of* $\widetilde{\varepsilon}$ *more risky (respect. less risky).*

The first part of Proposition 6 states that the optimal parametric cover of the prudent policyholder is written as a straight deductible contract, in which the trigger is an adjusted random loss $\widehat{Z}(Y)$ larger than the expected loss $Z(Y)$. The proof of the proposition shows that the parametric indemnity $I(Y) = \max\{0, \widehat{Z}(Y) - \widehat{z}_0\}$ coincides with the coverage that would be optimal if the true loss exposure were $\widehat{Z}(Y)$ without basis risk $\widetilde{\varepsilon}$, with unchanged utility function $u(.)$.In other words, taking into account the basis risk $\widetilde{\varepsilon} = X - Z(Y)$ is equivalent to considering a traditional insurance problem without residual component $\widetilde{\varepsilon}$, where the policyholder would face an insurable loss exposure $\widehat{Z}(Y)$ larger than the conditional expected loss $Z(Y)$. Further characterizing the indemnity schedule $I(Y)$ thus requires to be more specific about the relationship between the loss adjustment $\widehat{Z}(Y) - Z(Y)$ and the parameter vector Y. The second part of the proposition shows that $\widehat{Z}(Y) - Z(Y)$ is related to the relation between Y and the size of the basis risk. Considering two parameter vectors y_1 and y_2 in S, if the conditional distribution of the basis risk is more risky when $Y = y_2$ than when $Y = y_1$, then the loss adjustment is larger in the first case than in the second. In that sense, and perhaps paradoxically, the existence of the basis risk stimulates the demand for parametric insurance.[8] Corollary 1 states a direct consequence of this second part of Proposition 6. If the conditional distribution of the basis risk is more risky when the expected loss is larger, then the increase in the insurance payout is larger than the

[8]Note however that we cannot conclude that for all Y the optimal parametric insurance contract provides a larger indemnity than what would be optimal under risk exposure $Z(Y)$ without basis risk, because the increase of risk exposure from $Z(Y)$ to $\widehat{Z}(Y)$ simultaneously affects the optimal deductible. See Eeckhoudt et al. (1991) on the effect of an increase in risk on optimal insurance with deductible.

increase in expected loss, which corresponds to a vanishing deductible, as established by Gollier (1996) in his study of optimal indemnity insurance with basis risk. In the opposite case, the indemnity schedule entails an increasing deductible.[9]

Assuming that the basis risk only depends on expected loss is very restrictive and, in general, the optimal parametric indemnity schedule $I(Y)$ cannot be written as a function of $Z(Y)$. In other words, in general the optimal parametric insurance contract is not index-based and, in that case, characterizing this indemnity schedule requires additional assumptions about the informational content of the parameter vector. To do so, we may consider the case where parameter vector Y is splitted in two components, one affecting the expected loss and the other being related with the basis risk. Let us write $Y = (Y_a, Y_b)$, with $Y_a \in \mathcal{S}_a \subset \mathbb{R}^{n_a}, Y_b \in \mathcal{S}_b \subset \mathbb{R}^{n_b}, n_a + n_b = n$ and $\mathcal{S} = \mathcal{S}_a \times \mathcal{S}_b$. We assume that component Y_a is a sufficient statistic for the expected loss $Z(Y)$, while only component Y_b may be correlated with the basis risk $\tilde{\varepsilon}$. Proposition 7 considers this case and shows how the two components of the parameter vector should be combined in order to provide the optimal coverage.

Proposition 7 *If $Z(y_a) = \mathbb{E}[X \mid Y = (y_a, y_b)]$ for all $y_a \in \mathcal{S}_a, y_b \in \mathcal{S}_b$ and Y_a and $\tilde{\varepsilon}$ are independently distributed, then the optimal parametric insurance is written as $I(Y) = \max\{0, Z(Y_a) - z_0(Y_b)\}$, where payout $I(Y)$ is equal to the excess of the expected loss $Z(Y_a)$ above a deductible $z_0(Y_b)$ that depends on component Y_b of the parameter vector. Furthermore, when $u''' > 0$, if $\widetilde{\varepsilon}_1$ and $\widetilde{\varepsilon}_2$ are random variables distributed as $\widetilde{\varepsilon}$ given $\widetilde{y}_b = y_{b1}$ and y_{b2}, respectively, and if $\widetilde{\varepsilon}_2$ is more risky than $\widetilde{\varepsilon}_1$, then $z_0(y_{b2}) < z_0(y_{b1})$.*

Proposition 6 provides conditions under which the optimal parametric insurance contract takes the form of a conditional deductible. There is full coverage of conditional expected losses above a deductible that depends on the basis risk, and the larger

[9] When the basis risk and the parameter vector are independently distributed, we are at the frontier between these two cases, with full coverage of expected loss above a deductible, as shown in Proposition 3.

the basis risk, the lower the deductible. The intuition of this result is clear if one keeps in mind the fact that the policyholder's risk exposure X is made up of two additive components: the expected loss $Z(Y)$ and the basis risk $\widetilde{\varepsilon}$. Under the assumptions of Proposition 6, these components can be expressed as functions of subvectors Y_a and Y_b extracted from Y: the conditional expected loss depends on $Z(Y_a)$, while only Y_b affects the conditional probability distribution of $\widetilde{\varepsilon}$. In that case, conditionally on $Y_b = y_b$, the basis risk $\widetilde{\varepsilon}$ and the expected loss $Z(Y_a)$ are independently distributed and, as in Proposition 3, the optimal cover entails full coverage of the expected loss above a deductible. This deductible depends on Y_b and is written as $z_0(Y_b)$. Under downside risk aversion, when $Y_b = y_b$, the larger the conditional basis risk $\widetilde{\varepsilon}_{|y_b}$, the lower the deductible $z_0(y_b)$, and thus the larger the coverage of expected losses $Z(Y_a)$. Interestingly, this is reminiscent of the precautionary motive of the prudent insured highlighted by Schlesinger (2013), whose intuition was provided by Eeckhoudt & Schlesinger (2006), and which states that uncertainty about uninsurable losses exacerbates insurance demand.

Example 1 *For illustrative purposes, let us consider the case of a risk-averse firm facing a double risk of property loss and price uncertainty. To be concrete, assume that the firm is an electrical energy supplier with normal output q in kWh per year, sold at unit price p. We consider that q and p have been determined and specified in long-term contracts with customers.*[10] *Accidents due to meteorological uncertainty may induce repair costs, as for example when electricity pylons are blown over or offshore windmills are damaged when a hurricane hits power plants. For simplicity, it is assumed that these property damages do not affect the firm's yearly output (i.e., repair does not entail significant production delay) and we denote $\widetilde{\ell}$ the repair costs, with $\widetilde{\ell} = \overline{\ell}(Y_a) + \widetilde{\eta}_a$ where $Y_a \in \mathcal{S}_a$ is a random vector of publicly observable meteorological data and $\widetilde{\eta}_a$ is a zero-*

[10] It is presumed that the sales contract with fixed values for price and quantity reflects purchasers' risk-aversion. Another version of this example would consists in assuming that electricity output is sold at spot price, the electricity supplier being able to hedge its price risk at actuarial price through forward exchange contracts.

mean random variable, Y_a and $\widetilde{\eta}_a$ being independently distributed. The actual output is $q(1+\widetilde{\eta}_b)$ where $\widetilde{\eta}_b$ is a zero-mean random variable, pair-wise independent from Y_a and $\widetilde{\eta}_a$. The difference $q\widetilde{\eta}_b$ between actual and normal outputs results from all factors that may affect electricity production for privately observed reasons, such as technological failures, delivery delays by subcontractors or wind speed outside accident risk, in the case of a wind farm. It is sold or purchased in a centralized spot market, at publicly observable price $Y_b \in \mathcal{S}_b = \mathbb{R}_+$, Y_b and $\widetilde{\eta}_b$ being also independent, with zero-mean net proceeds $q\widetilde{\eta}_b Y_b$. Production costs (apart from repair costs) are fixed and denoted C. With these notations, the firm's profit is written as

$$q(p + \widetilde{\eta}_b Y_b) - \overline{\ell}(Y_a) - \widetilde{\eta}_a - C,$$

which may be reformulated with previous notations by denoting $w_0 = qp - C$ and $X = \overline{\ell}(Y_a) + \widetilde{\eta}_a - q\widetilde{\eta}_b Y_b$. In other words, initial wealth w_0 is the difference between normal turnover and fixed cost, while losses X is the sum of repair cost and net purchases in the spot market. We have $\mathbb{E}[\widetilde{\eta}_b Y_b] = 0$ because $\widetilde{\eta}_b$ and Y_b are independently distributed with $\mathbb{E}\widetilde{\eta}_b = 0$, which gives $\mathbb{E}[X \mid Y_a, Y_b] = \overline{\ell}(Y_a) = Z(Y_a)$ and $\widetilde{\varepsilon} = X - Z(Y_a) = \widetilde{\eta}_a + q\widetilde{\eta}_b Y_b$. By applying Proposition 7, we deduce that the optimal parametric cover is a straight deductible contract, where the trigger is the expected repair cost $\overline{\ell}(Y_a)$ under meteorological data Y_a and the deductible depends on the electricity spot price Y_b. Furthermore, $\widetilde{\varepsilon}_{|Y_b=y_{b2}}$ is more risky than $\widetilde{\varepsilon}_{|Y_b=y_{b1}}$ if y_{b2} is larger than y_{b1}. Hence, if the electricity supplier is prudent, the larger the spot price, the lower the deductible.

4.2.2 Case where utility is CARA and basis risk is normal

For illustrative purpose, consider the case where the policyholder displays constant absolute risk aversion with a normally distributed basis risk. We denote $u(w) = -\exp(-\gamma w)$ and $\widetilde{\varepsilon}(y) \mapsto \mathcal{N}(0, \sigma_\varepsilon(y)^2)$ for all $y \in \mathcal{S}$, where $\widetilde{\varepsilon}(y)$ is the basis risk conditionally distributed on $Y = y$, with standard deviation $\sigma_\varepsilon(y)$. In that case, simple calculations allow us to write the adjusted risk as[11]

$$\widehat{Z}(y) = Z(y) + \frac{\gamma \sigma_\varepsilon(y)^2}{2} \text{ for all } y \in \mathcal{S}, \tag{8}$$

When Y and $\widetilde{\varepsilon}$ are independent, we have $\sigma_\varepsilon(y) = \sigma_\varepsilon$ for all y, and Proposition 5 is equivalent to Proposition 3 with

$$\widehat{z}_0 = z_0 + \frac{\gamma {\sigma_\varepsilon}^2}{2},$$

which is the case where the optimal contract is index-based, with indemnity equal to the excess of the expected loss $Z(y)$ over deductible z_0.

When Y and $\widetilde{\varepsilon}$ are not independent, (8) shows that the risk adjustment $\widehat{Z}(y) - Z(y)$ is proportional to the conditional variance $\sigma_\varepsilon(y)^2$. In the case of the normal law, a larger risk in the sense of Rothschild-Stiglitz (1970) corresponds to a larger variance, which illustrates the second part of Proposition 6.

When the distribution of $\widetilde{\varepsilon}(y)$ only depends on $Z(y)$, there exists a function $\overline{\sigma}_\varepsilon(.) : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\sigma_\varepsilon(y) = \overline{\sigma}_\varepsilon(Z(y))$ for all y. In this case, the optimal contract is index-based and, with the notation of Corollary 1, we may write

$$\xi(Z(y)) = Z(y) + \frac{\gamma \overline{\sigma}_\varepsilon(Z(y))^2}{2},$$

with $\xi' > 1$ if $\overline{\sigma}'_\varepsilon > 0$ and $\xi' < 1$ if $\overline{\sigma}'_\varepsilon < 0$: there is a vanishing deductible in the first case, and an increasing deductible in the second..

Finally, under the assumptions and notations of Proposition 6, we may write $Z(y) = Z(y_a)$ and $\sigma_\varepsilon(y) = \sigma_\varepsilon(y_b)$ for all $y = (y_a, y_b)$ and

$$I(y_a, y_b) = \max\{0, Z(y_a) - z_0(y_b)\},$$

with

$$z_0(y_b) = k - \frac{\gamma \sigma_\varepsilon(y_b)^2}{2}.$$

Hence the optimal insurance payout is the excess of the expected loss over a conditional deductible $z_0(y_b)$ that depends linearly and decreasingly on the variance of the basis risk.

[11] See the appendix for details.

4.2.3 Binary model

As Clarke (2016), we may also consider the case where the possible loss is single-valued $X \in \{0, L\}$ with $\mathcal{S} = \{0, 1\}$, and joint probabilities as follows:

X/Y	0	1
0	π_{00}	π_{01}
L	π_{10}	π_{11}

with probability of loss L equal to $\pi_{10} + \pi_{11}$, and conditional probability

$$\begin{aligned} \mathbb{P}[X &= L \mid Y = 1] = \frac{\pi_{11}}{\pi_{01} + \pi_{11}}, \\ \mathbb{P}[X &= L \mid Y = 0] = \frac{\pi_{10}}{\pi_{00} + \pi_{10}}. \end{aligned}$$

We express the fact that $Y = 1$ is an informative signal about the realization of the loss $X = L$, by assuming

$$\mathbb{P}[X = L \mid Y = 1] > \mathbb{P}[X = L \mid Y = 0],$$

which holds if

$$\frac{\pi_{11}}{\pi_{01}} > \frac{\pi_{10}}{\pi_{00}}. \tag{9}$$

The index insurance provides a payout I when $y = 1$. Clarke (2016) considers the CARA and CRRA classes of utility functions. When $\sigma > 0$, the optimal indemnity $I^*(\gamma)$ may be non-monotonic with respect to γ, which denotes the coefficient of absolute (CARA case) or relative (CRRA) risk aversion. More precisely, he shows that either $I^*(\gamma) = 0$ for all $\gamma \in (0, \infty)$, or $I^*(\gamma) = 0$ for all $\gamma < \gamma_1$, $I^*(\gamma)$ is strictly increasing for all $\gamma_1 < \gamma < \gamma_2$ and $I^*(\gamma)$ and strictly decreasing for all $\gamma_2 < \gamma < \infty$ for some $\gamma_1 < \gamma_2 < \infty$. In words, the optimal coverage is increasing and then decreasing with risk aversion, so the most risk averse individual does not necessarily purchase more insurance.

We may write

$$Z(0) = L\frac{\pi_{10}}{\pi_{10} + \pi_{00}} \qquad Z(1) = L\frac{\pi_{11}}{\pi_{11} + \pi_{01}},$$

with $Z(1) > Z(0)$ from (4), and

$$\tilde{\varepsilon}|_{Y=0} = \begin{cases} -L\frac{\pi_{10}}{\pi_{10}+\pi_{00}} & \text{with probability} \quad \frac{\pi_{00}}{\pi_{10}+\pi_{00}} \\ L\frac{\pi_{00}}{\pi_{10}+\pi_{00}} & \text{with probability} \quad \frac{\pi_{10}}{\pi_{10}+\pi_{00}} \end{cases},$$

and

$$\tilde{\varepsilon}|_{Y=1} = \begin{cases} -L\frac{\pi_{11}}{\pi_{11}+\pi_{01}} & \text{with probability} \quad \frac{\pi_{01}}{\pi_{11}+\pi_{01}} \\ L\frac{\pi_{01}}{\pi_{11}+\pi_{01}} & \text{with probability} \quad \frac{\pi_{11}}{\pi_{11}+\pi_{01}} \end{cases}.$$

By construction, we have $\mathbb{E}(\tilde{\varepsilon}|Y = y) = 0$ for $y = 0$ and 1, but the distribution of $\tilde{\varepsilon}$ differs according to whether $Y = 0$ or $Y = 1$: hence Y and $\tilde{\varepsilon}$ are not independent.[12] Because of that, Proposition 4 is not valid, and the optimal insurance payout may not be increasing with respect to risk aversion because of the interaction with prudence. Some intuition of how risk aversion and prudence interact may be obtained as follows. Conditionally on $X = 0$, the net expected transfer from the insurer to the policyholder is

$$\overline{T}_0 = \frac{\pi_{01}}{\pi_{00} + \pi_{01}} I - P,$$

and the actual net transfer is

$$\widetilde{T}_0 = \begin{cases} I - P \text{ with prob. } \frac{\pi_{01}}{\pi_{00}+\pi_{01}} \\ -P \quad \text{with prob. } \frac{\pi_{00}}{\pi_{00}+\pi_{01}} \end{cases},$$

with $P = (1 + \sigma)(\pi_{01} + \pi_{11})$. Similarly, in state $X = L$, the net expected and actual transfers to the policyholder are

$$\overline{T}_1 = \frac{\pi_{11}}{\pi_{10} + \pi_{11}} I - P,$$

and

$$\widetilde{T}_1 = \begin{cases} I - P \text{ with prob. } \frac{\pi_{11}}{\pi_{10}+\pi_{11}} \\ -P \quad \text{with prob. } \frac{\pi_{10}}{\pi_{10}+\pi_{11}} \end{cases},$$

[12] This is an intrinsic property of the binary model, since this dependence follows from Assumption (5) which reflects the fact that signal $Y = 1$ is informative about the occurence of the loss $X = L$.

respectively. We have $\overline{T}_1 > \overline{T}_0$ because of inequality (5), meaning that, on average, the payment received by the policyholder is larger when $X = L$ that when $X = 0$. Because of this first effect, as in a standard insurance demand problem, the larger the degree of risk-aversion, the larger the optimal average indemnity, obtained through an increase in I. However, the actual transfer $\widetilde{T}_0$ or $\widetilde{T}_1$(conditionally on $X = 0$ or L, respectively) is uncertain, and we have $\mathrm{Var}(\widetilde{T}_1) > \mathrm{Var}(\widetilde{T}_0)$. Because this uncertainty on the conditional payment (as measured by its variance) is larger in the loss state than in the no-loss state, downside risk-aversion creates a countervailing effect that reduces insurance demand. Under CARA and CRRA preferences, when parameter γ increases, the coefficient of absolute prudence also increases. The countervailing effect reflecting prudence becomes stronger and it may dominate the risk aversion effect, hence a possible decrease in insurance demand.

This interaction between risk aversion and prudence may be further illustrated through simple examples. The optimal parametric insurance contract maximizes

$$\mathbb{E}u = \pi_{00}u(w_0 - P) + \pi_{01}u(w_0 - P + I) + \pi_{10}u(w_0 - P - L) + \pi_{11}u(w_0 - P - L + I),$$

with respect to I, P, subject to $P = (1+\sigma)(\pi_{01} + \pi_{11})I$. Consider first the case where u is quadratic

$$u(w) = w - bw^2,$$

where $b < 1/2w$ parametrizes risk aversion. In that case, $u''' = 0$, so the policyholder does not display downside risk aversion. Straightforward calculations show that the optimal level of coverage is

$$I^* = \frac{2bL[\pi_{11} - (1+\sigma)\pi p] - \sigma\pi(1 - 2bw_0)}{2bw_0[1 - (1-\sigma^2)\pi_0]},$$

where $p = \pi_{10} + \pi_{11}$ is the probability of loss and $\pi = \pi_{01} + \pi_{11}$ is the probability that the index triggers the indemnity payment. In the quadratic case, the optimal coverage I^* is therefore always increasing in b. The non-monotonicity of insurance coverage with respect to risk aversion does not arise in this case because the decision maker does not display dowside risk aversion.

Figure 1 represents the optimal parametric insurance I^* on the vertical axis, as a function of risk-aversion parameter γ on the horizontal axis, for several values of the loading factor σ, when

$$u(w) = -\frac{1}{\alpha}\exp(-\alpha w) - \gamma w,$$

where $\alpha = -u'''(w)/u''(w)$ is the coefficient of absolute prudence.[13] The index of absolute risk aversion

$$A(w) = \frac{\alpha \exp(-\alpha w)}{\exp(-\alpha w) - \gamma}$$

is increasing in γ. Increasing γ therefore increases risk aversion at all wealth levels without modifying the prudence of the agent. When $\sigma > 0$, the standard risk aversion comparative statics applies : at a given level of absolute prudence α, an increase in the risk aversion parameter γ results in an increased coverage I^*. An increase in σ shifts the curve downward, and insurance is not affected by risk aversion when $\sigma = 0$.

Remark 1 *It is tempting to draw a parallel between the parametric insurance problem in the binary model and Jewitt's (1988) analysis of portfolio choices. He considers a standard portfolio problem with one risky asset and an additive background risk affecting wealth, in which the returns of the risky asset and of the background wealth are affiliated random variables. He shows that the usual relationship between absolute risk aversion and portfolio choices still holds, under the additional assumption of decreasing absolute risk aversion. More precisely, if two investors display DARA preferences, the one with the larger absolute risk aversion has the smaller demand for the risky asset.*[14] *When the insurance payout is assumed to be a fraction of the expected loss (i.e.,* $I(Y) = \alpha Z(Y)$*, with* $\alpha > 0$*) - which is not a loss of generality in the binary model*

[13] Figure 1 corresponds to $\pi_{00} = 0.55, \pi_{01} = 0.1, \pi_{10} = 0.1 = 0.25, w_0 = 2,000, L = 1,000, \alpha = 0.0005$, $\sigma \in \{0, 0.1, 0.2, 0.3\}$ with $\gamma \in [0.001, 0.025]$, with $u' > 0$ for all relevant value of final wealth.

[14] See Proposition 26 in Gollier (2004).

-, then choosing coefficient α, is formally equivalent to the standard portfolio problem. In that case, if the loss $Z(Y)$ and the background risk $\tilde{\varepsilon}$ are affiliated, and if two individuals have DARA preferences, then the more risk-averse has a higher demand for insurance, meaning that his optimal coinsurance coefficient α is larger. However, one can check that $\tilde{\varepsilon}$ and $Z(Y)$ are not affiliated in the binary model, which explains why the static comparative analysis of Jewitt (1988) does not apply in that case.

5 Conclusion

Reframing the parametric insurance problem in an asymmetric information setting brings about new insights on the design of optimal coverage. When the actual loss incurred by the policyholder is his private information, the insurance indemnity depends only on the publicly-observed parameter vector. The information structure is finer when this vector provides a more accurate information on the loss incurred, and in that case the policyholder reaches a higher expected utility and the basis risk is lower. The most important conclusion that emerge in this context is the fact that optimal parametric insurance depends on the stochastic relationship between the loss index and the basis risk, i.e. between the parameter-based best estimate of the loss on one side, and the residual unobserved risk on the other. If these two components of the risk exposure are independently distributed, with some caveats such as the strong comparability criterion used in Proposition 4, important results of insurance demand theory extend to the parametric insurance setting. As we have seen, this follows from the similarity with the insurance demand problem under independent background risk. Under constant loading, a straight deductible contract triggered by the loss index is optimal. Furthermore, the amount of insurance demand of two individuals who face the same risk exposure depend on their respective degrees of risk aversion.

Conclusions are far less simple when the parameter vector and the basis risk are not independently distributed, a case that may be more relevant in many concrete situations. The reason is simple: if two parameter vectors leading to the same loss index correspond to different distributions of the basis risk, then they provide different information on the loss, and this should be reflected in the optimal insurance coverage. In that case, the optimal parametric insurance is not index-based. In other words, the insurance payout should depend on the parameter vector itself, and not only on the best estimate of the loss that can be infered from this information. Once said that, this raises questions about, at least, two issues: the structure of the optimal indemnity schedule and the relationship between the attitude toward risk and the demand for insurance. With respect to the first question, we have shown that the optimal indemnity schedule corresponds to a straight deductible contract applied to an adjusted expected loss exposure. Under downward risk aversion, the larger the basis risk conditionally on the parameter vector, the larger the risk adjustment. In other words, if the conditional basis risk increases when we move from a parameter vector to another one, then the risk adjustment should be larger in the second case than in the first one. This adjustment takes a more simple form when the parameter vector can be splitted in two independently distributed subvectors, affecting the expected loss and the conditional basis risk, respectively. In that case, the optimal indemnity schedule takes the form of a conditional deductible, and the larger the conditional basis risk, the lower the conditional deductible. Concerning the relationship between the attitude toward risk and the demand for parametric insurance, risk aversion and prudence codetermine the demand for parametric insurance, and, contrary to the case where the basis risk and the loss index are independtly distributed, a lower degree of risk aversion does not necessarily means a lower demand for insurance with a lower premium, if this risk aversion effect is more than compensated by a larger degree of prudence. In other words, risk aversion and downside risk aversion may go in opposite direction, which invalidates the usual comparative static analysis of insurance choices.

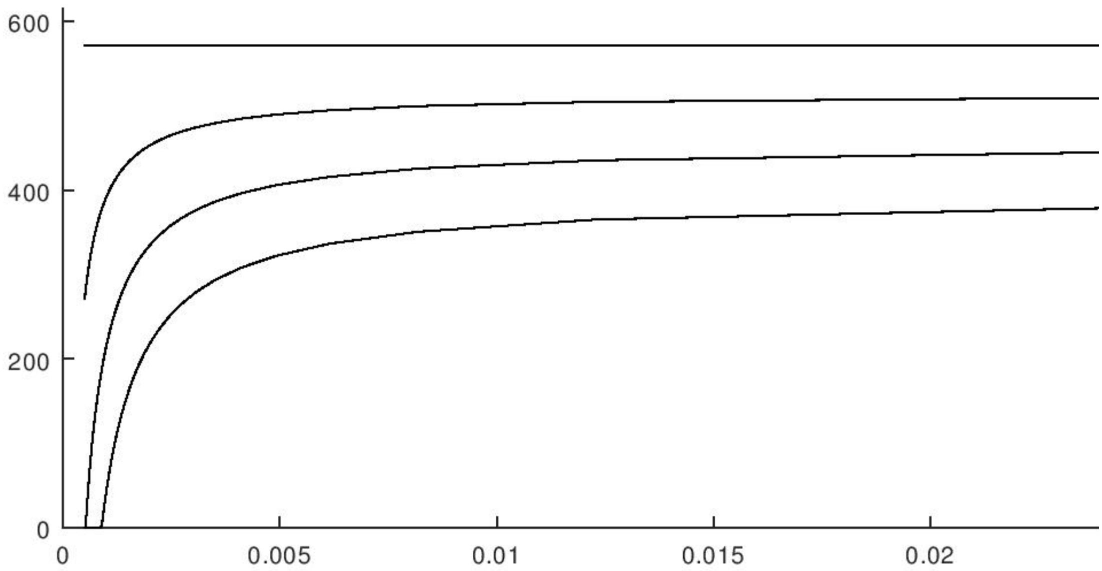

Figure 1: Optimal insurance demand I^* as a function of absolute risk-aversion parameter γ for given absolute prudence α and various values of the loading factor σ.

6 Proofs

6.1 Proof of Lemma 1

Assume $Y_2(\omega) = \Phi(Y_1(\omega))$ for all ω. Let $y_2 \in \mathcal{S}_2$. We have

$$\begin{aligned}
\mathcal{O}_2(y_2) &= \{\omega \in \Omega \text{ s.t. } Y_2(\omega) = y_2\} \\
&= \{\omega \in \Omega \text{ s.t. } \Phi(Y_1(\omega)) = y_2\} \\
&= \{\omega \in \Omega \text{ s.t. } Y_1(\omega) \in \Phi^{-1}(y_2)\} \\
&= \bigcup_{y_1 \in \mathcal{K}(y_2)} \mathcal{O}_1(y_1),
\end{aligned}$$

with $\mathcal{K}(y_2) = \{y_1 \in \mathcal{S}_1 \text{ s.t. } y_2 = \Phi(y_1)\} = \Phi^{-1}(y_2)$. We also have $\mathcal{K}(y_2) \cap \mathcal{K}(y_2') = \phi$ if $y_2 \neq y_2'$ and

$$\bigcup_{y_2 \in \mathcal{S}_2} \mathcal{K}(y_2) = \bigcup_{y_2 \in \mathcal{S}_2} \{y_1 \in Y_1 \text{ s.t. } y_2 = \Phi(y_1)\} = Y_1.$$

Hence $\{\mathcal{K}(y_2), y_2 \in \mathcal{S}_2\}$ is a partition of $\mathcal{S}_1$, and thus $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$.

Conversely, assume that $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$, i.e.

$$\mathcal{O}_2(y_2) = \bigcup_{y_1 \in \mathcal{Z}(y_2)} \mathcal{O}_1(y_1),$$

with $\{\mathcal{K}(y_2), y_2 \in Y_2\}$ a partition of Y_1. Let $\Phi(.) : \mathcal{S}_1 \to \mathcal{S}_2$ defined by $\Phi(y_1) = y_2$ if $y_1 \in \mathcal{Z}(y_2)$ with $y_2 \in Y_2$. For all ω there exists $y_1 \in Y_1$ such that $Y_1(\omega) = y_1$, i.e. $\omega \in \mathcal{O}_1(y_1)$. We have $\mathcal{O}_1(y_1) \subset \mathcal{O}_2(y_2)$ with $y_1 \in \mathcal{Z}(y_2)$, and thus $\mathcal{O}_1(y_1) \subset \mathcal{O}_2(\Phi(y_1))$ from the definition of $\Phi(.)$, which implies $Y_2(\omega) = \Phi(y_1) = \Phi(Y_1(\omega))$.

6.2 Proof of Proposition 1

Under information structure $(\mathcal{S}_i, Y_i)$, for $i = 1$ or 2, the optimal contract $\{P_i^*, I_i^*(.)\}$ maximizes

$$\mathbb{E}u(w_0 - X + I_i(Y_i) - P_i),$$

with respect to P_i and $I_i(.) : \mathcal{S}_i \to \mathbb{R}_+$ subject to

$$P_i = (1+\sigma)\mathbb{E}I_i(Y_i),$$

with optimal expected utility

$$\overline{u}_i^* = \mathbb{E}u(w_0 - X + I_i^*(Y_i) - P_i^*).$$

Consider the indemnity schedule $I_1(.) : \mathcal{S}_1 \to \mathbb{R}_+$ defined by $I_1(y_1) = I_2^*(\Phi(y_1))$ for all $y_1 \in \mathcal{S}_1$. In any state of nature ω, the insurance payout is the same for $I_1(.)$ and $I_2^*(.)$. Hence, the contract $\{P_2^*, I_1(.)\}$ is feasible under information structure $(\mathcal{S}_1, Y_1)$ with expected utility $\overline{u}_1 = \overline{u}_2^*$, and thus we have $\overline{u}_1^* \geq \overline{u}_2^*$.

Consider the indemnity schedule $I_1(.) = I_2^*(\Phi(.)) : \mathcal{S}_1 \to \mathbb{R}_+$ which is feasible under information structure $(\mathcal{S}_1, Y_1)$. Replacing $(\mathcal{S}_2, Y_2)$ by $(\mathcal{S}_1, Y_1)$ allows us to increase the insurance indemnity $I_1(y_1)$ above $I_2^*(\Phi(y_1))$ when $y_1 \in \mathcal{A}_1^1(y_2)$ and simultaneously to decrease $I_1(y_1)$ under $I_2^*(\Phi(y_1))$ when $y_1 \in \mathcal{A}_1^2(y_2)$. This can be done for all $y_2 \in \mathcal{A}_2$ in such a way that the expected insurance payout is unchanged when $Y^1(\omega) \in \mathcal{A}_1^1(y_2) \cup \mathcal{A}_1^1(y_2)$. Furthermore $I_1(y_1)$ is kept equal to $I_2^*(\Phi(y_1))$ if $y_1 \notin \{\mathcal{A}_1^1(y_2) \cup \mathcal{A}_1^1(y_2), y_2 \in \mathcal{A}_2\}$.This change increases the insurance payout in states with higher losses and reduces this payout in states with lower losses, starting from an initial solution $I_1(.)$ where these payouts are equal, and the insurance premium is unchanged. Because of the concavity of the utility function, this induces an increase in expected utility, hence the dominance of $(\mathcal{S}_1, Y_1)$.

6.3 Proof of Proposition 2

Assume that $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$. We have

$$\begin{aligned} \widetilde{\varepsilon}_2(\omega) &= X(\omega) - Z_2(Y_2(\omega)) \\ &= \widetilde{\varepsilon}_1(\omega) + \widetilde{\eta}_\varepsilon(\omega), \end{aligned} \tag{10}$$

for all $\omega \in \Omega$, where

$$\widetilde{\eta}_\varepsilon(\omega) = Z_1(Y_1(\omega)) - Z_2(Y_2(\omega)).$$

Using Lemma 1 yields

$$Z_2(Y_2(\omega)) = \mathbb{E}[Z_1(y_1) \mid y_1 \in \mathcal{K}(Y_2(\omega))] \text{ for all } \omega \in \Omega,$$

which gives

$$\begin{aligned}\mathbb{E}[\widetilde{\eta}(\omega) \mid \widetilde{\varepsilon}_1(\omega) = \varepsilon_1] = \mathbb{E}[Z_1(Y_1(\omega)) \mid \widetilde{\varepsilon}_1(\omega) = \varepsilon_1] \\ -\mathbb{E}[Z_1(y_1) \mid y_1 \in \mathcal{Z}(Y_2(\omega))], \widetilde{\varepsilon}_1(\omega) = \varepsilon_1].\end{aligned} \tag{11}$$

Let $F_1(y_1 \mid y_2, \varepsilon_1)$ be the distribution function of $y_1 \in \mathcal{S}_1$ conditionally on $Y_2 = y_2 \in \mathcal{S}_2$ and $\widetilde{\varepsilon}_1 = \varepsilon_1$, and let

$$\overline{Z}_1(y_2, \varepsilon_1) = \int_{y_1 \in \mathcal{K}(y_2)} Z_1(y_1) dF_1(y_1 \mid y_2, \varepsilon_1) = 1.$$

Let $F_2(y_2 \mid \varepsilon_1)$ be the distribution function of $y_2 \in Y_2$ conditionally on $\widetilde{\varepsilon}_1 = \varepsilon_1$. Using (7) gives

$$\begin{aligned}\mathbb{E}[\widetilde{\eta}_\varepsilon \mid \widetilde{\varepsilon}_1 = \varepsilon_1] &= \int_{y_2 \in Y_2} \left\{ \int_{y_1 \in \mathcal{K}(y_2)} [Z_1(y_1) - \overline{Z}_1(y_2, \varepsilon_1)] dF_1(y_1 \mid y_2, \varepsilon_1) \right\} dF_2(y_2 \mid \varepsilon_1) \\ &= 0,\end{aligned}$$

for all ε_1, which shows that $\widetilde{\varepsilon}_2$ is more risky than $\widetilde{\varepsilon}_1$.

Furthermore, we have

$$\begin{aligned} Z_2(y_2) &= \mathbb{E}[X \mid Y_2 = y_2] \\ &= \mathbb{E}[X \mid Y_1 \in \mathcal{K}(y_2)] \\ &= \mathbb{E}[\mathbb{E}[X \mid Y_1] \mid Y_1 \in \mathcal{K}(y_2)] \\ &= \mathbb{E}[Z_1(Y_1) \mid Y_1 \in \mathcal{K}(y_2)] \end{aligned}$$

for all $y_2 \in \mathrm{im}(Y_2) \subset \mathcal{S}_2$. Consequently,

$$z_2 = \int_{y_1 \in \mathcal{K}(Z_2^{-1}(z_2))} Z_1(y_1) dF_1(y_1 \mid Y_1 \in \mathcal{K}(Z_2^{-1}(z_2))), \tag{12}$$

for all $z_2 \in \mathrm{im}(Z_2)$, where $F_1(y_1 \mid Y_1 \in \mathcal{K}(Z_2^{-1}(z_2)))$ is the distribution function of Y_1 conditionally on $Y_1 \in \mathcal{K}(Z_2^{-1}(z_2))$. We have

$$Y_1 \in \mathcal{K}(Z_2^{-1}(z_2))) \Leftrightarrow Z_2(\Phi(Y_1) = z_2.$$

Using $Y_2 = \Phi(Y_1)$ and (9) yields

$$\mathbb{E}[Z_1 \mid Z_2 = z_2] = z_2 \ \text{ for all } z_2,$$

Equivalently,

$$Z_1 = Z_2 + \widetilde{\eta}_Z,$$

with

$$\mathbb{E}[\widetilde{\eta}_Z \mid Z_2 = z_2] = 0 \ \text{ for all } z_2 \in \mathrm{im}(Z_2)$$

which shows that Z_1 is more risky than Z_2.

6.4 Proof of Proposition 3

Consider random variables $Z_1, Z_2, \widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2$ and X such that

$$Z_1 + \widetilde{\varepsilon}_1 \equiv Z_2 + \widetilde{\varepsilon}_2 \equiv X, \tag{13}$$

with $\widetilde{\varepsilon}_1, Z_1$ and $Z_2, \widetilde{\varepsilon}_2$ pairwise independent. Define random variable $\widetilde{\eta}$ by

$$\widetilde{\varepsilon}_2 \equiv \widetilde{\varepsilon}_1 + \widetilde{\eta},$$

and thus with

$$Z_1 \equiv Z_2 + \widetilde{\eta}.$$

Assume that random variables $Z_2, \widetilde{\eta}$ and $\widetilde{\varepsilon}_1$ are pairwise independent. Consider probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω includes all three-dimension states $\omega_z = (\omega_z^1, \omega_z^2, \omega_z^3)$ such that $z \in \mathbb{R}_+$, $\omega_z^1 \in$ supp $(Z_2), \omega_z^2 \in \mathrm{supp}(\widetilde{\eta}), \omega_z^3 \in \mathrm{supp}(\widetilde{\varepsilon}_1)$. Define random variable $\widetilde{\omega}(.) : \Omega \rightarrow \mathbb{R}_3$ by $\widetilde{\omega}(\omega_z) = \omega_z$ for all $z \in \mathbb{R}_+$ and choose $(\mathcal{F}, \mathbb{P})$ such that $\widetilde{\omega} \equiv (Z_2, \widetilde{\eta}, \widetilde{\varepsilon}_1)$. Define loss function and information structures by

$$
\begin{aligned}
& X(\omega_z) = \omega_z^1 + \omega_z^2 + \omega_z^3, \\
& \mathcal{S}_1 = \mathbb{R}_3 \text{ and } Y_1(\omega_z) = (\omega_z^1, \omega_z^2), \\
& \mathcal{S}_2 = \mathbb{R} \text{ and } Y_2(\omega_z) = \omega_z^1.
\end{aligned}
$$

for all $z \in \mathbb{R}_+$. We have $Y_2(\omega_z) = \Phi(Y_1(\omega_z))$ with $\Phi(\omega_z^1, \omega_z^2) = \omega_z^1$ and thus $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$.

Using $\mathbb{E}\widetilde{\varepsilon}_1 = 0$ and the fact that $Z_2, \widetilde{\eta}$ and $\widetilde{\varepsilon}_1$ are pairwise independent yields

$$
\begin{aligned}
\mathbb{E}[X(\omega_z) \quad | \quad & Y_1 = (\omega_z^1, \omega_z^2)] = \omega_z^1 + \omega_z^2 + \mathbb{E}[\widetilde{\omega}_z^3 \mid \omega_z^1, \omega_z^2] \\
= \quad & \omega_z^1 + \omega_z^2 + \mathbb{E}[\widetilde{\varepsilon}_1 \mid Z_2 = \omega_z^1, \widetilde{\eta} = \omega_z^2] \\
= \quad & \omega_z^1 + \omega_z^2 + \mathbb{E}\widetilde{\varepsilon}_1 \\
= \quad & \omega_z^1 + \omega_z^2
\end{aligned}
$$

We also have

$$
\begin{aligned}
\mathbb{E}[X(\omega_z) \quad | \quad & Y_2 = \omega_z^1] = \omega_z^1 + \mathbb{E}[\widetilde{\omega}_z^2 + \widetilde{\omega}_z^3 \mid \omega_z^1] \\
= \quad & \omega_z^1 + \mathbb{E}[\widetilde{\eta} + \widetilde{\varepsilon}_1 \mid Z_2 = \omega_z^1] \\
= \quad & \omega_z^1.
\end{aligned}
$$

This gives $Z_1(\omega_z) = \omega_z^1 + \omega_z^2$ and $Z_2(\omega_z) = \omega_z^1$. Hence, if probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is defined as above, then loss index Z_i and basis risk $\widetilde{\varepsilon}_i$ are induced by information structure $(\mathcal{S}_i, Y_i)$ for $i = 1$ and 2. Since $(\mathcal{S}_1, Y_1)$ is finer than $(\mathcal{S}_2, Y_2)$ we deduce from Proposition 1 that, for any risk-averse individual, the optimal expected utility is (weakly or strongly) larger under $(\mathcal{S}_1, Y_1)$ than under $(\mathcal{S}_2, Y_2)$. When the loss index and the basis risk are independently distributed (which is the case here), the optimal expected utility only depends on their probability distribution, independently from the underlying probability distribution of the parameter vector.[15] Hence, the fact that the

[15] This is intuitive but formally established in the proof of Proposition 4 below.

optimal expected utility is higher for $Z_1, \widetilde{\varepsilon}_1$ than for $Z_2, \widetilde{\varepsilon}_2$ holds for any probability space and information structure that sustains random variables $Z_1, Z_2, \widetilde{\varepsilon}_1, \widetilde{\varepsilon}_2$ and X.

6.5 Proof of Proposition 4

Let $I^*(.) : \mathcal{S} \rightarrow \mathbb{R}_+, P^*$ the optimal indemnity schedule and premium, with optimal expected utility

$$\overline{u}^* = \mathbb{E}_Y[v(w_0 - Z(Y) + I^*(Y) - P^*)].$$

Define $\overline{J}(.) : Z(Y(\Omega)) \rightarrow \mathbb{R}_+$ by

$$\overline{J}(z) = \mathbb{E}_Y[I^*(Y) \mid Z(Y) = z].$$

Using $u'' < 0$ allows us to write

$$\begin{aligned}
\overline{u}^* &= \mathbb{E}_Z[\mathbb{E}_Y[v(w_0 - Z + I^*(Y) - P^*) \mid Z = z]] \\
&\leq \mathbb{E}_Z[v(w_0 - Z + \mathbb{E}_Y[I^*(Y) \mid Z = z] - P^*] \mid Z = z] \\
&= \mathbb{E}[v(w_0 - Z + \overline{J}(Z) - P^*] \mid Z = z] \\
&= \overline{u},
\end{aligned}$$

with strict inequality if $I^*(.)$ is not index-based in a positive probability event. Furthermore, we have

$$\mathbb{E}\overline{J}(Z) = \mathbb{E}_Z[\mathbb{E}_Y[I^*(Y) \mid Z(Y) = z]] = \mathbb{E}I^*(Y),$$

and

$$P^* = (1 + \sigma)\mathbb{E}I^*(Y) = (1 + \sigma)\mathbb{E}\overline{J}(Z).$$

Thus, the index-based contract $\overline{J}(.), P^*$ is feasible, with higher expected utility than $I^*(.), P^*$, hence a contradiction.

The rest of the proof results from the optimality of a straight deductible contract with loss Z, under constant loading and utility function $v(.)$.

6.6 Proof of Proposition 5

The Proposition directly follows from the analysis of comparative risk aversion when there is an independent background risk: see Proposition 24 and 25 in Gollier (2004).

6.7 Proof of Proposition 6

Let λ be a Lagrange multiplier associated with constraint (2). The first-order optimality conditions are written as

$$\mathbb{E}[u'(w_0 - Z(y) - \widetilde{\varepsilon} + I(y) - P) \mid Y = y] - \lambda(1+\sigma) \begin{cases} \leq 0 \text{ for all } y \in \mathcal{S} \\ = 0 \text{ if } I(y) > 0 \end{cases}, \quad (14)$$

$$\mathbb{E}[u'(w_0 - Z(Y) - \widetilde{\varepsilon} + I(Y) - P)] = \lambda. \quad (15)$$

Using $\mathbb{E}[\widetilde{\varepsilon} \mid Y = y] = 0$ for all $y \in \mathcal{S}$ and $u''' > 0$ yields

$$\mathbb{E}[u'(w_0 - Z(y) - \widetilde{\varepsilon} + I(y) - P) \mid Y = y] > u'(w_0 - Z(y) + I(y) - P).$$

Let $\widehat{Z}(y)$ be defined for all $y \in Y$ by

$$\mathbb{E}[u'(w_0 - Z(y) - \widetilde{\varepsilon} + I(y) - P) \mid Y = y] = u'(w_0 - \widehat{Z}(y) + I(y) - P)$$

with $\widehat{Z}(y) > Z(y)$ from $u'' < 0$. Hence, the optimality conditions may be rewritten as

$$u'(w_0 - \widehat{Z}(y) + I(y) - P) - \lambda(1+\sigma) \begin{cases} \leq 0 \text{ for all } y \in \mathcal{S} \\ = 0 \text{ if } I(y) > 0 \end{cases},$$

$$\mathbb{E}[u'(w_0 - \widehat{Z}(Y) + I(Y) - P)] = \lambda.$$

These are the first-order optimality conditions of the optimization problem for an individual with risk exposure $\widehat{Z}(Y)$, in which the expected utility

$$\mathbb{E}[u(w_0 - \widehat{Z}(Y) + I(Y) - P)$$

is maximized with respect to $I(.) : \mathcal{S} \to \mathbb{R}_+$ and P, subject to constraint (2). We know that such conditions implies that there exists $\widehat{z}_0 \geq 0$ such that

$$I(y) = \max\{\widehat{Z}(y) - \widehat{z}_0, 0] \text{ for all } y \in \mathcal{S}$$

with $\widehat{z}_0 = 0$ if $\sigma = 0$ and $\widehat{z}_0 > 0$ if $\sigma > 0$.

Let $y_1, y_2 \in \mathcal{S}$ with $I(y_1), I(y_2) > 0$. Denote $\Delta Z_1 = \widehat{Z}(y_1) - Z(y_1) > 0$ and $\Delta Z_2 = \widehat{Z}(y_2) - Z(y_2) > 0$. Assume that the conditional distribution of $\widetilde{\varepsilon}$ is more risky in when $Y = y_2$ than when $Y = y_1$, and suppose $\Delta Z_2 \leq \Delta Z_1$. The optimality conditions give

$$u'(w_0 - \widehat{Z}(y_1) + I(y_1) - P) = u'(w_0 - \widehat{Z}(y_2) + I(y_2) - P) = \lambda(1+\sigma),$$

or, equivalently

$$\mathbb{E}[u'(w_0 - Z(y_1) - \widetilde{\varepsilon} + I(y_1) - P) \mid Y = y_1] = \mathbb{E}[u'(w_0 - Z(y_2) - \widetilde{\varepsilon} + I(y_2) - P) \mid Y = y_2].$$

In this first case, we have $I(y_1) = \widehat{Z}(y_1) - \widehat{z}_0 = Z(y_1) - \widehat{z}_0 + \Delta Z_1$ and $I(y_2) = \widehat{Z}(y_2) - \widehat{z}_0 = Z(y_2) - \widehat{z}_0 + \Delta Z_2$. Hence, the last equation may be rewritten as

$$\mathbb{E}[u'(w_0 + \Delta Z_1 - \widehat{z}_0 - \widetilde{\varepsilon} - P) \mid Y = y_1] = \mathbb{E}[u'(w_0 + \Delta Z_2 - \widehat{z}_0 - \widetilde{\varepsilon}_2 - P) \mid Y = y_2],$$

or

$$\mathbb{E}[u'(w + \Delta Z_1 - \Delta Z_2 - \widetilde{\varepsilon}) \mid Y = y_1] = \mathbb{E}[u'(w - \widetilde{\varepsilon}) \mid Y = y_2],$$

where $w = w_0 + \Delta Z_2 - \widehat{z}_0 - P$. However, we have

$$\begin{aligned} \mathbb{E}[u'(w + \Delta Z_1 - \Delta Z_2 - \widetilde{\varepsilon}) \quad &| \quad Y = y_1] \leq \mathbb{E}[u'(w - \widetilde{\varepsilon}) \mid Y = y_1] \\ &< \quad \mathbb{E}[u'(w - \widetilde{\varepsilon}) \mid Y = y_2], \end{aligned}$$

where the first inequality comes from $\Delta Z_1 \geq \Delta Z_2$ and $u'' < 0$, and the second from the fact that $\widetilde{\varepsilon}_{|Y=y_2}$ is more risky than $\widetilde{\varepsilon}_{|Y=y_1}$ and $u''' > 0$. This is a contradiction.

6.8 Proof of Proposition 7

The first-order optimality conditions are

$$\begin{aligned} &\mathbb{E}[u'(w_0 - Z(y_a) - \widetilde{\varepsilon} + I(y) - P) \mid Y_b = y_b] - \lambda(1+\sigma) \begin{cases} \leq 0 \text{ for all } y \in \mathcal{S}, \\ = 0 \text{ if } I(y) > 0, \end{cases}, \\ &\mathbb{E}[u'(w_0 - Z(Y_a) - \widetilde{\varepsilon} + I(Y) - P)] = \lambda. \end{aligned}$$

where $y = (y_a, y_b)$. Let us define

$$\mathcal{U}(w, y_b) \equiv \mathbb{E}_{\widetilde{\varepsilon}}[u(w - \widetilde{\varepsilon}) \mid Y_b = y_b],$$

with $\mathcal{U}'_w > 0, \mathcal{U}''_{w^2} < 0$. This allows us to rewrite the optimality conditions as

$$\begin{aligned} &\mathcal{U}'_w(w_0 - Z(y_a) + I(y) - P, y_b) - \lambda(1+\sigma) \begin{cases} \leq 0 \text{ for all } y \in Y \\ = 0 \text{ if } I(y) > 0 \end{cases} \\ &\mathbb{E}[\mathcal{U}'_w(w_0 - Z(Y_a) + I(Y) - P), Y_b] = \lambda. \end{aligned}$$

When $I(y) > 0$, we have

$$I(y) = Z(y_a) - z_0(y_b),$$

where $z_0(y_b)$ is defined by

$$\mathcal{U}'_w(w_0 - z_0(y_b) - P, y_b) = \lambda(1+\sigma).$$

Furthermore, when $I(y) = 0$ we have $Z(y_a) < z_0(y_b)$. Patching up these two cases yields

$$I(y) = \max\{0, Z(y_a) - z_0(y_b)\} \text{ for all } y = (y_a, y_b).$$

Let $\widetilde{\varepsilon}_1$ *and* $\widetilde{\varepsilon}_2$ *be random variables distributed as* $\widetilde{\varepsilon}$ *given* $\widetilde{y}_b = y_{b1}$ *and* y_{b2}*, respectively, and assume that* $\widetilde{\varepsilon}_2$ *is more risky than* $\widetilde{\varepsilon}_1$*.* We have

$$\begin{aligned} \mathcal{U}'_w(w_0 - z_0(y_{b1}) - P, y_{b1}) &= \lambda(1+\sigma), \\ \mathcal{U}'_w(w_0 - z_0(y_{b2}) - P, y_{b2}) &= \lambda(1+\sigma). \end{aligned}$$

The last equality may be rewritten as

$$\mathbb{E}u'(w_0 - z_0(y_{b2}) - \widetilde{\varepsilon}_2 - P) = \lambda(1+\sigma).$$

Since $\widetilde{\varepsilon}_2$ *is more risky than* $\widetilde{\varepsilon}_1$, we may write

$$\widetilde{\varepsilon}_2 \equiv \widetilde{\varepsilon}_1 + \widetilde{\eta},$$

where random variable $\widetilde{\eta}$ is such that $\mathbb{E}[\widetilde{\eta} \mid \widetilde{\varepsilon}_1 = \varepsilon_1] = 0$ for all ε_1. When $u''' > 0$, we have

$$\begin{aligned}
\mathbb{E}u'(w_0 - z_0(y_{b2}) - \widetilde{\varepsilon}_2 - P) &= \mathbb{E}_{\widetilde{\varepsilon}_1}[\mathbb{E}_{\widetilde{\eta}}[u'(w_0 - z_0(y_{b2}) - \varepsilon_1 - \widetilde{\eta} - P) \mid \widetilde{\varepsilon}_1 = \varepsilon_1]] \\
&> \mathbb{E}_{\widetilde{\varepsilon}_1} u'(w_0 - z_0(y_{b2}) - \widetilde{\varepsilon}_1 - P) \\
&= \mathcal{U}'_w(w_0 - z_0(y_{b2}) - P, y_{b1}).
\end{aligned}$$

We deduce

$$\mathcal{U}'_w(w_0 - z_0(y_{b2}) - P, y_{b1}) < \mathcal{U}'_w(w_0 - z_0(y_{b1}) - P, y_{b1}),$$

and using $\mathcal{U}''_{w^2} < 0$ gives $z_0(y_{b2}) < z_0(y_{b1})$.

6.9 Proof of results when utility is CARA and basis risk is normally distributed.

When $u(w) = -\exp(-\gamma w)$, optimality condition (14) yields

$$\mathbb{E}[\exp(\gamma(Z(y) - I(y) + \widetilde{\varepsilon}(y)] = \frac{\lambda(1+\sigma)\exp(\gamma(w_0 - P))}{\gamma}$$

if $I(y) > 0$, and thus

$$I(y) = Z(y) + \frac{1}{\gamma}\ln\{\mathbb{E}[\exp(\gamma\widetilde{\varepsilon}(y)]\} - k,$$

if $I(y) > 0$, where

$$k = w_0 - P + \frac{1}{\gamma}\ln\left[\frac{\lambda(1+\sigma)}{\gamma}\right].$$

When $\widetilde{\varepsilon}(y) \mapsto \mathcal{N}(0, \sigma_\varepsilon(y)^2)$, we have

$$\mathbb{E}[\exp(\gamma\widetilde{\varepsilon}(y)] = \exp\left(\frac{\gamma^2\sigma_\varepsilon(y)^2}{2}\right),$$

which gives

$$I(y) = Z(y) + \frac{\gamma\sigma_\varepsilon(y)^2}{2} - k,$$

for all y such that $I(y) > 0$. This gives $I(y) = \max\{\widehat{Z}(y) - \widehat{z}_0, 0\}$ with $\widehat{z}_0 = k$, and

$$\widehat{Z}(y) = Z(y) + \frac{\gamma\sigma_\varepsilon(y)^2}{2}.$$

References

Barnett, B. and O. Mahul, 2007, Weather index insurance for agricultural and rural areas in lower-income countries, *American Journal of Agricultural Economics*, 89(5): 1241-1247

Barnett, B., J.R. Black, Y. Hu and J. Skees, 2005, Is area yield insurance competitive with farm yield insurance?, *Journal of Agricultural and Resource Economics*, 30(2): 285-301.

Biffis, E. and E. Chavez, 2017, Satellite data and machine learning for weather risk management and food security, *Risk Analysis*, 37(8): 1508-1521.

Biffis, E., E. Chavez, A. Louaas and P. Picard, 2022, Parametric insurance and technology adoption in developing countries, *Geneva Risk and Insurance Review*, 47, 7-44.

Bryan, G., 2019, Ambiguity aversion decreases the impact of partial insurance: Evidence from African farmers, *Journal of the European Economic Association*, 17(5): 1428-1469.

Carter, M., A.D. Janvry, E. Sadoulet and A. Sarris, 2017, Index-insurance for developing country agriculture: a reassesment, *Annual Review of Resource Economics*, 9(1): 421-438.

Chantarat, S., C.B. Barrett, A.G. Mude and C.G. Turvey, 2007, Using weather index insurance to improve drought response for famine prevention, *American Journal of Agricultural Economics*, 89, 1262–1268.

Chantarat, S., A.G. Mude, C.B. Barrett and M.M. Carter, 2013, Designing index-based livestock insurance for managing asset risk in Northern Kenya, *Journal of Risk and Insurance*, 80, 205–237.

Clarke, D.J., 2016, A theory of rational demand for index insurance, *American Economic Journal: Microeconomics*, 8(1): 283-306.

Eeckhoudt, L., C. Gollier and H. Schlesinger, 1991, Increase in risk and deductible insurance, *Journal of Economic Theory*, 55: 435-440.

Eeckhoudt. L. and H. Schlesinger, 2006, Putting risk in its proper place, *American Economic Review*, 96: 280-289.

Emanuel, K., 2011, Global warming effects on US hurricane damage, *Weather, Climate, and Society*, 3(4): 261-268.

Gollier, C., 1996,Optimal insurance of approximate losses, *Journal of Risk and Insurance*, 63(3): 369-380.

Gollier, C., 2004, *The Economics of Risk and Time*, MIT Press.

Jewitt, I., 1988, Risk and risk aversion in the two risky asset portfolio problem, *mimeo*, University of Bristol.

Katz, R.W., 2002, Stochastic modeling of hurricane damage, *Journal of Applied Meteorology*, 41: 754-762.

Kusuma, A., B. Jackson, and J. Noy, 2018, A viable and cost-effective weather index insurance for rice in Indonesia, *Geneva Risk and Insurance Review*, 43:186-218.

Liao, S.-C., S.-C. Chang and T.-C. Cheng, 2021, Managing the volatility risk of renewable energy: Index insurance for off-shore wind farms in Taiwan, *Sustainability*, MDPI, 13, 8985.

Miranda, M., 1991, Area-yield crop insurance reconsidered, *American Journal of Agricultural Economics*, 73(2): 233-242.

Nordhaus, W.D., 2010, The economics of hurricanes and implications of global warming, *Climate Change Economics*, 1(01): 1-20.

Omondi, C. K.,T. Rientjes, M.J. Booij and A.D. Nelson, 2021, Satellite rainfall bias assessment for crop growth simulation: A case study of maize growth in Kenya, *Agricultural Water Management*, 258, 1-14.

Pielke, R.A., 2007, Future economic damage from tropical cyclones: sensitivities to societal and climate changes, *Philosophical Transactions of the Royal Society A: Mathematical, Physical, and Engineering Sciences*, 365(1860): 2717-2729.

Ross, S. A., 1981, Some stronger measures of risk aversion in the small and in the large with applications, *Econometrica*, 3, 621-638.

Rothschild, M. and J. Stiglitz, 1970, Increasing risk: I. A definition, *Journal of Economic Theory*, 2(3): 225-243.

Schlesinger, H, 2013, The theory of insurance demand, *Handbook of Insurance* (G. Dionne, Ed.), Springer, 167-184.

Skees, J.R., 2008, Innovation in index insurance for the poor in lower income countries, Agricultural Resource Economic Review, 37, 1–15.

Teh, T.-L. and C. Woolnough, 2019, A better trigger: Indices for insurance, *Journal of Risk and Insurance*, 86(4): 861-885.

Zhang, J., K.S. Tan, and C. Weng, 2018, Index insurance design, *Astin Bulletin: The Journal of the IAA*, 1-33.