# Risk Aversion and Insurance Propensity*

Fabio Maccheroni Massimo Marinacci Ruodu Wang Qinyu Wu

February 9, 2025

**Abstract**

We provide a new foundation of risk aversion by showing that this attitude is fully captured by the propensity to seize insurance opportunities. Our foundation, which applies to all probabilistically sophisticated preferences, well accords with the commonly held prudential interpretation of risk aversion that dates back to the seminal works of Arrow (1963) and Pratt (1964). In our main results, we first characterize the Arrow-Pratt risk aversion in terms of propensity to *full* insurance and the stronger notion of risk aversion of Rothschild and Stiglitz (1970) in terms of propensity to *partial* insurance. We then extend the analysis to comparative risk aversion by showing that the notion of Yaari (1969) corresponds to comparative propensity to full insurance, while the stronger notion of Ross (1981) corresponds to comparative propensity to partial insurance.

The seminal works of Arrow (1963), Pratt (1964), and Rothschild and Stiglitz (1970) provide different behavioral notions of risk aversion that under expected utility amount to the familiar concavity property of utility functions. All these papers relate their notions of risk aversion with insurance choices, arguably the most basic domain of application of any risk analysis. In this paper, we show how insurance choice behavior can be used to *define* risk attitudes. At a theoretical level, our results provide compelling economic explanations of risk attitudes, thus clarifying their economic appeal and normative status. At an empirical level, they show how these attitudes may motivate the most common features of marketed insurance policies.

As both Arrow (1963) and Pratt (1964) observe, risk aversion can be characterized as preference for full insurance over no insurance at an actuarially fair premium. To formalize this claim, consider an agent who, say because of real or financial assets held, faces a random wealth change $w$ that we call *risk*.[1] A full insurance for risk $w$ at premium $\pi$ is a contract with random payoff $-w-\pi$ that eliminates all uncertainty by replacing the random loss $-w$ with a fixed cost $\pi$. The actuarially fair premium is, by definition, the expected loss $\mathbb{E}[-w]$. Thus, the agent prefers to sign up an actuarially fair full insurance rather than facing the risk if

$$w + \underbrace{(-w - \mathbb{E}[-w])}_{\text{full insurance at fair premium}} \gtrsim w + \underbrace{0}_{\text{no insurance}} \tag{1}$$

that is, if $\mathbb{E}[w] \gtrsim w$ for all risks $w$. In words, it amounts to a preference for a sure amount over a random one with the same expectation. This is the classical Arrow-Pratt notion of risk aversion.

While the concept of full insurance is natural and is, since a long time, common in the insurance practice, in reality no insurance comes at fair premium. For instance, insurance companies have operating expenses

*Maccheroni (fabio.maccheroni@unibocconi.it) and Marinacci (massimo.marinacci@unibocconi.it) are with the Department of Decision Sciences, Università Bocconi; Wang (ruodu.wang@uwaterloo.ca) and Wu (q35wu@uwaterloo.ca) are with the Department of Statistics and Actuarial Science, University of Waterloo. All authors contributed equally to this work. For helpful comments that substantially improved the paper, we thank the editors, three anonymous referees, numerous seminar and conference audiences, as well as Alain Chateauneuf, Paul Embrechts, Martin Larsson, Marcel Nutz, Luciano Pomatto, Marco Scarsini, Peter Wakker, Jingni Yang, and Bin Wang. This work was supported by the Italian Ministry of University and Research [Grant PRIN-2017CY2NCA] and the Natural Sciences and Engineering Research Council of Canada [Grants RGPIN-2018-03823 and RGPAS-2018-522590].

[1]For instance, an agent who owns a car with market value $\nu$ faces a wealth change $w$ equal to $-\nu$ if the car is stolen, and to 0 otherwise. As it is typically the case, here the risk $w$ is a negative random variable, with $-w$ representing the agent loss. The 'risk' terminology that we adopt comes from Pratt (1964) and is also used by Gollier (2001).

that affect premia.[2] But then the connection of expression (1) to actual insurance choices becomes a weak one – we may want to define individuals who pick insurances at higher premia as risk averse, but little can be said of individuals who refuse them.

Our first contribution is to show that there is, indeed, a strong connection. We provide an equivalent definition of risk aversion that is purely based on insurance concepts and does not rely on expectations and fair premia. Consider an agent who, at the same price $\pi$, can either buy full insurance $-w$ or make another investment $h$ that has the same distribution of $-w$. Agent's payoff is $w + (-w - \pi)$ when full insurance is purchased, while it is $w + (h - \pi)$ when investment $h$ is purchased. Arrow-Pratt's risk aversion implies that

$$w + \underbrace{(-w - \pi)}_{\text{full insurance at price } \pi} \succsim w + \underbrace{(h - \pi)}_{h \text{ distributed as } -w \text{ at price } \pi} \tag{2}$$

for all $w$ and all $\pi$. Indeed, the sure payoff $-\pi$ on the left-hand side is the expectation of the random payoff $w + (h - \pi)$ on the right-hand side (because $h$ is distributed as $-w$). Preference pattern (2) has a clear meaning of *propensity to full insurance*, does not make use of expected values, and applies to actual insurance choices dealing with insurance premia that are not actuarially fair. In terms of dependence, a full insurance is perfectly negatively correlated with risk $w$ and therefore guarantees a constant payoff. In contrast, the equally distributed $h$ might well have a different correlation structure with risk $w$, as a simple example in Section II.A illustrates. Arrow-Pratt's risk aversion thus manifests itself into different attitudes towards identically distributed modifications of $-w$ based on their correlation with risk $w$, perfect negative correlation being favored because it eliminates wealth variability.

Our first main result, Theorem 1, shows that this propensity to full insurance is equivalent to Arrow-Pratt's risk aversion for every transitive preference $\succsim$ over random payoffs that depends only on payoffs' distributions. It thus applies to expected utility, where it provides a novel underpinning for concave utility, but goes well beyond it. In particular, it applies to all probabilistically sophisticated preferences in the sense of Machina and Schmeidler (1992) such as cumulative prospect theory with probability weighting (Tversky and Kahneman, 1992).[3] It also applies to preferences that do not satisfy stochastic dominance, like the original prospect theory of Kahneman and Tversky (1979), and that might not even be complete, like the mean-variance preferences of Markowitz (1952).

One may then argue that many insurance contracts do not provide full coverage. Some of them, like most health insurance policies, have a proportional form as they reimburse only a fraction of the loss. Others, like many property insurance policies, have a deductible-limit form as they impose a deductible and a policy limit. The resulting notions of propensity to partial insurance are natural extensions of (2). For instance, *propensity to proportional insurance* requires

$$w + \underbrace{(-\alpha w - \pi)}_{\text{proportional insurance at price } \pi} \succsim w + \underbrace{(h - \pi)}_{h \text{ distributed as } -\alpha w \text{ at price } \pi} \tag{3}$$

for all $w$, all $\pi$, and all $\alpha \in (0, 1]$.[4] The definition of *propensity to deductible-limit insurance* is analogous.

Our second main result, Theorem 2, shows that propensity to proportional insurance and propensity to deductible-limit insurance are both equivalent to the Rothschild-Stiglitz notion of risk aversion

$$\mathbb{E}[\varphi(f)] \geq \mathbb{E}[\varphi(g)] \text{ for all concave } \varphi : \mathbb{R} \to \mathbb{R} \text{ implies } f \succsim g \tag{4}$$

Like the equivalence of Arrow-Pratt's risk aversion (1) and propensity to full insurance (2), also the equivalence between Rothschild-Stiglitz's risk aversion (4) and propensity to proportional insurance (3) is based

[2]For a textbook treatment see, e.g., Dickson (2017), who writes 'It is unlikely that an insurer who calculates premiums by this [fair premium] principle will remain in business very long'.

[3] This class also includes the preferences introduced by Machina (1982), rank-dependent utility (Quiggin, 1982, Yaari, 1987), betweenness preferences (Dekel, 1986, Chew, 1989, Gul, 1991), multiplier preferences (Hansen and Sargent, 2008), quantile preferences (Rostek, 2010), and cautious expected utility (Cerreia-Vioglio, Dillenberger, and Ortoleva, 2015).

[4]The *percentage excess* $1 - \alpha$ is the fraction of the loss not covered by the insurance policy. When $\alpha = 1$ proportional coverage corresponds to full coverage. As a consequence, propensity to proportional insurance is a stronger requirement than propensity to full insurance.

on basic insurance concepts that do not rely on expectations. At a theoretical level, our findings provide new definitions of risk aversion that are economically founded and may clarify its normatively appeal. At an empirical level, they show how Rothschild-Stiglitz's risk aversion may underlie two important market phenomena: (i) the prevalence of proportional and deductible-limit policies in insurance markets, (ii) the fact that insurance policyholders typically have both kinds of contracts in their portfolios.

In our analysis we also consider more general, yet standard, definitions of partial insurance that only require coverage to increase with loss.[5] We show that the resulting notions of propensity to partial insurance correspond again to Rothschild-Stiglitz risk aversion, thus providing further support to this popular concept.

We then extend the analysis to comparative risk attitudes. We show that comparative risk aversion in the sense of Yaari (1969) corresponds to comparative propensity to full insurance, while the stronger notion of comparative risk aversion due to Ross (1981) corresponds to comparative propensity to partial insurance, in its various forms (proportional, deductible-limit, and so on). These comparative results complete our analysis, which thus provides a unified economic perspective on weak and strong notions of absolute and comparative risk aversion in terms of insurance choices, as Figure 2 shows in the conclusion.

Finally, we relate our results to the ones on correlation aversion of Epstein and Tanny (1980), and to the ones on expected-value preferences of the classical de Finetti (1931) and of the recent Pomatto, Strack, and Tamuz (2020).

# I Preliminaries

## A Risk setting

We study an agent who has to choose, at time 0, among actions that yield, at time 1, monetary payoffs that depend on uncertain contingencies outside the agent control. Uncertainty resolves at time 1 and is represented by a probability space $(S, \Sigma, P)$, where $S$ is a space of payoff-relevant states (the contingencies), $\Sigma$ is a $\sigma$-algebra of events in $S$, and $P$ is the probability measure on $\Sigma$ that governs states' realizations.

Each action corresponds to a random variable $f : S \to \mathbb{R}$ with $f(s)$ interpreted as the, positive or negative, monetary payoff obtained in state $s$ when the action is taken.

The probability measure $P$ is given and, in the tradition of Savage (1954), it is assumed throughout to be *adequate*, that is, either nonatomic on $\Sigma$ or uniform on a finite partition that generates $\Sigma$. Nonatomicity is a standard divisibility assumption requiring that, for each event A with $P(A) > 0$, there exists an event $B \subseteq A$ such that $0 < P(B) < P(A)$; it amounts to the existence of a random variable with continuous distribution (e.g., a normal distribution).

We restrict our attention to random variables that admit all moments. We call them *random payoffs* and denote their collection by $\mathcal{F}$, with typical elements $f$, $g$, and $h$. Formally, we denote by $\mathcal{L}^0$ the space of all random variables and by $\mathcal{L}^\infty$ the subspace of $\mathcal{L}^0$ that consists of all (almost surely) bounded random variables. Moreover, for each $p \in [1, \infty)$ we denote by $\mathcal{L}^p$ the subspace of all elements $f$ of $\mathcal{L}^0$ with finite absolute $p$-th moment $\mathbb{E}[|f|^p]$. With this, we consider

$$\text{either } \mathcal{F} = \mathcal{L}^\infty \text{ or } \mathcal{F} = \mathcal{M}^\infty$$

where $\mathcal{M}^\infty = \bigcap_{p \in \mathbb{N}} \mathcal{L}^p$ is the space of all random variables with finite moments of all orders (as usual, $\mathbb{N} = \{0, 1, ...\}$ is the set of all natural numbers). The space $\mathcal{M}^\infty$ contains $\mathcal{L}^\infty$, the usual setting of decision theory under risk, yet it allows for random variables that are commonly used in applications – like normals, log-normals, and gammas – with distributions that admit all moments, but may have unbounded support. All the results in the main text, with the exception of Proposition 3, hold for both spaces. Also, in the main text, we consider convergence of random payoffs in $\mathcal{F}$ with respect to all integer $p$-norms, that is, $f_n \to f$ whenever $\mathbb{E}[|f_n - f|^p] \to 0$ for all $p \in \mathbb{N}$. In Appendix B.1 we detail how this mode of convergence can be weakened.

[5] They include proportional and deductible-limit insurances as special cases since their payoffs are, indeed, (weakly) increasing functions of the loss (see Figure 1).

Each random payoff $f$ induces a distribution $P_f = P \circ f^{-1}$ of deterministic payoffs, called 'lottery' in the decision theory jargon. In particular, $P_f(B)$ is the probability that $f$ yields an outcome in the Borel subset $B$ of the real line.

**Definition 1.** *Two random payoffs $f$ and $g$ are* equally distributed*, written $f \overset{d}{=} g$, when $P_f = P_g$.*

Equally distributed random payoffs generate the same lottery, but they may have different realizations in the same state, as the example in Section II.A illustrates.

## B Risk preferences

The agent preferences are represented by a binary relation $\succsim$ on the space $\mathcal{F}$ of random payoffs. We read $f \succsim g$ as 'the agent prefers $f$ to $g$'. As usual, $\sim$ and $\succ$ denote the indifference and strict preference relations.

**Definition 2.** *A binary relation $\succsim$ on $\mathcal{F}$ is a* risk preference *when it is both transitive and law invariant, that is,*

$$f \overset{d}{=} g \Longrightarrow f \sim g$$

*for all random payoffs $f$ and $g$.*

Besides the standard assumption of transitivity, the definition of risk preference assumes law invariance, which requires the agent to be indifferent between equally distributed random payoffs. The fact that only the lottery $P_f$ induced by $f$ matters to the agent is what characterizes choice under risk, hence the name risk preferences. Law invariance guarantees reflexivity, which is thus automatically satisfied by a risk preference.

As previously mentioned, risk preferences include all probabilistically sophisticated preferences, like the preferences introduced by Machina (1982), rank-dependent utility (Quiggin, 1982, Yaari, 1987), betweenness preferences (Dekel, 1986, Chew, 1989, Gul, 1991), cumulative prospect theory (Tversky and Kahneman, 1992), multiplier preferences (Hansen and Sargent, 2008), quantile preferences (Rostek, 2010), and cautious expected utility (Cerreia-Vioglio, Dillenberger, and Ortoleva, 2015). Risk preferences also include some classes of incomplete preferences, like the expected multi-utility of Dubra, Maccheroni, and Ok (2004), classes of preferences that do not satisfy stochastic dominance, like the original prospect theory of Kahneman and Tversky (1979), and classes with both of these features, like the mean-variance preferences of Markowitz (1952) defined by $f \succsim_{\mathrm{MV}} g$ if and only if $\mathbb{E}[f] \geq \mathbb{E}[g]$ and $\mathbb{V}[f] \leq \mathbb{V}[g]$.

**Definition 3.** *A risk preference is* continuous *when*

$$f_n \succsim g_n \text{ for all } n \quad \Longrightarrow \quad \lim_n f_n \succsim \lim_n g_n$$

*for all convergent sequences $\{f_n\}$ and $\{g_n\}$ of random payoffs.*

This assumption is weaker than continuity in distribution because our notion of convergence implies convergence in distribution.

## C Classical risk attitudes

As mentioned in the introduction, there are two classical approaches to risk attitudes. One approach is due to Arrow (1963) and Pratt (1964). It is based on the observation that a random payoff $f$ is 'risky' when it is not constant, that is, when $f \neq \mathbb{E}[f]$. This leads to the definition of *weak* risk attitudes.

**Definition 4.** *A risk preference $\succsim$ is:*

*(i)* weakly risk averse *when, for all random payoffs $f$,*

$$\mathbb{E}[f] \succsim f$$

*(ii)* weakly risk propense *when, for all random payoffs $f$,*

$$\mathbb{E}[f] \precsim f$$

*(iii)* weakly risk neutral *when it is both weakly risk averse and propense.*

The other approach is due to Rothschild and Stiglitz (1970). They show that the relation $\geq_{\text{cv}}$ on $\mathcal{F}$ defined by

$$f \geq_{\text{cv}} g \iff \mathbb{E}\left[\varphi\left(f\right)\right] \geq \mathbb{E}\left[\varphi\left(g\right)\right] \text{ for all concave } \varphi : \mathbb{R} \to \mathbb{R}$$

meaningfully captures the idea that '$f$ is less risky than $g$' (for example, in terms of mean preserving spreads). This leads to the definition of *strong* risk attitudes.

**Definition 5.** *A risk preference* $\succsim$ *is:*

*(i)* strongly risk averse *when, for all random payoffs* $f$ *and* $g$*,*

$$f \geq_{\text{cv}} g \Longrightarrow f \succsim g$$

*(ii)* strongly risk propense *when, for all random payoffs* $f$ *and* $g$*,*

$$f \geq_{\text{cv}} g \Longrightarrow f \precsim g$$

*(iii)* strongly risk neutral *when it is both strongly risk averse and propense.*

Clearly, strong risk aversion (propensity) implies weak risk aversion (propensity). As well-known, these two notions are equivalent for expected utility preferences, but not in general.[6] In contrast, the strong and weak notions of risk neutrality always coincide, so we can talk of 'risk neutrality' without further qualification.

# II Absolute attitudes

## A Insurance contracts and attitudes

As discussed in the introduction, our main objective is to characterize classical risk attitudes in terms of insurance choices. To tackle this problem we need to answer two questions:

- Which random payoffs can be seen as insurances for a given risk $w$?
- How can we describe the attitudes towards insurance of an agent facing a risk $w$?

Let us identify, as in Arrow (1974), insurance policies (also called contracts) with the random payoffs detailing their state-contingent net payments to the agent. Formally, an insurance policy that pays $h(s)$ in each state $s$ and has premium $\pi$ corresponds to the random payoff $f = h - \pi$. We consider an agent who, *before* an insurance policy is chosen, faces a risk $w \in \mathcal{F}$, so a random loss $-w$. Therefore, *after* the policy $f \in \mathcal{F}$ is chosen, the risk changes to $w + f$.

In our static analysis, the policy is chosen at time 0 and uncertainty resolves at time 1. For our purposes, it is immaterial whether we interpret the risk $w$ as the random wealth change over the considered period (as we maintain throughout to ease exposition) or, rather, as the final random wealth of the agent. Indeed, in our setting the insurance problem of an agent with initial wealth $w_0 \in \mathbb{R}$ who confronts risk $w \in \mathcal{F}$ is equivalent to that of an agent with initial wealth 0 who confronts risk $w_0 + w \in \mathcal{F}$. By purchasing insurance, agents seek protection against payoff variability, which is unaffected by the addition of constants. For instance, to fully insure risk $w$ at premium $\pi$, thus receiving $-w$ at cost $\pi$, is equivalent to fully insure $w_0 + w$ at premium $\pi - w_0$. Formally, this equivalence corresponds to the accounting identity $-w - \pi = -\left(w_0 + w\right) - \left(\pi - w_0\right)$.[7] For this reason, in our analysis, random payoffs can take both positive and negative values.

Next we introduce a basic taxonomy of insurance policies.

[6] Yaari (1987), Wakker (1994), Cohen (1995), and Schmidt and Zank (2008) study several notions of risk aversion for rank-dependent and cumulative prospect theory preferences. For instance, in the dual model of Yaari (1987) weak risk aversion corresponds to a probability weighting function that is dominated by the identity function, while strong risk aversion to a convex probability weighting function.

[7] Similar identities hold for partial insurances, which thus feature analogous equivalences. Appendix C.1 discusses these properties and their behavioral implications in more detail.

**Definition 6.** *Given any risk $w$, a random payoff $f$ is:*

*(i)* *a* full insurance *for $w$, written $f \in \mathcal{I}^{\text{fi}}(w)$, when*

$$f = -w - \pi$$

*for some premium $\pi \in \mathbb{R}$;*

*(ii)* *a* proportional insurance *for $w$, written $f \in \mathcal{I}^{\text{pr}}(w)$, when*

$$f = -(1-\varepsilon)\, w - \pi$$

*for some premium $\pi \in \mathbb{R}$ and percentage excess $\varepsilon \in [0,1)$;*

*(iii)* *a* deductible-limit insurance *for $w$, written $f \in \mathcal{I}^{\text{dl}}(w)$, when*

$$f = \min\left\{(-w-\delta)^{+}, \lambda\right\} - \pi$$

*for some premium $\pi \in \mathbb{R}$, deductible $\delta \in \mathbb{R}$, and limit $\lambda \in [0,\infty)$.*[8]

Full insurances completely cover the agent position by neutralizing, at a cost, the uncertainty that the agent faces. Instead, proportional and deductible-limit insurances provide only partial cover: they reimburse either a proportion $1-\varepsilon$ of the loss or the part of the loss exceeding $\delta$, up to $\lambda$. They are the most common and simplest kinds of insurance contracts. Health insurance contracts are typically proportional, while property insurance ones have a deductible-limit form.

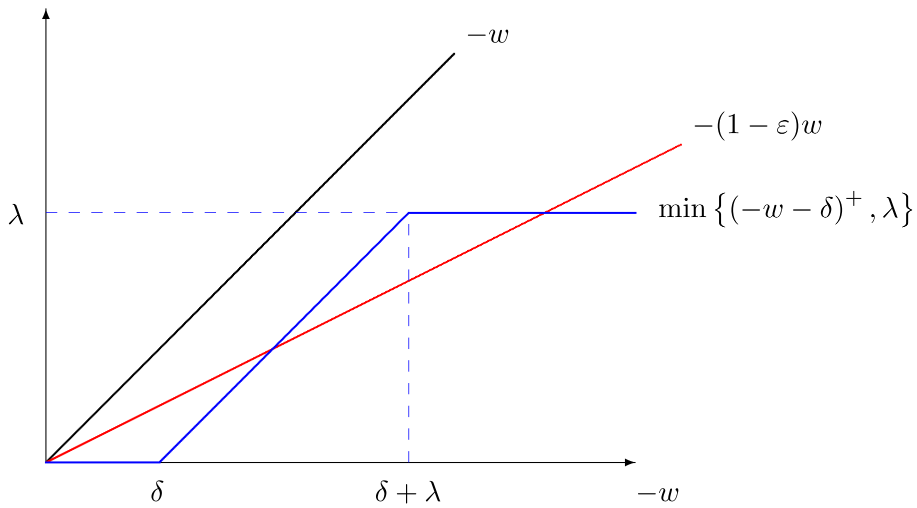


Figure 1: Proportional insurance (in red) and deductible-limit insurance (in blue) for loss $-w$

Next we introduce attitudes towards insurance using the types of contracts that we just presented. Recall that an agent facing risk $w$ who purchases insurance $f$ ends up with $w+f$.

**Definition 7.** *A risk preference $\succsim$ is:*

*(i)* propense to full insurance *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \in \mathcal{I}^{\text{fi}}(w) \implies w + f \succsim w + g$$

*(ii)* propense to proportional insurance *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \in \mathcal{I}^{\text{pr}}(w) \implies w + f \succsim w + g$$

[8] As usual, $(-w-\delta)^{+}$ denotes the positive part of $-w-\delta$.

*(iii)* propense to deductible-limit insurance *when, for all* $w, f, g \in \mathcal{F}$ *with* $g \overset{d}{=} f$,

$$f \in \mathcal{I}^{\text{dl}}(w) \implies w + f \succsim w + g$$

Intuitively, insurance has the benefit of being negatively correlated to the risk that the agent is facing: being insured flattens the agent payoff. Other contracts may have the same distribution, but they may not be correlated in the same way. Our whole point is that an insurance-propense agent prefers a contract that, being negatively correlated with the risk, reduces payoff variability. In particular, an agent who is propense to full insurance favors a perfect negative correlation that altogether eliminates variability. When the agent also values a milder negative correlation that partially reduces variability, the stronger notion of propensity to partial insurance takes the stage.

These definitions address the initial questions of this section by relying on a common principle: once a kind of insurance is defined for risk $w$ (Definition 6), propensity to insurance of that kind means that the agent prefers to purchase these insurances $f$ over other equally distributed random payoffs $g$ (Definition 7). Equidistribution is a *ceteris paribus* assumption that disciplines comparisons by ensuring, for example, that neither of the random payoffs at hand be statewise dominated (with dominance considerations then confounding insurance motives).[9]

For a simple illustration, consider two important extreme weather events like 'excess rainfall' and 'drought'. Wine grapes are an example of crop much more vulnerable to excess rainfall than to drought, while the opposite is true for rice. If the two extreme events are equally likely, the rain insurance $f$ and the drought insurance $g$ paying, respectively, 1 in case of excess rainfall and 0 otherwise, and 1 in case of drought and 0 otherwise, have the same distribution. For a viticulturist growing wine grapes – with, say, revenue $w_{\text{grapes}}$ equal to 0 in case of excessive rainfall and to 1 otherwise – the rain insurance $f$ is a full insurance policy, while the equally distributed drought insurance $g$ is not. Representing these in a table, we have,

| | excess rainfall | drought | other weather conditions |
|---|---|---|---|
| $w_{\text{grapes}} + f$ | 1 | 1 | 1 |
| $w_{\text{grapes}} + g$ | 0 | 2 | 1 |

In contrast, for a rice farmer – with, say, revenue $w_{\text{rice}}$ equal to 0 in case of drought and to 1 otherwise – it is the drought insurance $g$ that becomes a full insurance policy, while the equally distributed rain insurance $f$ is not. Indeed,

| | excess rainfall | drought | other weather conditions |
|---|---|---|---|
| $w_{\text{rice}} + f$ | 2 | 0 | 1 |
| $w_{\text{rice}} + g$ | 1 | 1 | 1 |

In conclusion, equally distributed random payoffs can be vastly different when viewed as possible insurance policies for a given agent, depending on their correlation with the risk that the agent is facing. Agents' behavior will differ accordingly: when both farmers are propense to full insurance, the viticulturist will prefer the acquisition of $f$, the rice farmer that of $g$.

There is a natural hierarchy among the insurance attitudes introduced in Definition 7. When $w$ is bounded, it can be shown that,

$$\mathcal{I}^{\text{fi}}(w) = \mathcal{I}^{\text{pr}}(w) \cap \mathcal{I}^{\text{dl}}(w) \tag{5}$$

Thus, propensity to full insurance is weaker than propensity to either proportional or deductible-limit insurance. The next results show that this hierarchy in insurance attitudes corresponds to the hierarchy in risk attitudes.

[9]The notions presented in Definition 7 are equivalent to the ones discussed in the introduction because the equidistribution relation $\overset{d}{=}$ is invariant under the addition of constants. For instance, Definition 7-(i) is just a theoretically convenient rewriting of (2) because, when $g \overset{d}{=} f = -w - \pi \in \mathcal{I}^{\text{fi}}(w)$, by setting $h = g + \pi$ we have $h \overset{d}{=} -w$ as well as $w + f = w + (-w - \pi)$ and $w + g = w + (h - \pi)$. See also Appendix B.2.

## B Weak risk aversion

We start with weak risk aversion.

**Theorem 1.** *The following properties are equivalent for a risk preference:*

*(i) weak risk aversion;*

*(ii) propensity to full insurance.*

This theorem provides a novel foundation of weak risk aversion for all risk preferences. It shows how the traditional notion of 'preference for the expectation of a random payoff over the random payoff itself' emerges from a minimal requirement of propensity to insurance. It is minimal because only the purchase of full insurance is required to be preferred over the purchase of other equally distributed random payoffs; in other words, because weak risk aversion is silent about attitudes towards partial insurance. Furthermore, by making no use of expectations, the equivalence presented in Theorem 1 also addresses the normative critique of weak risk aversion that hinges on the seemingly *ad hoc* use of expectations over other possible statistics (such as the median).

Theorem 1 relies on a novel result in probability theory of some independent interest.

**Lemma 1.** *The following properties are equivalent for* $f \in \mathcal{F}$*:*

*(i)* $\mathbb{E}[f] = 0$*;*

*(ii) there exist* $h, h' \in \mathcal{F}$ *such that* $h \stackrel{d}{=} h'$ *and* $f \stackrel{d}{=} h - h'$*.*

The nontrivial part is that (i) implies (ii). Yet, a simple explanation is possible in the finite uniform case when $S = \{1, 2, ..., n\}$ and $P(s) = 1/n$ for all $s \in S$. In this case,

$$\mathbb{E}[f] = 0 \iff \sum_{s=1}^{n} f(s) = 0$$

Define the random payoffs $h$ and $h'$ by $h(s) = \sum_{i=1}^{s} f(i)$ and $h'(s) = \sum_{i=1}^{s-1} f(i)$ for all $s \in S$, with the convention $h'(1) = 0$. Diagrammatically,

| | 1 | 2 | $\cdots$ | $n-1$ | $n$ |
|---|---|---|---|---|---|
| $h$ | $f(1)$ | $f(1)+f(2)$ | $\cdots$ | $\sum_{i=1}^{n-1} f(i)$ | $\sum_{i=1}^{n} f(i) = 0$ |
| $h'$ | $0$ | $f(1)$ | $\cdots$ | $\sum_{i=1}^{n-2} f(i)$ | $\sum_{i=1}^{n-1} f(i)$ |

Therefore, $f(s) = h(s) - h'(s)$ for all $s \in S$, that is, $f = h - h'$. Moreover, it is easy to see that $h \stackrel{d}{=} h'$ since all states are equally probable. Thus, $h$ and $h'$ are the sought-after random payoffs showing that (i) implies (ii).

The general nonatomic case cannot be directly tackled through a limit argument building upon the finite uniform case because the cumulant random payoffs that we constructed above may lose boundedness or integrability when passing to the limit. Different techniques are needed. Interestingly, the probabilistic Lemma 1 and the decision-theoretic Theorem 1 turn out to be mathematically equivalent, as detailed in Appendix B.7. If one were able to prove directly Theorem 1 (something that eluded us), the lemma would follow.

To see how Lemma 1 implies Theorem 1, assume propensity to full insurance. For each random payoff $f$, by the lemma there exist two equally distributed risks $w$ and $w'$ such that $f - \mathbb{E}[f] \stackrel{d}{=} w - w'$. Now, $-w + \mathbb{E}[f]$ is a full insurance for risk $w$ that is equally distributed with $-w' + \mathbb{E}[f]$. Propensity to full insurance then implies

$$w + (-w + \mathbb{E}[f]) \succsim w + \left(-w' + \mathbb{E}[f]\right)$$

Hence, $\mathbb{E}[f] \succsim w - w' + \mathbb{E}[f] \stackrel{d}{=} f$ and so, by the law invariance of $\succsim$, we have $w - w' + \mathbb{E}[f] \sim f$. By transitivity, we conclude that $\mathbb{E}[f] \succsim f$, i.e., weak risk aversion holds, as desired. The easy converse was explained in the introduction.

## C Strong risk aversion

We now move to strong risk aversion.

**Theorem 2.** *The following properties are equivalent for a continuous risk preference:*

*(i) strong risk aversion;*

*(ii) propensity to proportional insurance;*

*(iii) propensity to deductible-limit insurance.*

This result shares the same features as the previous one in terms of economic scope, technical accessibility, and normative soundness. Furthermore, it justifies the use of the concave order instead of other dispersion orders (such as the one of Bickel and Lehmann, 1976) to define strong risk aversion.

Theorem 2 also has clear empirical relevance because proportional and deductible-limit insurances are the most commonly held and legally disciplined insurance policies. On the one hand, it shows that the strong risk aversion of agents may motivate the demand for these two types of insurance contracts. On the other hand, the prevalence of these two contracts in the insurance practice may support the hypothesis of strong risk aversion of most policyholders. Moreover, the *a priori* non-obvious equivalence between propensity to proportional insurance (ii) and to deductible-limit insurance (iii) is consistent with the fact that policyholders often have both kinds of contracts in their insurance portfolios.[10]

Finally, by expressing both concepts in the same language (that of insurance), Theorems 1 and 2 jointly provide a new perspective on the well-known fact that weak risk aversion is implied by strong risk aversion. When offered equally distributed payoffs, a weakly risk-averse agent only favors full insurance, whereas a strongly risk-averse one also favors some forms of partial insurance. In the next section we show that strongly risk-averse agents actually favor any kind of partial insurance.

## D More insurances

To further develop our analysis, and make it more realistic, we consider more general forms of partial insurance. A first principle of insurance theory is that an insurance policy pays more when the incurred loss is larger. There are two similar ways to formalize this principle, depending on whether we require the insurance payment to be a function of the realized loss. We regroup them in the following definition.

**Definition 8.** *Given any risk $w$, a random payoff $f$ is:*

*(iv) an* indemnity-schedule insurance *for $w$, written $f \in \mathcal{I}^{\mathrm{is}}(w)$, when*

$$f = I(-w)$$

*for some real-valued increasing map $I$ defined on the image of $-w$;*

*(v) a* contingency-schedule insurance *for $w$, written $f \in \mathcal{I}^{\mathrm{cs}}(w)$, when*

$$-w(s) > -w(s') \implies f(s) \geq f(s')$$

*for almost all states $s$ and $s'$.*[11]

Again, these two notions have a common basic meaning: greater losses cannot lead to smaller insurance payments. This is best seen by writing condition (iv), due to Arrow (1963), as

$$-w(s) \geq -w(s') \implies f(s) \geq f(s')$$

[10] Different forms of insurance policies address various issues in the insurance market. For example, deductible-limit insurance reduces labor costs in damage assessment for auto insurance, where damage verification is costly and moral hazard is a concern. Proportional insurance, on the other hand, is common in health insurance, where claim assessment is simpler.

[11] That is, almost surely with respect to the product probability measure $P \times P$.

for all states $s$ and $s'$. Thus, (iv) is obtained by (v) under the additional requirement that equal losses must lead to equal insurance payments.

With this, (v) is the most general notion of insurance that we consider.[12] It embodies a strong form of positive correlation between insurance $f$ and loss $-w$, known as comonotonicity (see Schmeidler, 1989). This property is what ultimately characterizes insurances, among all possible random payoffs, for an agent confronting risk $w$. We can now enrich relation (5) by adding the inclusions:

$$\mathcal{I}^{\text{pr}}(w) \cup \mathcal{I}^{\text{dl}}(w) \subseteq \mathcal{I}^{\text{is}}(w) \subseteq \mathcal{I}^{\text{cs}}(w)$$

The next definition is based on a different notion: rather than defining insurance for $w$, it describes different degrees of coverage for the loss $-w$ provided by two different policies $f$ and $g$. Yet, as it will be seen momentarily, this concept naturally connects to the previous ones.

**Definition 9.** *Given any risk $w$, a random payoff $f$ is a* better hedge *for $w$ than a random payoff $g$, written $f \trianglerighteq_w g$, when $f \overset{d}{=} g$ and*

$$P(f \leq \tau \mid w \leq \lambda) \leq P(g \leq \tau \mid w \leq \lambda)$$

*for all payments $\tau \in \mathbb{R}$ and risk levels $\lambda \in \mathbb{R}$.*

This means that $f$ first-order stochastically dominates $g$ on the left tails of $w$. In the language of Epstein and Tanny (1980, p. 18), $f \trianglerighteq_w g$ if and only if $f$ is *less correlated* (or *less concordant*) with $w$ than $g$. The next proposition connects the concepts of insurance and hedge.

**Proposition 1.** *Given any risk $w$, a random payoff $f$ is a contingency-schedule insurance for $w$ if and only if it is a best hedge for $w$, that is,*

$$\mathcal{I}^{\text{cs}}(w) = \left\{ f \in \mathcal{F} : f \trianglerighteq_w g \text{ for all } g \overset{d}{=} f \right\}$$

In other words, contingency-schedule insurances for $w$ are the policies that are less correlated to $w$ within any given distribution class. Next we introduce the definitions of propensity to insurance and to hedging relevant here, which are completely analogous to the ones given before.

**Definition 10.** *A risk preference $\succsim$ is:*

*(iv)* propense to indemnity-schedule insurance *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \in \mathcal{I}^{\text{is}}(w) \implies w + f \succsim w + g$$

*(v)* propense to contingency-schedule insurance *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \in \mathcal{I}^{\text{cs}}(w) \implies w + f \succsim w + g$$

*(vi)* propense to hedging *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \trianglerighteq_w g \implies w + f \succsim w + g$$

We are now ready for an omnibus result on the equivalence of strong risk aversion and propensity to partial insurance.

**Theorem 3.** *The following properties are equivalent for a continuous risk preference:*

*(i) strong risk aversion;*

*(ii) propensity to proportional insurance;*

[12] Proportional insurances with state-dependent percentage excesses might not be contingency-schedule insurances. This is also the case for deductible limit insurances with state-dependent deductibles.

*(iii) propensity to deductible-limit insurance;*

*(iv) propensity to indemnity-schedule insurance;*

*(v) propensity to contingency-schedule insurance;*

*(vi) propensity to hedging.*

Some implications easily follow from our earlier analysis, others are less obvious. Some attitudes, like (ii) and (iii), seem mild, easy to understand, and normatively compelling. Others, like (i) and (vi), seem instead more demanding and theoretically sophisticated. Be that as it may, they are all equivalent. In particular, as points (ii)-(v) embody different forms of propensity to partial insurance, we can summarize this result as:

$$\text{strong risk aversion} \iff \text{propensity to partial insurance} \iff \text{propensity to hedging}$$

To the best of our knowledge, the only precursor of this result is the equivalence between (i) and (vi) for expected utility preferences that can be derived from the findings of Epstein and Tanny (1980). Their results connect risk aversion and hedging propensity for expected utility preferences, but remain silent about insurance choice behavior, which is the lens that we adopt here to analyze risk aversion.

Finally, let us recall that both Theorems 1 and 3 (which subsumes Theorem 2) are valid for all preferences on $\mathcal{F}$ that are transitive, law invariant, and continuous. Therefore, the applicability of our results goes well beyond expected utility. This makes the present analysis relevant for popular models of risk behavior in psychology (such as the prospect theory of Kahneman and Tversky, 1979) and allows us to account for robustness concerns in economics and finance (as captured, e.g., by the multiplier preferences of Hansen and Sargent, 2008, or by the expected shortfall criterion of Artzner, Delbaen, Eber, and Heath, 1999). Our analysis shows that insurance propensity characterizes risk aversion irrespective of whether preferences abide to the expected utility model or violate it.

# III Neutrality

The definitions of *aversion* to the different kinds of insurance and to hedging are obtained from those of *propensity* by replacing $\succsim$ with $\precsim$. As usual, *neutrality* is then defined as simultaneous propensity and aversion. With this, the counterparts of Theorems 1, 2, and 3 hold as expected. In particular, all definitions of insurance neutrality coincide both with risk neutrality and with hedging neutrality.

The concept of neutrality is important because it serves as a benchmark to connect the absolute attitudes that we studied in the previous section and the comparative ones that we will analyze in the next section. With this motivation we go a bit deeper in its study. To this end, we introduce two more notions.

**Definition 11.** *A risk preference $\succsim$ is:*

- monotone *when, for all $w \in \mathcal{F}$ and all $\varepsilon \in (0, \infty)$,*

$$w + \varepsilon \succ w$$

- dependence neutral *when, for all $w, f, g \in \mathcal{F}$,*

$$g \overset{d}{=} f \implies w + f \sim w + g$$

Monotonicity just requires that the addition of a sure positive payoff is always preferred, a natural assumption when monetary outcomes are considered. Dependence neutrality means that preferences are unaffected by the possible correlation between risk $w$ and two identically distributed investments $f$ and $g$. It strengthens the requirement of law invariance, which corresponds to $w = 0$, to situations where risk is present.

**Proposition 2.** *The following conditions are equivalent for a risk preference $\succsim$:*

*(i) risk neutrality;*

*(ii) neutrality to full insurance;*

*(iii) neutrality to hedging;*

*(iv) dependence neutrality.*

*Moreover, $\succsim$ is monotone and satisfies any of the equivalent conditions above if and only if*

$$f \succsim g \iff \mathbb{E}[f] \geq \mathbb{E}[g] \tag{6}$$

*for all random payoffs $f$ and $g$.*

This proposition characterizes risk neutrality and makes explicit its relation with expected-value preferences. Its most innovative contribution is the characterization of these preferences based on dependence neutrality, which shows how law invariance irrespective of the outstanding risk leads to expected value maximization.

A fundamental feature of these preferences is their consistency with first-order stochastic dominance, $\geq_{\text{fsd}}$. This consistency is crucial in the existing characterizations of expected-value preferences, in particular the classic one of de Finetti (1931) and the more recent one of Pomatto, Strack, and Tamuz (2020). In our result, consistency with $\geq_{\text{fsd}}$ is implicit because, in the derivation, it follows from monotonicity and dependence neutrality. Yet, to better connect the approaches, next we provide a characterization of expected-value preferences that makes explicit the role of first-order stochastic dominance.

**Proposition 3.** *Let $\mathcal{F} = \mathcal{M}^\infty$ and $P$ be nonatomic. The following conditions are equivalent for a monotone risk preference $\succsim$:*

*(i) $\succsim$ admits an expected-value representation (6);*

*(ii) for all $w, f, g \in \mathcal{F}$,*

$$f \geq_{\text{fsd}} g \implies w + f \succsim w + g$$

*(iii) for all $w, f, g \in \mathcal{F}$,*

$$f \succsim g \implies w + f \succsim w + g$$

*(iv) $\succsim$ is complete and*

$$f \succ g \Longrightarrow w + \tilde{f} >_{\text{fsd}} w + \tilde{g}$$

*for some $w, \tilde{f}, \tilde{g} \in \mathcal{F}$ such that $f \overset{d}{=} \tilde{f}$, $g \overset{d}{=} \tilde{g}$ and $w$ is independent of both $\tilde{f}$ and $\tilde{g}$.*

The equivalence of conditions (i) and (ii) is the sought-after characterization of expected-value preferences in terms of first-order stochastic dominance. For perspective, Proposition 3 also reports the earlier characterizations of de Finetti (1931), which in preferential form corresponds to the equivalence of points (i) and (iii), and of Pomatto, Strack, and Tamuz (2020), which corresponds to the equivalence of points (i) and (iv).

In comparing condition (ii) with (iii), it is important to contrast the objective premise $f \geq_{\text{fsd}} g$ of the implication in (ii) with the subjective premise $f \succsim g$ of the one in (iii). In comparing (ii) with (iv), it is important to observe that the former is not the contrapositive of the latter. Indeed, the equivalence between (i), (ii) and (iii) continues to hold when $\mathcal{F} = \mathcal{L}^\infty$, while expected-value preferences on $\mathcal{L}^\infty$ fail to satisfy (iv).

## IV Comparative attitudes

We have shown how absolute risk attitudes – both strong and weak – can be characterized in terms of insurance behavior, without recurring to the concept of expectation, and how this leads to novel insights on old and recent results about risk preferences. It is then natural to wonder whether the same exercise can be performed for comparative attitudes.

## A Classical comparative risk attitudes

As it is the case for absolute risk attitudes, also comparative attitudes have a weak and a strong form. According to Yaari (1969), agent B (Bob) is *weakly more risk averse than* agent A (Ann) if whenever Ann prefers a sure payoff to a random one, so does Bob. Formally,

$$\gamma \succsim_{\mathrm{A}} f \implies \gamma \succsim_{\mathrm{B}} f$$

for all $f \in \mathcal{F}$ and $\gamma \in \mathbb{R}$. Ross (1981) introduces a stronger notion: B is *strongly more risk averse than* A if

$$\left.\begin{array}{l} f \geq_{\mathrm{cv}} g \\ g \sim_{\mathrm{A}} f - \rho_{\mathrm{A}} \\ g \sim_{\mathrm{B}} f - \rho_{\mathrm{B}} \end{array}\right\} \implies \rho_{\mathrm{B}} \geq \rho_{\mathrm{A}}$$

for all $f, g \in \mathcal{F}$ and $\rho_{\mathrm{A}}, \rho_{\mathrm{B}} \in \mathbb{R}$. The interpretation becomes transparent once one observes that $\rho_{\mathrm{A}}$ (resp. $\rho_{\mathrm{B}}$) is the amount of money Ann (resp. Bob) is willing to pay to replace $g$ with the less risky $f$. For the ease of exposition, next we introduce a class of risk preferences for which this amount always exists.

**Definition 12.** *A risk preference $\succsim$ is* secular *when, for all $f, g \in \mathcal{F}$, there exists $\rho \in \mathbb{R}$ such that $g \sim f - \rho$.*

When $\succsim$ is monotone, $\rho$ is the largest scalar $r$ such that $f - r \succsim g$, that is, the highest amount of money that the agent is willing to pay to trade $g$ with $f$. Equivalently, $-\rho$ is the smallest compensation for which the agent accepts this trade. Secularity, implicit in Ross (1981), thus requires that the agent is willing to trade any random payoff with another one for some suitable compensation. Briefly, 'every risk has its price' (see Gollier, 2001). This notion allows us to extend the observation of Ross, who studies the monotone and strictly concave expected utility case, that his definition is stronger than the one of Yaari.

**Lemma 2.** *The following conditions are equivalent for two monotone and secular risk preferences $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$:*

*(i)* B *is weakly more risk averse than* A*;*

*(ii) for all $f, g \in \mathcal{F}$ and $\rho_{\mathrm{A}}, \rho_{\mathrm{B}} \in \mathbb{R}$,*

$$\left.\begin{array}{l} f = \mathbb{E}[g] \\ g \sim_{\mathrm{A}} f - \rho_{\mathrm{A}} \\ g \sim_{\mathrm{B}} f - \rho_{\mathrm{B}} \end{array}\right\} \implies \rho_{\mathrm{B}} \geq \rho_{\mathrm{A}}$$

*In particular, if* B *is strongly more risk averse than* A*, then* B *is weakly more risk averse than* A.

This lemma also shows how Yaari's and Ross' notions are the comparative counterparts of the ones of Arrow-Pratt and Rothschild-Stiglitz. Indeed, in both the absolute and comparative cases, the weak notion corresponds to preference for expectation, the strong one to preference for less risky payoffs in general. The parallel does not stop here: the absolute risk aversion notions can be obtained from the comparative ones by assuming agent A to be risk neutral, as next we show. The result is known for the definition of Yaari (we report it for the sake of completeness), while it seems novel for the one of Ross.

**Lemma 3.** *Let $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$ be monotone and secular risk preferences. If* A *is risk neutral, then:*

1. B *is weakly more risk averse than* A *if and only if* B *is weakly risk averse.*
2. B *is strongly more risk averse than* A *if and only if* B *is strongly risk averse.*

To further elaborate, observe that when a risk preference is monotone and secular, given any $g$ and $f$ in $\mathcal{F}$ the sure amount $\rho = \rho(g, f)$ such that $g \sim f - \rho$ exists and is unique. So, the function

$$(g, f) \mapsto \rho(g, f) \tag{7}$$

is well defined. Intuitively, the greater $\rho(g, f)$ is, the more $f$ is preferred over $g$. With this, we can interpret the function (7) as a measure of the strength of preference. In view of Lemma 2, this function permits to reformulate the comparative notions of Yaari and Ross as follows:

- B is weakly more risk averse than A when $f = \mathbb{E}[g]$ implies $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$ for all $f, g \in \mathcal{F}$.
- B is strongly more risk averse than A when $f \geq_{\mathrm{cv}} g$ implies $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$ for all $f, g \in \mathcal{F}$.

The difference in the definitions is now evident. Not only $\mathbb{E}[g] \geq_{\mathrm{cv}} g$, but we also have $\mathbb{E}[g] \geq_{\mathrm{cv}} h$ for all $h \geq_{\mathrm{cv}} g$. In words, $\mathbb{E}[g]$ is the least risky among the random payoffs that are less risky than $g$. The sure payoff $\mathbb{E}[g]$ completely eliminates the risk involved in $g$, while a generic payoff $h \geq_{\mathrm{cv}} g$ only reduces it. Thus, Yaari compares the strength of preferences only when risk is eliminated, while Ross compares it also when risk is just reduced.

## B Comparative insurance propensity

In light of the previous analysis, the formalization of the concept of comparative propensity to full insurance is now natural:

**Definition 13.** *Let $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$ be monotone and secular risk preferences. We say that* B *is* more propense to full insurance than A *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \in \mathcal{I}^{\mathrm{fi}}(w) \implies \rho_{\mathrm{B}}(w+g, w+f) \geq \rho_{\mathrm{A}}(w+g, w+f)$$

In words, Bob is 'more willing to pay than' Ann in order to achieve full insurance for the risk $w$ that he faces. We can now state the comparative version of Theorem 1.

**Theorem 4.** *The following properties are equivalent for two monotone and secular risk preferences $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$:*

*(i)* B *is weakly more risk averse than* A*;*

*(ii)* B *is more propense to full insurance than* A*.*

To move to strong comparative attitudes, first observe that the definitions of comparative propensity to proportional insurance, deductible-limit insurance, indemnity-schedule insurance, and contingency-schedule insurance can be obtained by replacing $\mathcal{I}^{\mathrm{fi}}(w)$ with $\mathcal{I}^{\mathrm{pr}}(w)$, $\mathcal{I}^{\mathrm{dl}}(w)$, $\mathcal{I}^{\mathrm{is}}(w)$, and $\mathcal{I}^{\mathrm{cs}}(w)$ in Definition 13. Also the comparative version of propensity to hedging yields no surprises.

**Definition 14.** *Let $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$ be monotone and secular risk preferences. We say that* B *is* more propense to hedging than A *when, for all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$,*

$$f \trianglerighteq_w g \implies \rho_{\mathrm{B}}(w+g, w+f) \geq \rho_{\mathrm{A}}(w+g, w+f)$$

We can now state the comparative version of Theorem 3.

**Theorem 5.** *The following properties are equivalent for two continuous, monotone, and secular risk preferences $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$:*

*(i)* B *is strongly more risk averse than* A*;*

*(ii)* B *is more propense to proportional insurance than* A*;*

*(iii)* B *is more propense to deductible-limit insurance than* A*;*

*(iv)* B *is more propense to indemnity-schedule insurance than* A*;*

*(v)* B *is more propense to contingency-schedule insurance than* A*;*

*(vi)* B *is more propense to hedging than* A*.*

The interpretations and implications of these comparative results are similar to the absolute ones we discussed in Section II. In particular,

$$\text{stronger risk aversion} \iff \text{higher propensity to partial insurance} \iff \text{higher propensity to hedging}$$

In view of the fact that $f$ is a better hedge than $g$ for $w$ if and only if $f$ is more correlated than $g$ with $-w$ (in the sense of Epstein and Tanny, 1980), these equivalences confirm the classical intuition that agents are more risk averse if and only if they exhibit a stronger preference for insurance contracts that are more correlated with losses.

# V Conclusion

We have shown how the classic, weak and strong, absolute and comparative notions of risk aversion can be completely characterized through insurance choice behavior. Our analysis thus provides a unified economic perspective on these all-important attitudes. Figure 2 summarizes. In the tables, the superscript 'pi' (partial insurance) stands for any one of 'pr, dl, is, cs', with the set $\mathcal{I}^{\mathrm{pi}}(w)$ describing the corresponding class of partial insurance contracts for $w$.

ABSOLUTE ATTITUDES

Arrow (1963), Pratt (1964)

**Weakly risk averse**

$$f = \mathbb{E}[g] \implies f \succsim g$$

risk elimination is preferred

$\iff$

This paper

**Propense to full insurance**

$$f \overset{d}{=} g;\ f \in \mathcal{I}^{\mathrm{fi}}(w) \implies w + f \succsim w + g$$

full insurance acquisition is preferred

$\Uparrow$ $\Uparrow$

Rothschild and Stiglitz (1970)

**Strongly risk averse**

$$f \geq_{\mathrm{cv}} g \implies f \succsim g$$

risk reduction is preferred

$\iff$

This paper

**Propense to partial insurance**

$$f \overset{d}{=} g;\ f \in \mathcal{I}^{\mathrm{pi}}(w) \implies w + f \succsim w + g$$

partial insurance acquisition is preferred

COMPARATIVE ATTITUDES

Yaari (1969)

**Weakly more risk averse**

$$f = \mathbb{E}[g] \implies \rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$$

risk elimination is more valued

$\iff$

This paper

**More propense to full insurance**

$$f \overset{d}{=} g;\ f \in \mathcal{I}^{\mathrm{fi}}(w) \implies \rho^{w}_{\mathrm{B}}(g, f) \geq \rho^{w}_{\mathrm{A}}(g, f)$$

full insurance acquisition is more valued

$\Uparrow$ $\Uparrow$

Ross (1981)

**Strongly more risk averse**

$$f \geq_{\mathrm{cv}} g \implies \rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$$

risk reduction is more valued

$\iff$

This paper

**More propense to partial insurance**

$$f \overset{d}{=} g;\ f \in \mathcal{I}^{\mathrm{pi}}(w) \implies \rho^{w}_{\mathrm{B}}(g, f) \geq \rho^{w}_{\mathrm{A}}(g, f)$$

partial insurance acquisition is more valued

Figure 2: Summary tables, where $\rho^{w}(g, f)$ stands for $\rho(w + g, w + f)$

In sum, our unified analysis of the classical notions of risk aversion in the expectation-free language of insurance contracts roots these concepts into basic economic objects, thus improving their economic appeal. It also makes it possible to talk of risk attitudes for random variables with an infinite first moment, like those with some Pareto or Cauchy distributions, something that the traditional expectational analysis is unable to do. The study of these extended notions is an object of future study.

A secondary contribution of our analysis is to highlight the potential advantages of a state-space approach, based on random variables, for studying risk attitudes and decisions under risk more broadly. Lotteries may fully describe random variables that are considered in isolation. This is the case, for instance, in the preferential rankings over pairs of random variables – with either of them being possibly chosen – that underlie the classical von Neumann-Morgenstern axiomatization of expected utility. However, insurance inherently involves multiple interacting random variables, where correlations play a central role. A state-space framework then provides a natural way to address interdependence. In contrast, a purely lottery-type distributional analysis would require more intricate tools, such as multivariate distributions or copulas, which our approach does not need.

Finally we remark that our results can be readily extended to $\mathcal{L}^p$ spaces featuring any scalar $p \in [1, \infty)$, integer or not, as well as to the space $\mathcal{F}_0$ of simple random payoffs. Moreover, the conclusions of our theorems continue to hold under weaker definitions of propensity to insurance when premium calculation principles are explicitly specified. See Appendices C.2 and C.3 for these extensions.

## Appendix A Outlines of the proofs of Theorems 3, 4, and 5

The proofs of Theorem 1 and Lemma 1 are outlined in Section II.B. As Theorem 2 is implied by Theorem 3, here we will consider Theorems 3, 4, and 5.

**Proof sketch of Theorem 3.** Property (i) implies the other properties because, for any $w, f, g \in \mathcal{F}$, we have $w + f \geq_{\mathrm{cv}} w + g$ if $f \overset{d}{=} g$ and either $f \in \mathcal{I}^{\mathrm{pi}}(w)$ – where 'pi' denotes any of 'pr, dl, is, cs' – or the condition $f \trianglerighteq_w g$ is satisfied. This is a standard result in stochastic orders (see Müller and Stoyan, 2002). Property (vi) implies (v) by Proposition 1. Since $\mathcal{I}^{\mathrm{pr}}(w)$ and $\mathcal{I}^{\mathrm{dl}}(w)$ are subsets of $\mathcal{I}^{\mathrm{is}}(w) \subseteq \mathcal{I}^{\mathrm{cs}}(w)$ for all $w \in \mathcal{F}$, it is easy to see that (v) implies (iv), and that (iv) implies both (ii) and (iii).

The most challenging parts are that (ii) implies (i) and that (iii) implies (i). We first focus on the latter implication. Let us start with the finite uniform case, where $S = \{1, \ldots, n\}$ and $P(s) = 1/n$ for all $s \in S$. The proof is based on constructing mean preserving spreads. The first and most critical step is to verify that (iii) implies that $f \succsim g$ when $g$ is a mean preserving spread of $f$. For such $f$ and $g$, our purpose is to construct $\tilde{f}, \tilde{g}, \tilde{w} \in \mathcal{F}$ such that

$$\tilde{f} \overset{d}{=} \tilde{g},\ \tilde{f} \in \mathcal{I}^{\mathrm{dl}}(\tilde{w}),\ f = \tilde{w} + \tilde{f} \text{ and } g = \tilde{w} + \tilde{g} \tag{8}$$

Indeed, this yields $f \succsim g$ by (iii). An explicit construction of $\tilde{f}$, $\tilde{g}$, and $\tilde{w}$ is provided in the proof of Lemma 10 in Appendix B.3. Further, it is well-known that when $f \geq_{\mathrm{cv}} g$ there exists a sequence $h_0, , h_1, \ldots, h_m$ such that $f = h_0$, $g = h_m$ and each $h_{k+1}$ is a mean preserving spread of $h_k$. Therefore, transitivity and the existence of $\tilde{f}, \tilde{g}, \tilde{w} \in \mathcal{F}$ in (8) yield (i). The general nonatomic case can be directly tackled through a limiting argument building upon the finite uniform case.

The proof that (ii) implies (i) is similar, with 'dl' replaced by 'pr' in (8). The additional technical complexity is that $\tilde{f}$, $\tilde{g}$, and $\tilde{w}$ may not exist for all pairs $f, g$ with a mean preserving relationship. They do exist, however for a sufficiently large subset of such pairs, from which all pairs $f, g$ with a mean preserving relationship can be approximated as limits of sequences within this subset. This, combined with continuity, confirms that (ii) implies (i).

**Proof sketch of Theorem 4.** It is easy to see that (i) implies (ii) by noting that $w + f$ is constant when $f \in \mathcal{I}^{\mathrm{fi}}(w)$. Conversely, Lemma 1 plays an important role. To be specific, let $\gamma \succsim_{\mathrm{A}} h$ with $h \in \mathcal{F}$ and $\gamma \in \mathbb{R}$. Lemma 1 implies that there exist $w, w' \in \mathcal{F}$ such that $w \overset{d}{=} w'$ and $h - \mathbb{E}[h] \overset{d}{=} w - w'$. Thus, we can construct two random payoffs $f = -w + \mathbb{E}[h]$ and $g = -w' + \mathbb{E}[h]$ satisfying $f \overset{d}{=} g$, $f \in \mathcal{I}^{\mathrm{fi}}(w)$, $w + g \overset{d}{=} h$ and $w + f$ constant. Set

$$\eta_{\mathrm{A}} = (w + f) - \rho_{\mathrm{A}}(w + g, w + f) \quad \text{and} \quad \eta_{\mathrm{B}} = (w + f) - \rho_{\mathrm{B}}(w + g, w + f)$$

They are both constant. It follows from (ii) that $\rho_{\mathrm{A}}(w+g, w+f) \leq \rho_{\mathrm{B}}(w+g, w+f)$, which implies $\eta_{\mathrm{A}} \geq \eta_{\mathrm{B}}$. Note that

$$\eta_{\mathrm{A}} \sim_{\mathrm{A}} w+g \stackrel{d}{=} h \precsim_{\mathrm{A}} \gamma \quad \text{and} \quad \eta_{\mathrm{B}} \sim_{\mathrm{B}} w+g \stackrel{d}{=} h$$

By law invariance, $\eta_{\mathrm{A}} \sim_{\mathrm{A}} h$ and $\eta_{\mathrm{B}} \sim_{\mathrm{B}} h$. Monotonicity of $\succsim_{\mathrm{A}}$ yields $\gamma \geq \eta_{\mathrm{A}}$. As $\eta_{\mathrm{A}} \geq \eta_{\mathrm{B}}$, we get $\gamma \geq \eta_{\mathrm{B}}$. Further, we have $\gamma \succsim_{\mathrm{B}} \eta_{\mathrm{B}} \sim_{\mathrm{B}} h$, where the '$\succsim_{\mathrm{B}}$' step is due to the monotonicity of $\succsim_{\mathrm{B}}$. Transitivity shows that B is weakly more risk averse than A.

**Proof sketch of Theorem 5.** Using arguments similar to those in the proof of Theorem 3, we can demonstrate that property (i) implies the other properties and establish the sequence of implications from (vi) to (v), from (v) to (iv), and from (iv) to both (ii) and (iii).

We now address the more challenging implications (ii) $\Rightarrow$ (i) and (iii) $\Rightarrow$ (i). We focus on the latter implication as the former is similar but, like in Theorem 3, involves additional technical complexities requiring some standard convergence arguments. As in the proof of Theorem 3, the technique of mean preserving spreads is central. In the finite uniform case, we recall that if $g$ is a mean preserving spread of $f$, there exist $\tilde{f}, \tilde{g}, \tilde{w} \in \mathcal{F}$ such that (8) holds. By (iii), we have $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$. To extend the result for all $f, g \in \mathcal{F}$ with $f \geq_{\mathrm{cv}} g$, we can assume that $f = h_0$, $g = h_m$, and $h_{k+1}$ is a mean preserving spread of $h_k$ for $k = 0, \ldots, m-1$. In this step, we use some standard analysis to prove by induction that for every $x \in \mathbb{R}$ and for each $j = 1, 2, \ldots, m$,

$$\rho_{\mathrm{B}}(h_j - x, h_0 - x) \geq \rho_{\mathrm{A}}(h_j - x, h_0 - x)$$

This establishes a result stronger than $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$. The extension to the general nonatomic case can be directly tackled through a limit argument. In particular, by noting that the risk preference in Theorem 5 is continuous, we demonstrate that $\rho : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ is (jointly) sequentially continuous by Lemma 12 in Appendix B.14.

# Appendix B Proofs and related analysis

## B.1 Preamble

Recall that $(S, \Sigma, P)$ is an adequate probability space. We denote by $\mathcal{L}^0 = \mathcal{L}^0(S, \Sigma, P)$ the space of all measurable functions $f : S \to \mathbb{R}$, by $\mathcal{L}^\infty = \mathcal{L}^\infty(S, \Sigma, P)$ the space of all almost surely (a.s.) bounded elements of $\mathcal{L}^0$, and by $\mathcal{L}^p = \mathcal{L}^p(S, \Sigma, P)$ the space of all elements of $\mathcal{L}^0$ which admit finite absolute $p$-th moment (for $p \in (0, \infty)$). For $p \in [1, \infty]$, $\|\cdot\|_p$ is the usual (semi-)norm of $\mathcal{L}^p$. By *convergence in* $\mathcal{L}^p$, we mean convergence in this norm. By *convergence in* $\mathcal{M}^\infty = \bigcap_{p \in \mathbb{N}} \mathcal{L}^p$, we mean convergence in all of the $\|\cdot\|_p$ norms (for $p \in \mathbb{N}$). By *bounded a.s. convergence in* $\mathcal{L}^\infty$, we mean almost sure convergence of a sequence which is bounded in $\|\cdot\|_\infty$ norm. By the Dominated Convergence Theorem, bounded a.s. convergence implies convergence in $\mathcal{M}^\infty$.

When we say that a risk preference $\succsim$ is continuous on $\mathcal{F}$, we consider bounded a.s. convergence of sequences if $\mathcal{F} = \mathcal{L}^\infty$, and convergence of sequences in $\mathcal{M}^\infty$ otherwise.

We denote by $L^p = L^p(S, \Sigma, P)$ the quotient of $\mathcal{L}^p = \mathcal{L}^p(S, \Sigma, P)$ when almost surely equal measurable functions are identified (e.g. Pollard, 2002). Analogously, $M^\infty = M^\infty(S, \Sigma, P)$ is the quotient of $\mathcal{M}^\infty = \mathcal{M}^\infty(S, \Sigma, P)$.

Given any $f \in \mathcal{L}^0$, the cumulative distribution function $F : \mathbb{R} \to [0,1]$ of $f$ is defined by $F(x) = P(f \leq x)$ for all $x \in \mathbb{R}$. The function $F$ is increasing and right-continuous, with $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$. Its left-continuous inverse $F^{-1} : (0,1) \to \mathbb{R}$ is defined by $F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) \geq t\}$, also denoted by $q_f^-(t)$ or $F_f^{-1}(t)$ when its dependence on $f$ needs to be emphasized. The function $F^{-1}$ is always increasing, and it belongs to $\mathcal{L}^p(\lambda)$ if and only if $f \in \mathcal{L}^p(P)$ (for all $p \in [0, \infty]$), where as usual $\lambda$ is the Lebesgue measure on $(0,1)$.

Let $\mathcal{U}$ be the collection of all $v \in \mathcal{L}^0$ having a uniform distribution on $(0,1)$, i.e. $P(v \leq t) = t$ for all $t \in (0,1)$. It is without loss to assume $v(S) = (0,1)$ for all $v \in \mathcal{U}$. For each $f \in \mathcal{L}^0$, $f_v \in \mathcal{L}^0$ is defined by $f_v = q_f^- \circ v$.

**Lemma 4.** *Let $P$ be nonatomic and $f \in L^0$. Then:*

*(i) for each $v \in \mathcal{U}$, it holds $f_v \overset{d}{=} f$;*

*(ii) there exists $v \in \mathcal{U}$ such that $f_v = f$ a.s.*

**Proof.** See, e.g., Lemmas A.23 and A.32 of Föllmer and Schied (2016). ∎

In what follows, for each $n \in \mathbb{N}$ we denote by

$$\Psi_n = \left\{ \left( \frac{0}{2^n}, \frac{1}{2^n} \right], \left( \frac{1}{2^n}, \frac{2}{2^n} \right], \dots, \left( \frac{2^n - 1}{2^n}, \frac{2^n}{2^n} \right) \right\}$$

the partition of $(0,1)$ into segments of equal length $2^{-n}$. If $P$ is nonatomic and $v \in \mathcal{U}$, for each $n \in \mathbb{N}$,

$$\Pi_n^v = v^{-1}(\Psi_n)$$

is a partition of $S$ in $\Sigma$ such that $P(E) = 1/2^n$ for all $E \in \Pi_n^v$. By setting $\Sigma_n^v = \sigma(\Pi_n^v) = v^{-1}(\sigma(\Psi_n))$ for all $n \in \mathbb{N}$, we have a filtration $\{\Sigma_n^v\}_{n \in \mathbb{N}}$ in $\Sigma$. As usual, $\Sigma_\infty^v = \sigma\left(\bigcup_{n\in\mathbb{N}} \Sigma_n^v\right)$.

**Lemma 5.** *Let $P$ be nonatomic and $p \in [1, \infty]$. For each $v \in \mathcal{U}$,*

$$\Sigma_\infty^v = \sigma(v)$$

*and, for each $f \in \mathcal{L}^p$ (resp. $f \in \mathcal{M}^\infty$),*

$$\mathbb{E}[f_v \mid \Sigma_n^v] \to f_v$$

*almost surely, in $\mathcal{L}^p$ if $p < \infty$, and in bounded a.s. convergence if $p = \infty$ (resp. in $\mathcal{M}^\infty$). In particular, by choosing $v$ such that $f = f_v$ a.s., it follows that*

$$\mathbb{E}[f \mid \Sigma_n^v] \to f$$

*in the above senses. Moreover, for each $v \in \mathcal{U}$ and each $f \in \mathcal{L}^p$,*

$$q^-_{\mathbb{E}(f_v \mid \Sigma_n^v)} = \mathbb{E}_\lambda \left[ q_f^- \mid \sigma(\Psi_n) \right] \qquad \lambda\text{-a.s.}$$

*for all $n \in \mathbb{N}$.*

**Proof.** Note that the $\sigma$-algebra $\sigma\left(\bigcup_{n\in\mathbb{N}} \Psi_n\right)$ is the Borel $\sigma$-algebra $\mathcal{B}(0,1)$ on $(0,1)$ because $\bigcup_{n\in\mathbb{N}} \Psi_n$ is countable and separates the points of $(0,1)$ (see, e.g., Mackey, 1957, Theorem 3.3). Then,

$$\begin{aligned} \sigma(v) = v^{-1}(\mathcal{B}(0,1)) = v^{-1}\left(\sigma\left(\bigcup_{n\in\mathbb{N}} \Psi_n\right)\right) = \sigma\left(v^{-1}\left(\bigcup_{n\in\mathbb{N}} \Psi_n\right)\right) = \sigma\left(\bigcup_{n\in\mathbb{N}} v^{-1}(\Psi_n)\right) \\ = \sigma\left(\bigcup_{n\in\mathbb{N}} \Pi_n^v\right) = \sigma\left(\bigcup_{n\in\mathbb{N}} \Sigma_n^v\right) = \Sigma_\infty^v \end{aligned}$$

By the Martingale Convergence Theorem on $\mathcal{L}^1$ and on $\mathcal{L}^p$, $p \in (1, \infty)$ (see Theorems 4.2.11 and 4.4.6 of Durrett, 2019, respectively),

$$\mathbb{E}[f_v \mid \Sigma_n^v] \to \mathbb{E}[f_v \mid \Sigma_\infty^v]$$

both almost surely, and in $\mathcal{L}^p$ if $p < \infty$ (in particular, if $f \in \mathcal{M}^\infty$, convergence in $\mathcal{M}^\infty$ follows). In case $p = \infty$, we have bounded a.s. convergence because $f$ is a.s. bounded. But $\Sigma_\infty^v = \sigma(v)$ and $f_v$ is $\sigma(v)$-measurable, and so, almost surely

$$f_v = \mathbb{E}[f_v \mid \sigma(v)] = \mathbb{E}[f_v \mid \Sigma_\infty^v]$$

This proves the first part of the statement.

For each $n \in \mathbb{N}$. Define $G_n = \mathbb{E}_\lambda \left[ q_f^- \mid \sigma(\Psi_n) \right]$ on $(0,1)$, and observe that $G_n$ is an increasing function. Moreover, by the change of variable formula (Lemma 6 below),

$$\mathbb{E}_P [f_v \mid \Sigma_n^v] = \mathbb{E}_P \left[ q_f^- \circ v \mid v^{-1} (\sigma(\Psi_n)) \right] = \mathbb{E}_\lambda \left[ q_f^- \mid \sigma(\Psi_n) \right] \circ v = G_n \circ v$$

almost surely. By Lemma A.27 of Föllmer and Schied (2016), we then have, $\lambda$-a.s.,

$$q^-_{\mathbb{E}[f_v | \Sigma_n^v]} = q^-_{G_n \circ v} = G_n \circ q_v^- = G_n$$

as desired. ■

We close with two technical results.

**Lemma 6.** *Let $(X, \Sigma_X, P)$ be a probability space, $(Y, \Sigma_Y)$ be a measurable space, $T : X \to Y$ be a measurable function, $g : Y \to \mathbb{R}$ be a measurable function such that $g \circ T$ is $P$-summable. Then $g$ is $P \circ T^{-1}$-summable and, for every sub-$\sigma$-algebra $\mathcal{A}$ of $\Sigma_Y$,*

$$\mathbb{E}_P \left[ g \circ T \mid T^{-1}(\mathcal{A}) \right] = \mathbb{E}_{P \circ T^{-1}} [g \mid \mathcal{A}] \circ T$$

*Moreover, for all $A \in \Sigma_Y$, it holds $\mathbb{E}_P \left[ g \circ T \mid T^{-1}(A) \right] = \mathbb{E}_{P \circ T^{-1}} [g \mid A]$.*

**Proof.** The proof is standard. ■

**Lemma 7.** *Let $f, g, f', g' \in \mathcal{L}^1$. If $P(f \leq x, g \leq y) \leq P(f' \leq x, g' \leq y)$ for all $x, y \in \mathbb{R}$, then $f + g \geq_{\mathrm{cv}} f' + g'$.*

**Proof.** See, e.g., Theorem 3.8.2 of Müller and Stoyan (2002). ■

### B.2 On equivalent definitions of insurance propensity

The definitions of propensity to full (resp. proportional) insurance that we provide in the introduction are equivalent to those appearing in Section II. Indeed, $f \in \mathcal{I}^{\mathrm{pr}}(w)$ if and only if $f = -(1-\varepsilon) w - \pi$ for some $\pi \in \mathbb{R}$ and some $\varepsilon \in [0,1)$. Thus, propensity to proportional insurance, as defined by point (ii) of Definition 7, requires

$$w - (1-\varepsilon) w - \pi \succsim w + g$$

for all $w \in \mathcal{F}$, $\varepsilon \in [0,1)$, $\pi \in \mathbb{R}$ and $g \overset{d}{=} -(1-\varepsilon) w - \pi$, that is,

$$w - (1-\varepsilon) w - \pi \succsim w + h - \pi$$

for all $w \in \mathcal{F}$, $\varepsilon \in [0,1)$, $\pi \in \mathbb{R}$ and $h \overset{d}{=} -(1-\varepsilon) w$. The latter is the definition of propensity to proportional insurance proposed in the introduction.

The case of full insurance is obtained by considering only the case $\varepsilon = 0$.

### B.3 On mean preserving spreads

In this section, we assume that $\Sigma$ is generated by a partition $\mathcal{S}$ of equiprobable events (called cells), and we fix a risk preference $\succsim$ on $\mathcal{F}$.

**Definition 15.** *Given $f, g \in \mathcal{F}$, we say that $g$ is a* mean preserving spread *of $f$ when there exist $\delta \geq 0$ and two distinct cells $S_1$ and $S_2$ in $\mathcal{S}$, with $f(S_1) \leq f(S_2)$ such that*[13]

$$g = f - \delta 1_{S_1} + \delta 1_{S_2}$$

[13]Clearly, $f$ is constant on cells, so $f(S_i)$ is the constant value of $f$ on $S_i$, for $i = 1, 2$.

**Lemma 8.** *Let $f, g \in \mathcal{F}$ be such that $g$ is a mean preserving spread of $f$ satisfying $g = f - \delta 1_{S_1} + \delta 1_{S_2}$ with $f(S_1) < f(S_2)$ and $\delta > 0$. Then there exist $\tilde{f}, \tilde{g}, \tilde{w} \in \mathcal{F}$ such that*

$$\tilde{f} \stackrel{d}{=} \tilde{g},\ \tilde{f} = \eta \tilde{w} \text{ with } \eta \in (-1, 0)\,, \text{ and } f = \tilde{w} + \tilde{f} \text{ and } g = \tilde{w} + \tilde{g} \tag{9}$$

*in particular $\tilde{f} \in \mathcal{I}^{\mathrm{pr}}(\tilde{w})$, and so $f \succsim g$ if the risk preference $\succsim$ is propense to proportional insurance.*

**Proof.** Denote by $m_i = f(S_i)$, $i = 1, 2$. Let $a = (m_1 - m_2)/\delta - 1 < -1$, and define

$$\tilde{f} = f/(a+1),\ \tilde{g} = \tilde{f} 1_{S \setminus \{S_1, S_2\}} + \tilde{f}(S_1) 1_{S_2} + \tilde{f}(S_2) 1_{S_1},\ \tilde{w} = a\tilde{f}.$$

We aim to show that $\tilde{f}, \tilde{g}, \tilde{w}$ satisfy all conditions in (9). It is straightforward to see $\tilde{f} \stackrel{d}{=} \tilde{g}$. Moreover, it is easy to verify that

$$\tilde{f} = \frac{1}{a}\tilde{w} \quad \text{with} \quad \frac{1}{a} \in (-1, 0)$$

because $a < -1$, and

$$\tilde{w} + \tilde{f} = a\tilde{f} + \tilde{f} = (a+1)\,\frac{f}{a+1} = f$$

and

$$\begin{aligned}
&\tilde{w} + \tilde{g} = a\tilde{f} + \tilde{f} 1_{S \setminus \{S_1, S_2\}} + \tilde{f}(S_1) 1_{S_2} + \tilde{f}(S_2) 1_{S_1} \\
&= a\tilde{f} 1_{S \setminus \{S_1, S_2\}} + a\tilde{f}(S_2) 1_{S_2} + a\tilde{f}(S_1) 1_{S_1} + \tilde{f} 1_{S \setminus \{S_1, S_2\}} + \tilde{f}(S_1) 1_{S_2} + \tilde{f}(S_2) 1_{S_1} \\
&= f 1_{S \setminus \{S_1, S_2\}} + (a\tilde{f}(S_2) + \tilde{f}(S_1)) 1_{S_2} + (a\tilde{f}(S_1) + \tilde{f}(S_2)) 1_{S_1} \\
&= f 1_{S \setminus \{S_1, S_2\}} + \left(a\frac{f(S_2)}{a+1} + \frac{f(S_1)}{a+1}\right) 1_{S_2} + \left(a\frac{f(S_1)}{a+1} + \frac{f(S_2)}{a+1}\right) 1_{S_1} \\
&= f 1_{S \setminus \{S_1, S_2\}} + \underbrace{\left(\left(\frac{m_1 - m_2}{\delta} - 1\right)\frac{m_2}{\frac{m_1 - m_2}{\delta}} + \frac{m_1}{\frac{m_1 - m_2}{\delta}}\right)}_{= m_2 + \delta = f(S_2) + \delta} 1_{S_2} + \underbrace{\left(\left(\frac{m_1 - m_2}{\delta} - 1\right)\frac{m_1}{\frac{m_1 - m_2}{\delta}} + \frac{m_2}{\frac{m_1 - m_2}{\delta}}\right)}_{= m_1 - \delta = f(S_1) - \delta} 1_{S_1} \\
&= g
\end{aligned}$$

as desired. ■

**Lemma 9.** *Let $f, g \in \mathcal{F}$ be such that $g$ is a mean preserving spread of $f$. If the risk preference $\succsim$ is continuous and propense to proportional insurance, then $f \succsim g$.*

**Proof.** Let $f, g \in \mathcal{F}$ be such that $g$ is a mean preserving spread of $f$. Then there exist $\delta \geq 0$ and two distinct cells $S_1$ and $S_2$ in $\mathcal{S}$, with $f(S_1) \leq f(S_2)$ such that

$$g = f - \delta 1_{S_1} + \delta 1_{S_2}$$

If $\delta = 0$, then $f = g$ and reflexivity of $\succsim$ yields $f \succsim g$. If $\delta > 0$ and $f(S_1) < f(S_2)$, the previous lemma yields $f \succsim g$. If $\delta > 0$ and $f(S_1) = f(S_2)$, define $f_\varepsilon = f - \varepsilon 1_{S_1} + \varepsilon 1_{S_2}$ with $\varepsilon \in (0, \delta)$. Note that

$$\begin{aligned}
f_\varepsilon &= f - \varepsilon 1_{S_1} + \varepsilon 1_{S_2} = f 1_{S \setminus \{S_1, S_2\}} + (f(S_1) - \varepsilon)\, 1_{S_1} + (f(S_2) + \varepsilon)\, 1_{S_2} \\
g &= f - \delta 1_{S_1} + \delta 1_{S_2} = f - (\varepsilon + (\delta - \varepsilon))\, 1_{S_1} + (\varepsilon + (\delta - \varepsilon))\, 1_{S_2} = f_\varepsilon - (\delta - \varepsilon)\, 1_{S_1} + (\delta - \varepsilon)\, 1_{S_2}
\end{aligned}$$

Thus $g$ is a mean preserving spread of $f_\varepsilon$ with $f_\varepsilon(S_1) < f_\varepsilon(S_2)$ and $\delta - \varepsilon > 0$. By the previous argument $f_\varepsilon \succsim g$ for all $\epsilon \in (0, \delta)$. Letting $\varepsilon_n = \delta/2^n \to 0$, we have $f_{\varepsilon_n} \to f$, and continuity implies $f \succsim g$. ■

**Lemma 10.** *Let $f, g \in \mathcal{F}$ be such that $g$ is a mean preserving spread of $f$. Then there exist $\tilde{f}, \tilde{g}, \tilde{w} \in \mathcal{F}$ such that*

$$\tilde{f} \stackrel{d}{=} \tilde{g},\ \tilde{f} \in \mathcal{I}^{\mathrm{dl}}(\tilde{w}),\ f = \tilde{w} + \tilde{f} \text{ and } g = \tilde{w} + \tilde{g}$$

*and so $f \succsim g$ if the risk preference $\succsim$ is propense to deductible-limit insurance.*

**Proof.** For a mean preserving spread $g$ of $f$, we can write

$$g = f - 2\delta 1_{S_1} + 2\delta 1_{S_2}$$

where $\delta \geq 0$ and $f(S_1) \leq f(S_2)$. Define the events

$$E_1 = \{f \leq f(S_1)\} \setminus S_1 \qquad E_2 = \{f(S_1) < f < f(S_2)\} \qquad E_3 = \{f \geq f(S_2)\} \setminus S_2$$

The events $S_1$, $S_2$, $E_1$, $E_2$, and $E_3$ form a measurable partition of $S$. Define $\tilde{f}, \tilde{g}, \tilde{w}$ by the following table:

| | $E_1$ | $S_1$ | $E_2$ | $S_2$ | $E_3$ |
|---|---|---|---|---|---|
| $\tilde{f}$ | $\delta$ | $\delta$ | $\delta$ | $-\delta$ | $-\delta$ |
| $\tilde{g}$ | $\delta$ | $-\delta$ | $\delta$ | $\delta$ | $-\delta$ |
| $\tilde{w}$ | $f-\delta$ | $f-\delta$ | $f-\delta$ | $f+\delta$ | $f+\delta$ |

Write $\xi = -f(S_2) - \delta$. One can check $\tilde{f} \overset{d}{=} \tilde{g}$ and

$$\tilde{f} = (-\tilde{w} - \xi)^+ \wedge (2\delta) - \delta$$

in fact

- if $s \in E_1 \cup S_1 \cup E_2$, then $f(s) \leq f(S_2)$, and

  $$-\tilde{w}(s) - \xi = -f(s) + \delta + f(S_2) + \delta = f(S_2) - f(s) + 2\delta \geq 2\delta \geq 0$$

  so

  $$(-\tilde{w}(s) - \xi)^+ = f(S_2) - f(s) + 2\delta \geq 2\delta$$

  and

  $$(-\tilde{w}(s) - \xi)^+ \wedge (2\delta) - \delta = 2\delta - \delta = \delta = \tilde{f}(s)$$

- else $s \in S_2 \cup E_3$, then $f(s) \geq f(S_2)$, and

  $$-\tilde{w}(s) - \xi = -f(s) - \delta + f(S_2) + \delta = f(S_2) - f(s) \leq 0$$

  so

  $$(-\tilde{w}(s) - \xi)^+ = 0$$

  and

  $$(-\tilde{w}(s) - \xi)^+ \wedge (2\delta) - \delta = -\delta = \tilde{f}(s)$$

This implies $\tilde{f} \in \mathcal{I}^{\mathrm{dl}}(\tilde{w})$. On the other hand, it is easy to see $\tilde{w} + \tilde{f} = f$ and $\tilde{w} + \tilde{g} = g$, as wanted. ■

**Lemma 11.** *If the risk preference $\succsim$ is continuous, and propense to either proportional or deductible-limit insurance, then*

$$f \geq_{\mathrm{cv}} g \implies f \succsim g$$

**Proof.** If $f \geq_{\mathrm{cv}} g$ in $\mathcal{F}$, then there exists a sequence $h_0, h_1, \ldots, h_m$ such that $f = h_0$, $g = h_m$ and each $h_{k+1}$ is either a mean preserving spread of $h_k$ or it is obtained by $h_k$ through the permutation of the values that $h_k$ takes on two cells. In the first case, $h_k \succsim h_{k+1}$ by what we just proved. In the second, $h_k \sim h_{k+1}$ because $\succsim$ is law invariant. By the transitivity of $\succsim$, we conclude that $f \succsim g$. ■

## B.4 A deus ex machina

In what follows, for $f \in L^\infty$, let $u_f$ be the essential supremum of $f$ and $\ell_f$ be the essential infimum of $f$, defined by $u_f = \inf\{x \in \mathbb{R} : P(f \leq x) = 1\}$ and $\ell_f = \sup\{x \in \mathbb{R} : P(f \geq x) = 1\}$.

**Theorem 6.** *Let $k \geq 1$ and $f \in L^k$. Then $\mathbb{E}[f] = 0$ if and only if there exist $g, g' \in L^{k-1}$ such that $g \stackrel{d}{=} g'$ and $g - g' \stackrel{d}{=} f$. Moreover,*

*(i) if $f \in L^\infty$, it is possible to choose $g, g' \in L^\infty$ so that $\ell_f \leq g, g' \leq u_f$;*

*(ii) if $f \in M^\infty$, it is possible to choose $g, g' \in M^\infty$;*

*(iii) if the probability space is finite, it is possible to choose $g$ and $g'$ so that $g - g' = f$.*

To prove Theorem 6, we first note that the "if" direction can be verified in a straightforward manner. Suppose that $f \stackrel{d}{=} g - g'$ for some $g \stackrel{d}{=} g'$. If $g, g' \in L^1$ then it is obvious that $\mathbb{E}[g - g'] = 0$. In general, Simons (1977) showed that $\mathbb{E}[g - g'] = 0$ even if $g, g'$ are not in $L^1$, as long as the mean $\mathbb{E}[g - g']$ is well defined, justified by $f \in L^k$. Therefore, $\mathbb{E}[f] = \mathbb{E}[g - g'] = 0$.

Next, we focus on the more important "only if" direction of Theorem 6. For this, we first prove the case of $L^\infty$, and then the case of $L^k$, which is much more technically involved.

**Proof of Theorem 6 on finite spaces.** We begin with a finite state space $S = \{1, \ldots, n\}$ of equiprobable states. Let $f : S \to \mathbb{R}$ have mean 0, and set $x_i = f(i)$ for each $i = 1, \ldots, n$. If $f = 0$, there is nothing to prove. Otherwise choose $j_1 \in \{1, \ldots, n\}$ such that $x_{j_1} > 0$. Now

$$\min\{x_1, \ldots, x_n\} \leq \sum_{i=1}^{1} x_{j_i} \leq \max\{x_1, \ldots, x_n\}$$

Assume for some $1 \leq k < n$ to have found distinct $j_1, j_2, \ldots, j_k \in \{1, \ldots, n\}$ such that

$$\min\{x_1, \ldots, x_n\} \leq \sum_{i=1}^{m} x_{j_i} \leq \max\{x_1, \ldots, x_n\} \qquad \forall m = 1, \ldots, k$$

We next show that there is $j_{k+1} \in J_{k+1} := \{1, \ldots, n\} \setminus \{j_1, \ldots, j_k\}$ such that

$$\min\{x_1, \ldots, x_n\} \leq \sum_{i=1}^{m} x_{j_i} \leq \max\{x_1, \ldots, x_n\} \qquad \forall m = 1, \ldots, k, k+1$$

1. If $x_j = 0$ for some $j \in J_{k+1}$, set $j_{k+1} = j$.
2. If $\sum_{i=1}^{k} x_{j_i} = 0$, arbitrarily choose $j_{k+1} \in J_{k+1}$.
3. Else $x_j \neq 0$ for all $j \in J_{k+1}$ and $\sum_{i=1}^{k} x_{j_i} \neq 0$;

   (a) if $\sum_{i=1}^{k} x_{j_i} > 0$, it cannot be the case that $x_j \geq 0$ for all elements of $J_{k+1} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_k\}$, otherwise we would have

   $$0 < \sum_{i=1}^{k} x_{j_i} \leq \sum_{i=1}^{k} x_{j_i} + \sum_{j \in J_{k+1}} x_j = \sum_{j=1}^{n} x_j = 0$$

   then it is possible to choose $j_{k+1} \in J_{k+1}$ such that $x_{j_{k+1}} < 0$, and

   $$\min\{x_1, \ldots, x_n\} \leq x_{j_{k+1}} < \sum_{i=1}^{k} x_{j_i} + x_{j_{k+1}} < \sum_{i=1}^{k} x_{j_i} \leq \max\{x_1, \ldots, x_n\}$$

(b) else $\sum_{i=1}^{k} x_{j_i} < 0$, it cannot be the case that $x_j \leq 0$ for all elements of $J_{k+1} = \{1, \ldots, n\} \setminus \{j_1, \ldots, j_k\}$, otherwise we would have

$$0 > \sum_{i=1}^{k} x_{j_i} \geq \sum_{i=1}^{k} x_{j_i} + \sum_{j \in J_{k+1}} x_j = \sum_{j=1}^{n} x_j = 0$$

then it is possible to choose $j_{k+1} \in J_{k+1}$ such that $x_{j_{k+1}} > 0$, and

$$\min\{x_1, \ldots, x_n\} \leq \sum_{i=1}^{k} x_{j_i} < \sum_{i=1}^{k} x_{j_i} + x_{j_{k+1}} < x_{j_{k+1}} \leq \max\{x_1, \ldots, x_n\}$$

In exactly $n$ steps this produces a rearrangement $(x_{j_1}, \ldots, x_{j_n})$ of $(x_1, \ldots, x_n)$, which by construction satisfies

$$\min\{x_1, \ldots, x_n\} \leq \sum_{i=1}^{m} x_{j_i} \leq \max\{x_1, \ldots, x_n\} \qquad \forall m = 1, \ldots, n \tag{10}$$

Define $g, g' : S \to \mathbb{R}$ by $g(j_k) = \sum_{i=1}^{k} x_{j_i}$ and $g'(j_k) = \sum_{i=1}^{k-1} x_{j_i}$ for each $1 \leq k \leq n$, with the convention $g'(j_1) = 0$. Diagram $g$ and $g'$ as follows:

| | $j_1$ | $j_2$ | $\cdots$ | $j_{n-1}$ | $j_n$ |
|---|---|---|---|---|---|
| $g$ | $x_{j_1}$ | $x_{j_1} + x_{j_2}$ | $\cdots$ | $\sum_{i=1}^{n-1} x_{j_i}$ | $0 = \sum_{i=1}^{n} x_{j_i}$ |
| $g'$ | $0$ | $x_{j_1}$ | $\cdots$ | $\sum_{i=1}^{n-2} x_{j_i}$ | $\sum_{i=1}^{n-1} x_{j_i}$ |

We have that $g(j_i) - g'(j_i) = f(j_i)$ for all $i = 1, \ldots, n$ and hence

$$f = g - g'$$

and since states are equally probable, $g \overset{d}{=} g'$. In view of (10), we conclude that

$$f = g - g' \quad ; \quad g \overset{d}{=} g' \quad \text{and} \quad \min_S f \leq g, g' \leq \max_S f \tag{11}$$

This proves the statement for a finite state space.

**Proof of Theorem 6 on $L^\infty$.** Now, let $S$ be an infinite state space. Let $f \in L^\infty$. Choose $v \in \mathcal{U}$ such that $f = f_v$, by Lemma 5,

$$f_n := \mathbb{E}[f \mid \Sigma_n^v] \to f$$

both almost surely and in $L^1$. Moreover, for all $n \in \mathbb{N}$,

$$\ell_f \leq f \leq u_f$$

implies

$$\ell_f = \mathbb{E}[\ell_f \mid \Sigma_n^v] \leq f_n \leq \mathbb{E}[u_f \mid \Sigma_n^v] = u_f$$

In view of (11), by choosing the standard versions of the $f_n$, given by

$$f_n(s) = \frac{1}{2^n} \int_E f \mathrm{d}P \qquad \forall s \in E \in \Pi_n^v$$

there exist two sequences $\{g_n\}$ and $\{g_n'\}$ such that, for each $n \in \mathbb{N}$,

$$g_n \overset{d}{=} g_n' \quad , \quad \ell_f \leq g_n, g_n' \leq u_f \quad \text{and} \quad g_n - g_n' = f_n$$

Since $f_n \in L^\infty$, we have $g_n, g_n' \in L^\infty$ for all $n \in \mathbb{N}$. Moreover, by the almost sure convergence of $f_n$ to $f$, it follows that

$$g_n - g_n' \overset{d}{\to} f \tag{12}$$

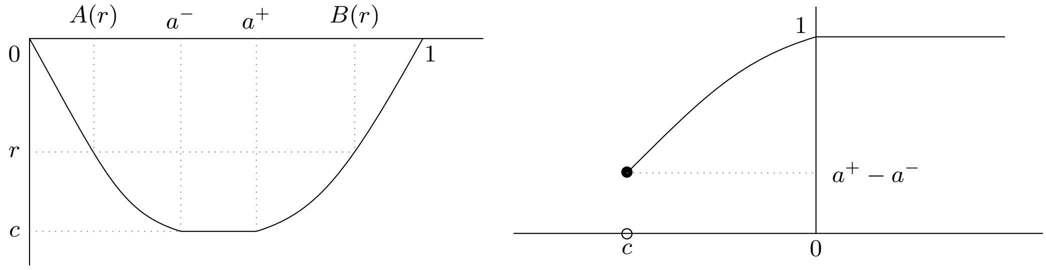


Figure 3: The functions $H$ (left panel) and $K$ (right panel).

Denote by $\mu_n$ the joint distribution of $(g_n, g'_n)$. The sequence $\{\mu_n\}$ is tight since is supported in the compact square

$$C = [\ell_f, u_f] \times [\ell_f, u_f]$$

of $\mathbb{R}^2$. By Prohorov's Theorem, there exists a subsequence $\{\mu_{n_k}\}$ that converges weakly to a probability measure $\mu$ on $\mathbb{R}^2$ with support in $C$. As $P$ is nonatomic, by a version of Skorokhod's Theorem there exists a random vector $(g, g') : S \to \mathbb{R}^2$ with joint distribution $\mu$.[14] By the Continuous Mapping Theorem,[15]

$$g_{n_k} \xrightarrow{d} g \quad , \quad g'_{n_k} \xrightarrow{d} g' \quad \text{and} \quad g_{n_k} - g'_{n_k} \xrightarrow{d} g - g'$$

Since $g_{n_k} \stackrel{d}{=} g'_{n_k}$ for all $k \geq 1$, we have $g \stackrel{d}{=} g'$. By (12), we also have $g_{n_k} - g'_{n_k} \xrightarrow{d} f$ and so $g - g' \stackrel{d}{=} f$. Note that $\ell_f \leq g, g' \leq u_f$ since $\mu$ is supported in $C$. ■

**Preparation for the proof on $L^k$.** We first present some preliminaries. For $f \in L^1$, denote by $\mu_f = P \circ f^{-1}$. Recall that the left quantile $q_f^- : (0,1) \to \mathbb{R}$ is defined as $q_f^-(p) = \inf\{x \in \mathbb{R} : P(f \leq x) \geq p\}$.

Let $f \in L^1$ with $\mathbb{E}[f] = 0$. We assume that $f$ is not constantly 0. Define

$$H(t) = \int_0^t q_f^- \mathrm{d}\lambda \qquad \forall t \in [0,1]$$

and denote by $a^- = \mu_f((-\infty, 0))$ and $a^+ = \mu_f((-\infty, 0])$. It is easy to see that $H$ is strictly decreasing on $[0, a^-]$ and strictly increasing on $[a^+, 1]$, $H(0) = H(1) = 0$, and the minimum value of $H$ is given by $c := H(a^-) = H(a^+) < 0$, which is attained by any point in $[a^-, a^+]$. Moreover, $H$ is convex because $q_f$ is increasing, and hence $H$ is almost everywhere differentiable on $[0,1]$. For $r \in [c, 0]$, define

$$A(r) = \inf\{t \in [0,1] : H(t) = r\} \quad \text{and} \quad B(r) = \sup\{t \in [0,1] : H(t) = r\} \tag{13}$$

Obviously, $A(r) \in [0, a^-]$, $B(r) \in [a^+, 1]$, $A(c) = a^-$, $B(c) = a^+$ and $H \circ A(r) = H \circ B(r) = r$. Moreover, $A(r)$ and $B(r)$ are also both continuous and strictly monotone as $H$ is so on $[0, a^-]$ and $[a^+, 1]$. Define

$$K(r) = \begin{cases} 1 & r > 0 \\ B(r) - A(r) & c \leq r \leq 0 \\ 0 & r < c \end{cases} \tag{14}$$

One can check that $K$ is right-continuous and increasing, with $\lim_{x \uparrow c} K(x) = 0$, $\lim_{x \downarrow c} K(x) = K(c) = a^+ - a^-$ and $K(0) = 1$. Hence, $K$ is a distribution function on $[c, 0]$. More precisely, $K$ is continuous and strictly increasing on $(c, 0]$ and has probability mass $a^+ - a^-$ at $c$. The functions $H$ and $K$ are plotted in Figure 3.

Define the function $\Phi$ on $[c, 0)$ by

$$\Phi = q_f^- \circ B - q_f^- \circ A \tag{15}$$

[14]See, e.g., Theorem 3.1 of Berti, Pratelli, and Rigo (2007).
[15]See, e.g., Theorem 4.27 of Kallenberg (2002).

It is easy to see that $\Phi$ is increasing. Note that $q_f^- \circ A(r) \le 0 \le q_f^- \circ B(r)$ for $r \in [c,0]$ with strict inequalities on $(c,0)$. It holds that $\Phi \ge \max\{|q_f^- \circ B|, |q_f^- \circ A|\}$. The functions $A, B, H, K, \Phi$ have been studied by Wang and Wang (2015) in a different context. An important technical tool that we will use is Lemma 2.4 of Wang and Wang (2015), which says that $\Phi(w)$ where $w \stackrel{d}{\sim} K$ is in $L^{k-1}$ (the distribution of $\Phi(w)$ is denoted by $\tilde{F}$ in that paper). That is,

$$\int_c^0 (\Phi(r))^{k-1} \mathrm{d}K(r) < \infty \tag{16}$$

**Proof of Theorem 6 on $L^k$.** Consider any $f \in L^k$ with $k \ge 1$ which satisfies $\mathbb{E}[f] = 0$. The case of $f = 0$ is trivial. We assume that $f$ is not constantly 0. Choose $v \in \mathcal{U}$ such that $f = f_v$. Recall the functions $A$, $B$ defined by (13) and $K$ by (14). We have that $K = B - A$ is continuous and strictly increasing on $(c, 0]$ and has probability mass $a^+ - a^-$ at $c$. Let $r_1 = c$, and define

$$r_n = \inf\left\{ r > r_{n-1} : K(r) - K(r_{n-1}) = \frac{1-(a^+ - a^-)}{2^{n-1}} \right\} \qquad \forall n \ge 2$$

It is easy to see that the sequence $\{r_n\}$ is contained in $[c,0)$, and increasing with $K(r_n) - K(r_{n-1}) = (1 - a^+ + a^-)/2^{n-1}$ for $n \ge 2$. Moreover, $r_n \to 0$ because $K$ is strictly increasing on $[c,0]$ with $K(0) = 1$, and

$$K(r_n) = K(r_1) + \sum_{i=2}^{n} K(r_i) - K(r_{i-1}) = 1 - \frac{1 - a^+ + a^-}{2^{n-1}} \to 1$$

Denote by $T_n = [A(r_n), B(r_n)]$. Further, write $T_0 = \emptyset$, $T_\infty = \lim_{n\to\infty} T_n = (0,1)$. For each $n \in \mathbb{N}$, define $\mu_n$ by

$$\mu_n(D) = P(f \in D \mid v \in T_n \setminus T_{n-1}) \qquad \forall D \in \mathcal{B}(\mathbb{R})$$

Note that $\{r_n\}_{n\in\mathbb{N}} \subseteq [c,0)$, and $A$ and $B$ are both strictly monotone on $[c,0]$ satisfying $A \le B$ and $A(0) = 1 - B(0) = 0$. It holds that $A(r_n), B(r_n) \in (0,1)$ for all $n \in \mathbb{N}$. Hence, $\mu_n$ is a compactly supported Borel probability measure. Below we will show $\int_{\mathbb{R}} x \mathrm{d}\mu_n(x) = 0$ for $n \in \mathbb{N}$, and

$$\mu_f = \sum_{n\in\mathbb{N}} (K(r_n) - K(r_{n-1}))\mu_n \tag{17}$$

where $r_0$ is any number in $(-\infty, c)$ so that $K(r_0) = 0$. To show the claim, using Lemma 6 and denoting by $m = K(r_n) - K(r_{n-1}) > 0$, we have

$$\begin{aligned}
m \int_{\mathbb{R}} x \mathrm{d}\mu_n(x) &= m \int_{\mathbb{R}} x \mathrm{d}\left(P_{\{v\in T_n\setminus T_{n-1}\}} \circ f_v^{-1}\right)(x) = m \int f_v \mathrm{d}P_{\{v\in T_n\setminus T_{n-1}\}} \\
&= m \int q_f^- \mathrm{d}\left(P \circ v^{-1}\right)_{T_n\setminus T_{n-1}} = m \int q_f^- \mathrm{d}\lambda_{T_n\setminus T_{n-1}} \\
&= \int_{T_n} q_f^- \mathrm{d}\lambda - \int_{T_{n-1}} q_f^- \mathrm{d}\lambda \\
&= \int_0^{B(r_n)} q_f^- \mathrm{d}\lambda - \int_0^{A(r_n)} q_f^- \mathrm{d}\lambda - \int_0^{B(r_{n-1})} q_f^- \mathrm{d}\lambda + \int_0^{A(r_{n-1})} q_f^- \mathrm{d}\lambda \\
&= H\circ B(r_n) - H \circ A(r_n) - H \circ B(r_{n-1}) + H \circ A(r_{n-1}) = 0
\end{aligned}$$

where the last step follows from $H \circ A(r) = H \circ B(r)$ for all $r \in [c,0]$. This implies that $\int_{\mathbb{R}} x \mathrm{d}\mu_n(x) = 0$ for $n \in \mathbb{N}$. To see (17), note that $P(v \in T_n \setminus T_{n-1}) = \lambda(T_n \setminus T_{n-1}) = \lambda(T_n) - \lambda(T_{n-1}) = K(r_n) - K(r_{n-1})$. Hence,

$$\begin{aligned}
\sum_{n\in\mathbb{N}} (K(r_n) - K(r_{n-1}))\mu_n(D) &= P\left( f \in D,\ v \in \bigcup_{n\in\mathbb{N}} (T_n \setminus T_{n-1}) \right) \\
&= P\left(f \in D,\ v \in T_\infty \setminus T_0\right) = \mu_f(D)
\end{aligned}$$

Therefore, we have verified (17). Take independent random variables $v : S \to (0,1)$ and $w : S \to \mathbb{N}$ with $v \in \mathcal{U}$ and $w$ such that $P(w = n) = K(r_n) - K(r_{n-1})$ for all $n \in \mathbb{N}$. By the construction of $\{r_n\}$, we have $2P(w = n+1) = P(w = n)$ for all $n \geq 2$. Using the result of Theorem 6 on bounded random variables, on the space $(S, \sigma(v), P)$ there exist identically distributed random variables $g_n$ and $g_n'$ such that $g_n - g_n' \overset{d}{\sim} \mu_n$ for each $n \in \mathbb{N}$. Moreover, $q_f^- \circ A(r_n) \leq g_n, g_n' \leq q_f^- \circ B(r_n)$ as the support of $\mu_n$ is contained in $[q_f^- \circ A(r_n), q_f^- \circ B(r_n)]$. Define the random variables $g$ and $g'$ by

$$g(s) = g_{w(s)}(s) \quad \text{and} \quad g'(s) = g'_{w(s)}(s) \qquad \forall s \in S$$

First observe that, for all $D \in \mathcal{B}(\mathbb{R})$,

$$\{g \in D\} = \bigcup_{n\in\mathbb{N}} \{g \in D,\ w = n\} = \bigcup_{n\in\mathbb{N}} \{g_n \in D,\ w = n\}$$

This shows that $g$ is measurable. Moreover, since $g_n$ and $w$ are independent we have that, for all $D \in \mathcal{B}(\mathbb{R})$,

$$\mu_g(D) = \sum_{n\in\mathbb{N}} P(w = n)\, \mu_{g_n}(D) = \sum_{n\in\mathbb{N}} (K(r_n) - K(r_{n-1}))\mu_{g_n}(D)$$

The same argument for $g'$ and the fact that $g_n \overset{d}{=} g_n'$ for all $n \in \mathbb{N}$ show that $g \overset{d}{=} g'$; the same argument for $g - g'$ and the fact that $g_n - g_n' \overset{d}{\sim} \mu_n$ for each $n \in \mathbb{N}$ combining with (17) yield $g - g' \overset{d}{\sim} \mu_f$. It remains to verify that the constructed $g$ is in $L^{k-1}$. Recall the definition of $\Phi = q_f^- \circ B - q_f^- \circ A$ in (15). We have $|g_n| \leq \Phi(r_n)$ because $q_f^- \circ A(r_n) \leq g_n \leq q_f^- \circ B(r_n)$ and $q_f^- \circ A \leq 0 \leq q_f^- \circ B$. Using (16), we obtain

$$\begin{aligned}
\infty > \int_c^0 (\Phi(r))^{k-1} \mathrm{d}K(r) &\geq \sum_{n\in\mathbb{N}} \int_{(r_n, r_{n+1}]} (\Phi(r))^{k-1} \mathrm{d}K(r) \\
&\geq \sum_{n\in\mathbb{N}} P(w = n+1)\, (\Phi(r_n))^{k-1} \geq \frac{1}{2} \sum_{n=2}^{\infty} P(w = n)\, (\Phi(r_n))^{k-1} \\
&\geq \frac{1}{2} \sum_{n=2}^{\infty} P(w = n)\, \mathbb{E}\left[|g_n|^{k-1}\right] = \frac{1}{2} \left( \mathbb{E}\left[|g|^{k-1}\right] - K(r_1)\mathbb{E}\left[|g_1|^{k-1}\right] \right)
\end{aligned}$$

Noting that $\mathbb{E}\left[|g_1|^{k-1}\right] < \infty$ as $g_1$ is bounded, we have $\mathbb{E}\left[|g|^{k-1}\right] < \infty$. This completes the proof of the necessity statement. ■

### B.5 Proof of Lemma 1

It is a direct consequence of Theorem 6, which we proved above. ■

### B.6 Proof of Theorem 1

(i) $\Longrightarrow$ (ii). Let $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$, if $f \in \mathcal{I}^{\text{fi}}(w)$, then $f = -w - \pi$ for some $\pi \in \mathbb{R}$, then

$$w + f = -\pi = \mathbb{E}[w] + \mathbb{E}[-w - \pi] = \mathbb{E}[w] + \mathbb{E}[g] = \mathbb{E}[w + g] \succsim w + g$$

where the third equality follows from $g \overset{d}{=} -w - \pi$, and the final preference follows from weak risk aversion. Thus propensity to full insurance holds.

(ii) $\Longrightarrow$ (i). For each $h \in \mathcal{F}$, by Lemma 1, there exist $w, w' \in \mathcal{F}$ such that $w \overset{d}{=} w'$ and $h - \mathbb{E}[h] \overset{d}{=} w - w'$. Let $f = -w + \mathbb{E}[h]$ and $g = -w' + \mathbb{E}[h]$, clearly $f \overset{d}{=} g$ and $f \in \mathcal{I}^{\text{fi}}(w)$. Propensity to full insurance implies that $w + f \succsim w + g$, which gives $\mathbb{E}[h] = w + f \succsim w + g = w - w' + \mathbb{E}[h] \overset{d}{=} h$. Law invariance of $\succsim$ yields $w - w' + \mathbb{E}[h] \sim h$, implying $\mathbb{E}[h] \succsim w - w' + \mathbb{E}[h] \sim h$ and transitivity implies $\mathbb{E}[h] \succsim h$. Thus weak risk aversion holds. ■

## B.7 On the relation between Lemma 1 and Theorem 1

We have just proved Theorem 1 by means of Lemma 1. Here we show how, if Theorem 1 could be proved *without* relying on Lemma 1, the lemma would actually result as a corollary of the theorem.

(i) $\implies$ (ii).[16] Consider, for each $c \in \mathbb{R}$, the set

$$\mathcal{G}_c = \{f \in \mathcal{F} : f \overset{d}{=} c + h - h' \text{ for some } h, h' \in \mathcal{F} \text{ with } h \overset{d}{=} h'\} \subseteq \{f \in \mathcal{F} : \mathbb{E}[f] = c\}$$

Now define a relation $\sim$ on $\mathcal{F}$ by

$$f \sim g \iff \text{either } f \overset{d}{=} g \text{ or } f, g \in \mathcal{G}_c \text{ for some } c \in \mathbb{R}$$

Clearly $\sim$ is law invariant (and symmetric).

Before proving transitivity note that

$$f, g \in \mathcal{G}_c \implies \mathbb{E}[f] = c = \mathbb{E}[g]$$

Now let $f_1 \sim f_2$ and $f_2 \sim f_3$, in order to prove $f_1 \sim f_3$, we consider the following four cases.

- If $f_1 \overset{d}{=} f_2$ and $f_2 \overset{d}{=} f_3$, then $f_1 \overset{d}{=} f_3$, and so $f_1 \sim f_3$.
- If $f_1 \overset{d}{=} f_2$ and [not $f_2 \overset{d}{=} f_3$], then there exists $c \in \mathbb{R}$ such that $f_2, f_3 \in \mathcal{G}_c$, that is, $f_2 \overset{d}{=} c + h_2 - h_2'$ for some $h_2, h_2' \in \mathcal{F}$ with $h_2 \overset{d}{=} h_2'$, and $f_3 \overset{d}{=} c + h_3 - h_3'$ for some $h_3, h_3' \in \mathcal{F}$ with $h_3 \overset{d}{=} h_3'$. But
$$f_1 \overset{d}{=} f_2 \overset{d}{=} c + h_2 - h_2'$$
thus $f_1 \in \mathcal{G}_c$, and so $f_1, f_3 \in \mathcal{G}_c$, which implies $f_1 \sim f_3$.
- If [not $f_1 \overset{d}{=} f_2$] and $f_2 \overset{d}{=} f_3$, the conclusion $f_1 \sim f_3$ is obtained as in the previous case.
- If [not $f_1 \overset{d}{=} f_2$] and [not $f_2 \overset{d}{=} f_3$], then there exist $c_{12}, c_{23} \in \mathbb{R}$ such that $f_1, f_2 \in \mathcal{G}_{c_{12}}$ and $f_2, f_3 \in \mathcal{G}_{c_{23}}$, but this implies $\mathbb{E}[f_2] = c_{12}$ and $\mathbb{E}[f_2] = c_{23}$. Therefore, $c_{12} = c_{23} = c$, and $f_1, f_3 \in \mathcal{G}_c$ implies $f_1 \sim f_3$.

Summing up, $\sim$ is a risk preference (indeed a law invariant equivalence relation) on $\mathcal{F}$. Next we show that $\sim$ is propense to full insurance. Take any $w, f, g \in \mathcal{F}$ such that $g \overset{d}{=} f$. If $f$ is a full insurance for $w$, then $f = -w - \pi$ for some $\pi \in \mathbb{R}$. It follows that:

- $w + f \in \mathcal{G}_{-\pi}$, because $w + f = -\pi = -\pi + 0 - 0$ with $0 \in \mathcal{F}$ and $0 \overset{d}{=} 0$;
- $w + g \in \mathcal{G}_{-\pi}$, because $w + g = -f - \pi + g = -\pi + g - f$ with $g, f \in \mathcal{F}$ and $g \overset{d}{=} f$;

therefore (by definition of $\sim$) $w + f \sim w + g$. By Theorem 1 (that we are *assuming to be true*), $\sim$ is weakly risk averse, that is, $\mathbb{E}[f] \sim f$ for all $f \in \mathcal{F}$.

We use the latter fact to show that, given any $f \in \mathcal{F}$, if $\mathbb{E}[f] = 0$, then $f \in \mathcal{G}_0$, that is, (i) $\implies$ (ii). If $\mathbb{E}[f] = 0$, since $f \sim \mathbb{E}[f]$, then $f \sim 0$.

- If $f$ is almost surely constant, then $\mathbb{E}[f] = 0$ implies that $f = 0$ almost surely, and so $f \overset{d}{=} 0 = 0 + 0 - 0$ with $0 \in \mathcal{F}$ and $0 \overset{d}{=} 0$, thus $f \in \mathcal{G}_0$.
- Else $f$ is not almost surely constant, and so it cannot be the case that $f \overset{d}{=} 0$. Then $f \sim 0$ implies that there exists $c \in \mathbb{R}$ such that $f, 0 \in \mathcal{G}_c$, but as observed, it must then be the case that $c = \mathbb{E}[f] = 0$, then $f \in \mathcal{G}_0$.

(ii) $\implies$ (i) of Lemma 1 is trivial. ■

[16] Of Lemma 1, assuming Theorem 1 to be true.

### B.8 Proof of Theorem 2

It is a direct consequence of Theorem 3, which we prove below. ■

### B.9 Proof of Proposition 1

The $\subseteq$ inclusion. As observed, $\mathcal{I}^{\mathrm{cs}}(w)$ is the set of all elements of $\mathcal{F}$ that are counter-monotonic with $w$, that is, such that

$$\left[f(s)-f\left(s'\right)\right]\left[w(s)-w\left(s'\right)\right]\leq 0$$

$P\times P$ almost surely. Thus, by Theorem 2.14 of Rüschendorf (2013),[17] if $f\in\mathcal{I}^{\mathrm{cs}}(w)$, then

$$F_{f,w}\leq G$$

for all joint distributions with marginals $F_f$ and $F_w$. In particular, if $g\overset{d}{=}f$, then $F_{f,w}\leq F_{g,w}$ which is equivalent to $f\trianglerighteq_w g$.

The $\supseteq$ inclusion. Assume that $f\trianglerighteq_w g$ for all $g\overset{d}{=}f$, that is, $F_{f,w}\leq F_{g,w}$ for all $g\overset{d}{=}f$. We want to show that $f$ is counter-monotonic with $w$. By Theorem 3.1 of Puccetti and Wang (2015), it suffices to show that

$$F_{f,w}(x,y)\leq\left(F_f(x)+F_w(y)-1\right)^+\qquad\forall(x,y)\in\mathbb{R}^2$$

since the opposite inequality is true for all joint distributions with marginals $F_f$ and $F_w$.

Let $g\in\mathcal{F}$ be such that $g\overset{d}{=}f$ and $g$ is counter-monotonic with $w$. If $\Sigma$ is generated by a finite partition of equiprobable cells, then such a $g$ can be constructed by rearranging the values of $f$ over the cells. Else, we can take $v\in\mathcal{U}$ such that a.s. $w=w_v=F_w^{-1}(v)$ and define $g=F_f^{-1}(1-v)$, now $g\overset{d}{=}f$ because $1-v\in\mathcal{U}$, and it is counter-monotonic with $w$ because

$$(w,g)=\left(F_w^{-1}(v),F_f^{-1}(1-v)\right)$$

$P\times P$ almost surely.

With this, for all $x,y\in\mathbb{R}^2$,

$$F_{f,w}(x,y)\leq F_{g,w}(x,y)=\left(F_g(x)+F_w(y)-1\right)^+=\left(F_f(x)+F_w(y)-1\right)^+$$

where the first equality follows from Theorem 3.1 of Puccetti and Wang (2015) and the counter-monotonicity of $g$ and $w$, the second from the fact that $g\overset{d}{=}f$. ■

### B.10 Proof of Theorem 3

(i) $\Longrightarrow$ (vi). Let $w,f,g\in\mathcal{F}$ with $g\overset{d}{=}f$ and $f\trianglerighteq_w g$. By Lemma 7, $w+f\geq_{\mathrm{cv}}w+g$, and strong risk aversion implies $w+f\succsim w+g$. Thus $\succsim$ is propense to hedging.

(vi) $\Longrightarrow$ (v). Let $w,f,g\in\mathcal{F}$ with $g\overset{d}{=}f$ and $f\in\mathcal{I}^{\mathrm{cs}}(w)$, by Proposition 1, it follows that $f\trianglerighteq_w g$, and propensity to hedging implies $w+f\succsim w+g$. Thus $\succsim$ is propense to contingency-schedule insurance.

(v) $\Longrightarrow$ (iv) because $\mathcal{I}^{\mathrm{is}}(w)\subseteq\mathcal{I}^{\mathrm{cs}}(w)$ for all $w\in\mathcal{F}$.

(iv) $\Longrightarrow$ (iii) and (iv) $\Longrightarrow$ (ii) because $\mathcal{I}^{\mathrm{dl}}(w),\mathcal{I}^{\mathrm{pr}}(w)\subseteq\mathcal{I}^{\mathrm{is}}(w)$ for all $w\in\mathcal{F}$.

(iii) $\Longrightarrow$ (i) and (ii) $\Longrightarrow$ (i). The case in which $\Sigma$ is generated by a finite partition of equiprobable events follows from Lemma 11. Now, let $P$ be nonatomic. Let $f,g\in\mathcal{F}$ be such that $f\geq_{\mathrm{cv}}g$. We want to show that $f\succsim g$. Let $v\in\mathcal{U}$. By Lemma 4-(i), $f_v\overset{d}{=}f$ and $g_v\overset{d}{=}g$. Consider the filtration $\{\Sigma_n^v:n\in\mathbb{N}\}$ that we built for Lemma 5 and note that

$$f_n:=\mathbb{E}\left[f_v\mid\Sigma_n^v\right]\to f_v\quad\text{and}\quad g_n:=\mathbb{E}\left[g_v\mid\Sigma_n^v\right]\to g_v$$

[17] There is a typo in both relation (2.39) and the last line of the mentioned theorem of Rüschendorf: the inequality on the left-hand side of the implication should be strict, in both cases.

in $\mathcal{L}^\infty$ with respect to bounded a.s. convergence if $\mathcal{F} = \mathcal{L}^\infty$ and in $\mathcal{M}^\infty$ if $\mathcal{F} = \mathcal{M}^\infty$. We want to show that, for each $n \in \mathbb{N}$, $f_n \geq_{\text{cv}} g_n$. To this end, let $F, G$ and $F_n, G_n$ be the distribution functions of $f, g$ and $f_n, g_n$, respectively. Define $\varphi, \gamma : [0,1] \to \mathbb{R}$ by

$$\varphi(p) = \int_0^p F^{-1}(t)\mathrm{d}\lambda \quad \text{and} \quad \gamma(p) = \int_0^p G^{-1}(t)\mathrm{d}\lambda$$

As well-known,[18] $f \geq_{\text{cv}} g$ is equivalent to $\varphi \geq \gamma$ with $\varphi(1) = \gamma(1)$. Arbitrarily choose $n \in \mathbb{N}$ and define $\varphi_n$ and $\gamma_n$ in a similar way. Now note that, by Lemma 5, we have $\lambda$-a.s.

$$F_n^{-1} = \mathbb{E}_\lambda \left[F^{-1} \mid \sigma(\Psi_n)\right]$$

Therefore, for each $i = 1, \ldots, 2^n$,

$$\varphi_n\left(\frac{i}{2^n}\right) = \int_0^{\frac{i}{2^n}} F_n^{-1}(t)\mathrm{d}\lambda = \int_0^{\frac{i}{2^n}} F^{-1}(t)\mathrm{d}\lambda = \varphi\left(\frac{i}{2^n}\right) \tag{18}$$

A similar argument holds for $g$ and $g_n$. Thus,

$$\varphi \geq \gamma \Longrightarrow \varphi_n\left(\frac{i}{2^n}\right) \geq \gamma_n\left(\frac{i}{2^n}\right) \qquad \forall i = 1, \ldots, 2^n$$

By definition $\varphi_n(0) = \gamma_n(0) = 0$. The functions $\varphi_n$ and $\gamma_n$ are absolutely continuous on $[0,1]$. Moreover, on each segment $[(i-1)/2^n, i/2^n]$, for each $p \in [(i-1)/2^n, i/2^n]$, we have

$$\begin{aligned}\varphi_n(p) &= \int_0^p F_n^{-1}(t)\mathrm{d}\lambda = \int_0^{\frac{i-1}{2^n}} F_n^{-1}(t)\mathrm{d}\lambda + \int_{\frac{i-1}{2^n}}^p \underbrace{F_n^{-1}(t)}_{=c_{i,n}\ \lambda\text{-a.s.}} \mathrm{d}\lambda \\ &= \varphi_n\left(\frac{i-1}{2^n}\right) + c_{i,n}\left(p - \frac{i-1}{2^n}\right)\end{aligned}$$

because $F_n^{-1}(t)$ is $\lambda$-a.s. constant on $((i-1)/2^n, i/2^n)$. But then $\varphi_n$ is affine on $[(i-1)/2^n, i/2^n]$, and the same is true for $\gamma_n$. Therefore, the inequality $\varphi_n \geq \gamma_n$ on the points $\{i/2^n : i = 0, \ldots, 2^n\}$ implies $\varphi_n \geq \gamma_n$ on $[0,1]$. As the equality $\varphi_n(1) = \gamma_n(1)$ follows from $\varphi(1) = \gamma(1)$, this proves that $f_n \geq_{\text{cv}} g_n$ in $\mathcal{L}^\infty(S, \Sigma, P)$, but then $f_n \geq_{\text{cv}} g_n$ in $\mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$.[19] As $n$ was chosen arbitrarily in $\mathbb{N}$, we conclude that, for each $n \in \mathbb{N}$, $f_n \geq_{\text{cv}} g_n$ in $\mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$. Now the restriction of $\succsim$ to $\mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$ is either propense to deductible-limit insurance or propense to proportional insurance because $\succsim$ satisfies either (iii) or (ii) on $\mathcal{F}$, and we can apply Lemma 11 to conclude that

$$f_n \succsim g_n \qquad \forall n \in \mathbb{N} \tag{19}$$

But, as observed, $f_n \to f_v$ and $g_n \to g_v$, thus the continuity of $\succsim$ guarantees that $f_v \succsim g_v$, and law invariance delivers $f \succsim g$, as wanted. ■

The conclusions of Theorem 3 hold also for risk preferences on $\mathcal{L}^p$ for $p \in [1, \infty)$ if continuity is formulated with respect to convergence in $\mathcal{L}^p$. This is because in Lemma 5, we proved that the convergence of $f_n \to f_v$ and $g_n \to g_v$ is in the corresponding sense.

## B.11 Weak monotonicity and weak secularity

Next we introduce weaker notions of monotonicity and secularity that are sufficient for some of the results that follow.

**Definition 16.** *A risk preference $\succsim$ is:*

[18] See, e.g., Theorem 3.A.5 of Shaked and Shanthikumar (2007).

[19] Since $\Sigma_n^v$ is finite, then $\mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right) = \mathcal{M}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$.

- weakly monotone *when, for all* $\eta, \gamma \in \mathbb{R}$,

$$\eta > \gamma \implies \eta \succ \gamma$$

- weakly secular *(or* solvable*) when, for all* $g \in \mathcal{F}$*, there exists* $\gamma \in \mathbb{R}$*, such that* $g \sim \gamma$*.*

As for the interpretation, weak monotonicity just requires that larger sure payoffs are preferred to smaller ones, weak secularity that every random payoff has a certainty equivalent.

## B.12 Proof of Proposition 2

This proof only requires weak monotonicity.

Clearly, (iv) $\implies$ (iii) $\implies$ (ii) $\implies$ (i). For the sake of brevity, call (v) the property

$$f \succsim g \iff \mathbb{E}[f] \geq \mathbb{E}[g]$$

for all $f, g \in \mathcal{F}$. Let $w, f, g \in \mathcal{F}$.

(i) $\implies$ (iv) If $g \overset{d}{=} f$, then $\mathbb{E}[w+f] = \mathbb{E}[w] + \mathbb{E}[f] = \mathbb{E}[w] + \mathbb{E}[g] = \mathbb{E}[w+g]$. Risk neutrality delivers

$$w + f \sim \mathbb{E}[w+f] = \mathbb{E}[w+g] \sim w + g$$

and transitivity implies $w + f \sim w + g$. Thus dependence neutrality holds.

(v) $\implies$ (i) Since $\mathbb{E}[f] = \mathbb{E}[\mathbb{E}[f]]$, condition (v) implies $f \sim \mathbb{E}[f]$. Thus risk neutrality holds.

(i) $\implies$ (v) If $f \succsim g$, then risk neutrality yields $\mathbb{E}[f] \sim f \succsim g \sim \mathbb{E}[g]$, and transitivity implies $\mathbb{E}[f] \succsim \mathbb{E}[g]$. If $\mathbb{E}[f] < \mathbb{E}[g]$, weak monotonicity would imply $\mathbb{E}[f] \prec \mathbb{E}[g]$, a contradiction, therefore it must be the case that $\mathbb{E}[f] \geq \mathbb{E}[g]$. Summing up: $f \succsim g \implies \mathbb{E}[f] \geq \mathbb{E}[g]$.

Conversely, if $\mathbb{E}[f] \geq \mathbb{E}[g]$, then:

- either $\mathbb{E}[f] = \mathbb{E}[g]$, then risk neutrality and reflexivity yield $f \sim \mathbb{E}[f] \sim \mathbb{E}[g] \sim g$, and transitivity implies $f \succsim g$;

- or $\mathbb{E}[f] > \mathbb{E}[g]$, then risk neutrality and weak monotonicity yield $f \sim \mathbb{E}[f] \succ \mathbb{E}[g] \sim g$, and transitivity implies $f \succsim g$.

Summing up: $\mathbb{E}[f] \geq \mathbb{E}[g] \implies f \succsim g$. Thus, (v) holds. ∎

## B.13 Proof of Proposition 3

This proof only requires weak monotonicity.

Let $w, f, g \in \mathcal{F}$.

(i) $\implies$ (ii). If $f \geq_{\text{fsd}} g$, then $\mathbb{E}[w+f] = \mathbb{E}[w] + \mathbb{E}[f] \geq \mathbb{E}[w] + \mathbb{E}[g] = \mathbb{E}[w+g]$, it follows that $\mathbb{E}[w+f] \geq \mathbb{E}[w+g]$ and, by (i), $w + f \succsim w + g$.

(ii) $\implies$ (i). If $f \overset{d}{=} g$, then $f \geq_{\text{fsd}} g \geq_{\text{fsd}} f$. By (ii), we have $w+f \succsim w+g \succsim w+f$ and so $w+f \sim w+g$. Hence, $\succsim$ is dependence neutral, and Proposition 2 implies that $\succsim$ admits an expected-value representation.

(i) $\implies$ (iii). Since $\succsim$ is represented by the expected value, (iii) follows immediately.

(iii) $\implies$ (i) If $f \overset{d}{=} g$, by law invariance, $f \sim g$, by (iii), $w + f \sim w + g$. Thus, (iii) yields dependence neutrality, and Proposition 2 implies that $\succsim$ admits an expected-value representation.

(i) $\implies$ (iv). Since $\succsim$ is represented by the expected value, it is complete. Also, if $f \succ g$, then $\mathbb{E}[f] > \mathbb{E}[g]$; Theorem 1 of Pomatto, Strack, and Tamuz (2020) implies that $w + \tilde{f} >_{\text{fsd}} w + \tilde{g}$ for some $w, \tilde{f}, \tilde{g} \in \mathcal{F}$ such that $f \overset{d}{=} \tilde{f}$, $g \overset{d}{=} \tilde{g}$, and $w$ is independent of both $\tilde{f}$ and $\tilde{g}$.

(iv) $\implies$ (i). If $f \succ g$, then $w + \tilde{f} >_{\text{fsd}} w + \tilde{g}$ for some $w, \tilde{f}, \tilde{g} \in \mathcal{F}$ such that $f \overset{d}{=} \tilde{f}$, $g \overset{d}{=} \tilde{g}$, and $w$ is independent of both $\tilde{f}$ and $\tilde{g}$. Thus, $f \succ g$ implies $\mathbb{E}[w + \tilde{f}] > \mathbb{E}[w + \tilde{g}]$, whence $\mathbb{E}[f] = \mathbb{E}[\tilde{f}] > \mathbb{E}[\tilde{g}] = \mathbb{E}[g]$ and $\mathbb{E}[f] > \mathbb{E}[g]$. Since $\succsim$ is complete, by contraposition, it follows that $\mathbb{E}[f] \leq \mathbb{E}[g]$ implies $f \precsim g$. In particular, $\mathbb{E}[f] = \mathbb{E}[g]$ implies $f \sim g$. Finally, $\mathbb{E}[f] = \mathbb{E}[\mathbb{E}[f]]$ implies $f \sim \mathbb{E}[f]$. Thus risk neutrality holds. Since $\succsim$ is a (weakly) monotone risk preference, by Proposition 2, it admits an expected-value representation.
∎

## B.14 Proofs of the results of Section IV

**Proof of Lemma 2.** This proof only requires weak monotonicity and weak secularity.

Let $f, g \in \mathcal{F}$, $\rho_{\mathrm{A}}, \rho_{\mathrm{B}}, \gamma \in \mathbb{R}$.

(i) $\Longrightarrow$ (ii). If $f = \mathbb{E}[g]$, then both $f - \rho_{\mathrm{A}}$ and $f - \rho_{\mathrm{B}}$ are sure payoffs. Since $f - \rho_{\mathrm{A}} \succsim_{\mathrm{A}} g$, by (i), $f - \rho_{\mathrm{A}} \succsim_{\mathrm{B}} g \sim_{\mathrm{B}} f - \rho_{\mathrm{B}}$. By weak monotonicity, $\rho_{\mathrm{A}} > \rho_{\mathrm{B}}$ would lead to the contradiction $f - \rho_{\mathrm{B}} \succ_{\mathrm{B}} f - \rho_{\mathrm{A}}$, then it must be the case that $\rho_{\mathrm{B}} \geq \rho_{\mathrm{A}}$.

(ii) $\Longrightarrow$ (i). If $\gamma \succsim_{\mathrm{A}} g$, then $\mathbb{E}[g] - (\mathbb{E}[g] - \gamma) \succsim_{\mathrm{A}} g$. Now, let $f = \mathbb{E}[g]$, if $g \sim_{\mathrm{A}} f - \rho_{\mathrm{A}}$ (and such a $\rho_{\mathrm{A}}$ exists by weak secularity), we have

$$f - (\mathbb{E}[g] - \gamma) = \mathbb{E}[g] - (\mathbb{E}[g] - \gamma) \succsim_{\mathrm{A}} g \sim_{\mathrm{A}} f - \rho_{\mathrm{A}}$$

By weak monotonicity, $\mathbb{E}[g] - \gamma > \rho_{\mathrm{A}}$ would lead to the contradiction $f - \rho_{\mathrm{A}} \succ_{\mathrm{A}} f - (\mathbb{E}[g] - \gamma)$, then it must be the case that $\mathbb{E}[g] - \gamma \leq \rho_{\mathrm{A}}$. Now let $\rho_{\mathrm{B}}$ be such that $g \sim_{\mathrm{B}} f - \rho_{\mathrm{B}}$ (and such a $\rho_{\mathrm{B}}$ exists by weak secularity). By (ii) and what we have just observed, $\rho_{\mathrm{B}} \geq \rho_{\mathrm{A}} \geq \mathbb{E}[g] - \gamma$, and $f - (\mathbb{E}[g] - \gamma) \geq f - \rho_{\mathrm{B}}$, by weak monotonicity (and reflexivity for the equality case)

$$f - (\mathbb{E}[g] - \gamma) \succsim_{\mathrm{B}} f - \rho_{\mathrm{B}}$$

but then $\gamma = f - (\mathbb{E}[g] - \gamma) \succsim_{\mathrm{B}} f - \rho_{\mathrm{B}} \sim_{\mathrm{B}} g$, so that $\gamma \succsim_{\mathrm{B}} g$.

The final part of the statement is a consequence of the fact that $\mathbb{E}[g]$ dominates any random payoff $g$ according to $\geq_{\mathrm{cv}}$. ■

**Lemma 12.** *Let $\succsim$ be a monotone and secular risk preference on $\mathcal{F}$. Then:*

1. *for all $f, g \in \mathcal{F}$, $f \succsim g \iff \rho(g, f) \geq 0$;*
2. *the certainty equivalent map $g \mapsto -\rho(g, 0)$ represents $\succsim$ on $\mathcal{F}$;*
3. *if $f, f', g, g' \in \mathcal{F}$, $f \overset{d}{=} f'$, and $g \overset{d}{=} g'$, then $\rho(g, f) = \rho(g', f')$.*

*If moreover $\succsim$ is continuous, then $\rho : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$ is (jointly) sequentially continuous.*

**Proof.** Let $f, g \in \mathcal{F}$.

1. By definition of $\rho : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$, $g \sim f - \rho(g, f)$. If $f \succsim g$, by transitivity $f \succsim f - \rho(g, f)$, monotonicity then excludes the case $\rho(g, f) < 0$. Conversely, if $\rho(g, f) \geq 0$, monotonicity and reflexivity imply

$$f = (f - \rho(g, f)) + \rho(g, f) \succsim f - \rho(g, f) + 0 \sim g$$

transitivity allows to conclude $f \succsim g$.

2. By definition of $\rho : \mathcal{F} \times \mathcal{F} \to \mathbb{R}$, $g \sim 0 - \rho(g, 0) = -\rho(g, 0)$, then $-\rho(g, 0)$ is the certainty equivalent of $\mathcal{F}$. With this, for all $f, g \in \mathcal{F}$

$$f \succsim g \iff -\rho(f, 0) \succsim -\rho(g, 0) \iff -\rho(f, 0) \geq -\rho(g, 0)$$

where the latter relation follows by monotonicity.

3. Note $f \overset{d}{=} f'$ implies $f - \rho(g, f) \overset{d}{=} f' - \rho(g, f)$, repeated application of law invariance yield

$$g' \sim g \sim f - \rho(g, f) \sim f' - \rho(g, f)$$

transitivity and the definition of $\rho$ yield $\rho(g, f) = \rho(g', f')$.

Finally, assume that $\succsim$ is continuous. Next we show that, if $k \in \mathbb{R}$, $f_n \to f$ in $\mathcal{F}$, $g_n \to g$ in $\mathcal{F}$, and $\rho(g_n, f_n) \leq k$ (resp. $\geq k$) for all $n \in \mathbb{N}$, then $\rho(g, f) \leq k$ (resp. $\geq k$). Indeed, for all $n \in \mathbb{N}$, $\rho(g_n, f_n) \leq k$ implies $-\rho(g_n, f_n) \geq -k$, by monotonicity,

$$g_n \sim f_n - \rho(g_n, f_n) \succsim f_n - k$$

by continuity

$$g \succsim f - k$$

but then $f - \rho(g, f) \sim g \succsim f - k$, and monotonicity again yields $\rho(g, f) \leq k$. Analogously, for all $n \in \mathbb{N}$, $\rho(g_n, f_n) \geq k$ implies $\rho(g, f) \geq k$.

Now assume that $f_n \to f$ in $\mathcal{F}$, and $g_n \to g$ in $\mathcal{F}$, and, per contra $\rho(g_n, f_n) \nrightarrow \rho(g, f)$. Then there exists $\eta > 0$ such that for all $m \in \mathbb{N}$ there exists $n_m > m$ such that $\rho(g_{n_m}, f_{n_m}) \notin (\rho(g, f) - \eta, \rho(g, f) + \eta)$. Therefore there exists a subsequence $\{(g_{n_l}, f_{n_l})\}_{l \in \mathbb{N}}$ of $\{(g_n, f_n)\}_{n \in \mathbb{N}}$ such that $\rho(g_{n_l}, f_{n_l}) \notin (\rho(g, f) - \eta, \rho(g, f) + \eta)$ for all $l \in \mathbb{N}$. But then, either $\rho(g_{n_l}, f_{n_l}) \leq \rho(g, f) - \eta$ for infinitely many $l$, or $\rho(g_{n_l}, f_{n_l}) \geq \rho(g, f) + \eta$ for infinitely many $l \in \mathbb{N}$. In the first case, there exists a subsequence $\{(g_{n_i}, f_{n_i})\}_{i \in \mathbb{N}}$ of $\{(g_{n_l}, f_{n_l})\}_{l \in \mathbb{N}}$ such that $\rho(g_{n_i}, f_{n_i}) \leq \rho(g, f) - \eta$ for all $i \in \mathbb{N}$, and by the previous observation $\rho(g, f) \leq \rho(g, f) - \eta$, a contradiction. In the second case, the contradiction $\rho(g, f) \geq \rho(g, f) + \eta$ is obtained. This yields the desired joint sequential continuity. ■

**Proof of Lemma 3.** Let A be risk neutral. Note that for $\succsim_{\mathrm{A}}$ the assumptions of monotonicity and secularity are implied by weak monotonicity. In fact, by Proposition 2, weak monotone and risk neutral risk preferences are represented by the expected value, so they are monotone. As to secularity, for all $f, g \in \mathcal{F}$,

$$f - (\mathbb{E}[f] - \mathbb{E}[g]) \sim_{\mathrm{A}} \mathbb{E}[f - (\mathbb{E}[f] - \mathbb{E}[g])] = \mathbb{E}[g] \sim_{\mathrm{A}} g$$

that is, $\rho_{\mathrm{A}}(g, f) = \mathbb{E}[f] - \mathbb{E}[g]$.

We only prove point 2 because point 1 is well known.

2. Let B be strongly more risk averse than A. If $f \geq_{\mathrm{cv}} g$, then $\mathbb{E}[f] = \mathbb{E}[g]$. Since A is risk neutral, as observed, $\rho_{\mathrm{A}}(g, f) = \mathbb{E}[f] - \mathbb{E}[g] = 0$. Since B is strongly more risk averse than A, then

$$\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f) = 0$$

Lemma 12 yields $f \succsim_{\mathrm{B}} g$, and so B is strongly risk averse.

Conversely, if B is strongly risk averse, then

$$f \geq_{\mathrm{cv}} g \implies f \succsim_{\mathrm{B}} g$$

Lemma 12 yields $\rho_{\mathrm{B}}(g, f) \geq 0$. But, as observed, since A is risk neutral, $\rho_{\mathrm{A}}(g, f) = \mathbb{E}[f] - \mathbb{E}[g] = 0$, and so $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$ which shows that B is strongly more risk averse than A. ■

**Proof of Theorem 4**. This proof only requires weak monotonicity and weak secularity.

(i) $\implies$ (ii). Let $w, f, g \in \mathcal{F}$, with $f \stackrel{d}{=} g$, if $f \in \mathcal{I}^{\mathrm{fi}}(w)$, then

$$(w + f) - \rho_{\mathrm{A}}(w + g, w + f) \sim_{\mathrm{A}} w + g \quad \text{and} \quad (w + f) - \rho_{\mathrm{B}}(w + g, w + f) \sim_{\mathrm{B}} w + g$$

but $\gamma = (w + f) - \rho_{\mathrm{A}}(w + g, w + f) \in \mathbb{R}$, because $f \in \mathcal{I}^{\mathrm{fi}}(w)$. By (i), $(w + f) - \rho_{\mathrm{A}}(w + g, w + f) \succsim_{\mathrm{B}} w + g \sim_{\mathrm{B}} (w + f) - \rho_{\mathrm{B}}(w + g, w + f)$, by weak monotonicity, $\rho_{\mathrm{A}}(w + g, w + f) \leq \rho_{\mathrm{B}}(w + g, w + f)$. This shows that B is more propense to full insurance than A.

(ii) $\implies$ (i). Let $h \in \mathcal{F}$ and $\gamma \in \mathbb{R}$ be such that $\gamma \succsim_{\mathrm{A}} h$. By weak secularity there exists $\eta \in \mathbb{R}$ such that $\gamma \succsim_{\mathrm{A}} h \sim_{\mathrm{A}} \eta$, and by weak monotonicity $\gamma \geq \eta$. By Lemma 1, there exist $w, w' \in \mathcal{F}$ such that $w \stackrel{d}{=} w'$ and $h - \mathbb{E}[h] \stackrel{d}{=} w - w'$. Let $f = -w + \mathbb{E}[h]$ and $g = -w' + \mathbb{E}[h]$, clearly $f \stackrel{d}{=} g$ and $f \in \mathcal{I}^{\mathrm{fi}}(w)$. By (ii), $\rho_{\mathrm{A}}(w + g, w + f) \leq \rho_{\mathrm{B}}(w + g, w + f)$, and by definition of $\rho$,

$$\underbrace{(w + f) - \rho_{\mathrm{A}}(w + g, w + f)}_{=\mathbb{E}[h] - \rho_{\mathrm{A}}(w+g, w+f)} \sim_{\mathrm{A}} \underbrace{w + g}_{\stackrel{d}{=} h} \quad \text{and} \quad (w + f) - \rho_{\mathrm{B}}(w + g, w + f) \sim_{\mathrm{B}} w + g$$

law invariance yields $(w + f) - \rho_{\mathrm{A}}(w + g, w + f) \sim_{\mathrm{A}} h$, but since $(w + f) - \rho_{\mathrm{A}}(w + g, w + f)$ is constant, then

$$\eta = (w + f) - \rho_{\mathrm{A}}(w + g, w + f) \geq (w + f) - \rho_{\mathrm{B}}(w + g, w + f) \sim_{\mathrm{B}} w + g \stackrel{d}{=} h$$

Weak monotonicity and law invariance yield $\eta \succsim_{\mathrm{B}} h$, and weak monotonicity again yields $\gamma \succsim_{\mathrm{B}} h$. This shows that B is weakly more risk averse than A. ■

In the following two lemmas, analogous to those of Appendix B.3, we assume that $\Sigma$ is generated by a partition $\mathcal{S}$ of equiprobable events (called cells), and we fix two continuous, monotone, and secular risk preferences $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$ on $\mathcal{F}$.

**Lemma 13.** *Let $f, g \in \mathcal{F}$ be such that $g$ is a mean preserving spread of $f$. If either (ii) or (iii) of Theorem 5 holds, then $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$.*

**Proof.** When (iii) of Theorem 5 holds, the results follows immediately from Lemma 10. Suppose now that (ii) of Theorem 5 holds. Let

$$g = f - \delta 1_{S_1} + \delta 1_{S_2}$$

with $\delta \geq 0$ and $S_1, S_2$ two distinct cells in $\mathcal{S}$ such that $f(S_1) \leq f(S_2)$. If $\delta = 0$, then $f = g$, and $\rho_{\mathrm{A}}(g, f) = \rho_{\mathrm{B}}(g, f) = 0$. If $\delta > 0$ and $f(S_1) < f(S_2)$, it follows from Lemma 8 that $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$. If $\delta > 0$ and $f(S_1) = f(S_2)$, define $f_\varepsilon = f - \varepsilon 1_{S_1} + \varepsilon 1_{S_2}$ with $\varepsilon \in (0, \delta)$. Note that

$$\begin{aligned} f_\varepsilon &= f - \varepsilon 1_{S_1} + \varepsilon 1_{S_2} = f 1_{S \setminus \{S_1, S_2\}} + (f(S_1) - \varepsilon) 1_{S_1} + (f(S_2) + \varepsilon) 1_{S_2} \\ g &= f - \delta 1_{S_1} + \delta 1_{S_2} = f - (\varepsilon + (\delta - \varepsilon)) 1_{S_1} + (\varepsilon + (\delta - \varepsilon)) 1_{S_2} = f_\varepsilon - (\delta - \varepsilon) 1_{S_1} + (\delta - \varepsilon) 1_{S_2} \end{aligned}$$

Thus $g$ is a mean preserving spread of $f_\varepsilon$ with $f_\varepsilon(S_1) < f_\varepsilon(S_2)$ and $\delta - \varepsilon > 0$. By the previous argument $\rho_{\mathrm{B}}(g, f_\varepsilon) \geq \rho_{\mathrm{A}}(g, f_\varepsilon)$ for all $\varepsilon \in (0, \delta)$. Let $\{\varepsilon_n\}_{n \in \mathbb{N}} \subseteq (0, \delta)$ be such that $\lim_{n \to \infty} \varepsilon_n = 0$. By Lemma 12 and continuity of both $\succsim_{\mathrm{A}}$ and $\succsim_{\mathrm{B}}$, it follows that $\rho_{\mathrm{A}}(g, f_{\varepsilon_n}) \to \rho_{\mathrm{A}}(g, f)$ and $\rho_{\mathrm{B}}(g, f_{\varepsilon_n}) \to \rho_{\mathrm{B}}(g, f)$, and so $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$. This completes the proof. ■

**Lemma 14.** *Let $f, g \in \mathcal{F}$ be such that $f \geq_{\mathrm{cv}} g$. If either (ii) or (iii) of Theorem 5 holds, then $\rho_{\mathrm{B}}(g, f) \geq \rho_{\mathrm{A}}(g, f)$.*

**Proof.** If $f \geq_{\mathrm{cv}} g$ in $\mathcal{F}$, then there exists a sequence $h_0, h_1, \ldots, h_m$ such that $f = h_0$, $g = h_m$ and each $h_{k+1}$ is either a mean preserving spread of $h_k$ or it is obtained by $h_k$ through the permutation of the values that $h_k$ takes on two cells. By the previous lemma, we have $\rho_{\mathrm{B}}(h_{k+1} - x, h_k - x) \geq \rho_{\mathrm{A}}(h_{k+1} - x, h_k - x)$ for all $x \in \mathbb{R}$ and $k = 0, 1, \ldots, m-1$ as either $h_{k+1} - x$ is a mean preserving spread of $h_k - x$ or $h_{k+1} - x \overset{d}{=} h_k - x$. Next, we prove by induction that, for all $x \in \mathbb{R}$ and $j = 1, 2, \ldots, m$,

$$\rho_{\mathrm{B}}(h_j - x, h_0 - x) \geq \rho_{\mathrm{A}}(h_j - x, h_0 - x)$$

As we have just observed, for $j = 1$, we have $\rho_{\mathrm{B}}(h_1 - x, h_0 - x) \geq \rho_{\mathrm{A}}(h_1 - x, h_0 - x)$ for all $x \in \mathbb{R}$. Suppose that, for $j = k$, $\rho_{\mathrm{B}}(h_k - x, h_0 - x) \geq \rho_{\mathrm{A}}(h_k - x, h_0 - x)$ for all $x \in \mathbb{R}$; it then suffices to verify that $\rho_{\mathrm{B}}(h_{k+1} - x, h_0 - x) \geq \rho_{\mathrm{A}}(h_{k+1} - x, h_0 - x)$ for all $x \in \mathbb{R}$. To see this, denote by $\eta_{\mathrm{A}} = \rho_{\mathrm{A}}(h_{k+1} - x, h_k - x)$ and $\eta_{\mathrm{B}} = \rho_{\mathrm{B}}(h_{k+1} - x, h_k - x)$. It holds that

$$h_k - x - \eta_{\mathrm{A}} \sim_{\mathrm{A}} h_{k+1} - x \quad \text{and} \quad h_k - x - \eta_{\mathrm{B}} \sim_{\mathrm{B}} h_{k+1} - x$$

As we have observed above, $\eta_{\mathrm{B}} \geq \eta_{\mathrm{A}}$ and since $\succsim_{\mathrm{A}}$ is monotone we have $h_k - x - \eta_{\mathrm{B}} \precsim_{\mathrm{A}} h_{k+1} - x$. Therefore

$$h_0 - x - \eta_{\mathrm{B}} - \rho_{\mathrm{A}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}}) \sim_{\mathrm{A}} h_k - x - \eta_{\mathrm{B}} \precsim_{\mathrm{A}} h_{k+1} - x \sim_{\mathrm{A}} h_0 - x - \rho_{\mathrm{A}}(h_{k+1} - x, h_0 - x)$$

and, by monotonicity, $\rho_{\mathrm{A}}(h_{k+1} - x, h_0 - x) \leq \eta_{\mathrm{B}} + \rho_{\mathrm{A}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}})$. Moreover,

$$h_0 - x - \eta_{\mathrm{B}} - \rho_{\mathrm{B}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}}) \sim_{\mathrm{B}} h_k - x - \eta_{\mathrm{B}} \sim_{\mathrm{B}} h_{k+1} - x$$

and so $\rho_{\mathrm{B}}(h_{k+1} - x, h_0 - x) = \eta_{\mathrm{B}} + \rho_{\mathrm{B}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}})$. By induction $\rho_{\mathrm{B}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}}) \geq \rho_{\mathrm{A}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}})$, and so

$$\begin{aligned} \rho_{\mathrm{A}}(h_{k+1} - x, h_0 - x) &\leq \eta_{\mathrm{B}} + \rho_{\mathrm{A}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}}) \\ &\leq \eta_{\mathrm{B}} + \rho_{\mathrm{B}}(h_k - x - \eta_{\mathrm{B}}, h_0 - x - \eta_{\mathrm{B}}) \\ &= \rho_{\mathrm{B}}(h_{k+1} - x, h_0 - x) \end{aligned}$$

as wanted. ■

**Proof of Theorem 5**. (i) $\implies$ (vi). Let $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$ and $f \trianglerighteq_w g$. By Lemma 7, $w+f \geq_{\text{cv}} w+g$, and (i) implies $\rho_{\text{B}}(w+g, w+f) \geq \rho_{\text{A}}(w+g, w+f)$. Thus (vi) holds.

(vi) $\implies$ (v). Let $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$ and $f \in \mathcal{I}^{\text{cs}}(w)$, by Proposition 1, it follows that $f \trianglerighteq_w g$, and (vi) implies $\rho_{\text{B}}(w+g, w+f) \geq \rho_{\text{A}}(w+g, w+f)$. Thus (v) holds.

(v) $\implies$ (iv) because $\mathcal{I}^{\text{is}}(w) \subseteq \mathcal{I}^{\text{cs}}(w)$ for all $w \in \mathcal{F}$.

(iv) $\implies$ (iii) and (iv) $\implies$ (ii) because $\mathcal{I}^{\text{dl}}(w), \mathcal{I}^{\text{pr}}(w) \subseteq \mathcal{I}^{\text{is}}(w)$ for all $w \in \mathcal{F}$.

(iii) $\implies$ (i) and (ii) $\implies$ (i). The case in which $\Sigma$ is generated by a finite partition of equiprobable events follows from Lemma 14. Now, let $P$ be nonatomic. Let $f, g \in \mathcal{F}$ be such that $f \geq_{\text{cv}} g$. We want to show that $\rho_{\text{B}}(g, f) \geq \rho_{\text{A}}(g, f)$.

The sequences $\{f_n\}_{n\in\mathbb{N}}$ and $\{g_n\}_{n\in\mathbb{N}}$ introduced in the proof of Theorem 3 in Appendix B.10 have the following properties:

- $f_n, g_n \in \mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$ where $\{\Sigma_n^v : n \in \mathbb{N}\}$ is the filtration that we built for Lemma 5;
- $f_n \geq_{\text{cv}} g_n$ for all $n \in \mathbb{N}$;
- $f_n \to f_v$ and $g_n \to g_v$ in $\mathcal{F}$, with $f_v \overset{d}{=} f$ and $g_v \overset{d}{=} g$.

The restrictions of $\succsim_{\text{A}}$ and $\succsim_{\text{B}}$ to $\mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$ are continuous, monotone, and secular risk preferences that satisfy either (ii) or (iii) in this theorem, and we can apply Lemma 14 to conclude

$$\rho_{\text{B}}(g_n, f_n) \geq \rho_{\text{A}}(g_n, f_n) \qquad \forall n \in \mathbb{N} \tag{20}$$

but, by Lemma 12, both $\rho_{\text{A}}$ and $\rho_{\text{B}}$ are law invariant and continuous, and hence

$$\rho_{\text{B}}(g, f) = \rho_{\text{B}}(g_v, f_v) \geq \rho_{\text{A}}(g_v, f_v) = \rho_{\text{A}}(g, f)$$

as wanted. ■

**Proposition 4.** *Let $\succsim_{\text{A}}$ and $\succsim_{\text{B}}$ be monotone and secular risk preferences.*

1. *If $\succsim_{\text{A}}$ is neutral to full insurance, then* B *is more propense to full insurance than* A *if and only if* B *is propense to full insurance.*

2. *If $\succsim_{\text{A}}$ is neutral to hedging, then* B *is more propense to hedging than* A *if and only if* B *is propense to hedging.*

**Proof.** Note that by Proposition 2, A is neutral to full insurance if and only if she is neutral to hedging if and only if she is risk neutral.

1. By Theorem 4, B is more propense to full insurance than A if and only if B is weakly more risk averse than A. By Lemma 3, B is weakly more risk averse than A if and only if B is weakly risk averse. By Theorem 1, B is weakly risk averse if and only if B is propense to full insurance.

2. Let $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$. Assume that B is more propense to hedging than A. If $f \trianglerighteq_w g$, then $\mathbb{E}[w+f] = \mathbb{E}[w+g]$ (because $f \overset{d}{=} g$). Since A is risk neutral, as observed in the proof of Lemma 3, $\rho_{\text{A}}(w+g, w+f) = \mathbb{E}[w+f] - \mathbb{E}[w+g] = 0$. Since B is more propense to hedging than A, then

$$\rho_{\text{B}}(w+g, w+f) \geq \rho_{\text{A}}(w+g, w+f) = 0$$

Lemma 12 yields $w+f \succsim_{\text{B}} w+g$, and so B is propense to hedging.

Conversely, if B is propense to hedging, then

$$f \trianglerighteq_w g \implies w+f \succsim_{\text{B}} w+g$$

Lemma 12 yields $\rho_{\text{B}}(w+g, w+f) \geq 0$. But, as observed in the proof of Lemma 3, since A is risk neutral, $\rho_{\text{A}}(w+g, w+f) = \mathbb{E}[w+f] - \mathbb{E}[w+g] = 0$, and so $\rho_{\text{B}}(w+g, w+f) \geq \rho_{\text{A}}(w+g, w+f)$, which shows that B is more propense to hedging than A. ■

# Appendix C Additional results and considerations

## C.1 Total wealth and wealth changes

In choice under risk, to each risk preference $\succsim$ on $\mathcal{F}$ and each initial wealth $w_0 \in \mathbb{R}$, another risk preference

$$f \succsim^{w_0} g \iff w_0 + f \succsim w_0 + g \tag{21}$$

is associated. In this perspective, random payoffs are interpreted as risks – that is, changes in wealth – relative to an initial wealth $w_0$. Accordingly, the ranking

$$f \succsim^{w_0} g$$

is interpreted as '$f$ is preferred to $g$, given $w_0$'. Obviously, the risk preference $\succsim^0$, corresponding to $w_0 = 0$, is nothing but $\succsim$ itself. This appendix shows that the study of risk attitudes – in its traditional form as well as in the insurance-based one of the current paper – is independent of whether we consider either a preference relation $\succsim$ over random final wealth levels or any preference relation $\succsim^{w_0}$ over risks.

**Proposition 5.** *The following properties are equivalent for a risk preference $\succsim$:*

*(i) $\succsim$ is propense to full insurance (weakly risk averse);*

*(ii) for some $w_0 \in \mathbb{R}$, $\succsim^{w_0}$ is propense to full insurance (weakly risk averse);*

*(iii) for every $w_0 \in \mathbb{R}$, $\succsim^{w_0}$ is propense to full insurance (weakly risk averse).*

**Proof.** We only prove that (ii) $\implies$ (iii), the rest being obvious. Assume that (ii) holds for a given $w_0$ and arbitrarily choose $w_0' \in \mathbb{R}$. For all $w, f, g \in \mathcal{F}$ with $g \overset{d}{=} f$, if $f$ is full insurance for $w$, then $f = -w - \pi$ for some $\pi \in \mathbb{R}$. But then $f$ is full insurance also for $y = w + w_0' - w_0$, in fact $f = -(w + w_0' - w_0) - (\pi - w_0' + w_0)$. Since $\succsim^{w_0}$ is propense to full insurance, then $y + f \succsim^{w_0} y + g$, explicitly

$$w_0 + \underbrace{w + w_0' - w_0}_{y} + f \succsim w_0 + \underbrace{w + w_0' - w_0}_{y} + g$$

and so $w_0' + (w + f) \succsim w_0' + (w + g)$, that is, $w + f \succsim^{w_0'} w + g$, as wanted. ■

In words, when a preference relation is propense to full insurance (weakly risk averse) at some initial wealth level, it remains so at any other initial level. Intuitively, propensity to full insurance (weak risk aversion) per se is a feature of a preference relation that depends only on the variability of payoffs and as such it is unaffected by the addition of a constant (i.e., by an initial wealth $w_0$). In contrast, the degree of propensity to full insurance (weak risk aversion) may well change with the initial wealth level as risk preferences are, in general, not invariant under the addition of constants (in the jargon, they are not translation invariant).

This intuition is confirmed by the main idea of the proof above: $f$ is full insurance for a risk $w$ if and only if it is full insurance for the corresponding final wealth $w_0 + w$, i.e., $\mathcal{I}^{\text{fi}}(w) = \mathcal{I}^{\text{fi}}(w_0 + w)$. This invariance is easily seen to hold for partial insurance as well, that is, $\mathcal{I}^{\text{pi}}(w) = \mathcal{I}^{\text{pi}}(w_0 + w)$ for each $\text{pi} \in \{\text{pr}, \text{dl}, \text{is}, \text{cs}\}$. Accordingly the last proposition continues to hold with 'partial' and 'strongly' in place of 'full' and 'weakly'.

## C.2 A generalization of Theorem 1

A possible issue in testing empirically our notion of propensity to full insurance is the universal quantification 'for all $\pi \in \mathbb{R}$' regarding insurance prices. Typically, in insurance markets there is only a finite number of insurance providers, each adopting a premium principle $\Pi$ that associates a premium $\pi = \Pi(h)$ to each insurance payoff $h$ in $\mathcal{F}$. In this appendix, we address this issue by providing a weaker version of our notion that still characterizes weak risk aversion. To this end, next we introduce a general class of pricing rules for insurance markets.

**Definition 17.** *A function* $\Pi : \mathcal{F} \to \mathbb{R}$ *is a* premium calculation principle *when there exists* $\theta > 0$ *such that*

$$\Pi(h+\gamma) = \Pi(h) + \theta\gamma$$

*for all* $h \in \mathcal{F}$ *and all* $\gamma \in \mathbb{R}$.

This notion includes most pricing rules used in the insurance industry, such as those presented by Dickson (2017, Chapter 3), and in particular the fair premium principle for which $\Pi$ is the expected value.

With a prespecified premium calculation principle $\Pi$ replacing the arbitrary premium $\pi \in \mathbb{R}$, propensity to full insurance takes the following weaker form.

**Definition 18.** *A risk preference* $\succsim$ *is* propense to full insurance at price $\Pi$ *when, for all* $w, f, g \in \mathcal{F}$ *with* $g \overset{d}{=} f$,

$$f = -w - \Pi(-w) \implies w + f \succsim w + g \tag{22}$$

The interpretation is analogous to the original Definition 7-(i), but without the quantification 'for all $\pi \in \mathbb{R}$' previously mentioned. With this amended definition, the conclusion of Theorem 1 continues to hold.

**Theorem 7.** *Let* $\Pi : \mathcal{F} \to \mathbb{R}$ *be a premium calculation principle. The following properties are equivalent for a risk preference:*

*(i) weak risk aversion;*

*(ii) propensity to full insurance;*

*(iii) propensity to full insurance at price* $\Pi$.

**Proof.** The implication (i) $\implies$ (ii) is part of Theorem 1 and (ii) $\implies$ (iii) is immediate. As for (iii) $\implies$ (i) observe that for each $h \in \mathcal{F}$ there exist, by Lemma 1, elements $z, z' \in \mathcal{F}$ such that $z \overset{d}{=} z'$ and $h - \mathbb{E}[h] \overset{d}{=} z - z'$. By (22), for

$$\gamma = \frac{\Pi(-z) + \mathbb{E}[h]}{\theta}$$

we have that $\Pi(-z-\gamma) = -\mathbb{E}[h]$. Let $w = z + \gamma$ and $w' = z' + \gamma$. Clearly, $w \overset{d}{=} w'$, $h - \mathbb{E}[h] \overset{d}{=} w - w'$, and $\mathbb{E}[h] = -\Pi(-z-\gamma) = -\Pi(-w)$. Let

$$f = -w + \mathbb{E}[h] = -w - \Pi(-w) \quad \text{and} \quad g = -w' + \mathbb{E}[h] = -w' - \Pi(-w)$$

Clearly, $f \overset{d}{=} g$. Propensity to full insurance at price $\Pi$, yields $w + f \succsim w + g$, that is,

$$\mathbb{E}[h] = w + f \succsim w + g = w - w' + \mathbb{E}[h] \overset{d}{=} h$$

Law invariance of $\succsim$ guarantees that $w - w' + \mathbb{E}[h] \sim h$, and so $\mathbb{E}[h] \succsim w - w' + \mathbb{E}[h] \sim h$. Transitivity then implies $\mathbb{E}[h] \succsim h$, showing that (i) holds. ■

Also the definitions of propensity to proportional and deductible-limit insurance can be weakened in the same manner, with the conclusions of Theorem 2 still holding true. The details are omitted for brevity.

We close by observing that another common feature of most pricing rules for insurance markets is *law invariance*,[20] with this (22) becomes

$$w + \underbrace{(-w - \Pi(-w))}_{\text{full insurance at its own price}} \succsim w + \underbrace{(h - \Pi(h))}_{\text{payoff } h \text{ at its own price}} \tag{23}$$

for all $h \overset{d}{=} -w$, because law invariance guarantees $\Pi(h) = \Pi(-w)$. Note that (23) is analogous to condition (2) in the introduction, again without the universal quantification.

[20] Law invariance of $\Pi$ means that $h \overset{d}{=} h'$ implies $\Pi(h) = \Pi(h')$.

### C.3 Extension to $\mathcal{L}^p$ spaces and to $\mathcal{F}_0$

In the following proofs, continuity for risk preferences on $\mathcal{L}^p$ spaces is with respect to $p$-norm convergence, continuity for risk preferences on the space $\mathcal{F}_0$ of *simple random payoffs* – those that take, almost surely, only finitely many values – is with respect to bounded a.s. convergence.

As discussed in the main text only the proofs of the results concerning propensity to full insurance (Theorems 1 and 4) need to be modified by adding the assumption of continuity, the ones regarding propensity about partial insurance remain unchanged.

**Proof of Theorem 1 for continuous risk preferences on $\mathcal{F}_0$ and $\mathcal{L}^p$, with $p \in [1, \infty)$.**

(i) $\implies$ (ii). The proof is the same as that in Appendix B.6.

(ii) $\implies$ (i). Let $f \in \mathcal{F}_0$ (resp. $\mathcal{L}^p$). Choosing $v$ and $\Sigma_n^v$ as in Lemma 5,

$$f_n := \mathbb{E}[f \mid \Sigma_n^v] \to f \tag{24}$$

in bounded a.s. convergence (resp. in $\mathcal{L}^p$). It is obvious to see that $f_n \in \mathcal{F}_0 \subseteq \mathcal{L}^p$ for all $n \in \mathbb{N}$. Theorem 1, applied to the restriction of $\succsim$ to $\mathcal{F}_0\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right) = \mathcal{L}^p\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right) = \mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$, yields $\mathbb{E}[f_n] \succsim f_n$ for all $n \in \mathbb{N}$. But $\mathbb{E}[f_n] \to \mathbb{E}[f]$ and $f_n \to f$, and the continuity of $\succsim$ implies $\mathbb{E}[f] \succsim f$. Thus weak risk aversion holds. ■

**Proof of Theorem 4 for continuous risk preferences on $\mathcal{F}_0$ and $\mathcal{L}^p$, with $p \in [1, \infty)$.**

(i) $\implies$ (ii). The proof is the same as that in Appendix B.14.

(ii) $\implies$ (i). Let $f \in \mathcal{F}_0$ (resp. $\mathcal{L}^p$) and $\gamma \in \mathbb{R}$ be such that $\gamma \succsim_{\mathrm{A}} f$. We want to show that $\gamma \succsim_{\mathrm{B}} f$. Define $\{f_n\}$ as in (24). Note that $\gamma \sim_{\mathrm{A}} f_n - \rho_{\mathrm{A}}(\gamma, f_n)$ for all $n \in \mathbb{N}$. Theorem 4, applied to the restriction of $\succsim$ to $\mathcal{F}_0\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right) = \mathcal{L}^p\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right) = \mathcal{L}^\infty\left(S, \Sigma_n^v, P_{|\Sigma_n^v}\right)$, yields $\gamma \succsim_{\mathrm{B}} f_n - \rho_{\mathrm{A}}(\gamma, f_n)$ for all $n \in \mathbb{N}$. But then

$$f_n - \rho_{\mathrm{B}}(\gamma, f_n) \sim_{\mathrm{B}} \gamma \succsim_{\mathrm{B}} f_n - \rho_{\mathrm{A}}(\gamma, f_n)$$

and, together with transitivity, monotonicity implies that $\rho_{\mathrm{B}}(\gamma, f_n) \leq \rho_{\mathrm{A}}(\gamma, f_n)$ for all $n \in \mathbb{N}$.[21] Since $\succsim$ is continuous and $f_n \to f$ suitably, it follows, by Lemma 12,[22] that $\rho_{\mathrm{A}}(\gamma, f_n) \to \rho_{\mathrm{A}}(\gamma, f)$ and $\rho_{\mathrm{B}}(\gamma, f_n) \to \rho_{\mathrm{B}}(\gamma, f)$, and so $\rho_{\mathrm{B}}(\gamma, f) \leq \rho_{\mathrm{A}}(\gamma, f)$. But

$$f - \rho_{\mathrm{A}}(\gamma, f) \sim_{\mathrm{A}} \gamma \succsim_{\mathrm{A}} f = f - 0$$

and another application of transitivity and monotonicity yields $\rho_{\mathrm{A}}(\gamma, f) \leq 0$, and so $\rho_{\mathrm{B}}(\gamma, f) \leq \rho_{\mathrm{A}}(\gamma, f) \leq 0$. With this (monotonicity again)

$$\gamma \sim_{\mathrm{B}} f - \rho_{\mathrm{B}}(\gamma, f) \succsim_{\mathrm{B}} f - 0 = f$$

and (transitivity again) $\gamma \succsim_{\mathrm{B}} f$, as desired. ■

## References


[1] Arrow, Kenneth J.. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 5: 941–973.

[2] Arrow, Kenneth J.. 1974. "Optimal Insurance and Generalized Deductibles." *Scandinavian Actuarial Journal* 1974: 1–42.

[3] Artzner, Philippe, Freddy Delbaen, Jean-Marc Eber, and David Heath. 1999. "Coherent Measures of Risk." *Mathematical Finance* 9: 203–228.


[21] Monotonicity is equivalent to

$$f - \zeta \succsim f - \xi \iff \zeta \leq \xi$$

whenever $f \in \mathcal{F}_0$ and $\zeta, \xi \in \mathbb{R}$.

[22] Adjusted for the corresponding type of convergence of sequences of random payoffs.

[4] Berti, Patrizia, Luca Pratelli, and Pietro Rigo. 2007. “Skorohod Representation on a Given Probability Space.” *Probability Theory and Related Fields* 137: 277–288.

[5] Bickel, Peter J., and Erich L. Lehmann. 1976. “Descriptive Statistics for Nonparametric Models III: Dispersion.” *Annals of Statistics* 4: 1139–1158.

[6] Billingsley, Patrick. 1995. *Probability and Measure*. 3rd ed. Wiley.

[7] Cerreia-Vioglio, Simone, David Dillenberger, and Pietro Ortoleva. 2015. “Cautious Expected Utility and the Certainty Effect.” *Econometrica* 83: 693–728.

[8] Chew, Soo Hong. 1989. “Axiomatic Utility Theories with the Betweenness Property.” *Annals of Operations Research* 19: 273–298.

[9] Cohen, Michele. 1995. “Risk-Aversion Concepts in Expected- and Non-Expected-Utility Models.” *The Geneva Papers on Risk and Insurance Theory* 20: 73–91.

[10] de Finetti, Bruno. 1931. “Sul concetto di media.” *Giornale dell’Istituto Italiano degli Attuari* 2(3): 369–396.

[11] Dekel, Eddie. 1986. “An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom.” *Journal of Economic Theory* 40: 304–318.

[12] Dickson, David C.M. 2017. *Insurance Risk and Ruin*. Cambridge University Press.

[13] Dubra, Juan, Fabio Maccheroni, and Efe A. Ok. 2004. “Expected Utility Theory without the Completeness Axiom.” *Journal of Economic Theory* 115: 118–133.

[14] Durrett, Rick. 2019. *Probability: Theory and Examples*. 5th ed. Cambridge University Press.

[15] Epstein, Larry G., and Stephen M. Tanny. 1980. “Increasing Generalized Correlation: A Definition and Some Economic Consequences.” *Canadian Journal of Economics* 13: 16–34.

[16] Föllmer, Hans, and Alexander Schied. 2016. *Stochastic Finance*. de Gruyter.

[17] Gollier, Christian. 2001. *The Economics of Risk and Time*. MIT Press.

[18] Gul, Faruk. 1991. “A Theory of Disappointment Aversion.” *Econometrica* 59: 667–686.

[19] Hansen, Lars P., and Thomas J. Sargent. 2008. *Robustness*. Princeton University Press.

[20] Kahneman, Daniel, and Amos Tversky. 1979. “Prospect Theory: An Analysis of Decision under Risk.” *Econometrica* 47: 263–292.

[21] Kallenberg, Olav. 2002. *Foundations of Modern Probability*. 2nd ed. Springer.

[22] Mackey, George W.. 1957. “Borel Structure in Groups and Their Duals.” *Transactions of the American Mathematical Society* 85: 134–165.

[23] Machina, Mark J.. 1982. “Expected Utility Analysis without the Independence Axiom.” *Econometrica* 50: 277–323.

[24] Machina, Mark J., and David Schmeidler. 1992. “A More Robust Definition of Subjective Probability.” *Econometrica* 60: 745–780.

[25] Markowitz, Harry M.. 1952. “Portfolio Selection.” *Journal of Finance* 7: 77–91.

[26] Müller, Alfred, and Dietrich Stoyan. 2002. *Comparison Methods for Stochastic Models and Risks*. Wiley.

[27] Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2020. “Stochastic Dominance under Independent Noise.” *Journal of Political Economy* 128: 1877–1900.

[28] Pollard, David. 2002. *A User's Guide to Measure Theoretic Probability.* Cambridge University Press.

[29] Pratt, John W.. 1964. "Risk Aversion in the Small and in the Large." *Econometrica* 32: 122–136.

[30] Puccetti, Giovanni, and Ruodu Wang. 2015. "Extremal Dependence Concepts." *Statistical Science* 30: 485–517.

[31] Quiggin, John. 1982. "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization* 3: 323–343.

[32] Rostek, Marzena. 2010. "Quantile Maximization in Decision Theory." *Review of Economic Studies* 77: 339–371.

[33] Rothschild, Michael, and Joseph E. Stiglitz. 1970. "Increasing Risk: I. A Definition." *Journal of Economic Theory* 2: 225–243.

[34] Ross, Stephen A.. 1981. "Some Stronger Measures of Risk Aversion in the Small and the Large with Applications." *Econometrica* 49: 621–638.

[35] Rüschendorf, Ludger. 2013. *Mathematical Risk Analysis.* Springer.

[36] Savage, Leonard. 1954. *The Foundations of Statistics.* Wiley.

[37] Schmeidler, David. 1989. "Subjective Probability and Expected Utility without Additivity." *Econometrica* 57: 571–587.

[38] Schmidt, Ulrich, and Horst Zank. 2008. "Risk Aversion in Cumulative Prospect Theory." *Management Science* 54: 208–216.

[39] Shaked, Moshe, and J. George Shanthikumar. 2007. *Stochastic Orders.* Springer.

[40] Simons, Gordon. 1977. "An Unexpected Expectation." *Annals of Probability* 5: 157–158.

[41] Tversky, Amos, and Daniel Kahneman. 1992. "Advances in Prospect Theory: Cumulative Representation of Uncertainty." *Journal of Risk and Uncertainty* 5: 297–323.

[42] Wakker, Peter. 1994. "Separating Marginal Utility and Probabilistic Risk Aversion." *Theory and Decision* 36: 1–44.

[43] Wang, Bin, and Ruodu Wang. 2015. "Extreme Negative Dependence and Risk Aggregation." *Journal of Multivariate Analysis* 136: 12–25.

[44] Yaari, Menahem E. 1969. "Some Remarks on Measures of Risk Aversion and on Their Uses." *Journal of Economic Theory* 1: 315–329.

[45] Yaari, Menahem E. 1987. "The Dual Theory of Choice under Risk." *Econometrica* 55: 95–115.