Quasi-convex Risk Measures and Acceptability Indices. Theory and Applications.

by

Ilaria Peri

MSc. in Economics and Finance (University of Milan - Bicocca) 2004

Dissertation submitted for the degree of
Doctor of Philosophy

in

Mathematics for the Analysis of Financial Markets (XXIV cycle)

in the
UNIVERSITY OF MILAN - BICOCCA

Supervisor:
Professor Marco Frittelli

Ph.D Coordinator:
Professor Giovanni Zambruno

January 2012

To Flavio,

for his love and respect.

Contents

Acknowledgments

I am deeply grateful to my supervisor, Professor Marco Frittelli, without his guidance and patience this thesis would not have been possible. His encouraging, understanding and precious suggestions have been of great value for me.

I would like to thank my tutor, Professor Fabio Bellini, for his continuous interest and support in my formation during the entire course of study. Interacting with him has contributed to my professional and personal growth.

A very special thank goes to my colleagues and friends of the Mathematics Department of the Milan University, in particular Marco Maggis, Lara Charawi and Francesca Tantardini, who have helped me in my work, supported and make me feel more "mathematician" than I thought.

I also wish to thank Ludovico Cavedon of UCSB, for his big help with Pyton and my colleagues and friends of the DIMEQUANT of Milano Bicocca University, that have enriched this experience.

My sincere thank goes to Flavio, who has always stood by me, with sincere love, patience and respect, during the moments of difficulty and in the most important choices of my life. Without his support I would have never taken the decision to quit my job and understood my true way. My special gratitude is due to my brother, Alessandro, who, despite of the distance, is always present when I need some suggestion, and to my family for their support and understanding.

Finally, I warmly thank Flavia Barsotti, who in a short while has truly understand me, and Sara, my closest friend from the high school time.

Chapter 1

Introduction

Risk measurement has always been a crucial topic involving both regulators and financial institutions. To this end, several measures were introduced, beginning with the variance, later replaced by the popular 'Value at Risk'. Only at the end of the twentieth century, Artzner, Delbaen, Eber and Heath (1999) [ADEH99] introduced the concept of a *coherent risk measure*. This seminal paper lighted up a debate on the set of axioms that a reasonable risk measure should have satisfied. So in the last ten years numerous mathematical extensions and variations were proposed.

One of the first generalization was the notion of *convex risk measure* introduced by Föllmer and Schied (2002) [FS02] and Frittelli and Rosazza (2002) [FR02], stressing the role of convexity as a counterpart of the fundamental diversification principle. Later, El Karoui and Ravanelli in 2009 [ER09], argued that one of the axioms of the convex risk measures, the cash-additivity, fails whenever there is any form of uncertainty about interest rates, for instance due to the lack of liquidity of such assets. Since this condition is quite common in the real financial markets, they suggested to replace cash-additivity with *cash-subadditivity*. Therefore, in a recent paper, Cerreia-Vioglio, Maccheroni, Marinacci, Montrucchio (2011) [CMMMa] have observed that if the cash-additivity is replaced with cash-subadditivity, then convexity should be relaxed in favour of *quasiconvexity,* in order to maintain the original interpretation in terms of diversification.

As well as the risk measurement, another important topic that has collected scientific interest is the assessment of the performance of the financial positions. Again in the spirit of an axiomatic approach, Cherny and Madan (2009) [CM09] proposed a class of quasiconcave performance measures, called acceptability index, pointing out a link to the coherent risk measures. For this reason we call this class of performance measures, '*coherent acceptability indices*'.

This thesis is organized in two parts: in the first one, we treat, in a unique chapter, risk measures and acceptability indices from a theoretical point of view, while in the second part, we propose two applications of the quasi-convex results.

The starting point of this thesis is an overview of the several risk measures introduced in the literature, with a particular regard to the axiomatic characterizations and the relation with the sets of acceptable positions (Section 2.1). One of the fundamental aspect of the risk measures concerns their dual representation (treated in the Section 2.2).

After the presentation of the dual results regarding the convex risk measures (Section 2.2.1), we focus our attention to the quasiconvex case (Section 2.2.2). In a general setting one important contribute has been provided by Penot and Volle [PV90] and Volle [Vo98]. However, this duality resulted incomplete. Hence, recently, Cerreia Vioglio et al.

[CMMMb] and later Drapeau and Kupper [DK10] have addressed this problem. Both of them obtained a complete duality involving monotone and quasi-convex real valued functions. [CMMMb] provided a solution under fairly general condition covering both the case of maps that are quasi-convex lower semicontinuous and quasi-convex upper semicontinuous maps, whereas [DK10] treated the case of quasi-convex lower semicontinuous maps under different assumptions on the vector space.

The first contribute of this thesis, reported in Subsection 2.2.2, has been to compare their representations and prove that they coincide. In particular, we explain how is possible to shift from the [CMMMb]'s representation to the [DK10]'s one. The key step is provided by the Proposition 34. On the light of this comparison, we also propose in Subsection 2.2.3 a new representation for the *quasiconcave* and *monotone acceptability indices*, as stated in the Theorem 40.

A particular class of risk measures, studied in literature, is represented by the *law invariant risk measures*. In order to facilitate the comparison with risk measures defined on distributions, proposed in the Chapter 3, we report, in the Subsection 2.2.3, the main results of Kusuoka [K01], Frittelli and Rosazza [FR05], Jouini et al. [JST06], Svindland [S10] and the recent dual representation for the quasi-convex case by Cerreia-Vioglio et al. [CMMMa] and later Drapeau et al. [DKR10].

In the second part of the thesis we propose two different applications of the quasi-convex analysis to different sectors. The common idea has been to build a quasiconvex risk measure defining a particular *family of acceptability sets*, taking inspiration from the papers of Cherny and Madan (2009) [CM09] and Drapeau and Kupper (2010) [DK10].

Cherny and Madan (2009) [CM09] pointed out that, for a cash additive coherent risk measure, all the positions can be split in two classes: acceptable and not acceptable; in contrast, for an acceptability index there is a whole continuum of degrees of acceptability defined by a system $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of sets. This formulation has been further investigated by Drapeau and Kupper (2010) [DK10] for the quasi convex case.

Adopting this approach, we introduce, in the Chapter 3, a new class of *law invariant risk measures,* directly defined on the set $\mathcal{P}(\mathbb{R})$ of probability measures on $\mathbb{R}$, that are monotone and quasi-convex on $\mathcal{P}(\mathbb{R})$. We build the maps $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ from a family $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of acceptance sets of distribution functions by defining:

$$\Phi(P) := -\sup\{m : P \in \mathcal{A}^m\}.$$

We study the properties of such maps, we provide some specific examples and in particular we propose a generalization of the classical notion of the Value at Risk, $V@R_\lambda$, that, in spite of its drawbacks, keeps on being used by many financial institutions.

The key idea of our proposal - the definition of the $\Lambda V@R$ - arises from the consideration that in order to assess the risk of a financial position it is necessary to consider not only the probability λ of the loss, as in the case of the $V@R_\lambda$, but the dependence between such *probability* λ and the *amount* of the loss. In other terms, we want to model the fact that a risk adverse agent is willing to accept greater losses only with smaller probabilities. Hence, we replace the constant λ with an increasing function $\Lambda : \mathbb{R} \to [0,1]$ defined on losses, which we call *Probability/Loss function.* The balance between the probability and the amount of the losses is incorporated in the definition of the family of acceptance sets

$$\mathcal{A}^m := \{Q \in \mathcal{P}(\mathbb{R}) \mid Q(-\infty, x] \leq \Lambda(x),\ \forall x \leq m\},\ m \in \mathbb{R}.$$

If P_X is the distribution function of the random variable X, our new measure is defined by:

$$\Lambda V@R(P_X) := -\sup\{m \mid P(X \le x) \le \Lambda(x),\ \forall x \le m\}.$$

This map is not any more translation invariant, but we obtain another similar property.

Furthermore we provide some interesting results about the dual representation.

We propose a further application of the quasiconcavity to a sector of particular interest for the academic community: the *scientific research evaluation.* In the recent years the evaluation of the scientists' performance has become increasingly important. In fact, most crucial decisions regarding faculty recruitment, accepting research projects, research time, academic positions, travel money, award of grants and promotions depend on great extent upon the scientific merits of the involved researchers.

In the Chapter 3 we introduce the issue of the evaluation of the scientific performance. We discuss about the different methodologies and analyze the several critics. A particular attention has been given to the recent document of the Italian ANVUR, which also proposes to use some bibliometric indices as selection parameters of the candidates, who aim to obtain the national qualification of full or associate professor, and their referees. In this chapter we also present an historical overview of the bibliometric indices and the axiomatic approach proposed in literature.

Differently from any existing approach, our formulation is clearly germinated from the theory of risk measures. We adopt the same approach of the seminal paper of [ADEH99] in order to determine a good class of scientific performance measures, that we call *Scientific*

Research Measures (SRM).

Also in this second application, proposed in the Chapter 5, the key idea underlying our definition of SRM is the representation of quasi-concave monotone maps in terms of a *family of acceptance sets*.

The *SRM* of an author is the map $\phi_{\mathbb{F}}$ with values in $[0, \infty]$ that associates to each author X a performance given by:

$$\begin{aligned}\phi_{\mathbb{F}}(X) \quad &: = \sup\{q \in \mathcal{I} \mid X \in \mathcal{A}_q\} \\ &= \sup\{q \in \mathcal{I} \mid X(x) \geq f_q(x) \text{ for all } x \in \mathbb{R}\}.\end{aligned}$$

where $\{\mathcal{A}_q\}_q$ is a *family of performance sets* associated to particular family of *performance curves* $\{f_q\}_q$ that has some specific properties. Different choices of the family $\{f_q\}_q$ lead to different SRM $\phi_{\mathbb{F}}$.

Through this approach, we propose a class of SRMs that are:

- *flexible* in order to fit peculiarities of different areas and ages;
- *inclusive*, as they comprehends several popular indices;
- *calibrated* to the particular scientific community;
- *coherent*, as they share the same structural properties - based on an axiomatic approach;
- *granular,* as they allow a more precise comparison between scientists and are based on *the whole citation curve* of a scientist.

A new interesting approach to the whole area of bibliometric indices is provided by the *dual representation* of a SRM.

In this chapter we also show the method to compute a particular SRM, called ϕ-index, and we report some empirical results obtained by *calibrating* the performance curves to a specific data set (built using Google Scholar).

Part I

Part: Theory

Chapter 2

Risk Measures and Acceptability Indices

The risk measure theory has been developed in order to find a reasonable assessment method for the riskiness of financial positions. For a long time this was a concern that involved both financial institutions and regulators, the first ones for the financial risk management and the second ones in order to safeguard the bank solvency and the overall economic stability.

The classical method of *financial risk evaluation* based on the variance was inadequate, as, for example, it did not keep into account the asymmetry of the financial positions. So in the second half of 90's, after the stock market crash of 1987, the popular Value at Risk $V@R$ was developed and diffused by several financial institutions in order to consider the downside risk, indeed it is based on a quantile of the lower tail of the profit and loss distribution. Nevertheless, the $V@R$ has some deficiencies and especially doesn't satisfy some natural requirements. This was the main reason that led several scientists to study a set of axioms that a *reasonable risk measure* has to satisfy, instead of analyzing each single measure of risk.

Hence, on one hand a risk measure is the mean by calculating the risk of a financial position, on the other hand it represents a *capital requirement*, namely the minimal amount of capital, required by the regulator, to put aside in order to guard against the risk of financial positions, or in other words, to make it acceptable.

The aim of this chapter is to recall some concepts, well known in literature, related to the risk measure theory (Section 2.1), with a particular focus on the quasi-convex risk measures and the acceptability indices, recently introduced. In Section 2.2 we report their dual representations and a first theoretical contribute of this thesis to the quasi-convex case. We conclude this chapter with a prior to the dynamic setting (Section 2.3).

2.1 Risk Measures and the Acceptability Indices

In this section we present an historical overview of the several risk measures, with particular attention to the interpretation of the axioms and the relation to the set of the acceptable positions. We prove only the main results and we suggest to refer to the text Föllmer and Shield [FS04] for a more detail study.

Let us give some notations. $(\Omega, \mathcal{F}, \mathbb{P})$ is the probability space. $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P})$ is the space of $\mathcal{F}$-measurable random variables defined on Ω and that take value in $\bar{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ and $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$ is its quotient space in respect to the $\sim$ $\mathbb{P}$-a.s. $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ is the space of $\mathcal{F}$-measurable random variables $\mathbb{P}$-a.s. bounded. Any equality and inequality in this section has to be considered $\mathbb{P}$-a.s. valid.

In these examples we consider a *static* approach to the risk, since we assumed that the risk depends only on the changing of the financial position value at the maturity T. Hence there exists only one period of uncertainty $(0, T)$.

A *financial position* is a random variable $X : (\Omega, \mathcal{F}) \to \mathbb{R}$ where $X(\omega)$ represents the discounted value of the position at the maturity T in case the status of nature $\omega \in \Omega$ happens. Let $\mathcal{X}$ be the set of the financial positions such that $L^\infty \subseteq \mathcal{X} \subseteq L^0$.

The first axiomatization of the risk measures has been provided by Artzner, Delbaen, Eber and Heath (1999) [ADEH99]. They introduced the notion of *coherent risk measure* as follows:

Definition 1 (Coherent RM) *A map $\rho : \mathcal{X} \to \mathbb{R}$ is a coherent risk measure if the following properties hold, for any $X, Y \in \mathcal{X}$:*

1. *decreasing monotonicity:* $X \leq Y \Rightarrow \rho(X) \geq \rho(Y)$*;*
2. *cash additivity:* $\rho(X + m) = \rho(X) - m, \quad \forall\ m \in \mathbb{R}$*;*
3. *positive homogeneity :* $\rho(\lambda X) = \lambda \rho(X), \quad \forall\ \lambda \geq 0$ *and* $\lambda \in \mathbb{R}$*;*
4. *subadditivity:* $\rho(X + Y) \leq \rho(X) + \rho(Y)$*.*

The financial meaning of these conditions is simple:

1. *decreasing monotonicity*: a financial position Y that in any state of nature assumes greater values than another one X rationally must have a lower risk;

2. *cash additivity* or *translation invariance*: an additional cash amount m to a financial position X makes its risk lower by exactly this amount. This property is strictly linked to the interpretation of $\rho(X)$ as a capital requirement. Indeed, considering a position X *acceptable* if its risk is such that $\rho(X) \leq 0$, the cash additivity axiom suggests we should interpret $\rho(X)$ as the cash amount that added to the position X makes it acceptable, since:

$$\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0.$$

3. *positive homogeneity*: scaling a financial position X scales the risk by the same factor λ;

4. *subadditivity*: the risk of the sum of two positions X and Y is smaller than the sum of their respective risks.

Later, Föllmer and Schied (2002) [FS02] and Frittelli and Rosazza (2002) [FR02] extended the concept of coherent risk measure by relaxing the subadditivity and positive homogeneity conditions in favour of the weaker **convexity** requirement, which allows to *control the risk* of a convex combination by the combination of each single risk:

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y), \quad \forall\lambda \in [0, 1]$$

This requirement clearly expresses the principle that "diversification should not increase the risk", supposing that $\lambda X + (1 - \lambda)Y$ represents the diversified position obtained investing the fraction λ of the initial wealth in the position X and the remaining part in the second alternative Y. This observation leads to define a new class of risk measure.

Definition 2 (Convex RM) *A map* $\rho : \mathcal{X} \to \mathbb{R}$ *is a convex risk measure if satisfies the conditions of decreasing monotonicity, cash additivity, convexity and normalization* ($\rho(0) = 0$)*.*

As anticipated:

$$\rho \textit{ coherent} \text{ risk measure} \Rightarrow \rho \textit{ convex} \text{ risk measure}$$

Now, we want to recall an important relation between the risk measures up to now introduced and the so called, *acceptance set*.

Remark 3 *Any* cash additive *map* $\rho : \mathcal{X} \to \mathbb{R}$ *can be written as the minimal capital requirement:*

$$\rho(X) := \inf \{ m \in \mathbb{R} \, | m + X \in \mathcal{A}_\rho \}$$

where the set

$$\mathcal{A}_\rho := \{ X \in \mathcal{X} \mid \rho(X) \leq 0 \ \} \, . \tag{2.1}$$

is called the acceptance set *of* ρ*, that is the set of all "acceptable positions" in the sense that they do not require additional capital in order to be acceptable in term of risk. Vice versa given an acceptance set* $\mathcal{A} \subset \mathcal{X}$ *it is possible to associate a cash additive map* $\rho_\mathcal{A}$

$$\rho_\mathcal{A}(X) := \inf \{ m \in \mathbb{R} \, | m + X \in \mathcal{A} \} \tag{2.2}$$

that evidently represents the minimal amount of capital m *that added to the position* X *makes it acceptable.*

Hence, the properties of a cash additive risk measure can be deduced by those of the acceptance set and viceversa. The following remark points out these considerations.

Remark 4 *Let $\rho : \mathcal{X} \to \mathbb{R}$ be a cash additive and monotone map and $\mathcal{A} \subseteq \mathcal{X}$.*

1. *If ρ is positively homogeneous. Then $\mathcal{A}_\rho$ defined as in (2.1) is a cone, i.e.*

$$X \in \mathcal{A} \Rightarrow \lambda X \in \mathcal{A} \quad \forall \lambda \geq 0.$$

Viceversa, if $\mathcal{A}$ is a cone, then $\rho_{\mathcal{A}}$ defined as in (2.1) is positively homogeneous.

2. *If ρ is convex. Then $\mathcal{A}_\rho$ defined as in (2.1) is convex, i.e.*

$$X, Y \in \mathcal{A} \Rightarrow \lambda X + (1-\lambda)Y \in \mathcal{A} \quad \forall \lambda \in (0,1).$$

Viceversa, if $\mathcal{A}$ is convex, then $\rho_{\mathcal{A}}$ defined as in (2.1) is convex.

Now, we provide some examples of coherent and convex risk measures and the case of the $V@R$.

Example 5 (Worst-Case risk measure) *The* worst-case risk measure *of the financial position $X \in \mathcal{X}$ is defined by*

$$\rho_w(X) := -ess \inf_{\Omega}(X) = ess \sup_{\Omega}(-X)$$

and represents the smaller capital amount that has to be added to X in order to make the losses null in the worst case scenario. It is the most conservative risk measure and the acceptance set associated to ρ_w is the set of all non-negative functions in $\mathcal{X}$

$$\mathcal{A}_{\rho_w} = \mathcal{X}_+ := \{X \in \mathcal{X} \mid X \geq 0\}$$

that is a convex cone. Hence, ρ_w is a coherent risk measure.

Example 6 (Entropic risk measure) *Given the utility function*

$$u_\gamma(x) := 1 - e^{-\gamma x}$$

for every coefficient of risk aversion $\gamma > 0$*, the* entropic risk measure *of* X *is the capital requirement associated to the acceptance set*

$$\mathcal{A}_{u_\gamma} := \{X \in \mathcal{X} \mid E[u_\gamma(X)] \geq u_\gamma(0) = 0\}$$

and it is defined by

$$\begin{aligned} \rho_{\mathcal{A}_{u_\gamma}}(X) \quad &: \quad = \inf\left\{m \in \mathbb{R} \,|m + X \in \mathcal{A}_{u_\gamma}\right\} \\ &= \quad \frac{1}{\gamma} \ln\left(E[e^{-\gamma X}]\right). \end{aligned}$$

It is a convex risk measure but it is not coherent since A_{u_γ} *is not a cone.*

Example 7 ($V@R_\lambda$**)** *The* Value at Risk *at level* $\lambda \in (0,1)$ *of a financial position* X*, is defined by*

$$\begin{aligned} V@R_\lambda(X) \quad &: \quad = \inf\{m \ \in \mathbb{R} \mid P[X + m < 0] \leq \lambda\} \\ &= \quad -\sup\{m \ \in \mathbb{R} \mid P[X \leq m] \leq \lambda\} \\ &= \quad -q_X^+(\lambda) \end{aligned}$$

where $q_X^+(\lambda)$ *is the upper quantile function of the distribution of* X*. It represents the the smallest amount of capital which, if added to* X *and invested in the risk-free asset, keeps the probability of a negative outcome below the level* λ*. However, Value at Risk only controls the probability of a loss; it does not capture the size of such a loss if it occurs. The* $V@R_\lambda$ *is decreasing monotone, cash additive and positively homogeneous, but in general it is not* *subadditive, hence it is not a coherent risk measure. Furthermore, it is not even a convex risk measure, indeed the acceptance set*

$$\mathcal{A}_{V@R_\lambda} = \{X \in \mathcal{X} \mid P[X < 0] \leq \lambda\}$$

is not convex. So the $V@R_\lambda$ *may penalize diversification instead of encouraging it. In some case, it penalizes the increase of the probability that something goes wrong, without rewarding the significant reduction of the expected loss conditional on the event of default. Nevertheless it is often used by banking institutions. They usually make the hypothesis that* X *is a Normal random variable with variance* σ^2. *In such case* $V@R_\lambda(X)$ *is convex and it is given by*

$$V@R_\lambda(X) = E[-X] + N^{-1}(1-\lambda)\sigma(X)$$

where N^{-1} *is the inverse of the distribution function* $N(0,\ 1)$.

Example 8 (WCE_λ**)** *The* worst conditional expectation *at level* $\lambda \in (0,1)$ *of a financial position* X *is defined by*

$$WCE_\lambda(X) := \sup\{E[-X \mid A] \mid A \in \mathcal{F} \text{ and } P(A) > \lambda\}$$

is a coherent risk measure on $\mathcal{X} := L^\infty$.

Example 9 ($AV@R_\lambda$**)** *The* Average Value at Risk *at level at level* $\lambda \in (0,1]$ *of a position* X *is defined by*

$$AV@R_\lambda(X) := -\frac{\int_0^\lambda q_X^+(s)ds}{\lambda} = \frac{\int_0^\lambda V@R_s(X)ds}{\lambda}. \tag{2.3}$$

It is a coherent risk measure on $\mathcal{X} := L^\infty$ *(see the proof in THM 4.47 of [FS04]) with representation*

$$AV@R_\lambda(X) = \sup_{Q \in \mathcal{Q}_\lambda} E_Q[-X] \tag{2.4}$$

where $\mathcal{Q}_\lambda := \left\{ Q \ll P \mid \frac{dQ}{dP} \leq \frac{1}{\lambda} \right\}$.

We highlight the following relations, in respect to:

- *the worst case measure:*

$$AV@R_0(X) = V@R_0(X) = -ess\inf(X) = \rho_w(X)$$

- *the mean of the losses:*

$$AV@R_1(X) = -\int_0^1 q_X^+(s)ds = E[-X]$$

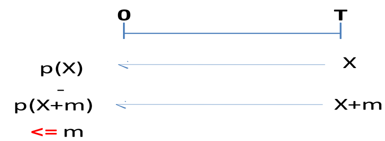
- *the worst conditional expectation:*

$$\begin{aligned} AV@R_\lambda(X) &\geq WCE_\lambda(X) \\ &\geq E[-X \mid -X \geq V@R_\lambda(X)] \\ &\geq V@R_\lambda(X) \end{aligned}$$

Moreover, if $P[X \leq q_X^+(\lambda)] = \lambda$*, namely* X *has a continuous distribution, the first two inequalities are in fact identities:*

$$AV@R_\lambda(X) = WCE_\lambda(X) = E[-X \mid -X \geq V@R_\lambda(X)].$$

Another progress in the risk measure theory has been done by El Karoui and Ravanelli in 2009 [ER09]. They criticized the cash additivity axiom since it requires that risky positions and risk measures (as reserve amounts) are expressed in the same numéraire ($\rho(X + \rho(X)) = 0$). This means that risky positions are discounted before applying the risk measure, assuming that the discounting process does not involve any additional risk. This is not realistic since the interest rates are uncertain, it is enough to think about the



illiquidity or the defaultability of zero coupon bonds. Thus, the strong assumption of the cash additivity is to consider the existence and the liquidity of a non-defaultable zero coupon bond with price $D \in (0,1]$, maturity T and nominal value 1 such that

$$\rho(X+m) = \rho(X) - Dm \quad \text{for any } m \in \mathbb{R}.$$

These considerations led to relax cash additivity in favour of the **cash subadditivity.** The meaning of this requirement is: "when m dollars are added to the future position X, the capital requirement today is reduced by less than m dollars", i.e.

$$\rho(X+m) \geq \rho(X) - m \quad \text{for any } m \in \mathbb{R}.$$

In other words the difference between the present capital requirement of the position X and that of the position X augmented by the cash amount m at the time T has to be less then m.

Therefore they proposed a new risk measure maintaining the convexity and the decreasing monotonicity but replacing cash-additivity with cash-subadditivity.

Recently Cerreia-Vioglio, Maccheroni, Marinacci, Montrucchio (2011) [CMMMa] have observed that if the cash-additivity is replaced with cash-subadditivity, then convexity should be relaxed in favour of quasiconvexity in order to maintain the original interpretation in terms of diversification. Indeed, this principle is literally expressed in this way: "if positions X and Y are less risky than Z, so it is any diversified position $\lambda X + (1-\lambda)Y$

within $\lambda \in (0,1)$"; that translated using the risk measure ρ is equal to

$$\text{if } \rho(X), \rho(Y) \leq \rho(Z) \text{ then } \rho(\lambda X + (1-\lambda)Y) \leq \rho(Z) \quad \forall \lambda \in (0,1)$$

This condition is equal to convexity only if we consider the cash-additivity assumption, otherwise in general it is only equal to the **quasiconvexity** requirement (also under cash-subadditivity).

Hence, they proposed the notion of *quasi-convex cash-subadditive risk measure.*

Definition 10 (quasi-convex cash-subadditive RM) *A map $\rho : \mathcal{X} \to \mathbb{R}$ is a* quasi-convex cash-subadditive risk measure *if satisfies the condition of decreasing monotonicity, cashsubadditivity, quasiconvexity and normalization ($\rho(0) = 0$).*

The first relevant mathematical findings on quasi-convex functions were provided by De Finetti [DeFin]. According to the classical mathematical representation, a map $\rho : \mathcal{X} \to \mathbb{R}$ is said to be quasi-convex if:

$$\textit{the lower level sets } \{X \in \mathcal{X} \mid \rho(X) \leq c\} \quad \forall c \in \mathbb{R} \textit{ are convex.}$$

Another characterization of the quasiconvexity condition is:

$$\rho(\lambda X + (1-\lambda)Y) \leq \max\{\rho(X); \rho(Y)\} \quad \forall \lambda \in [0,1] \ \text{ and } \lambda \in \mathbb{R} \tag{2.5}$$

Indeed, if the lower level sets are convex and $\overline{c} := \max[f(x); f(y)]$. Then $f(\lambda x + (1-\lambda)y) \leq \overline{\alpha} = \max[f(x); f(y)]$ for every $\lambda \in [0,1]$. Conversely, let $L(\rho, c)$ be any lower level set of ρ and $X, Y \in L(\rho, c)$. Then $\rho(X) \leq c$ and $\rho(Y) \leq c$ and by the 2.5 it follows that $\rho(\lambda X + (1-\lambda)Y) \leq c$ for every $\lambda \in [0,1]$. Hence $L(\rho, c)$ is a convex set and ρ is quasi-convex.

Remark 11 *It is clear that if ρ is convex it is also quasi-convex, indeed for any $\lambda \in (0,1)$*

$$\rho(\lambda X + (1-\lambda)Y) \leq \lambda\rho(X) + (1-\lambda)\rho(Y) \leq \max\{\rho(X); \rho(Y)\}$$

and it is easy to prove that

$$\text{quasiconvexity } \textit{and} \text{ cashadditivity} \Rightarrow \textit{convexity}$$

This is not true for quasiconvexity and cash-subadditivity.

Example 12 (Certainty Equivalent) *Let $\mathcal{X} = L^\infty$. The* certainty equivalent *of $X \in L^\infty$, and , defined as*

$$\rho(X) := l^{-1}(E[l(-X)])$$

where $l : \mathbb{R} \to \mathbb{R}$ is a continuous increasing loss function, is a quasi-convex risk measure.

In the mean time Cherny and Madan (2009) [CM09] proposed a class of performance measures of financial risk positions, called *acceptability index*, by formulating a set of axioms that such measure should satisfy. They defined it setting $\mathcal{X} = L^\infty$.

Definition 13 (Acceptability Index) *An* acceptability index *is a map $\alpha : L^\infty \to [0,\infty]$ that satisfies the following axioms, for any $X, Y \in L^\infty$:*

1. *increasing monotonicity: $X \leq Y \Rightarrow \alpha(X) \leq \alpha(Y)$;*

2. *quasi-concavity: $\alpha(\lambda X + (1-\lambda)Y) \geq \min\{\alpha(X),\ \alpha(Y)\}$ for all $\lambda \in [0,1]$;*

3. *scale invariance: $\alpha(X) = \alpha(\lambda X)$ for all $\lambda > 0$;*

4. *Fatou property: for any bounded sequence X_n which converges P-a.s. to some X, then*

$$\alpha(X) \leq \liminf_{n\to+\infty} \alpha(X_n).$$

The above axioms have a natural financial interpretation:

- the *increasing monotonicity* means that if the financial position Y dominates X in every states of nature then Y performs more or equal to X;
- the *quasiconcavity* states that a diversified portfolio $\lambda X + (1-\lambda)Y$ performs at higher level than its components X and Y;
- the *scale invariance* implies that the level of acceptance doesn't change if we scale the financial position;
- the *Fatou property* is a technical continuity property, which is used for constructing the duality between the acceptability indices and coherent risk measures.

A further characterization of the quasiconcavity requirement is given by the following remark.

Remark 14 *Properties of quasi-concave maps can be deduced by those of the quasi-convex maps, observing that*

$$\alpha \text{ quasi-concave} \Leftrightarrow -\alpha \text{ quasi-convex.} \tag{2.6}$$

Hence, like in the quasi-convex case, α is quasi-concave if

$$\text{the upper level sets } \{X \in \mathcal{X} \mid \alpha(X) \geq c\} \quad \forall c \in \mathbb{R} \text{ are convex.}$$

Cherny and Madan pointed out also that for a risk measure all the positions are split in two classes, acceptable and not acceptable, in contrast, for an acceptability index we have a whole continuum of degrees of acceptability defined by the system $\{\mathcal{A}^m\}_{m\in\mathbb{R}^+}$ and the index measures the degree of acceptability of a trade. Hence, an acceptability index can be built not just via one single acceptance set $\mathcal{A}$ but via an *acceptability system* $\{\mathcal{A}^m\}_{m\in\mathbb{R}^+}$ such that

$$\mathcal{A}^m := \{X \in L^\infty \mid \alpha(X) \geq m\}$$

that is the set of those positions which have acceptability level above m. This set is a convex cone since $\alpha(x)$ is quasi-concave and scale invariance.

In particular, Cherny and Madan also found a one to one correspondence between the acceptability index on L^∞ and a particular acceptability system $\{\mathcal{A}^m\}_{m\in\mathbb{R}^+}$ by mean of the following theorem.

Theorem 15 (Thm.1 [CM09]) *A map $\alpha : L^\infty \to [0, \infty]$ is an acceptability index if and only if there exists an increasing family $\{\mathcal{Q}_m\}_{m\in\mathbb{R}^+}$ of subsets of probabilities such that*

$$\begin{aligned} \alpha(X) &= \sup\left\{m \in \mathbb{R}^+ \mid X \in \mathcal{A}^m\right\} \\ \mathcal{A}^m &:= \left\{X \in L^\infty \mid \inf_{Q\in\mathcal{Q}_m} E_Q[X] \geq 0\right\} \end{aligned}$$

Let $\{\rho_m\}_{m\in\mathbb{R}^+}$ be a family of coherent risk measures increasing in m, then we can write

$$\mathcal{A}^m := \{X \in L^\infty \mid \rho_m(X) \leq 0\}$$

and the acceptability index $\alpha(X)$ can be represented as

$$\alpha(X) = \sup\left\{m \in \mathbb{R}^+ \mid \rho_m(X) \leq 0\right\}.$$

In this version, the acceptability index of a position X represents the greatest performance level m such that X is acceptable in terms of risk at that particular level q. For this reason we start to refer to this index with the term of *coherent acceptability index*.

In the following examples we show some performance measures, but only some of them are coherent acceptability indices.

Example 16 (SR) *The* Sharpe Ratio *of a financial position X, $SR(X)$, introduced by Sharpe (1964) [Sh64], is the ratio of the mean $E[X]$ to the standard deviation $\sigma(X)$:*

$$SR(X) := \begin{cases} \frac{E[X]}{\sigma(X)} & \text{if } E[X] > 0 \\ 0 & \text{otherwise} \end{cases}$$

We exclude the negative values in order to consider positive performance measures. It is easy to show that this measure is quasi-concave. However, the Sharpe Ratio does not satisfy the monotonicity property and hence is not a coherent acceptability index.

Example 17 (GLR) *The* Gain-Loss Ratio *of a financial position X, $GLR(X)$, introduced by Bernardo and Ledoit (2000) [BL00], is defined by the ratio of the mean to the expectation of the negative tail:*

$$GLR(X) := \begin{cases} \frac{E[X]}{E[X^-]} & \text{if } E[X] > 0 \\ 0 & \text{otherwise} \end{cases}$$

where $X^- = \max\{-X, 0\}$. This performance measure is a coherent acceptability index, since it is increasing monotone, quasi-concave, scale invariant and satisfies the Fatou Property.

In their work of 2009 Cherny and Madan also proposed some coherent acceptability indices, we recall the $RAROC(X)$ and the $AIT(X)$.

Example 18 (RAROC) *The* Coherent Risk-Adjusted Return on Capital *of a financial position X, $RAROC(X)$, is defined as the ratio of the mean to the coherent risk measure ρ:*

$$RAROC(X) := \begin{cases} \frac{E[X]}{\rho(X)} & \text{if } E[X] > 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.7}$$

By convention, if $\rho(X) \leq 0$ *then the* $RAROC(X) = +\infty$. *The* $RAROC$ *is a coherent acceptability index.*

Example 19 (AIT) *The* TVAR Acceptability Index *of* X, $AIT(X)$, *is defined by:*

$$AIT(X) := \sup\left\{m \in \mathbb{R}_+ : AV@R_{\frac{1}{1+m}}(X) \leq 0\right\}$$

where $AV@R_\lambda$ *is the Average Value at Risk at level* $\lambda \in (0,1]$ *as defined in (2.4) that is a coherent risk measure. If* X *has a continuous distribution, it is possible to obtain a more direct representation:*

$$AIT(X) = \left(\inf\left\{\lambda \in (0,1] \mid E[-X \mid -X \geq V@R_\lambda(X)] \leq 0\right\}\right)^{-1} - 1.$$

Based on [CM09] and on [CMMMa], Drapeau and Kupper (2010) [DK10] defined a quasi-convex risk measure via an **acceptability system**.

Theorem 20 (Thm. 1.7 [DK10]) *Given* $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$, *as collection of subsets* $\mathcal{A}^m \subseteq \mathcal{X}$ *such that*

1. $\mathcal{A}^m$ *is convex for any* m, *that is*

$$\text{if } Y, X \in \mathcal{A}^m \Rightarrow \lambda X + (1-\lambda)Y \in \mathcal{A}^m \quad \forall \lambda \in (0,1)$$

2. $\{\mathcal{A}^m\}$ *is monotone increasing with respect to* m, *that is*

$$\text{if } m \leq n \Rightarrow \mathcal{A}^n \subseteq \mathcal{A}^m$$

3. $\mathcal{A}^m$ *is monotone for any* m, *that is*

$$Y \geq X \in \mathcal{A}^m \Rightarrow Y \in \mathcal{A}^m$$

it is possible to associate a monotone decreasing and quasi-convex map $\rho : \mathcal{X} \to \mathbb{R}$ *defined by:*

$$\rho(X) := \inf \{m \in \mathbb{R} \mid X \in \mathcal{A}^m\} .$$

Vice versa, to any monotone decreasing and quasi-convex map $\rho : \mathcal{X} \to \mathbb{R}$ *we may associate a family* $\left\{\mathcal{A}_\rho^m\right\}_{m\in\mathbb{R}}$ *of* acceptance sets *(satisfying 1, 2, 3)*

$$\mathcal{A}_\rho^m := \{X \in \mathcal{X} \mid \rho(X) \leq m\}$$

such that:

$$\rho(X) := \inf \left\{m \in \mathbb{R} \mid X \in \mathcal{A}_\rho^m\right\} .$$

2.2 On the Dual Representations

Let us start this section with some notation and definition useful to our end.

Notation 21 $(\mathcal{X}, \tau)$ *is a locally convex topological vector and ordered space and* $(\mathcal{X}, \tau)'$ *its topological dual space, that is the vector space consisting of the* τ*-continuous linear functionals.*

In the dual pairing $(\mathcal{X}, \mathcal{X}')$*, the bilinear form* $\langle \cdot, \cdot \rangle : \mathcal{X} \times \mathcal{X}' \to \mathbb{R}$ *is given by* $\langle X, \mu \rangle$ *and the linear function* $X \mapsto \langle X, \mu \rangle = \mu(X)$*, with* $\mu \in \mathcal{X}'$*, is* τ*-continuous.*

Given a function $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$. The effective domain of f, denoted by $Dom(f)$, is defined as:

$$Dom(f) := \{X \in \mathcal{X} \mid f(X) < \infty\}$$

Definition 22 *A function $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ is said to be τ-lower semicontinuous (τ-l.s.c.) if the lower level sets*

$$\{X \in \mathcal{X} \mid f(X) \leq c\} \quad \forall c \in \mathbb{R}$$

are τ-closed.

If f is τ-lower semicontinuous (τ-l.s.c.), then $-f$ is τ-upper semicontinuous (τ-u.s.c.), i.e. the upper level sets

$$\{X \in \mathcal{X} \mid f(X) \geq c\} \quad \forall c \in \mathbb{R}$$

are τ-closed.

It is easy to prove that a further characterization of τ-l.s.c. functions is:

$$\forall X \in \mathcal{X} \text{ s.t. } X_n \xrightarrow{\tau} X \Rightarrow f(X) \leq \lim \inf_{n \to +\infty} f(X_n).$$

Naturally, f is a τ-u.s.c. function if and only if:

$$\forall X \in \mathcal{X} \text{ s.t. } X_n \xrightarrow{\tau} X \Rightarrow f(X) \geq \lim \sup_{n \to +\infty} f(X_n).$$

Any τ-continuous function is both τ-l.s.c. and τ-u.s.c.

Definition 23 *Let $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ be a function such that the $Dom(f) \neq \emptyset$. The conjugate function $f^* : \mathcal{X}' \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ is the function:*

$$f^*(\mu) := \sup_{X \in \mathcal{X}} \{\mu(X) - f(X)\}$$

and whenever $Dom(f^*) \neq \emptyset$, the biconjugate of f, f^{**}, is defined as the conjugate of f^*:

$$f^{**}(X) := \sup_{\mu \in \mathcal{X}'} \{\mu(X) - f^*(\mu)\}$$

2.2.1 Coherent and Convex Risk Measures

The fundamental theorem on which is founded the dual representation of the convex risk measure is the Fenchel-Moreau Theorem. For the proof we refer to the Theorem 5 in Rockafellar 1974 [ROC].

Theorem 24 (Fenchel-Moreau) *Let $f : \mathcal{X} \to \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ be convex, lower semi-continuous and $Dom(f) \neq \emptyset$, then f^{**} is well defined and*

$$f = f^{**}.$$

By this theorem, Frittelli Rosazza 2002 [FR02] derived some important consequences starting from the additional properties that a risk measure has to satisfy.

Theorem 25 ([FR02]) *Let $f : \mathcal{X} \to \mathbb{R} \cup \{\infty\}$ be convex, $\sigma(\mathcal{X}, \mathcal{X}')$-lower semicontinuous and $Dom(f) \neq \emptyset$. Then there exists a map $\alpha : \mathcal{X}' \to \mathbb{R} \cup \{\infty\}$ convex and $\sigma(\mathcal{X}, \mathcal{X}')$-lower semicontinuous such that*

$$f(X) = \sup_{\mu \in Dom(\alpha)} \{\mu(X) - \alpha(\mu)\}.$$

Suppose also that $f(0) = 0$. Furthermore:

a) *if f is monotone increasing $\Rightarrow Dom(\alpha) \subseteq \mathcal{X}'_+$*

b) *if f is such that $f(X + c) = f(X) + c$ for any $X \in \mathcal{X}$ and $c \in \mathbb{R} \Rightarrow Dom(\alpha) \subseteq \{\mu \in \mathcal{X}' \mid \mu(1_\Omega) = 1\}$*

c) *if f satisfies a) and b) $\Rightarrow Dom(\alpha) \subseteq \left\{\mu \in \mathcal{X}'_+ \mid \mu(1_\Omega) = 1\right\}$*

d) *if f is positively homogeneous* $\Rightarrow$

$$\alpha(\mu) = \begin{cases} 0 & \mu \in Dom(\alpha) \\ +\infty & \mu \in \mathcal{X}' \backslash Dom(\alpha) \end{cases}$$

Now we recall the dual representation of the convex risk measures in case of $\mathcal{X} = L^{\infty}(\Omega, \mathcal{F}, P)$, where $(\Omega, \mathcal{F}, P)$ is a non atomic probability space.

We denote with $ba := ba(\Omega, \mathcal{F}, P)$ the space of all the signed charge $\mu : \mathcal{F} \to \mathbb{R}$ [1]of bounded variation ($V_\mu < +\infty$) and absolutely continuous with respect to P (i.e. $\mu \ll P$)[2]:

$$ba(\mathbb{R}) := \{\mu \text{ signed charge} \mid V_\mu < +\infty \text{ and } \mu \ll P\}$$

where the variation of μ is

$$V_\mu = \sup \left\{ \sum_{i=1}^{n} |\mu(A_i)| \mid \{A_1, ..., A_n\} \text{ partition of } \Omega \right\}$$

Under the pointwise algebraic operations of addition and scalar multiplication:

$$(\mu + \nu)(A) := \mu(A) + \nu(A) \quad \text{and} \quad \alpha\mu(A) := (\alpha\mu)(A) \quad \forall A \in \mathcal{F}$$

the pointwise ordering, $\geq$:

$$\mu \geq \nu \quad \text{if} \quad \mu(A) \geq \nu(A) \quad \forall A \in \mathcal{F}$$

and the variation norm $\| \mu \| := V_\mu$, the space of charge ba is a Banach space (and a lattice) and the dual space of $(L^{\infty}(\Omega, \mathcal{F}, P), \| \cdot \|_\infty)$ (see the theorem 10.53 in [Ali]):

$$ba(\Omega, \mathcal{F}, P) = (L^{\infty}(\Omega, \mathcal{F}, P), \| \cdot \|_\infty)'.$$

[1] A *signed charge* $\mu : \mathcal{F} \to \mathbb{R}$ is any set function that is finitely additive and such that $\mu(\emptyset) = 0$. A *charge* is a nonnegative signed charge. Remember also that a *measure* is a charge that is countably additive.

[2] A signed charge ν is *absolutely continuous* with respect to another signed charge μ, written $\nu \ll \mu$, if for each $\varepsilon > 0$ there exists some $\delta > 0$ such that $A \in \mathcal{F}$ and $|\mu|(A) < \delta$ imply $|\nu|(A) < \varepsilon$.

Furthermore, it is easy to show that any convex risk measure ρ is $\|\cdot\|_\infty$-l.s.c., since it is $\|\cdot\|_\infty$-continuous, infact the decreasing monotonicity and the cash additivity properties imply that any convex risk measure is $\|\cdot\|_\infty$-Lipschitz continuous, i.e.:

$$\mid \rho(X) - \rho(X_n) \mid \leq \| X - X_n \|_\infty \quad \text{for any } n \in \mathbb{N}.$$

Under such assumptions the Fenchel Moreau Theorem 24 guarantees the dual representation of any convex risk measure ρ, by the simple substitution $\rho(X) = f(-X)$.

Theorem 26 ([FR02] and [FS02]) *A convex risk measure $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ admits the following representation*

$$\rho(X) = \sup_{\mu \in ba_+(1)} \{\mu(-X) - \alpha(\mu)\}$$

where

$$ba_+(1) = \{\mu \in ba_+ : \mu(1_\Omega) = 1\}$$

and $\alpha : ba_+(1) \to \mathbb{R} \cup \{+\infty\}$ is called "penalty function".

In particular, if ρ is a coherent risk measure then

$$\rho(X) = \sup_{\mu \in ba_+(1)} \{\mu(-X)\}.$$

On the other hand, if we endowed $L^\infty(\Omega, \mathcal{F}, P)$ with the weak topology $\sigma(L^\infty, L^1)$ we have that $L^1 := L^1(\Omega, \mathcal{F}, P)$ is the topological dual:

$$L^1 = (L^\infty, \sigma(L^\infty, L^1))'$$

Furthermore, the Radon-Nikodym theorem let us identify the set of the probability density with the set $\mathcal{Q}$ of the probability Q absolutely continuous with respect to P, so:

$$\mathcal{Q} := \{Q \ll P\} = \left\{ \frac{dQ}{dP} \in L^1_+ \mid E\left[\frac{dQ}{dP}\right] = 1 \right\}$$

In such case, it is possible to derive the dual representation of the convex risk measures by the Theorem 24 if, in addition, the $\sigma(L^\infty, L^1)$-l.s.c. holds.

Theorem 27 ([FR02] and [FS02]) *Any convex risk measure $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ that is* $\sigma(L^\infty, L^1)$-l.s.c. *admits the following representation:*

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{E_Q(-X) - \alpha(Q)\}$$

where $\mathcal{Q} := \{Q \ll P\}$ and $\alpha : \mathcal{Q} \to \mathbb{R} \cup \{+\infty\}$ is the "penalty function".

In particular, if ρ is a coherent risk measure and $\sigma(L^\infty, L^1)$-l.s.c. then:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} E_Q(-X)$$

Before moving to the dual representation of the quasi-convex risk measures we highlight a useful characterization of the $\sigma(L^\infty, L^1)$-lower semicontinuity for the convex risk measures (for the proof we refer to [FS04]).

Theorem 28 *Let $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \mathbb{R}$ be a convex risk measure. $X_n, X \in L^\infty(\Omega, \mathcal{F}, P)$ for any $n \in \mathbb{N}$. TFAE:*

i) *ρ is $\sigma(L^\infty, L^1)$-l.s.c;*

ii) *$\mathcal{A}_\rho$ is $\sigma(L^\infty, L^1)$-closed;*

iii) *ρ satisfies the* Fatou property*: that is $\|X_n\|_\infty \leq k \in \mathbb{R}$ for all $n \in \mathbb{N}$ and $X_n \to X$ P-a.s., then*

$$\rho(X) \leq \liminf_{n \to +\infty} \rho(X_n)$$

iv) *ρ is continuous from above:*

$$X_n \downarrow X \quad P\text{-a.s. then } \rho(X_n) \uparrow \rho(X) \quad P\text{-a.s.}$$

Remark 29 (Rmk 4.23 [FS04]) *Let ρ be a convex measure of risk which is continuous from below. Then ρ is also continuous from above.*

2.2.2 Quasi-convex Risk Measures

In literature we find several contributes to the dual representation of quasi-convex functions. In a general setting, such representation was provided by Penot and Volle [PV90] and later reformulated by Volle [Vo98] in Th.3.4.

Penot-Volle Duality

Theorem 30 ([Vo98], [PV90]) *Let $f : \mathcal{X} \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ be quasi-convex and lower semicontinuous. Then*

$$f(X) = \sup_{\mu \in \mathcal{X}'} F(\mu, \mu(X)) \tag{2.8}$$

where $F : \mathcal{X}' \times \mathbb{R} \to \overline{\mathbb{R}}$ is defined by

$$F(\mu, t) := \inf_{\xi \in \mathcal{X}} \{f(\xi) \mid \mu(\xi) \geq t\}.$$

This duality is, however, incomplete. In fact, there is no uniqueness: to any quasi-convex function f is possible to associate multiple functions $F(\mu, t)$. As a result, the duality is only one directional: to a function f we can associate a function like $F(\mu, t)$, but not vice versa.

More recently, Cerreia Vioglio et al. [CMMMb] and later Drapeau and Kupper [DK10] addressed this problem. In both case the main result is a complete duality involving monotone and quasi-convex real valued functions. [CMMMb] provided a solution under fairly general conditions covering both the case of maps that are quasi-convex lower semicontinuous and quasi-convex upper semicontinuous, whereas [DK10] treated the case of quasi-convex lower semicontinuous maps under different assumptions on the vector space. Since both of them found a unique representation, we have compared their representations and proved that they coincide. We begin by reporting the result in [CMMMb], then we explain how it is possible to move from this representation to the one presented in [DK10]. We report the assumption under which Drapeau and Kupper obtained their result and we conclude with a new theorem for the complete monotone duality of the quasi-convex and monotone maps.

Cerreia-Vioglio, Maccheroni, Marinacci, and Montrucchio Duality

Assumptions on the vector space in [CMMMb]

If $(\mathcal{X}, \geq)$ is an ordered normed space, we denote with $\mathcal{X}_+$ its positive cone $\{X \in \mathcal{X} \mid X \geq 0\}$ and by $\mathcal{X}'_+$ the set of all positive functionals in $\mathcal{X}'$. We also set

$$\Delta := \left\{\mu \in \mathcal{X}'_+ : \|\mu\| = 1\right\}$$

In the sequel, $\mathcal{X}'$ and any of its subsets will be always equipped with the weak* topology.

Let $\mathcal{X}$ be a M space with unit[3] e. We recall that an M-space is a normed Riesz

[3] A positive element e is a *unit* if for all $X \in \mathcal{X}$ there is some $\lambda \geq 0$ such that $|X| \leq \lambda e$.

space[4] equipped with an M-norm:

$$\|X \vee Y\| = \max\{\|X\|, \|Y\|\} \quad \forall X, Y \in \mathcal{X}_+.$$

We refer to [Ali], ch.9 for a detail study of the M-space. We only remember that any normed Riesz space with order unit e can be turned into an M-space, provided e is interior to the positive cone $\mathcal{X}'_+$. The sup norm $\|X\|_e = \inf\{\lambda > 0 \mid |X| \leq \lambda e\}$ generated by e is actually an equivalent M-norm.

If $\mathcal{X}$ is an M-space with unit, its closed unit ball is $[-e, e] = \{X \in \mathcal{X} \mid -e \leq X \leq e\}$. Hence $\|\mu\| = \mu(e)$ for all $\mu \in \mathcal{X}'_+$, and so

$$\Delta_e = \left\{\mu \in \mathcal{X}'_+ : \mu(e) = 1\right\},$$

which is therefore a convex and weak* compact set. We denote e with 1 and

$$\Delta := \Delta_1 = \left\{\mu \in \mathcal{X}'_+ : \mu(1) = 1\right\}.$$

The authors provided the dual representation of a more general class of function, the so-called *evenly quasi-concave* functions. The first notion of even convexity and its basic properties are due to Fenchel [Fe]. We need to recall some definitions.

Definition 31 *A subset C of $\mathcal{X}$ is* evenly convex *if it is the intersection of a family of open half spaces. With the convention that such intersection is $\mathcal{X}$ if the family is empty.*

It is also well known that a set C is evenly convex if and only if for each $\bar{X} \notin C$ there is $\bar{\mu} \in \mathcal{X}' \diagdown \{0\}$ such that

$$\bar{\mu}(\bar{X}) < \bar{\mu}(X) \quad \forall X \in C$$

[4] A Riesz space is an ordered vector space that is also a lattice.

By standard separation results, both open convex sets and closed convex sets are evenly convex.

Set $\mathbb{R}^{\Diamond} := \mathbb{R} \setminus \{0\}$, a subset C of $\Delta \times \mathbb{R}$ is said to be $\Diamond$-*evenly convex* if and only if for each $(\bar{\mu}, \bar{t}) \in \Delta \times \mathbb{R} \setminus C$ there exists $(X, s) \in \mathcal{X} \times \mathbb{R}^{\Diamond}$ such that

$$\bar{\mu}(X) + \bar{t}s < \mu(X) + ts \quad \forall(\mu, t) \in C.$$

Here is required that s is nonzero, which is stronger than requiring that both s and X are nonzero. As a result, $\Diamond$-even convexity is slightly more than an extension to products of topological vector spaces of the notion of even convexity. Clearly, $\Diamond$-evenly convex sets are evenly convex.

Definition 32 *A function $f : \mathcal{X} \to \overline{\mathbb{R}}$ is:*

i) evenly quasi-convex *if all the lower sets $\{X \in \mathcal{X} | f(X) \leq c\}$ are evenly convex $\forall c \in \mathbb{R}$;*

ii) evenly quasi-concave *if all the upper sets $\{X \in \mathcal{X} | f(X) \geq c\}$ are evenly convex $\forall c \in \mathbb{R}$;*

Similarly, a function f defined on $\Delta \times \mathbb{R}$ is:

$\Diamond$***i)*** $\Diamond$-evenly quasi-convex *if its lower level sets are $\Diamond$-evenly convex;*

$\Diamond$***ii)*** $\Diamond$-evenly quasi-concave *if its upper level sets are $\Diamond$-evenly convex.*

Note that $\Diamond$-evenly quasi-convex (quasi-concave) functions on $\Delta \times \mathbb{R}$ are evenly quasi-convex (quasi-concave) on $\mathcal{X}' \times \mathbb{R}$.

The following relations are easy to show:

$$f \text{ evenly quasi-convex} \Rightarrow f \text{ quasi-convex}$$

$$f \text{ quasi-convex and u.s.c} \;\Rightarrow\; f \text{ evenly quasi-convex}$$

$$f \text{ quasi-convex and l.s.c} \;\Rightarrow\; f \text{ evenly quasi-convex}$$

The same results hold for the quasi-concave case (observing (2.6)).

It is clear that, under their assumptions, the authors covered both case of:

- quasi-convex risk measures ρ *upper semicontinuous*;
- quasi-convex risk measures ρ *lower semicontinuous.*

On the class of dual functions

Let $\mathcal{H} := M^{\Diamond}_{qcx}(\Delta\times\mathbb{R})$ be the class (defined in [CMMMb]) of functions $H : \Delta\times\mathbb{R} \to \overline{\mathbb{R}}$ such that:

H1) $H(\mu, \cdot)$ is increasing over $\mathbb{R}$, for all $\mu \in \Delta$;

H2) for all $\mu, \nu \in \Delta$

$$\sup_{t\in\mathbb{R}} H(\mu, t) = \sup_{t\in\mathbb{R}} H(\nu, t)$$

H3) the function $(\mu, t) \to H(\mu, t)$ is $\Diamond-$evenly quasi-convex on $\Delta \times \mathbb{R}$;

H4)

$$\inf_{\mu\in\Delta} H^{+}(\mu, \mu(X)) = \inf_{\mu\in\Delta} H(\mu, \mu(X)).$$

where

$$H^{+}(\mu, t) := \inf_{p>t} H(\mu, p)$$

is the right continuous version of $H(\mu, \cdot)$, which coincides with the upper semicontinuous envelope of $H(\mu, \cdot)$.

Theorem 33 ([CMMMb]) *If $g : \mathcal{X} \to \overline{\mathbb{R}}$ is quasi-concave monotone increasing and upper semicontinuous then the only $H \in \mathcal{H}$ such that*

$$g(X) = \inf_{\mu \in \Delta} H(\mu, \mu(X)) \tag{2.9}$$

is given by

$$H(\mu, t) := \sup_{\xi \in \mathcal{X}} \{g(\xi) \mid \mu(\xi) \leq t\}. \tag{2.10}$$

Conversely, for every $H \in \mathcal{H}$ there is a unique quasi-concave monotone increasing upper semicontinuous $g : \mathcal{X} \to \overline{\mathbb{R}}$ such that (2.10) holds and g is given by (2.9).

From the [CMMMb] Duality to the [DK10] Duality

Our contribute to the dual representation of quasi-convex risk measure is given by the following proposition, that allow to pass from the [CMMMb] to the [DK10] complete duality.

Proposition 34 *Let $g : \mathcal{X} \to \overline{\mathbb{R}}$ and $H : \Delta \times \mathbb{R} \to \overline{\mathbb{R}}$ be defined by*

$$H(\mu, t) := \sup_{\xi \in \mathcal{X}} \{g(\xi) \mid \mu(\xi) \leq t\}. \tag{2.11}$$

The right continuous version of $H(\mu, \cdot)$ can be written as:

$$H^{+}(\mu, t) := \inf_{p > t} H(\mu, p) = \sup \{\beta \in \mathbb{R} \mid \gamma(\mu, \beta) \leq t\}, \tag{2.12}$$

where $\gamma : \Delta \times \mathbb{R} \to \overline{\mathbb{R}}$ is given by:

$$\gamma(\mu, \beta) := \inf \{\mu(X) \mid g(X) \geq \beta\}, \quad \beta \in \mathbb{R}.$$

Proof. Let the RHS of equation (2.12) be denoted by

$$S(\mu, t) := \sup \{\beta \in \mathbb{R} \mid \gamma(\mu, \beta) \leq t\}, \quad (\mu, t) \in \Delta \times \mathbb{R},$$

and note that $S(\mu, \cdot)$ is the right inverse of the increasing function $\gamma(\mu, \cdot)$ and therefore $S(\mu, \cdot)$ is right continuous (see [FS04]).

Step I. To prove that $H^{+}(\mu, t) \leq S(\mu, t)$ it is sufficient to show that for all $p > t$ we have:

$$H(\mu, p) \leq S(\mu, p), \tag{2.13}$$

Indeed, if (2.13) is true

$$H^{+}(\mu, t) = \inf_{p>t} H(\mu, p) \leq \inf_{p>t} S(\mu, p) = S(\mu, t),$$

as both H^{+} and S are right continuous in the second argument.

Writing explicitly the inequality (2.13)

$$\sup_{\xi \in \mathcal{X}} \{g(\xi) \mid \mu(\xi) \leq p\} \leq \sup \{\beta \in \mathbb{R} \mid \gamma(\mu, \beta) \leq p\}$$

and letting $\xi \in \mathcal{X}$ satisfying $\mu(\xi) \leq p$, we see that it is sufficient to show the existence of $\beta \in \mathbb{R}$ such that $\gamma(\mu, \beta) \leq p$ and $\beta \geq g(\xi)$. If $g(\xi) = \infty$ then $\gamma(\mu, \beta) \leq p$ for any β and therefore $S(\mu, p) = H(\mu, p) = \infty$.

Suppose now that $\infty > g(\xi) > -\infty$ and define $\beta := g(\xi)$. As $\mu(X) \leq p$, we have:

$$\gamma(\mu, \beta) := \inf \{\mu(X) \mid g(X) \geq \beta\} \leq p$$

Then $\beta \in \mathbb{R}$ satisfies the required conditions.

Step II : To obtain $H^{+}(\mu, t) := \inf_{p>t} H(\mu, p) \geq S(\mu, t)$ it is sufficient to prove that, for all $p > t$, $H(\mu, p) \geq S(\mu, t)$, that is :

$$\sup_{\xi \in \mathcal{X}} \{g(\xi) \mid \mu(\xi) \leq p\} \geq \sup \{\beta \in \mathbb{R} \mid \gamma(\mu, \beta) \leq t\}. \tag{2.14}$$

Fix any $p > t$ and consider any $\beta \in \mathbb{R}$ such that $\gamma(\mu, \beta) \leq t$. By the definition of γ, for all $\varepsilon > 0$ there exists $\xi_\varepsilon \in \mathcal{X}$ such that $g(\xi_\varepsilon) \geq \beta$ and $\mu(\xi_\varepsilon) \leq t + \varepsilon$. Take ε such that $0 < \varepsilon < p - t$. Then $\mu(\xi_\varepsilon) \leq p$ and $g(\xi_\varepsilon) \geq \beta$ and (2.14) follows. ∎

Fix a quasi-convex monotone decreasing and lower semicontinuous map $\rho : \mathcal{X} \to \overline{\mathbb{R}}$. Define

$$\rho(X) := -g(X).$$

Then $g : \mathcal{X} \to \overline{\mathbb{R}}$ is a quasi-concave monotone increasing and upper semicontinuous map. Applying Theorem 33 to the function g and applying the property $H4$ we deduce:

$$\begin{aligned} \rho(X) &= -\inf_{\mu \in \Delta} H^+(\mu, \mu(X))) = \sup_{\mu \in \Delta} \left\{ -H^+(\mu, \mu(X) \right\} \\ &= \sup_{\mu \in \Delta} R(\mu, -\mu(X)), \end{aligned}$$

where $R : \Delta \times \mathbb{R} \to \overline{\mathbb{R}}$ is defined by

$$R(\mu, t) := -H^+(\mu, -t),$$

H^+ is given in (2.12), H in (2.10) and is unique in the class $\mathcal{H}$. Applying (2.12) we get:

$$\begin{aligned} R(\mu, t) &= -H^+(\mu, -t) = \inf \{-\beta \in \mathbb{R} \mid \gamma(\mu, \beta) \leq -t\} \\ &= \inf \{\beta \in \mathbb{R} \mid -\gamma(\mu, -\beta) \geq t\} \\ &= \inf \{\beta \in \mathbb{R} \mid \alpha(\mu, \beta) \geq t\} \end{aligned}$$

where:

$$\begin{aligned} \alpha(\mu, \beta) := \quad -\gamma(\mu, -\beta) = \quad & -\inf \{\mu(X) \mid g(X) \geq -\beta\} \\ = \quad & \sup \{-\mu(X) \mid -g(X) \leq \beta\} = \sup \{\mu(-X) \mid \rho(X) \leq \beta\}. \end{aligned}$$

Let $\mathcal{H}^{lsc} := \left\{R : \Delta \times \mathbb{R} \to \overline{\mathbb{R}} \text{ such that } R(\mu,t) = -H^{+}(\mu,-t),\ H \in \mathcal{H}\right\}$. The following result is then an immediate corollary of Theorem 33.

Corollary 35 *If $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is quasi-convex monotone decreasing and lower semicontinuous then the only $R \in \mathcal{H}^{lsc}$ such that*

$$\rho(X) = \sup_{\mu \in \Delta} R(\mu, \mu(-X)), \tag{2.15}$$

is given by

$$R(\mu,t) := \inf \left\{\beta \in \mathbb{R} \mid \alpha(\mu,\beta) \geq t\right\}, \tag{2.16}$$

with

$$\alpha(\mu,\beta) = \sup \left\{\mu(-X) \mid \rho(X) \leq \beta\right\}.$$

Conversely, for every $R \in \mathcal{H}^{lsc}$ there is a unique quasi-convex monotone decreasing lower semicontinuous $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ such that (2.16) holds and ρ is given by (2.15).

Notice that if $R \in \mathcal{H}^{lsc}$ then R is left continuous (as H^{+} is right continuous) and increasing in the second argument (from assumption $H1$).

Drapeau and Kupper [DK10] Duality

Assumptions on the vector space in [DK10]

Let $\mathcal{X}$ be a locally convex topological vector space and $\mathcal{X}'$ be its dual space. Let $\succeq$ be a total preorder[5] on $\mathcal{X}$. Set the positive cone:

$$\mathcal{X}_{+} := \{X \in \mathcal{X} \mid X \succeq 0\}$$

[5] A total preorder is a transitive and complete binary relation. A binary relation $\succeq$ on $\mathcal{X}$ is transitive if $x \succeq y$ and $y \succeq z$ implies $x \succeq z$, and is complete if $x \succeq y$ or $y \succeq x$ for any $x, y \in \mathcal{X}$.

and remember that the bipolar theorem states that $X \succeq Y$ when $\mu(X-Y) \geq 0$ for all μ in the polar cone:

$$\mathcal{X}'_+ := \left\{\mu \in \mathcal{X}' \mid \mu(X) \geq 0 \text{ for all } X \in \mathcal{X}_+\right\}.$$

Assume that:

1. $\mathcal{X}$ is endowed with the $\sigma(\mathcal{X}, \mathcal{X}')$ topology;

2. $\mathcal{X}_+$ is $\sigma(\mathcal{X}, \mathcal{X}')$-closed;

3. the set

   $$Units := \left\{X \in \mathcal{X}_+ \mid \mu(X) > 0 \text{ for all } \mu \in \mathcal{X}'_+ \setminus \{0\}\right\}$$

   is not empty.

Define the normalized polar set

$$\Delta_e := \left\{\mu \in \mathcal{X}'_+ : \mu(e) = 1\right\}, \quad \text{for } e \in Units$$

As e will be fixed, we denote $e = 1$ and

$$\Delta := \Delta_1$$

On the class of dual functions

Let $\mathcal{R}_0 = \left\{R : \Delta \times \mathbb{R} \to \overline{\mathbb{R}} \text{ such that } R(\mu, \cdot) \text{ is increasing and left-continuous}\right\}$.

Let $\mathcal{R} := \mathcal{R}^{\max}$ be the class (defined in [DK10]) of functions $R : \Delta \times \mathbb{R} \to \overline{\mathbb{R}}$ such that:

R1) $R(\mu, \cdot)$ is increasing over $\mathbb{R}$, for all $\mu \in \Delta$;

R2) for all $\mu, \nu \in \Delta$

$$\inf_{t \in \mathbb{R}} R(\mu, t) = \inf_{t \in \mathbb{R}} R(\nu, t);$$

R3) the function $(\mu, t) \to R(\mu, t)$ is quasi-concave on $\Delta \times \mathbb{R}$;

R4) $R^+(\mu, t) := \inf_{p>t} R(\mu, p)$ is upper semicontinuous in the first argument;

R5) $R(\lambda\mu, t) = R(\mu, t/\lambda)$ for any $\mu \in \Delta$, $t \in \mathbb{R}$ and $\lambda > 0$;

R6) $R(\mu, \cdot)$ is left-continuous over $\mathbb{R}$, for all $\mu \in \Delta$.

Clearly $\mathcal{R} \subset \mathcal{R}_0$.

Theorem 36 ([DK10] Theorem 2.7) *If $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is quasi-convex monotone decreasing and lower semicontinuous then the only $R \in \mathcal{R}$ such that*

$$\rho(X) = \sup_{\mu \in \Delta} R(\mu, \mu(-X)) \tag{2.17}$$

is given by

$$R(\mu, t) := \inf \{\beta \in \mathbb{R} \mid \alpha_{\min}(\mu, \beta) \geq t\}, \tag{2.18}$$

where

$$\alpha_{\min}(\mu, \beta) := \sup \{\mu(-X) \mid \rho(X) \leq \beta\}, \quad \beta \in \mathbb{R}.$$

Conversely, for every $R \in \mathcal{R}$ the function ρ defined in (2.17) is quasi-convex monotone decreasing and lower semicontinuous.

Comparison between [CMMMb] and [DK10] duality and conclusion

Assume that the space $\mathcal{X}$ satisfies the assumptions in both [CMMMb] and [DK10] (for example take: $\mathcal{X} = (L^\infty(\Omega, \mathcal{F}, \mathbb{P}), \|\cdot\|_\infty, \geq)$. From Theorem 36 and Corollary 35 we have:

$$\rho(X) = \sup_{\mu\in\Delta} R(\mu, \mu(-X))$$

where R is given in (2.18) or (2.16) and is unique in the class $\mathcal{R}$ and in the class $\mathcal{H}^{lsc}$. Therefore, R given in (2.18) or (2.16) is unique in the intersection of the two class, $\mathcal{R} \cap \mathcal{H}^{lsc}$.

Conclusion 37 *Let $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ be a quasi-convex monotone decreasing and lower semicontinuous map. Then H given by*

$$H(\mu, t) := \sup_{\xi\in\mathcal{X}} \{-\rho(\xi) \mid \mu(\xi) \le t\} \tag{2.19}$$

belongs to $\mathcal{H}$, the map $\overline{H}$ defined by

$$\overline{H}(\mu, t) := -H^{+}(\mu, -t) \tag{2.20}$$

belongs to $\mathcal{R}$ and satisfies

$$\overline{H}(\mu, t) = \inf \{\beta \in \mathbb{R} \mid \alpha_{\min}(\mu, \beta) \ge t\} \tag{2.21}$$

$$\rho(X) := \sup_{\mu\in\Delta} \overline{H}(\mu, \mu(-X)) = \sup_{\mu\in\Delta} \left\{-H^{+}(\mu, \mu(X)\right\}$$

where

$$\alpha_{\min}(\mu, \beta) := \sup \{\mu(-X) \mid \rho(X) \le \beta\}.$$

The following simple proposition shows that the equation (2.21) holds true.

Proposition 38 *Suppose that H is given in (2.19), $\overline{H}$ in (2.20) and R is given in (2.18) and is left continuous in the second argument. Then*

$$\overline{H}(\mu, t) = R(\mu, t)$$

for all $(\mu, t) \in \Delta \times \mathbb{R}$.

Proof. Step I. To prove that $R(\mu,t) \leq \overline{H}(\mu,t) := -H^{+}(\mu,-t) = \sup_{p<t}[-H(\mu,-p)]$ it is sufficient to show that for all $p < t$ we have: $R(\mu,p) \leq -H(\mu,-p)$, i.e.:

$$\inf\{\beta \in \mathbb{R} \mid \alpha_{\min}(\mu,\beta) \geq p\} \leq \inf_{\xi\in\mathcal{X}}\{\rho(\xi) \mid \mu(-\xi) \geq p\}]. \tag{2.22}$$

Indeed, if (2.22) is true

$$R(\mu,t) = \sup_{p<t} R(\mu,p) \leq \sup_{p<t} -H(\mu,-p) = -\inf_{p<t} H(\mu,-p) = -H^{+}(\mu,-t),$$

as R is left continuous in the second argument.

Let ξ satisfy $\mu(-\xi) \geq p$. Then to prove (2.22) it is sufficient to show that there exists $\beta \in \mathbb{R}$ such that $\alpha_{\min}(\mu,\beta) \geq p$ and $\beta \leq \rho(\xi)$. If $\rho(\xi) = -\infty$ then $\alpha_{\min}(\mu,\beta) \geq p$ for any β and therefore $R(\mu,p) = -H(\mu,-p) = -\infty$.

Suppose now that $\infty > \rho(\xi) > -\infty$ and define $\beta := \rho(\xi)$. As $\mu(-X) \geq p$ we have:

$$\alpha_{\min}(\mu,\beta) := \sup\{\mu(-X) \mid \rho(X) \leq \beta\} \geq p.$$

Then $\beta \in \mathbb{R}$ satisfies the required conditions.

Step II : To obtain $R(\mu,t) \geq \overline{H}(\mu,t) := -H^{+}(\mu,-t) = \sup_{p<t}[-H(\mu,-p)]$ it is sufficient to prove that, for all $p < t$, $R(\mu,t) \geq -H(\mu,-p)$, that is :

$$\inf\{\beta \in \mathbb{R} \mid \alpha_{\min}(\mu,\beta) \geq t\} \geq \inf_{\xi\in\mathcal{X}}\{\rho(\xi) \mid \mu(-\xi) \geq p\}].$$

Fix any $p < t$ and consider any $\beta \in \mathbb{R}$ such that $\alpha_{\min}(\mu,\beta) \geq t$. Then for all $\varepsilon > 0$ there exists $\xi_\varepsilon \in \mathcal{X}$ such that $\rho(\xi_\varepsilon) \leq \beta$ and $\mu(-\xi_\varepsilon) \geq t - \varepsilon$. Take ε such that $0 < \varepsilon < t - p$. Then $\mu(-\xi_\varepsilon) \geq p$ and $\rho(\xi_\varepsilon) \leq \beta$ and the inequality follows. ∎

2.2.3 Quasi-concave Acceptability Indices

In this subsection we provide the dual representation of quasi-concave and monotone increasing maps, on the light of the result in [CMMMb], reported in [33], and the proposition [34]. So that we cover the case of *quasi-concave acceptability indices.*

Definition 39 (Quasi-concave Acceptability Index) *A* quasi-concave acceptability index *is a map* $\alpha : \mathcal{X} \to [0, \infty]$ *that it is increasing monotone and quasi-concave.*

The Fatou property mentioned in the definition [13] is here replaced by another appropriate continuity condition, i.e. the continuous from above (CFA). For the assumption on the space $\mathcal{X}$ we use the Notations 21.

Theorem 40 *Let* $\alpha : \mathcal{X} \to \overline{\mathbb{R}}$ *be a quasi-concave acceptability index, that is continuous from above. Then*

$$\alpha(X) = \inf_{\mu \in \Delta} H(\mu, \mu(X)) = \inf_{\mu \in \Delta} H^{+}(\mu, \mu(X)) \quad \text{for all } X \in \mathcal{X}$$

where

$$\Delta := \left\{ \mu \in \mathcal{X}'_{+} \mid \mu(1) = 1 \right\}$$

$H : \mathcal{X}' \times \mathbb{R} \to \overline{\mathbb{R}}$ *is defined by*

$$H(\mu, t) := \sup_{\xi \in \mathcal{X}} \{\alpha(\xi) \mid \mu(\xi) \leq t\} \quad \text{for all } (\mu, t) \in \mathcal{X}' \times \mathbb{R}$$

and its right continuous version $H^{+}(\mu, \cdot)$*:*

$$H^{+}(\mu, t) := \inf_{s > t} H(\mu, s) = \sup\{\beta \in \mathbb{R} \mid t \geq \gamma(\mu, \beta)\}$$

where $\gamma : \mathcal{X}' \times \mathbb{R} \to \overline{\mathbb{R}}$ *is given by:*

$$\gamma(\mu, \beta) := \inf_{X \in \mathcal{X}} \{\mu(X) \mid \alpha(X) \geq \beta\}, \quad \beta \in \mathbb{R}.$$

Proof. Step 1: $\alpha(X) = \inf_{\mu\in\Delta} H(\mu, \mu(X))$.

Fix $X \in \mathcal{X}$. As $X \in \{\xi \in \mathcal{X} \mid \mu(\xi) \le \mu(X)\}$, by the definition of $H(\mu, t)$ we deduce that, for all $\mu \in \mathcal{X}'$,

$$H(\mu, \mu(X)) \ge \alpha(X)$$

hence

$$\inf_{\mu\in\mathcal{X}'} H(\mu, \mu(X)) \ge \alpha(X).$$

We prove the opposite inequality. Let $\varepsilon > 0$ and define the set

$$C_\varepsilon := \{\xi \in \mathcal{X} \mid \alpha(\xi) \ge \alpha(X) + \varepsilon\}$$

As α is quasi-concave and $\sigma(\mathcal{X}, \mathcal{X}')$-upper semicontinuous, C is convex and $\sigma(X, \mu)-$closed. Since $X \notin C_\varepsilon$, the Hahn Banach theorem implies the existence of a continuous linear functional that strongly separates X and C_ε, that is there exist $k \in \mathbb{R}$ and $\mu_\varepsilon \in \mathcal{X}'$ such that

$$\mu_\varepsilon(\xi) > k > \mu_\varepsilon(X) \text{ for all } \xi \in C_\varepsilon.$$

Hence

$$\{\xi \in \mathcal{X} \mid \mu_\varepsilon(\xi) \le \mu_\varepsilon(X)\} \subseteq C_\varepsilon^c := \{\xi \in \mathcal{X} \mid \alpha(\xi) < \alpha(X) + \varepsilon\}$$

and

$$\begin{aligned}\alpha(X) &\le \inf_{\mu\in\mathcal{X}'} H(\mu, \mu(X)) \le H(\mu_\varepsilon, \mu_\varepsilon(X)) \\ &= \sup\{\alpha(\xi) \mid \xi \in \mathcal{X} \text{ and } \mu_\varepsilon(\xi) \le \mu_\varepsilon(X)\} \\ &\le \sup\{\alpha(\xi) \mid \xi \in \mathcal{X} \text{ and } \alpha(\xi) < \alpha(X) + \varepsilon\} \le \alpha(X) + \varepsilon.\end{aligned}$$

Therefore, $\alpha(X) = \inf_{\mu \in \mathcal{X}'} H(\mu, \mu(X))$. To show that the inf can be taken over the positive cone $\mathcal{X}'_+$, it is sufficient to prove that $\mu_\varepsilon \subseteq \mathcal{X}'_+$. Let $Y \in \mathcal{X}_+$ and $\xi \in C_\varepsilon$. Given that α is monotone increasing, $\xi + nY \in C_\varepsilon$ for every $n \in \mathbb{N}$ and we have:

$$\mu_\varepsilon(\xi + nY) > k > \mu_\varepsilon(X) \Rightarrow \mu_\varepsilon(Y) > \frac{\mu_\varepsilon(X - \xi)}{n} \to 0, \text{ as } n \to \infty.$$

As this holds for any $Y \in \mathcal{X}_+$ we deduce that $\mu_\varepsilon \subseteq \mathcal{X}'_+$. Therefore, $\alpha(X) = \inf_{\mu \in \mathcal{X}'_+} H(\mu, \mu(X))$.

By definition of $H(\mu, t)$,

$$H(\mu, \mu(X)) = H(\lambda\mu, E[X(\lambda\mu)]) \quad \forall \mu \in \mathcal{X}' \text{ and } \lambda \neq 0.$$

Hence we deduce

$$\alpha(X) = \inf_{\mu \in \mathcal{X}'} H(\mu, \mu(X)) = \inf_{\mu \in \cdot} H(\mu, \mu(X)).$$

Step 2: $\alpha(X) = \inf_{\mu \in \cdot} H^+(\mu, \mu(X))$.

Since $H(\mu, \cdot)$ is increasing and $\mu \in \mathcal{X}'_+$ we obtain

$$H^+(\mu, \mu(X)) := \inf_{s > E[\mu X]} H(\mu, s) \leq \lim_{X_m \downarrow X} H(\mu, \mu(X_m)),$$

$$\begin{aligned}
\alpha(X) = \inf_{\mu \in \mathcal{X}'_+} H(\mu, \mu(X)) &\leq \inf_{\mu \in \mathcal{X}'_+} H^+(\mu, \mu(X)) \leq \inf_{\mu \in \mathcal{X}'_+} \lim_{X_m \downarrow X} H(\mu, \mu(X_m)) \\
&= \lim_{X_m \downarrow X} \inf_{\mu \in \mathcal{X}'_+} H(\mu, \mu(X_m)) = \lim_{X_m \downarrow X} \alpha(X_m) \overset{(CFA)}{=} \alpha(X).
\end{aligned}$$

Step 3: $H^+(\mu, t) := \inf_{s > t} H(\mu, s) = \sup \{\beta \in \mathbb{R} \mid t \geq \gamma(\mu, \beta)\}$ follows from 34. ■

2.2.4 Law Invariant Risk Measures

In order to facilitate a comparison with a particular class of risk measures that we introduce in the Chapter 3, we highlight the main results on the dual representation of *law-invariance* risk measures.

In the furthering, we assume that the probability space $(\Omega, \mathcal{F}, P)$ is rich enough to support random variables with continuous distribution. This condition is equivalent to require that $(\Omega, \mathcal{F}, P)$ is atomless (see Prop. A.27 in [FS04]).

We denote with F_X the distribution function of the random variable X with respect to P and we write

$$X \sim_{\mathcal{D}} Y \Leftrightarrow F_X \equiv F_Y$$

The right-continuous inverse of F_X is defined as

$$q^+(t) := \inf\{x \in \mathbb{R} \mid F_X(x) > t\} \qquad \forall t \in [0,1)$$

and it is called the *right-quantile* of X. Further details on quantile functions can be found in the Appendix A.3. of [FS04]. Here we just recall the Remark A.16. of [FS04] for which holds that

$$q^+(t) = \sup\{x \in \mathbb{R} \mid F_X(x) \leq t\}.$$

Definition 41 *A risk measure $\rho : \mathcal{X} \to \overline{\mathbb{R}}$ is law invariant under P if*

$$X \sim_{\mathcal{D}} Y \Rightarrow \rho(X) = \rho(Y)$$

Namely, it assigns the same riskiness to two financial positions $X, Y \in \mathcal{X}$ that are identically distributed with respect to the probability P given a priori.

Here we study the law invariant risk measures defined on the space $\mathcal{X} = L^\infty(\Omega, \mathcal{F}, P)$.

We begin by reporting the result of Kusuoka 2001 [K01] that characterized the class of *law invariant coherent risk measures* with the Fatou property.

Theorem 42 (Thm. 4 [K01]) *A map* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant* coherent risk measure *with the Fatou property if and only if*

$$\rho(X) = \sup_{\mu \in \mathcal{P}((0,1])} \int_{(0,1]} AV@R_s(X)\mu(ds)$$

where $\mathcal{P}((0,1])$ *is the compact convex set of probability measures on* $(0,1]$ *and* $AV@R_s$ *is defined as in [2.3].*

Frittelli and Rosazza 2005 [FR05] generalized the previous result to the case of convex risk measures *continuous from above.* Remember that in the case of convex risk measures on $L^\infty(\Omega, \mathcal{F}, P)$ the Fatou property coincides with the continuity from above (see Lemma 28).

Theorem 43 (Thm.7 [FR05]) *A map* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant* convex risk measure *continuous from above if and only if*

$$\rho(X) = \sup_{\mu \in \mathcal{P}((0,1])} \left(\int_{(0,1]} AV@R_s(X)\mu(ds) - \alpha_{\min}(\mu) \right)$$

where $\mathcal{P}((0,1])$ *is the compact convex set of probability measures on* $(0,1]$, $AV@R_s$ *is defined as in [2.3] and*

$$\begin{aligned} \alpha_{\min}(\mu) &= \sup_{X \in \mathcal{A}_\rho} \int_{(0,1]} AV@R_s(X)\mu(ds) \\ &= \sup_{X \in L^\infty} \left(\int_{(0,1]} AV@R_s(X)\mu(ds) - \rho(X) \right) \end{aligned}$$

The following theorem represents another version of the dual representation of law invariant convex risk measures, that points out the one-to-one correspondence between laws of probability measures μ on $(0,1]$ and the Radon-Nykodim densities $\frac{dQ}{dP}$. Infact, it easy to show (see [FS04] p184) that for any probability measure $\mu \in \mathcal{P}((0,1])$ there is a probability $Q \in \mathcal{Q} := \{Q \ll P\}$ such that

$$\int_{(0,1]} AV@R_s(X)\mu(ds) = \int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t)dt.$$

holds, by defining

$$q_{\frac{dQ}{dP}}(t) := \int_{(1-\lambda,1]} \frac{\mu(ds)}{\lambda}$$

and observing that $q_{-X}(t) = V@R_{1-s}(X)$.

Theorem 44 (Thm.12 [FR05]) *A map* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant* convex risk measure *continuous from above if and only if*

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \left(\int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t)dt - \alpha_{\min}(Q) \right)$$

and the minimal penalty function

$$\begin{aligned} \alpha_{\min}(Q) &= \sup_{X \in \mathcal{A}_\rho} \int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t)dt \\ &= \sup_{X \in L^\infty} \left(\int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t)dt - \rho(X) \right) \end{aligned}$$

An important result for the law invariant convex risk measures has been obtained by Jouini, Schachermayer and Touzi, (2006) [JST06]. They showed that the Fatou property may actually be dropped as it is automatically implied by the hypothesis of law invariance, in other words every law invariant convex risk measure is already weakly lower semicontinuous.

Theorem 45 (Thm.2.2 [JST06]) *If* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant convex risk measure. Then* ρ *is* $\sigma(L^\infty, L^1)-$*l.s.c. and has the Fatou property.*

Recently, Cerreia-Vioglio, Maccheroni, Marinacci and Montrucchio (2011) [CMMMa] have provided a robust dual representation for law invariant quasi-convex risk measures that are *continuous from below.* They also proved that for a quasi-convex risk measure such condition is equal to the weakly upper semicontinuity ($\sigma(L^\infty, L^1)-$u.s.c.).

Theorem 46 (Thm.10 [CMMMa]) *A map* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant* $\sigma(L^\infty, L^1)-u.s.c.$ quasi-convex risk measure *if and only if*

$$\rho(X) = \max_{Q \in \mathcal{Q}} R\left(Q, \int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t) dt\right) \qquad \forall X \in L^\infty$$

where

$$R(Q, s) = \inf\left\{\rho(Y) \mid \int_0^1 q_{\frac{dQ}{dP}}(t) q_Y(1-t) dt = -s\right\} \qquad \forall (Q, s) \in \mathcal{Q} \times \mathbb{R}.$$

In a recent note, Svindland (2010) [S10] has shown that also quasi-convex $\|\cdot\|_\infty -$l.s.c. law invariant risk measures are already weakly lower semicontinuous.

Remark 47 ([S10]) *If* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant* $\|\cdot\|_\infty -l.s.c.$ *quasi-convex risk measure. Then* ρ *is* $\sigma(L^\infty, L^1)-l.s.c.$

The last result we recall is by Drapeau, Kupper and Reda (2010) [DKR10]. They provided a robust representation of law invariant quasi-convex risk measures that now are $\|\cdot\|_\infty -$l.s.c.

Theorem 48 (Thm.3.2 [DKR10]) *A map* $\rho : L^\infty(\Omega, \mathcal{F}, P) \to \overline{\mathbb{R}}$ *is a law invariant and* $\|\cdot\|_\infty -l.s.c.$ quasi-convex risk measure *if and only if there exists a unique* $R \in \mathcal{R}$ *such that*

$$\rho(X) = \max_{q_{\frac{dQ}{dP}} \in \Delta} R\left(q_{\frac{dQ}{dP}}, \int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t) dt\right) \qquad \forall X \in L^\infty$$

where $\Delta := \left\{ q_{\frac{dQ}{dP}} : (0,1] \to [0,+\infty) \mid q_{\frac{dQ}{dP}} \text{ is non decreasing, right-continuous and } \int_0^1 q_{\frac{dQ}{dP}}(t)dt = 1 \right\}$

and

$$R(q_{\frac{dQ}{dP}}, s) = \sup\{m \in \mathbb{R} \mid \alpha_{\min}(Q,m) < s\} \qquad \forall (q_{\frac{dQ}{dP}}, s) \in \Delta \times \mathbb{R}$$

for

$$\alpha_{\min}(q_{\frac{dQ}{dP}}, m) = \sup_{X \in \mathcal{A}^m} \int_0^1 q_{-X}(t) q_{\frac{dQ}{dP}}(t)dt.$$

2.3 Elements of Dynamic Risk Measures and Acceptability Indices

Thus far we have dealt with the risk measures in a static context, arguing the problem of quantifying *today* the riskiness of financial positions with a future maturity at time T. In this section we provide a brief introduction to the concept of *conditional* and *dynamic risk measure.* The main idea is "monitoring" the riskiness of financial positions at different times t between today and the maturity T.

We start by introducing some notation on the basis of this *conditional setting.* Let $(\Omega, \mathcal{F}_T, P)$ be the probability space. The information available to the agent who is assessing the riskiness of the financial position X at the time t is described by a sub σ-algebra $\mathcal{F}_t$ of the total information $\mathcal{F}_T$, so that:

$$\mathcal{F}_t \subseteq \mathcal{F}_T.$$

We denote with $L^p_{\mathcal{F}_T} := L^p(\Omega, \mathcal{F}_T, P)$ for $p > 1$ the set of all real-valued, $\mathcal{F}_T$-measurable and p-integrable random variables, where each element $X \in L^p_{\mathcal{F}_T}$ represents the random payoff to be delivered to an agent at a fixed future date T. While, $L^0_{\mathcal{F}_t} :=$

$L^0(\Omega, \mathcal{F}_t, P)$ is the set of all random variables defined on the probability space $(\Omega, \mathcal{F}_t, P)$, where each element represents the market value of the same financial positions at time t.

Definition 49 ([DS05]) *Given a time $t \in [0, T]$. A conditional risk measure is a map $\rho_t : L^p_{\mathcal{F}_T} \to L^0_{\mathcal{F}_t}$.*

This definition is intuitive, since it is natural that the conditional risk measure is a map assigning to every $\mathcal{F}_T$-measurable random variable X, representing a final payoff, a $\mathcal{F}_t$-measurable random variable $\rho_t(X)$, representing the riskiness of the financial position at time t, .i.e. "conditionally" to the information available in t.

Detlefsen and Scandolo (2004) [DS05] proved that the conditional convex risk measures satisfy a suitable property, called *regularity property.*

Definition 50 *A map $\rho_t : L^p_{\mathcal{F}_T} \to L^0_{\mathcal{F}_t}$ is regular (REG) if $\forall X, Y \in L^p_{\mathcal{F}_T}$ and $\forall A \in \mathcal{F}_t$*

$$\rho_t(X 1_A + Y 1_A^C) = \rho_t(X) 1_A + \rho_t(Y) 1_A^C \quad P-a.s.$$

If ρ_t is such that $\rho_t(0) = 0$, we have the following further characterizations of the regularity property $\forall X, Y, X_n \in L^p_{\mathcal{F}_T}$ and $\forall A \in \mathcal{F}_t$:

i) $\rho_t(X 1_A) = \rho_t(X) 1_A \quad P-a.s.$

ii) $\rho_t(X 1_A + Y 1_A^C) = \rho_t(X) 1_A + \rho_t(Y) 1_A^C \quad P-a.s.$

This property means that if we know that an event $A \in \mathcal{F}_t$ is prevailing, then the riskiness of X should depend only on what is really possible to happen, i.e. on the restriction of X to A.

We also refer to Detlefsen and Scandolo (2004) [DS05] for the definition of conditional convex and coherent risk measures, that in general extend the notion we already know to the conditional case. We only recall that:

Remark 51 *Every conditional convex risk measure is regular.*

In order to represent the continuous assessment in the time interval $[0, T]$ of the riskiness of a final payoff X occurring in T, Wang [W99] introduced the notion of *dynamic risk measure* as a collection of conditional risk measures. If $(\mathcal{F}_t)_{t\in[0,T]}$ is the filtration of σ-algebras describing how the total information $\mathcal{F}_T$ is disclosed through time, such that $(\Omega, \mathcal{F}_T, (\mathcal{F}_t)_{t\in[0,T]}, P)$ is a filtered probability space, then it is natural to define a *dynamic risk measure* as follows:

Definition 52 ([FR04]) *A dynamic risk measure (DRM) is a family* $(\rho_t)_{t\in[0,T]}$ *such that:*

i) $\rho_t : L^p_{\mathcal{F}_T} \to L^0_{\mathcal{F}_t}$ *for all* $t \in [0, T]$*;*

ii) ρ_0 *is a static risk measure;*

iii) $\rho_T(X) = -X \quad P-a.s. \ \forall X \in L^p_{\mathcal{F}_T}$.

Definition 53 *A DRM* $(\rho_t)_{t\in[0,T]}$ *is called CONVEX if it satisfies* $\forall X, Y \in L^p_{\mathcal{F}_T}$*:*

M) *decreasing monotonicity:* $X \geq Y \Rightarrow \rho_t(X) \leq \rho_t(Y) \quad P-a.s. \ \forall\, t \in [0, T]$*;*

TI) *translation invariance:* $\forall\, Z \quad \mathcal{F}_t-$*measurable in* $L^p_{\mathcal{F}_T} \quad \rho_t(X+Z) = \rho_t(X)-Z \quad P-a.s.$ $\forall\, t \in [0, T]$*;*

C) convexity: if $\forall \Lambda \in L^0_{\mathcal{F}_t}$ *s.t.* $0 \leq \Lambda \leq 1$,

$$\rho_t(\Lambda X + (1-\Lambda)Y) \leq \Lambda \rho_t(X) + (1-\Lambda)\rho_t(Y)$$

$P-$*a.s. for any* $t \in [0,T]$;

N) $\rho_t(0) = 0$.

A DRM $(\rho_t)_{t\in[0,T]}$ *is called COHERENT if it satisfies M), TI) and* $\forall X, Y \in L^p_{\mathcal{F}_T}$:

SA) *subadditivity:* $\rho_t(X+Y) \leq \rho_t(X) + \rho_t(Y) \quad P-a.s.\ \forall\ t \in [0,T]$;

PH) *positive homogeneity:* $\forall\ \lambda \geq 0,\ \rho_t(\lambda X) = \lambda\rho_t(X) \quad P-a.s.\ \forall\ t \in [0,T]$;

Risk measurements of the same payoff at different dates t should be related in some way. The issue of time-consistency was addressed by Artzner et al. [ADHEK], for the coherent case, and Detlefsen and Scandolo [DS05] generalized this result to the dynamic convex risk measures.

Definition 54 *A DRM* $(\rho_t)_{t\in[0,T]}$ *is time-consistent if* $\forall t \in [0,T]$ *and* $\forall X, Y \in L^p_{\mathcal{F}_T}$

$$\rho_t(X) = \rho_t(Y) \Rightarrow \rho_0(X) = \rho_0(Y) \quad P-a.s.$$

or equivalently $\forall t \in [0,T]$, $\forall X \in L^p_{\mathcal{F}_T}$ *and* $\forall A \in \mathcal{F}_t$

$$\rho_0(X1_A) = \rho_0(-\rho_t(X)1_A) \quad P-a.s.$$

The financial meaning of time consistency is that if two payoffs will have the same riskiness at time t in every state of nature, then the same conclusion should be valid *today*. If $\rho_t(X)$ is the conditional capital requirement that has to be set aside at date t in view of the final payoff X, then the risky position is equivalently described *today*, by the payoff $-\rho_t(X)$ occurring in t.

As in the static case, the immediate generalization of the convexity requirement is the quasiconvexity also in the conditional setting.

Definition 55 *A map* $\rho_t : L^p_{\mathcal{F}_T} \to L^0_{\mathcal{F}_t}$ *is quasi-convex (QCONV) if* $\forall X, Y \in L^p_{\mathcal{F}_T}$ *and* $\forall \Lambda \in L^0_{\mathcal{F}_t}$ *s.t.* $0 \leq \Lambda \leq 1$,

$$\rho_t(\Lambda X + (1-\Lambda)Y) \leq \rho_t(X) \vee \rho_t(Y);$$

or equivalently if all the lower level sets

$$\mathcal{A}(Z) = \left\{ X \in L^p_{\mathcal{F}_T} : \rho_t(X) \leq Z \right\} \qquad \forall Z \in L^0_{\mathcal{F}_t}$$

are conditionally convex, i.e.

$$\forall X, Y \in \mathcal{A}(Z) \Rightarrow \Lambda X + (1-\Lambda)Y \in \mathcal{A}(Z)$$

Recently, Frittelli and Maggis (2010) [FM11] have provided the dual representation of conditional quasi-convex maps in a more general setting. We start by reporting the notations and the assumptions on the spaces.

N*otations:*

- $L_{\mathcal{F}_T} := L(\Omega, \mathcal{F}_T, P) \subseteq L^0(\Omega, \mathcal{F}_T, P)$ is a lattice of $\mathcal{F}_T$-measurable random variables;
- $L_{\mathcal{F}_t} := L(\Omega, \mathcal{F}_t, P) \subseteq L^0(\Omega, \mathcal{F}_t, P)$ is a lattice of $\mathcal{F}_t$-measurable random variables;
- $L^*_{\mathcal{F}_T} := (L_{\mathcal{F}_T}, \geq)^*$ is the order continuous dual of $(L_{\mathcal{F}_T}, \geq)$ which is also a lattice.

Assumptions on the spaces:

- $L_{\mathcal{F}_T}$ (resp. $L_{\mathcal{F}_t}$) satisfies the property $\mathbf{1}_{\mathcal{F}}$ (resp. $\mathbf{1}_{\mathcal{F}_t}$):

$$X \in L_{\mathcal{F}_T} \text{ and } A \in \mathcal{F}_T \Rightarrow (X\mathbf{1}_A) \in L_{\mathcal{F}_T}$$

- $(L_{\mathcal{F}_T}, \sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T}))$ is a locally convex topological vector space;

- $L^*_{\mathcal{F}_T} \hookrightarrow L^1_{\mathcal{F}_T}$

- $L^*_{\mathcal{F}_T}$ satisfies the property $\mathbf{1}_{\mathcal{F}_T}$.

For example, the L^p spaces satisfies these assumptions: $L_{\mathcal{F}_T} := L^p_{\mathcal{F}_T}$ with $p \in [1, \infty]$ and $L^*_{\mathcal{F}_T} = L^q_{\mathcal{F}_T} \hookrightarrow L^1_{\mathcal{F}_T}$ (with $q = 1$ when $p = \infty$).

Frittelli and Maggis [FM11] provided a complete duality and a unique representation for the conditional quasi-convex maps ρ_t under the additional condition that either ρ_t is $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$l.s.c., i.e. all the lower level sets

$$\mathcal{A}(Z) = \{X \in L_{\mathcal{F}_T} : \rho_t(X) \le Z\} \qquad \forall Z \in L_{\mathcal{F}_t}$$

are $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$closed, or alternatively ρ_t is $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$u.s.c..

Furthermore, under very weak assumptions on the space $L_{\mathcal{F}_T}$ and when $\rho_t : L_{\mathcal{F}_T} \to L_{\mathcal{F}_t}$ is M and QCONV, the condition that ρ_t is $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$l.s.c. is equivalent to ρ_t is continuous from above (CFA). Thus in the following results, we may replace $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$l.s.c. with CFA.

Theorem 56 ([FM11]) *A $\rho_t : L_{\mathcal{F}_T} \to L_{\mathcal{F}_t}$ is M, QCONV, REG and either $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$l.s.c. or $\sigma(L_{\mathcal{F}_T}, L^*_{\mathcal{F}_T})-$u.s.c. if and only if there exists $S \in \mathcal{S}$ such that*

$$\rho_t(X) = ess \sup_{Q \in L^*_{\mathcal{F}_T} \cap \mathcal{P}} S(Q, E_Q[-X|\mathcal{F}_t])$$

where

$$\mathcal{P} := \left\{ \frac{dQ}{dP} \mid Q \ll P \text{ and } Q \text{ probability} \right\}.$$

$$S(Q,Y) := ess \inf_{\xi \in L_{\mathcal{F}_T}} \{\rho_t(\xi) \mid E_Q[-\xi|\mathcal{F}_t] \geq Y\}, \quad Y \in L_{\mathcal{F}_t}$$

and the class is defined

$$\mathcal{S} := \left\{ S : L^*_{\mathcal{F}_T} \times L^0_{\mathcal{F}_t} \to \bar{L}^0_{\mathcal{F}_t} \text{ such that } S(Q,\cdot) \text{ is M, REG and CFA} \right\}$$

Note that any map $S : L^*_{\mathcal{F}_T} \times L^0_{\mathcal{F}_t} \to \bar{L}^0_{\mathcal{F}_t}$ *such that* $S(Q,\cdot)$ *is M and REG is automatically QCONV in the first component.*

Observe that the dual representation of (QCONV) conditional maps turns out to have the same structure of the real valued case, but the expectations are conditional on the available information $\mathcal{F}_t$.

The following corollary provides the robust representation of the conditional convex maps, that was proved by [DS05].

Corollary 57 ([DS05]) *Suppose that the assumptions of the previous theorem hold true. Suppose that for every* $Q \in L^*_{\mathcal{F}_T} \cap \mathcal{P}_{\mathcal{F}_t}$, *where*

$$\mathcal{P}_{\mathcal{F}_t} := \left\{ \frac{dQ}{dP} \mid Q \in \mathcal{P} \text{ and } Q = P \text{ on } \mathcal{F}_t \right\},$$

and for any $\xi \in L_{\mathcal{F}_T}$ *we have that* $E_Q[-\xi|\mathcal{F}_t] \in L_{\mathcal{F}_T}$. *If* $\rho_t : L_{\mathcal{F}_T} \to L_{\mathcal{F}_t}$ *satisfies in addition* TI *then*

$$S(Q, E_Q[-X|\mathcal{F}_t]) = E_Q[-X|\mathcal{F}_t] - \alpha(Q)$$

and

$$\rho_t(X) = ess \sup_{Q \in L^*_{\mathcal{F}_T} \cap \mathcal{P}_{\mathcal{F}_t}} \{E_Q[-X|\mathcal{F}_t] - \alpha(Q)\}$$

where $\alpha : \mathcal{P}_{\mathcal{F}_t} \to \bar{L}^0_{\mathcal{F}_t}$ *is the random penalty function.*

Note that the additional information $\mathcal{F}_t$ allows to a-priori exclude some probabilistic models. Infact, only $\mathcal{P}_{\mathcal{F}_t} \subseteq \mathcal{P}$ enter the representation. The interpretation is simple: smaller is the information $\mathcal{F}_t$, larger is the subset $\mathcal{P}_{\mathcal{F}_t}$ of probabilistic models which can be considered in the worst case representation.

In a recent work Bielecki, Cialenco and Zhang (2011) [BCZ11] argued the acceptability indices in the dynamic case. Now we just recall the representation result in order to stress the link to the coherent dynamic risk measures.

Theorem 58 ([BCZ11]) *A map* $\alpha_t : L^\infty_{\mathcal{F}_T} \to [0, \infty]$ *is a dynamic coherent acceptability index if and only if there exists an increasing family of dynamic coherent risk measures* $\{\rho^m_t\}_{m \in \mathbb{R}^+}$ *such that*

$$\alpha(X) = \sup \left\{ m \in \mathbb{R}^+ \mid \rho^m_t(X) \leq 0 \right\}$$

or alternatively, if and only if there exists an increasing family $\{\mathcal{Q}_m\}_{m \in \mathbb{R}^+}$ *of dynamic subsets of probabilities such that*

$$\alpha(X) = \sup \left\{ m \in \mathbb{R}^+ \mid \inf_{Q \in \mathcal{Q}_m} E_Q[X|\mathcal{F}_t] \geq 0 \right\}.$$

Part II

Part: Applications

Chapter 3

Value At Risk with Probability/Loss function

In this chapter we introduce a new class of *law invariant risk measures* $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ that are directly defined on the set $\mathcal{P}(\mathbb{R})$ of probability measures on $\mathbb{R}$ and are monotone and quasi-convex on $\mathcal{P}(\mathbb{R})$.

As Cherny and Madan (2009) [CM09] pointed out, for a (*translation invariant)* coherent risk measure defined on random variables, all the positions can be split in two classes: acceptable and not acceptable; in contrast, for an *acceptability index* there is a whole continuum of degrees of acceptability defined by a system $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of sets. This formulation has been further investigated by Drapeau and Kupper (2010) [DK10] for the quasi convex case.

We adopt this approach and we build the maps Φ from a family $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of acceptance sets of distribution functions by defining:

$$\Phi(P) := -\sup\{m \in \mathbb{R} \mid P \in \mathcal{A}^m\}.$$

In Section 3.2 we study the properties of such maps, we provide some specific examples and in particular we propose an interesting generalization of the classical notion of $V@R_\lambda$.

The key idea of our proposal - the definition of the $\Lambda V@R$ in Section 3.3 - arises from the consideration that in order to assess the risk of a financial position it is necessary to consider not only the probability λ of the loss, as in the case of the $V@R_\lambda$, but the dependence between such *probability* λ and the *amount* of the loss. In other terms, a risk adverse agent is willing to accept greater losses only with smaller probabilities. Hence, we replace the constant λ with a (increasing) function $\Lambda : \mathbb{R} \rightarrow [0,1]$ defined on losses, which we call *Probability/Loss function.* The balance between the probability and the amount of the losses is incorporated in the definition of the family of acceptance sets

$$\mathcal{A}^m := \{Q \in \mathcal{P}(\mathbb{R}) \mid Q(-\infty, x] \leq \Lambda(x),\ \forall x \leq m\},\ m \in \mathbb{R}.$$

If P_X is the distribution function of the random variable X, our new measure is defined by:

$$\Lambda V@R(P_X) := -\sup\{m \mid P(X \leq x) \leq \Lambda(x),\ \forall x \leq m\}.$$

As a consequence, the acceptance sets $\mathcal{A}^m$ are not obtained by the translation of $\mathcal{A}^0$ which implies that the map is not any more translation invariant. However, the similar property

$$\Lambda V@R(P_{X+\alpha}) = \Lambda^\alpha V@R(P_X) - \alpha,$$

where $\Lambda^\alpha(x) = \Lambda(x + \alpha)$, holds true and is discussed in Section 3.3.

The $V@R_\lambda$ and the worst case risk measure are special cases of the $\Lambda V@R$.

In Section 3.4 we address the dual representation of these maps. We choose to define the risk measures on the entire set $\mathcal{P}(\mathbb{R})$ and not only on its subset of probabilities having compact support. We endow $\mathcal{P}(\mathbb{R})$ with the $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$ topology. The selection of this topology is also justified by the fact (see Proposition 64) that for monotone maps $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$-l.s.c. is equivalent to continuity from below.

Except for $\Phi = +\infty$, we show that there are no *convex*, $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R})) - lsc$ translation invariant maps $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$. But there are many quasi-convex and $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$-l.s.c. maps $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ that in addition are monotone and translation invariant, as for example the $V@R_\lambda$, the entropic risk measure and the worst case risk measure. This is another good motivation to adopt quasi convexity versus convexity.

Finally we provide the dual representation of quasi-convex, monotone and $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$-l.s.c. maps $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ - *defined on the entire set* $\mathcal{P}(\mathbb{R})$ - and compute the dual representation of the risk measures associated to families of acceptance sets and consequently of the $\Lambda V@R$.

3.1 Risk Measures defined on distributions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a non atomic probability space and we denote with $L^0 =: L^0(\Omega, \mathcal{F}, \mathbb{P})$ the space of $\mathcal{F}$ measurable random variables that are $\mathbb{P}$ almost surely finite defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and that take values in $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, where $\mathcal{B}_{\mathbb{R}}$ is the σ-field of the Borel sets.

Any random variable $X \in L^0$ induces a *distribution*, that is a probability measure

P_X on ($\mathbb{R}$, $\mathcal{B}_{\mathbb{R}}$) defined by

$$P_X(B) := \mathbb{P}(X^{-1}(B)) \quad \text{for any } B \in \mathcal{B}_{\mathbb{R}}$$

We refer to a distribution also with the term *lottery*.

Let us denote with $\mathcal{P} =: \mathcal{P}(\mathbb{R})$ the convex set of the distributions and $F_X(x) := P_X(-\infty, x]$ is the *distribution function* of X. We refer to [Ali] Chapter 15 for a detailed study of $\mathcal{P}$. Here we just recall some basic notions: for any $X \in L^0$ we have $P_X \in \mathcal{P}$ so that we will associate to any random variable a unique element in $\mathcal{P}$. If $\mathbb{P}(X = x) = 1$ for some $x \in \mathbb{R}$ then P_X turns out to be the delta distribution δ_x that concentrates the mass in the point x.

According to the definition 41, we remember that a risk measure $\rho : L \to \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$, defined on given subset $L \subset L^0$, is *law invariant* if for any $X, Y \in L$

$$P_X = P_Y \Leftrightarrow \rho(X) = \rho(Y).$$

As underlined by Weber 2006 [W06], since these risk measures depend only on the distribution of financial positions, they can either be considered as functionals on spaces of random variables or spaces of distributions. In the first case, the quasiconvexity of the risk measures leads to refine the robust representation, as recalled in the Theorem 46.

Here we consider the second perspective, that is, the interpretation of risk measures as functionals of distributions. To this end, we shift the problem to the set $\mathcal{P}$ by defining the new map $\Phi : \mathcal{P} \to \overline{\mathbb{R}}$ as

$$\Phi(P_X) = \rho(X).$$

This map Φ is well defined on the entire $\mathcal{P}$, since there exists a bi-injective relation between $\mathcal{P}$ and the quotient space $\frac{L^0}{\sim_{\mathcal{D}}}$, where the equivalence is given by $X \sim_{\mathcal{D}} Y \Leftrightarrow P_X = P_Y$.

However, $\mathcal{P}$ is only a convex set and the usual operations on $\mathcal{P}$ are not induced by those on L^0, namely

$$(P_X + P_Y)(A) = P_X(A) + P_Y(A) \neq P_{X+Y}(A), \quad A \in \mathcal{B}_{\mathbb{R}}.$$

We will also consider the first order stochastic dominance on $\mathcal{P}$ defined by:

$$Q \preccurlyeq_{mon} P \Leftrightarrow F_P(x) \leq F_Q(x) \quad \text{for all } x \in \mathbb{R},$$

where $F_P(x) = P(-\infty, x]$ and $F_Q(x) = Q(-\infty, x]$ are the distribution functions of $P, Q \in \mathcal{P}$. It will be more convenient to adopt on $\mathcal{P}$ the opposite order relation:

$$P \preccurlyeq Q \quad \Leftrightarrow \quad Q \preccurlyeq_{mon} P \quad \Leftrightarrow \quad F_P(x) \leq F_Q(x) \quad \text{for all } x \in \mathbb{R}.$$

The financial interpretation is natural: the financial position X has a lower level of risk with respect to $\preccurlyeq$ since the probability induced by X, F_X, concentrates more probability on higher values of x, indeed $F_X(x)$, goes faster to zero as $x \to -\infty$ and slower to one as $x \to +\infty$.

Definition 59 *A Risk Measure on $\mathcal{P}$ is a map $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ such that:*

(Mon) *Φ is monotone: $P \preccurlyeq Q$ implies $\Phi(P) \leq \Phi(Q)$;*

(QCo) *Φ is quasi-convex: $\Phi(\lambda P + (1-\lambda)Q) \leq \Phi(P) \vee \Phi(Q)$, $\lambda \in [0,1]$.*

The increasing (instead of decreasing) monotonicity assumption can be deduced by the case of risk measures on random variables, simply observing that $X \geq Y$ $\mathbb{P}$-a.s. implies $P_X \preccurlyeq P_Y$.

Also in this case, quasiconvexity can be equivalently reformulated in terms of sublevel sets.

Remark 60 *A map* $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ *is quasi-convex if and only if*

$$\text{every lower set } \mathcal{A}_c = \{P \in \mathcal{P} \mid \Phi(P) \leq c\} \quad \forall c \in \mathbb{R} \text{ is convex.}$$

As recalled in [W06] this notion of quasiconvexity is different from the one given for random variables (as in [FS04]) because it does not concern diversification of financial positions. In this context the risk perception is focused on the probability of a loss rather than in terms of values and the diversification principle is naturally in terms of compound lotteries. A natural interpretation of this principle is the following: whenever two probability measures P and Q are acceptable at some level c and $\lambda \in [0,1]$ is a probability, then the compound lottery $\lambda P + (1-\lambda)Q$, which randomizes over P and Q, is also acceptable at the same level. In terms of random variables (namely X, Y which induce P_X, P_Y) the randomized probability $\lambda P_X + (1-\lambda)P_Y$ will correspond to some random variable $Z \neq \lambda X + (1-\lambda)Y$ so that the diversification is realized at the level of distribution and not at the level of portfolio selection.

As suggested by [W06], we define the *translation operator* T_m on the set $\mathcal{P}$ by:

$$T_m P(-\infty, x] = P(-\infty, x-m], \quad \text{for every } m \in \mathbb{R}.$$

Equivalently, if P_X is the probability distribution of the random variable X we define the translation operator as

$$T_m P_X = P_{X+m}, \quad m \in \mathbb{R}.$$

As a consequence we map the distribution $F_X(x)$ into $F_X(x-m)$. Notice that $T_m P \preccurlyeq P$ for any $m > 0$.

Definition 61 *If* $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ *is a risk measure on* $\mathcal{P}$*, we say that*

(TrI) *Φ is translation invariant if*

$$\Phi(T_m P) = \Phi(P) - m, \quad m \in \mathbb{R}$$

Notice that *(TrI)* corresponds exactly to the notion of *cash additivity* for risk measures defined on a space of random variables as introduced in [ADEH99] (see Def. 1).

It is well known (see [CMMMa]) that for maps defined on random variables, quasiconvexity and cash additivity imply convexity. However, in the context of distributions *(QCo)* and *(TrI)* do not imply convexity of the map Φ, as can be shown with the simple examples of the $V@R_\lambda$ and the worst case risk measure ρ_w (see the examples in Section 3.2.1).

Additional topological conditions and results

It's well known that $\mathcal{P}(\mathbb{R})$ spans the space of all signed measures of bounded variations on $\mathbb{R}$,

$$ca(\mathbb{R}) := \{\mu \text{ signed measure} \mid V_\mu < +\infty\}$$

where

$$V_\mu = \sup\left\{\sum_{i=1}^{n} |\mu(A_i)| \mid \{A_i, ..., A_n\} \text{ partition of } \mathbb{R}\right\}$$

and $ca(\mathbb{R})$ (or simply ca) endowed with the norm V_μ is norm complete and an AL-space [1].

Let $C_b(\mathbb{R})$ (or simply C_b) be the space of bounded continuous function $f : \mathbb{R} \to \mathbb{R}$. We endow $ca(\mathbb{R})$ with the weak* topology $\sigma(ca, C_b)$. The dual pairing $\langle \cdot, \cdot \rangle : C_b \times ca \to \mathbb{R}$

[1] A normed Riesz space equipped with an L-norm (i.e a lattice norm s.t $x, y \geq 0 \Rightarrow \|x + y\| = \|x\| + \|y\|$) is an L-space. In case the space is also complete, it is called AL-space (see [Ali] for further details).

is given by

$$\langle f, \mu \rangle = \int f d\mu$$

and the function $\mu \mapsto \int f d\mu$ ($\mu \in ca$) is $\sigma(ca, C_b)$ continuous.

Notice that $\mathcal{P}$ is a $\sigma(ca, C_b)$-closed convex subset of ca (p. 507 in [Ali]) so that $\sigma(\mathcal{P}, C_b)$ is the relativization of $\sigma(ca, C_b)$ to $\mathcal{P}$ and any $\sigma(\mathcal{P}, C_b)$-closed subset of $\mathcal{P}$ is also $\sigma(ca, C_b)$-closed.

Even though $(ca, \sigma(ca, C_b))$ is not metrizable in general, its subset $\mathcal{P}$ is separable and metrizable (see [Ali], Th.15.12) and therefore when dealing with convergence in $\mathcal{P}$ we may work with sequences instead of nets.

For every real function F we denote by $\mathcal{C}(F)$ the set of points in which the function F is continuous.

Theorem 62 ([Shi] Theorem 2, p.314) *Suppose that P_n, $P \in \mathcal{P}$. Then*

$$P_n \overset{\sigma(\mathcal{P},C_b)}{\longrightarrow} P \quad \Leftrightarrow \quad F_{P_n}(x) \to F_P(x) \quad \forall x \in \mathcal{C}(F_P).$$

A sequence of probabilities $\{P_n\} \subset \mathcal{P}$ is *increasing*, denoted with $P_n \uparrow$, if

$$F_{P_n}(x) \le F_{P_{n+1}}(x) \quad \forall x \in \mathbb{R} \text{ and } \forall n.$$

Definition 63 *Suppose that P_n, $P \in \mathcal{P}$. We say that $P_n \uparrow P$ whenever*

$$P_n \uparrow \text{ and } F_{P_n}(x) \uparrow F_P(x) \text{ for every } x \in \mathcal{C}(F_P).$$

(CfB) *Φ is continuous from below if*

$$P_n \uparrow P \text{ implies } \Phi(P_n) \uparrow \Phi(P).$$

The selection of the topology $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$ is also justified by the fact that for monotone maps $\sigma(\mathcal{P}(\mathbb{R}), C_b(\mathbb{R}))$-lower semicontinuity. is equivalent to continuity from below. This result is shown in the following proposition.

Proposition 64 *Let $\Phi : \mathcal{P} \to \overline{\mathbb{R}}$ be (Mon). Then the following are equivalent:*

Φ is $\sigma(\mathcal{P}, C_b)$-lower semicontinuous

Φ is continuous from below.

Proof. Let Φ be $\sigma(\mathcal{P}, C_b)$-lower semicontinuous and suppose that $P_n \uparrow P$. Then $F_{P_n}(x) \uparrow F_P(x)$ for every $x \in \mathcal{C}(F_P)$ and we deduce from Theorem 62 that $P_n \overset{\sigma(\mathcal{P},C_b)}{\longrightarrow} P$. (Mon) implies $\Phi(P_n) \uparrow$ and $k := \lim_n \Phi(P_n) \leq \Phi(P)$. The lower level set $A_k = \{Q \in \mathcal{P} \mid \Phi(Q) \leq k\}$ is $\sigma(\mathcal{P}, C_b)$ closed and, since $P_n \in A_k$, we also have $P \in A_k$, i.e. $\Phi(P) = k$, and Φ is continuous from below.

Conversely, suppose that Φ is continuous from below. As $\mathcal{P}$ is metrizable we may work with sequences instead of nets. For $k \in \mathbb{R}$ consider $A_k = \{P \in \mathcal{P} \mid \Phi(P) \leq k\}$ and a sequence $\{P_n\} \subseteq A_k$ such that $P_n \overset{\sigma(\mathcal{P},C_b)}{\longrightarrow} P \in \mathcal{P}$. We need to show that $P \in A_k$. Lemma 65 shows that each $F_{Q_n} := (\inf_{m \geq n} F_{P_m}) \wedge F_P$ is the distribution function of a probability measure $Q_n \in \mathcal{P}$. Notice that $F_{Q_n} \leq F_{P_n}$ and $Q_n \uparrow$. From $P_n \overset{\sigma(\mathcal{P},C_b)}{\longrightarrow} P$ and the definition of Q_n, we deduce that $F_{Q_n}(x) \uparrow F_P(x)$ for every $x \in \mathcal{C}(F_P)$ so that $Q_n \uparrow P$. From (Mon) and $Q_n \preccurlyeq P_n$, we get $\Phi(Q_n) \leq \Phi(P_n)$. From (CfB) then: $\Phi(P) = \lim_n \Phi(Q_n) \leq \liminf_n \Phi(P_n) \leq k$. Thus $P \in A_k$. ■

Lemma 65 *For every $P_n \overset{\sigma(\mathcal{P},C_p)}{\longrightarrow} P$ we have that*

$$F_{Q_n} := \inf_{m \geq n} F_{P_m} \wedge F_P, \; n \in \mathbb{N},$$

is a distribution function associated to a probability measure $Q_n \in \mathcal{P}$.

Proof. For each n, F_{Q_n} is increasing and $\lim_{x\to-\infty} F_{Q_n}(x) = 0$. Moreover for real valued maps right continuity and upper semicontinuity are equivalent. Since the inf-operator preserves upper semicontinuity we can conclude that F_{Q_n} is right continuous for every n. Now we have to show that for each n, $\lim_{x\to+\infty} F_{Q_n}(x) = 1$. By contradiction suppose that, for some n, $\lim_{x\to+\infty} F_{Q_n}(x) = \lambda < 1$. We can choose a sequence $\{x_k\}_k \subset \mathbb{R}$ with $x_k \in \mathcal{C}(F_P)$, $x_k \uparrow +\infty$. In particular $F_{Q_n}(x_k) \leq \lambda$ for all k and $F_P(x_k) > \lambda$ definitively, say for all $k \geq k_0$. We can observe that since $x_k \in \mathcal{C}(F_P)$ we have, for all $k \geq k_0$, $\inf_{m\geq n} F_{P_m}(x_k) < \lim_{m\to+\infty} F_{P_m}(x_k) = F_P(x_k)$. This means that the infimum is attained for some index $m(k) \in \mathbb{N}$, i.e. $\inf_{m\geq n} F_{P_m}(x_k) = F_{P_{m(k)}}(x_k)$, for all $k \geq k_0$. Since $P_{m(k)}(-\infty, x_k] = F_{P_{m(k)}}(x_k) \leq \lambda$ then $P_{m(k)}(x_k, +\infty) \geq 1 - \lambda$ for $k \geq k_0$. We have two possibilities. Either the set $\{m(k)\}_k$ is bounded or $\overline{\lim}_k m(k) = +\infty$. In the first case, we know that the number of $m(k)$'s is finite. Among these $m(k)$'s we can find at least one $\overline{m}$ and a subsequence $\{x_h\}_h$ of $\{x_k\}_k$ such that $x_h \uparrow +\infty$ and $P_{\overline{m}}(x_h, +\infty) \geq 1 - \lambda$ for every h. We then conclude that

$$\lim_{h\to+\infty} P_{\overline{m}}(x_h, +\infty) \geq 1 - \lambda$$

and this is a contradiction. If $\overline{\lim}_k m(k) = +\infty$, fix $\overline{k} \geq k_0$ such that $P(x_{\overline{k}}, +\infty) < 1 - \lambda$ and observe that for every $k > \overline{k}$

$$P_{m(k)}(x_{\overline{k}}, +\infty) \geq P_{m(k)}(x_k, +\infty) \geq 1 - \lambda.$$

Take a subsequence $\{m(h)\}_h$ of $\{m(k)\}_k$ such that $m(h) \uparrow +\infty$. Then:

$$\lim_{h\to\infty} \inf P_{m(h)}(x_{\overline{k}}, +\infty) \geq 1 - \lambda > P(x_{\overline{k}}, +\infty),$$

which contradicts the weak convergence $P_n \overset{\sigma(\mathcal{P},C_b)}{\longrightarrow} P$. ■

Example 66 (The certainty equivalent) *It is very simple to build risk measures on $\mathcal{P}(\mathbb{R})$. Take any continuous, bounded from below and strictly decreasing function $f:\mathbb{R}\to\mathbb{R}$. Then the map $\Phi_f:\mathcal{P}\to\mathbb{R}\cup\{+\infty\}$ defined by:*

$$\Phi_f(P) := -f^{-1}\left(\int f dP\right)$$

is a Risk Measure on $\mathcal{P}(\mathbb{R})$. It is also easy to check that Φ_f is (CFB) and therefore $\sigma(\mathcal{P},C_b)-$l.s.c. Notice that Proposition 81 will then imply that Φ_f can not be convex. By selecting the function $f(x)=e^{-x}$ we obtain $\Phi_f(P)=\ln\left(\int \exp(-x)dF_P(x)\right)$, which is in addition (TrI). Its associated risk measure $\rho: L^0\to\mathbb{R}\cup\{+\infty\}$ defined on random variables, $\rho(X)=\Phi_f(P_X)=\ln\left(Ee^{-X}\right)$, is the Entropic Risk Measure. In Section 3.4 we will see more examples based on this construction.

3.2 A remarkable class of risk measures on $\mathcal{P}(\mathbb{R})$

As Cherny and Madan (2009) [CM09] pointed out, for a (*translation invariant)* coherent risk measure defined on random variables, all the positions can be split in two classes: acceptable and not acceptable; in contrast, for an *acceptability index* there is a whole continuum of degrees of acceptability defined by a system $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of sets. This formulation has been further investigated by Drapeau and Kupper (2010) [DK10] for the quasi convex case.

We adopt this approach and we build the maps Φ from a family $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of *acceptance sets* of distribution functions. Given a family $\{F_m\}_{m\in\mathbb{R}}$ of functions $F_m:\mathbb{R}\to$

$[0,1]$, we consider the associated sets of probability measures

$$\mathcal{A}^m := \{Q \in \mathcal{P} \mid F_Q \leq F_m\} \tag{3.1}$$

and the associated map $\Phi : \mathcal{P} \to \overline{\mathbb{R}}$ defined by

$$\Phi(P) := -\sup\{m \mid P \in \mathcal{A}^m\}. \tag{3.2}$$

Hereafter we assume that for each $P \in \mathcal{P}$ there exists m such that $P \notin \mathcal{A}^m$ so that $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$.

Definition 67 *A monotone decreasing family of sets $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ contained in $\mathcal{P}$ is* left continuous *in m if*

$$\mathcal{A}^m =: \bigcap_{\varepsilon>0} \mathcal{A}^{m-\varepsilon}$$

In particular it is left continuous *if it is* left continuous *in m for every $m \in \mathbb{R}$.*

Lemma 68 (Relations between $\{F_m\}_{m\in\mathbb{R}}$ and $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$) *Let $\{F_m\}_{m\in\mathbb{R}}$ be a family of functions $F_m : \mathbb{R} \to [0,1]$ and $\mathcal{A}^m$ be the set defined in (3.1). Then:*

1. *If, for every $x \in \mathbb{R}$, $F_\cdot(x)$ is decreasing (w.r.t. m) then the family $\{\mathcal{A}^m\}$ is monotone decreasing:*

   $$\mathcal{A}^m \subseteq \mathcal{A}^n \text{ for any level } m \geq n;$$

2. *For any m, $\mathcal{A}^m$ is convex and satisfies:*

   $$Q \preceq P \in \mathcal{A}^m \Rightarrow Q \in \mathcal{A}^m;$$

3. *If, for every $m \in \mathbb{R}$, $F_m(x)$ is right continuous w.r.t. x then*

   $$\mathcal{A}^m \text{ is } \sigma(\mathcal{P}, C_b)\text{-closed;}$$

4. Suppose that, for every $x \in \mathbb{R}$, $F_m(x)$ is decreasing w.r.t. m.

If $F_m(x)$ is left continuous w.r.t.$m \Rightarrow$ the family $\{\mathcal{A}^m\}$ is left continuous.

5. Suppose that, for every $x \in \mathbb{R}$, $F_m(x)$ is decreasing w.r.t. m and that, for every $m \in \mathbb{R}$, $F_m(x)$ is right continuous and increasing w.r.t. x and $\lim_{x\to+\infty} F_m(x) = 1$.

If the family $\{\mathcal{A}^m\}$is left continuous in m $\Rightarrow F_m(x)$ is left continuous in m.

Proof. 1. If $Q \in \mathcal{A}^m$ and $m \geq n$ then $F_Q \leq F_m \leq F_n$, i.e. $Q \in \mathcal{A}^n$.

2. Let $Q, P \in \mathcal{A}^m$ and $\lambda \in [0,1]$. Consider the convex combination $\lambda Q + (1-\lambda)P$ and notice that

$$F_{\lambda Q+(1-\lambda)P} \leq F_Q \vee F_P \leq F_m,$$

as $F_P \leq F_m$ and $F_Q \leq F_m$. Then $\lambda Q + (1-\lambda)P \in \mathcal{A}^m$.

3. Let $Q_n \in A^m$ and $Q \in \mathcal{P}$ satisfy $Q_n \overset{\sigma(\mathcal{P},C_b)}{\to} Q$. By Theorem 62 we know that $F_{Q_n}(x) \to F_Q(x)$ for every $x \in \mathcal{C}(F_Q)$. For each n, $F_{Q_n} \leq F_m$ and therefore $F_Q(x) \leq F_m(x)$ for every $x \in \mathcal{C}(F_Q)$. By contradiction, suppose that $Q \notin \mathcal{A}^m$. Then there exists $\bar{x} \notin \mathcal{C}(F_Q)$ such that $F_Q(\bar{x}) > F_m(\bar{x})$. By right continuity of F_Q for every $\varepsilon > 0$ we can find a right neighborhood $[\bar{x}, \bar{x} + \delta(\varepsilon))$ such that

$$|F_Q(x) - F_Q(\bar{x})| < \varepsilon \quad \forall\, x \in [\bar{x}, \bar{x} + \delta(\varepsilon))$$

and we may require that $\delta(\varepsilon) \downarrow 0$ if $\varepsilon \downarrow 0$.Notice that for each $\varepsilon > 0$ we can always choose an $x_\varepsilon \in (\bar{x}, \bar{x} + \delta(\varepsilon))$ such that $x_\varepsilon \in \mathcal{C}(F_Q)$. For such an x_ε we deduce that

$$F_m(\bar{x}) < F_Q(\bar{x}) < F_Q(x_\varepsilon) + \varepsilon \leq F_m(x_\varepsilon) + \varepsilon.$$

This leads to a contradiction since if $\varepsilon \downarrow 0$ we have that $x_\varepsilon \downarrow \bar{x}$ and thus by right continuity of F_m

$$F_m(\bar{x}) < F_Q(\bar{x}) \leq F_m(\bar{x}).$$

4. By assumption we know that $F_{m-\varepsilon}(x) \downarrow F_m(x)$ as $\varepsilon \downarrow 0$, for all $x \in \mathbb{R}$. By item 1, we know that $\mathcal{A}^m \subseteq \bigcap_{\varepsilon>0} \mathcal{A}^{m-\varepsilon}$. By contradiction we suppose that

$$\bigcap_{\varepsilon>0} \mathcal{A}^{m-\varepsilon} \supsetneqq \mathcal{A}^m,$$

so that there will exist $Q \in \mathcal{P}$ such that $F_Q \leq F_{m-\varepsilon}$ for every $\varepsilon > 0$ but $F_Q(\overline{x}) > F_m(\overline{x})$ for some $\overline{x} \in \mathbb{R}$. Set $\delta = F_Q(\overline{x}) - F_m(\overline{x})$ so that $F_Q(\overline{x}) > F_m(\overline{x}) + \frac{\delta}{2}$. Since $F_{m-\varepsilon} \downarrow F_m$ we may find $\overline{\varepsilon} > 0$ such that $F_{m-\overline{\varepsilon}}(\overline{x}) - F_m(\overline{x}) < \frac{\delta}{2}$. Thus $F_Q(\overline{x}) \leq F_{m-\varepsilon}(\overline{x}) < F_m(\overline{x}) + \frac{\delta}{2}$ and this is a contradiction.

5. Assume that $\mathcal{A}^{m-\varepsilon} \downarrow \mathcal{A}^m$. Define $F(x) := \lim_{\varepsilon\downarrow 0} F_{m-\varepsilon}(x) = \inf_{\varepsilon>0} F_{m-\varepsilon}(x)$ for all $x \in \mathbb{R}$. Then $F : \mathbb{R} \to [0,1]$ is increasing, right continuous (since the inf preserves this property). Notice that for every $\varepsilon > 0$ we have $F_{m-\varepsilon} \geq F \geq F_m$ and then $\mathcal{A}^{m-\varepsilon} \supseteq \{Q \in \mathcal{P} \mid F_Q \leq F\} \supseteq \mathcal{A}^m$ and $\lim_{x\to+\infty} F(x) = 1$. Necessarily we conclude $\{Q \in \mathcal{P} \mid F_Q \leq F\} = \mathcal{A}^m$. By contradiction we suppose that $F(\overline{x}) > F_m(\overline{x})$ for some $\overline{x} \in \mathbb{R}$. Define $F_{\overline{Q}} : \mathbb{R} \to [0,1]$ by: $F_{\overline{Q}}(x) = F(x)\mathbf{1}_{[\overline{x},+\infty)}(x)$. The above properties of F guarantees that $F_{\overline{Q}}$ is a distribution function of a corresponding probability measure $\overline{Q} \in \mathcal{P}$, and since $F_{\overline{Q}} \leq F$, we deduce $\overline{Q} \in \mathcal{A}^m$, but $F_{\overline{Q}}(\overline{x}) > F_m(\overline{x})$ and this is a contradiction. ■

Lemma 69 (Relations between $\{F_m\}_{m\in\mathbb{R}}$ and Φ) *Let $\{F_m\}_{m\in\mathbb{R}}$ be a family of functions $F_m : \mathbb{R} \to [0,1]$ and Φ be the associated map defined in (3.2). Then:*

1. *The map Φ is (Mon) on $\mathcal{P}$.*

2. *If, for every $x \in \mathbb{R}$, $F_{\bullet}(x)$ is decreasing (w.r.t. m) then Φ is (QCo) on $\mathcal{P}$.*

3. *If, for every $x \in \mathbb{R}$, $F_{\bullet}(x)$ is left continuous and decreasing (w.r.t. m) and if, for every $m \in \mathbb{R}$, $F_m(\bullet)$ is right continuous (w.r.t. x) then*

$$A_m := \{Q \in \mathcal{P} \mid \Phi(Q) \leq m\} = \mathcal{A}^{-m},\ \forall m, \tag{3.3}$$

and Φ is $\sigma(\mathcal{P}, C_b)-$lower-semicontinuous.

Proof. 1. From $Q \preceq P$ we have $F_Q \leq F_P$ and

$$\{m \in \mathbb{R} \mid F_P \leq F_m\} \subseteq \{m \in \mathbb{R} \mid F_Q \leq F_m\},$$

which implies $\Phi(Q) \leq \Phi(P)$.

2. We show that $Q_1, Q_2 \in \mathcal{P}$, $\Phi(Q_1) \leq n$ and $\Phi(Q_2) \leq n$ imply that $\Phi(\lambda Q_1 + (1-\lambda)Q_2) \leq n$, that is

$$\sup\left\{m \in \mathbb{R} \mid F_{\lambda Q_1 + (1-\lambda)Q_2} \leq F_m\right\} \geq -n.$$

By definition of the supremum, $\forall \varepsilon > 0\ \exists m_i$ s.t. $F_{Q_i} \leq F_{m_i}$ and $m_i > -\Phi(Q_i) - \varepsilon \geq -n - \varepsilon$. Then $F_{Q_i} \leq F_{m_i} \leq F_{-n-\varepsilon}$, as $\{F_m\}$ is a decreasing family. Therefore $\lambda F_{Q_1} + (1-\lambda)F_{Q_2} \leq F_{-n-\varepsilon}$ and $-\Phi(\lambda Q_1 + (1-\lambda)Q_2 \lambda) \geq -n - \varepsilon$. As this holds for any $\varepsilon > 0$, we conclude that Φ is quasi-convex.

3. The fact that $\mathcal{A}^{-m} \subseteq A_m$ follows directly from the definition of Φ, as if $Q \in \mathcal{A}^{-m}$

$$\Phi(Q) := -\sup\{n : Q \in \mathcal{A}^n\} = \inf\left\{n : Q \in \mathcal{A}^{-n}\right\} \leq m.$$

We have to show that $A_m \subseteq \mathcal{A}^{-m}$. Let $Q \in A_m$. Since $\Phi(Q) \leq m$, for all $\varepsilon > 0$ there exists m_0 such that $m + \varepsilon > -m_0$ and $F_Q \leq F_{m_0}$. Since $F_{\bullet}(x)$ is decreasing (w.r.t. m) we have that $F_Q \leq F_{-m-\varepsilon}$, therefore $Q \in \mathcal{A}^{-m-\varepsilon}$ for any $\varepsilon > 0$. By the left continuity in m of $F_{\cdot}(x)$, we know that$\{\mathcal{A}^m\}$ is left continuous (Lemma 68, item 4) and so: $Q \in \bigcap_{\epsilon>0} \mathcal{A}^{-m-\varepsilon} = \mathcal{A}^{-m}$.

From the assumption that $F_m(\cdot)$ is right continuous (w.r.t. x) and Lemma 68 item 3, we already know that $\mathcal{A}^m$ is $\sigma(\mathcal{P}, C_b)-$closed, for any $m \in \mathbb{R}$, and therefore the lower level sets $A_m = \mathcal{A}^{-m}$ are $\sigma(\mathcal{P}, C_b)-$closed and Φ is $\sigma(\mathcal{P}, C_b)-$lower-semicontinuous. ■

Definition 70 (feasible $\{F_m\}_{m\in\mathbb{R}}$ family) *A family $\{F_m\}_{m\in\mathbb{R}}$ of functions $F_m : \mathbb{R} \to [0, 1]$ is* feasible *if*

- *For any $P \in \mathcal{P}$ there exists m such that $P \notin \mathcal{A}^m$*
- *For every $m \in \mathbb{R}$, $F_m(\cdot)$ is right continuous (w.r.t. x)*
- *For every $x \in \mathbb{R}$, $F_{\cdot}(x)$ is decreasing and left continuous (w.r.t. m).*

From Lemmas 68 and 69 we immediately deduce:

Proposition 71 *Let $\{F_m\}_{m\in\mathbb{R}}$ be a feasible family. Then the associated family $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ is monotone decreasing and left continuous and each set $\mathcal{A}^m$ is convex and $\sigma(\mathcal{P}, C_b)-$closed. The associated map $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ is well defined, (Mon), (Qco) and $\sigma(\mathcal{P}, C_b)$-l.s.c.*

Remark 72 *Let $\{F_m\}_{m\in\mathbb{R}}$ be a feasible family. If there exists an $\overline{m}$ such that $\lim_{x\to+\infty} F_{\overline{m}}(x) < 1$ then $\lim_{x\to+\infty} F_m(x) < 1$ for every $m \geq \overline{m}$ and then $\mathcal{A}^m = \emptyset$ for every $m \geq \overline{m}$. Obviously if an acceptability set is empty then it does not contribute to the computation of the risk measure defined in (3.2). For this reason we will always consider w.l.o.g. a class $\{F_m\}_{m\in\mathbb{R}}$ such that $\lim_{x\to+\infty} F_m(x) = 1$ for every m.*

3.2.1 Examples

As explained in the introduction, we define a family of risk measures employing a Probability/Loss function Λ. Fix the *right continuous* function $\Lambda : \mathbb{R} \to [0,1]$ and define the family $\{F_m\}_{m\in\mathbb{R}}$ of functions $F_m : \mathbb{R} \to [0,1]$ by

$$F_m(x) := \Lambda(x)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x). \tag{3.4}$$

It is easy to check that if $\sup_{x\in\mathbb{R}} \Lambda(x) < 1$ then the family $\{F_m\}_{m\in\mathbb{R}}$ is feasible and therefore, by Proposition 71, the associated map $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ is well defined, *(Mon)*, *(Qco)* and $\sigma(\mathcal{P}, C_b)$−l.s.c.

Example 73 *When* $\sup_{x\in\mathbb{R}} \Lambda(x) = 1$, Φ *may take the value* $-\infty$. *The extreme case is when, in the definition of the family (3.4), the function* Λ *is equal to the constant one,* $\Lambda(x) = 1$, *and so:* $\mathcal{A}^m = \mathcal{P}$ *for all* m *and* $\Phi = -\infty$.

Example 74 (Worst case measure: $\Lambda(x) = 0$) *Take in the definition of the family (3.4) the function* Λ *to be equal to the constant zero:* $\Lambda(x) = 0$. *Then:*

$$\begin{aligned} F_m(x) \ &: = \mathbf{1}_{[m,+\infty)}(x) \\ \mathcal{A}^m \ &: = \{Q \in \mathcal{P} \mid F_Q \leq F_m\} = \{Q \in \mathcal{P} \mid Q \preccurlyeq \delta_m\} \\ \Phi_w(P) \ &: = -\sup\{m \mid P \in \mathcal{A}^m\} = -\sup\{m \mid P \preccurlyeq \delta_m\} = -\inf_{x\in\mathbb{R}}(F_P(x)) \end{aligned}$$

so that, if $X \in L^0$ *has distribution* P_X,

$$\Phi_w(P_X) = -\sup\{m \mid P_X \preccurlyeq \delta_m\} = -ess\inf(X) := \rho_w(X)$$

coincide with the worst case risk measure ρ_w.

As the family $\{F_m\}$ *is feasible,* $\Phi_w : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{+\infty\}$ *is* (Mon), (Qco) *and* $\sigma(\mathcal{P}, C_b)-$*l.s.c, hence it is a risk measure on* $\mathcal{P}$*. In addition, it also satisfies* (TrI)*.*

Even though $\rho_w : L^0 \to \mathbb{R}\cup\{+\infty\}$ *is convex, as a map defined on random variables, the corresponding* $\Phi_w : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$*, as a map defined on distribution functions, is not convex, but it is quasi-convex and concave. Indeed, let* $P \in \mathcal{P}$ *and, since* $F_P \geq 0$*, we set:*

$$-\Phi_w(P) = \inf(F_P) := \sup\{x \in \mathbb{R} : F_P(x) = 0\}.$$

If F_1*,* F_2 *are two distribution functions corresponding to* P_1*,* $P_2 \in \mathcal{P}$ *then for all* $\lambda \in (0,1)$ *we have:*

$$\inf(\lambda F_1 + (1-\lambda)F_2) = \min(\inf(F_1), \inf(F_2)) \leq \lambda \inf(F_1) + (1-\lambda)\inf(F_2)$$

and therefore, for all $\lambda \in [0,1]$

$$\min(\inf(F_1), \inf(F_2)) \leq \inf(\lambda F_1 + (1-\lambda)F_2) \leq \lambda \inf(F_1) + (1-\lambda)\inf(F_2).$$

Example 75 (Value at Risk $V@R_\lambda$: $\Lambda(x) := \lambda \in (0,1)$) *Take in the definition of the family (3.4) the function* Λ *to be equal to the constant* λ*,* $\Lambda(x) = \lambda \in (0,1)$*. Then*

$$\begin{aligned} F_m(x) \quad &: \quad = \lambda \mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x) \\ \mathcal{A}^m \quad &: \quad = \{Q \in \mathcal{P} \mid F_Q \leq F_m\} \\ \Phi_{V@R_\lambda}(P) \quad &: \quad = -\sup\{m \mid P \in \mathcal{A}^m\} \end{aligned} \tag{3.5}$$

If the random variable $X \in L^0$ *has distribution* P_X *and distribution function* F_X*, notice that:*

$$\sup\{m \mid F_X(m) \leq \lambda\} = \sup\{m \mid F_X(x) \leq \lambda \; \forall x < m\}.$$

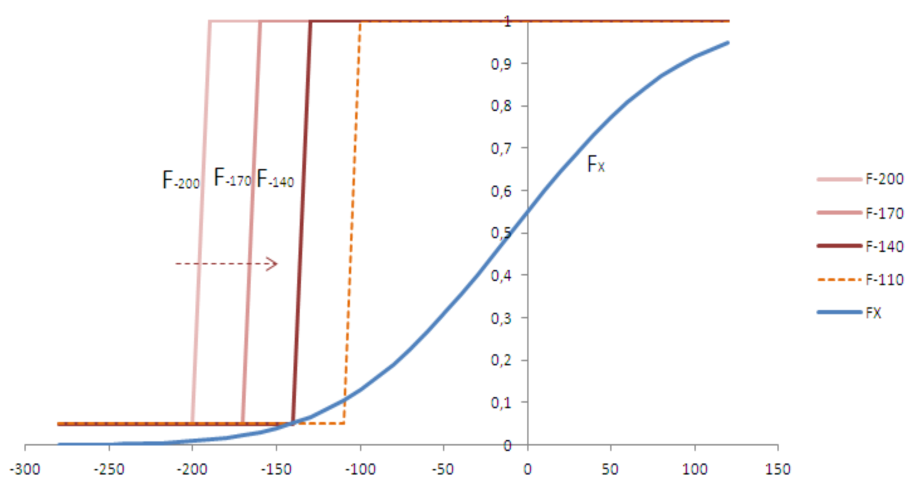
and

$$q_X^+(\lambda) = \sup\{x \in \mathbb{R} \mid \mathbb{P}(X \le x) \le \lambda\}$$

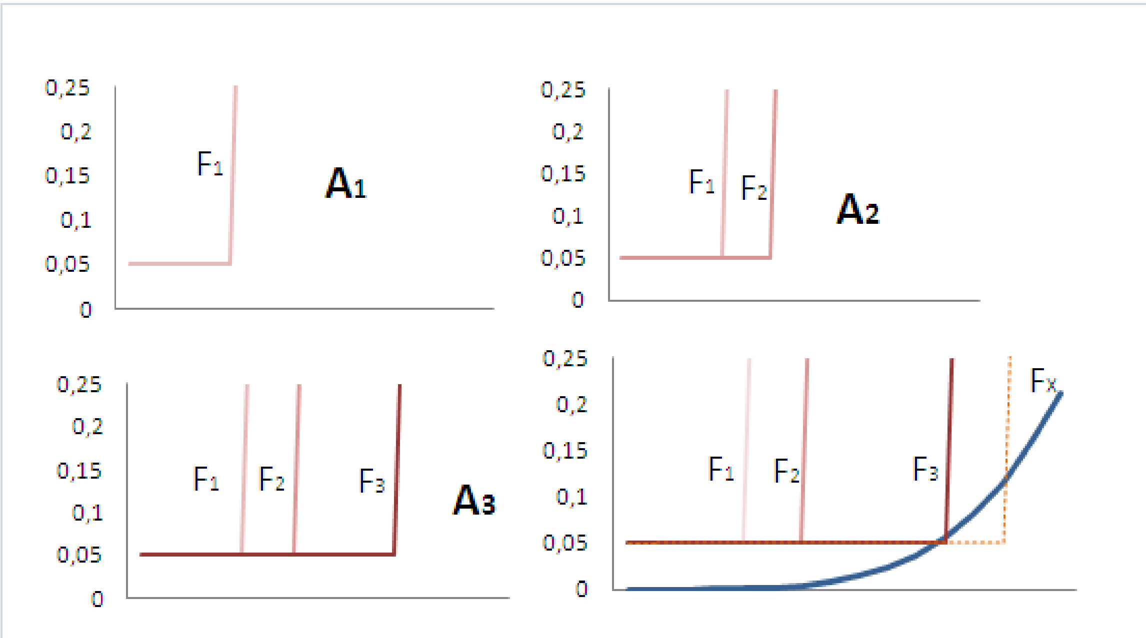
is the right continuous inverse of P_X, i.e. the right quantile of X, then

$$\begin{aligned}\Phi_{V@R_\lambda}(P_X) \quad : &= -\sup\{m \mid P_X \in \mathcal{A}^m\} \\ &= -\sup\{m \mid \mathbb{P}(X \le x) \le \lambda \; \forall x < m\} \\ &= -\sup\{m \mid \mathbb{P}(X \le m) \le \lambda\} \\ &= -q_X^+(\lambda) := V@R_\lambda(X)\end{aligned}$$

coincides with the Value At Risk of level $\lambda \in (0,1)$. The following figure provides a graphical interpretation of the $V@R_\lambda(X)$ as map defined on distributions (in the example $V@R_{0.05}(X) = 140$).



As the family $\{F_m\}$ is feasible, $\Phi_{V@R_\lambda} : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ is (Mon), (Qco), *$\sigma(\mathcal{P}, C_b)$-l.s.c, hence it is a risk measure on $\mathcal{P}$. In addition, it also satisfies* (TrI)*. The following figure shows the decreasing monotonicity over m of the family $\{\mathcal{A}^m\}_{m\in\mathbb{R}}$ of acceptance sets (the* *distributions acceptable at level* 1*,* $P \in \mathcal{A}^1$*, are those below* F_1 *and so on, the blue one is acceptable at level* 3 *but not* 4*).*



As well known, $V@R_\lambda : L^0 \to \mathbb{R} \cup \{+\infty\}$ *is not quasi-convex, as a map defined on random variables, even though the corresponding* $\Phi_{V@R_\lambda} : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$*, as a map defined on distribution functions, is quasi-convex (see [DK10] for a discussion on this issue).*

Example 76 *Fix the family* $\{\Lambda_m\}_{m\in\mathbb{R}}$ *of functions* $\Lambda_m : \mathbb{R} \to [0,1]$ *such that for every* $m \in \mathbb{R}$*,* $\Lambda_m(\boldsymbol{.})$ *is right continuous (w.r.t. x) and for every* $x \in \mathbb{R}$*,* $\Lambda_{\boldsymbol{.}}(x)$ *is decreasing and left continuous (w.r.t. m). Define the family* $\{F_m\}_{m\in\mathbb{R}}$ *of functions* $F_m : \mathbb{R} \to [0,1]$ *by*

$$F_m(x) := \Lambda_m(x)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x). \tag{3.6}$$

It is easy to check that if $\sup_{x\in\mathbb{R}} \Lambda_{m_0}(x) < 1$*, for some* $m_0 \in \mathbb{R}$*, then the family* $\{F_m\}_{m\in\mathbb{R}}$ *is feasible and therefore the associated map* $\Phi : \mathcal{P} \to \mathbb{R} \cup \{+\infty\}$ *is well defined,* (Mon)*,* (Qco)*,* $\sigma(\mathcal{P}, C_b)$*-l.s.c.*

3.3 On the $\Lambda V@R_\lambda$

We now propose a generalization of the $V@R_\lambda$ which appears useful for possible applications whenever an agent is facing some ambiguity on the parameter λ, namely λ is given by some uncertain value in a confidence interval $[\lambda^m, \lambda^M]$, with $0 \leq \lambda^m \leq \lambda^M \leq 1$.

Substantially, our generalization of the $V@R_\lambda$ replaces, in the formulation based on acceptance sets (formula 3.5 in the example 75), the

constant λ

with a

function $\Lambda : \mathbb{R} \to [\lambda^m, \lambda^M]$

In particular, $V@R_\lambda$ corresponds to case $\lambda^m = \lambda^M$ and one typical value is $\lambda^M = 0.05$.

We will distinguish two possible classes of agents:

Risk prudent Agents The main idea is to model the fact that, for a *risk prudent agent*, large losses are acceptable only under smaller probabilities. So that, we fix the *increasing* right continuous function $\Lambda : \mathbb{R} \to [\lambda^m, \lambda^M]$ and we choose, as in (3.4):

$$F_m(x) = \Lambda(x)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x).$$

and set $\lambda^m := \inf \Lambda \geq 0$, $\lambda^M := \sup \Lambda \leq 1$. Since Λ is increasing, the risk adverse agent will reserve more capital for larger losses.

Furthermore, we can also distinguish between several level of risk aversion: given two possible choices Λ_1, Λ_2 for two different agents, the condition $\Lambda_1 \leq \Lambda_2$ means that the agent 1 is more risk averse than agent 2.

Now we set, as in (3.1), $\mathcal{A}^m = \{Q \in \mathcal{P} \mid F_Q \leq F_m\}$ and define as in (3.2):

$$\Lambda V@R(P) := -\sup\{m \mid P \in \mathcal{A}^m\}.$$

Thus, in case of a random variable X

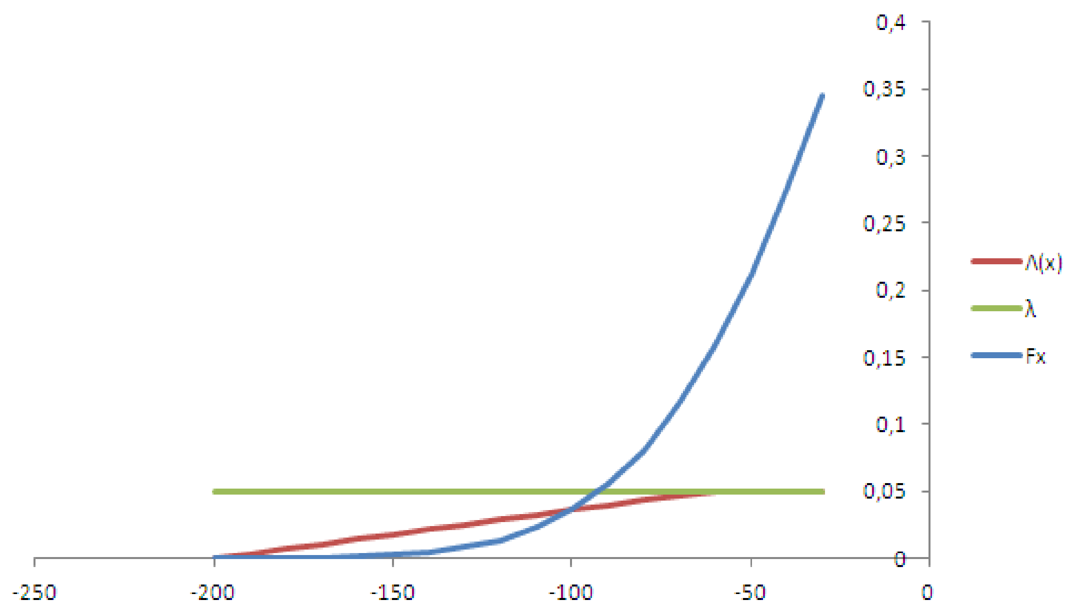
$$\Lambda V@R(P_X) := -\sup\{m \mid \mathbb{P}(X \leq x) \leq \Lambda(x),\ \forall x \leq m\}.$$

In particular it can be rewritten as

$$\Lambda V@R(P_X) = -\inf\{x \in \mathbb{R} \mid \mathbb{P}(X \leq x) > \Lambda(x)\}.$$

as the case of the right quantile of X where $\Lambda(x)$ is substituted to λ.

If both F_X and Λ are continuous $\Lambda V@R$ corresponds to the smallest intersection between the two curves.



The risk adverse agent requires smaller probabilities of greater losses

In this section, we assume that

$$\lambda^M < 1.$$

Besides its obvious financial motivation, this request implies that the corresponding family F_m is feasible and so $\Lambda V@R(P) > -\infty$ for all $P \in \mathcal{P}$.

By an appropriate choice of the function Λ it is possible to provide a model to detect the probability (hence the capital requirement). We need to define for an amount x of money the probability $\Lambda(x)$ that we accept to loose it. Whereas the $V@R_\lambda$ impose the probability λ indipendently by the propension to loose that amount.

Example 77 *One possible simple choice of the function Λ is represented by the step function:*

$$\Lambda(x) = \lambda^m \mathbf{1}_{(-\infty,\bar{x})}(x) + \lambda^M \mathbf{1}_{[\bar{x},+\infty)}(x)$$

The idea is that with a probability of λ^M we are accepting to loose at most $\bar{x}$. In this case we observe that:

$$\Lambda V@R(P) = \begin{cases} V@R_{\lambda^M}(P) & \text{if } V@R_{\lambda^m}(P) \leq -\bar{x} \\ V@R_{\lambda^m}(P) & \text{if } V@R_{\lambda^m}(P) > -\bar{x}. \end{cases}$$

Even though the $\Lambda V@R$ is continuous from below (proposition 71 and 64), it may not be continuous from above, as this example shows. For instance take $\bar{x} = 0$ and P_{X_n} induced by a sequence of uniformly distributed random variables $X_n \sim U\left[-\lambda^m - \frac{1}{n}, 1 - \lambda^m - \frac{1}{n}\right]$. We have $P_{X_n} \downarrow P_{U[-\lambda^m,1-\lambda^m]}$ but $\Lambda V@R(P_{X_n}) = -\frac{1}{n}$ for every n and $\Lambda V@R(P_{U[-\lambda^m,1-\lambda^m]}) = \lambda^M - \lambda^m$.

Remark 78 *(i) If $\lambda^m = 0$ the domain of $\Lambda V@R(P)$ is not the entire convex set $\mathcal{P}$. We have two possible cases*

- *supp*$(\Lambda) = [x^*, +\infty)$*: in this case* $\Lambda V@R(P) = -\inf supp(F_P)$ *for every* $P \in \mathcal{P}$ *such that* $supp(F_P) \supset supp(\Lambda)$*.*

- *supp*$(\Lambda) = (-\infty, +\infty)$*: in this case*

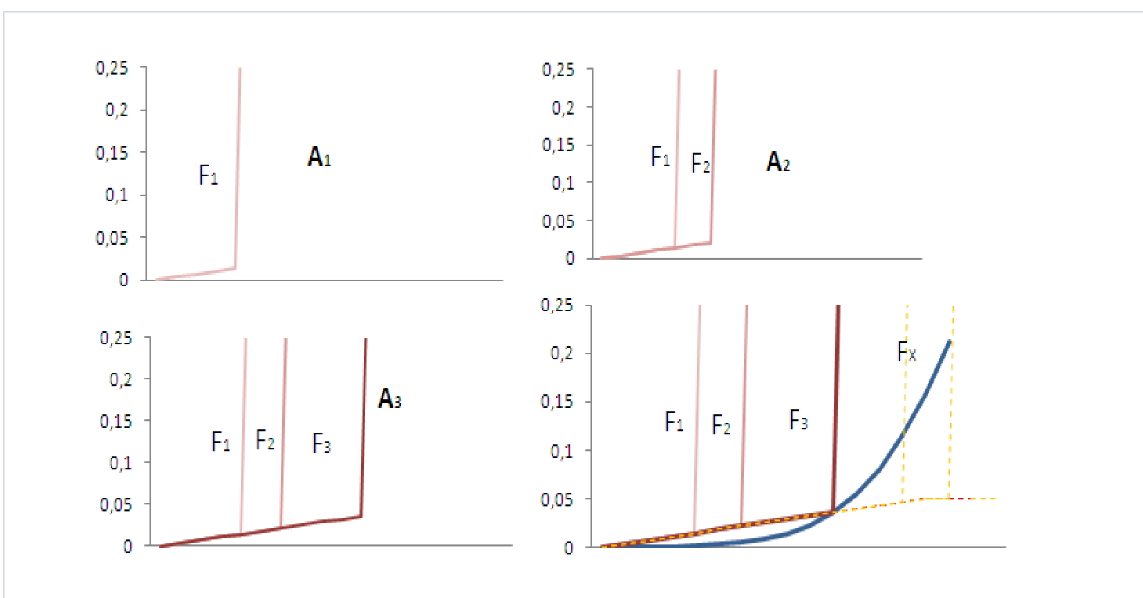
$$\Lambda V@R(P) = +\infty \qquad \text{for all } P \text{ such that } \lim_{x\to-\infty} \frac{F_P(x)}{\Lambda(x)} > 1$$

$$\Lambda V@R(P) < +\infty \qquad \text{for all } P \text{ such that } \lim_{x\to-\infty} \frac{F_P(x)}{\Lambda(x)} < 1$$

In the case $\lim_{x\to-\infty} \frac{F_P(x)}{\Lambda(x)} = 1$ *both the previous behaviors might occur.*

(ii) In case that $\lambda^m > 0$ *then* $\Lambda V@R(P) < +\infty$ *for all* $P \in \mathcal{P}$*, so that* $\Lambda V@R$ *is finite valued.*

The feasibility of the family $\{F_m\}$ implies that the $\Lambda V@R : \mathcal{P} \to \mathbb{R}$ is a well defined, *(Mon)*, *(QCo)* and *(CfB)* (or equivalently $\sigma(\mathcal{P}, C_b)$-l.s.c.) map. The following figure shows the evolution of the acceptance sets when m increases. The acceptance sets $\mathcal{A}_m$ are not anymore the translation of the acceptance set $\mathcal{A}_0$ ($\mathcal{A}_m \neq \mathcal{A}_0 - m$).



We drop in this way cash additivity *(TrI)*, but we obtain another similar property, which is the counterpart of *(TrI)* for the $\Lambda V@R$. Let $\alpha \in \mathbb{R}$ any cash amount

$$\begin{aligned}
\Lambda V@R(P_{X+\alpha}) &= -\sup\{m \mid \mathbb{P}(X+\alpha \le x) \le \Lambda(x),\ \forall x \le m\} \\
&= -\sup\{m \mid \mathbb{P}(X \le x-\alpha) \le \Lambda(x),\ \forall x \le m\} \\
&= -\sup\{m \mid \mathbb{P}(X \le y) \le \Lambda(y+\alpha),\ \forall y \le m-\alpha\} \\
&= -\sup\{m+\alpha \mid \mathbb{P}(X \le y) \le \Lambda(y+\alpha),\ \forall y \le m\} \\
&= \Lambda^{\alpha} V@R(P_X) - \alpha
\end{aligned}$$

where $\Lambda^{\alpha}(x) = \Lambda(x+\alpha)$. We may conclude that if we add a sure positive (resp. negative) amount α to a risky position X then the risk decreases (resp. increases) of the value $-\alpha$, constrained to lower (resp. higher) the level of risk aversion described by $\Lambda^{\alpha} \ge \Lambda$ (resp. $\Lambda^{\alpha} \le \Lambda$). For an arbitrary $P \in \mathcal{P}$ this property can be written as

$$\Lambda V@R(T_{\alpha}P) = \Lambda^{\alpha} V@R(P) - \alpha, \quad \forall\, \alpha \in \mathbb{R},$$

where $T_{\alpha}P(-\infty, x] = P(-\infty, x-\alpha]$.

Risk Seeking Agents Fix the *decreasing* right continuous function $\Lambda : \mathbb{R} \to [0,1]$, with $\inf \Lambda < 1$. Similarly as above, we define

$$F_m(x) = \Lambda(x)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x)$$

and the (Mon), (QCo) and (CfB) map

$$\Lambda V@R(P) := -\sup\{m \in \mathbb{R} \mid F_P \le F_m\} = -\sup\{m \in \mathbb{R} \mid \mathbb{P}(X \le m) \le \Lambda(m)\}.$$

In this case, for eventual huge losses we are allowing the highest level of probability. As in the previous example let $\alpha \in \mathbb{R}$ and notice that

$$\Lambda V@R(P_{X+\alpha}) = \Lambda^{\alpha} V@R(P_X) - \alpha.$$

where $\Lambda^{\alpha}(x) = \Lambda(x+\alpha)$. The property is exactly the same as in the former example but here the interpretation is slightly different. If we add a sure positive (resp. negative) amount α to a risky position X then the risk decreases (resp. increases) of the value $-\alpha$, constrained to a lower (resp. higher) level of risk seeking since $\Lambda^{\alpha} \leq \Lambda$ (resp. $\Lambda^{\alpha} \geq \Lambda$).

Remark 79 *For a decreasing Λ, there is a simpler formulation - which will be used in Section 3.4.3 - of the $\Lambda V@R$ that is obtained replacing in F_m the function Λ with the line $\Lambda(m)$ for all $x < m$. Let*

$$\tilde{F}_m(x) = \Lambda(m)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x).$$

This family is of the type (3.6) and is feasible, provided the function Λ is continuous. For a decreasing Λ, it is evident that

$$\Lambda V@R(P) = \Lambda \widetilde{V}@R(P) := -\sup\left\{m \in \mathbb{R} \mid F_P \leq \tilde{F}_m\right\},$$

as the function Λ lies above the line $\Lambda(m)$ for all $x \leq m$.

3.4 Quasi-convex Duality

As seen in the previous chapter, in literature we find several results about the dual representation of law invariant risk measures. Kusuoka [K01] contributed to the coherent case, whereas Frittelli and Rosazza [FR05] extended this result to the convex case. Jouini et al. [JST06], in the convex case, and Svindland (2010) [S10] in the quasi-convex case, showed that every law invariant risk measure is already weakly lower semicontinuous. Recently, Cerreia-Vioglio et al. [CMMMa] provided a robust dual representation for law invariant quasi-convex risk measures.

In Sections 3.4.1 and 3.4.2 we will treat the general case of maps defined on $\mathcal{P}$, while in Section 3.4.3 we specialize these results to show the dual representation of maps associated to feasible families.

3.4.1 Reasons of the failure of the convex duality for Translation Invariant maps on $\mathcal{P}$

As previously seen the classical convex duality provided by the Fenchel-Moreau theorem (see Thm.24), useful for the dual representation of convex risk measures (see [FR02]), guarantees the representation of any convex and lower semicontinuous functions. Therefore, for any map $\Phi : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$ let Φ^* be the convex conjugate:

$$\Phi^*(f) := \sup_{Q \in \mathcal{P}} \left\{ \int f dQ - \Phi(Q) \right\}, \quad f \in C_b.$$

Applying the fact that $\mathcal{P}$ is a $\sigma(ca, C_b)$ closed convex subset of ca, one can easily check that the following version of the Fenchel-Moreau Theorem holds true for maps defined on $\mathcal{P}$.

Proposition 80 (Fenchel-Moreau) *Suppose that* $\Phi : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$ *is* $\sigma(\mathcal{P}, C_b)$*-l.s.c. and convex. If* $Dom(\Phi) \neq \varnothing$ *then* $Dom(\Phi^*) \neq \varnothing$ *and*

$$\Phi(Q) = \sup_{f \in C_b} \left\{ \int f dQ - \Phi^*(f) \right\}.$$

One trivial example of a proper $\sigma(\mathcal{P}, C_b)$-l.s.c. and convex map on $\mathcal{P}$ is given by $Q \to \int f dQ$, for some $f \in C_b$. But this map does not satisfy the *(TrI)* property. Indeed, we show that in the setting of risk measures defined on $\mathcal{P}$, weakly lower semicontinuity and convexity are incompatible with translation invariance.

Proposition 81 *For any map $\Phi : \mathcal{P} \to \overline{\mathbb{R}}$, if there exists a sequence $\{Q_n\}_n \subseteq \mathcal{P}$ such that $\lim_n \Phi(Q_n) = -\infty$ then $Dom(\Phi^*) = \varnothing$. Thus the only $\sigma(\mathcal{P}, C_b)-$lsc, convex and (TrI) map $\Phi : \mathcal{P} \to \overline{\mathbb{R}}$ is $\Phi = +\infty$.*

Proof. For any $f \in C_b(\mathbb{R})$

$$\Phi^*(f) = \sup_{Q\in\mathcal{P}} \left\{ \int f dQ - \Phi(Q) \right\} \geq \int f d(Q_n) - \Phi(Q_n) \geq \inf_{x\in\mathbb{R}} f(x) - \Phi(Q_n) \uparrow \infty.$$

Observe that a translation invariant map satisfies $\lim_n \Phi(T_n Q) = \lim_n \{\Phi(Q) - n\} = -\infty$, for any $Q \in Dom(\Phi)$. The thesis follows from Proposition 80 and what just proved, replacing Q_n with $T_n Q$. ■

3.4.2 Quasi-convex duality

As described in the Examples in Section 3.2, the $\Phi_{V@R_\lambda}$ and Φ_w are proper, $\sigma(ca, C_b)$-l.s.c., quasi-convex, (Mon) and (TrI) maps $\Phi : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$. Therefore, the negative result outlined in Proposition 81 for the convex case can not be true in the quasi-convex setting.

We recall that one of the main contribution to quasi-convex duality comes from the dual representation by Volle [Vo98].

Here we replicate this result and provide the dual representation of a *(QCo)* and $\sigma(\mathcal{P}, C_b)$-l.s.c. maps defined on the entire set $\mathcal{P}$. The main difference is that our map Φ is defined on a convex subset of ca and not a vector space. But since $\mathcal{P}$ is $\sigma(ca, C_b)$-closed, the first part of the proof will match very closely the one given by Volle. In order to achieve the dual representation of $\sigma(\mathcal{P}, C_b)$ lsc risk measures $\Phi : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$ we will impose the monotonicity assumption of Φ and deduce that in the dual representation the supremum can be restricted to the set

$$C_b^- = \{f \in C_b \mid f \text{ is decreasing}\}.$$

This is natural as the first order stochastic dominance implies (see Th. 2.70 [FS04]) that

$$C_b^- = \left\{ f \in C_b \mid Q, P \in \mathcal{P} \text{ and } Q \preceq P \Rightarrow \int f dQ \leq \int f dP \right\}. \tag{3.7}$$

Notice that differently from [DK10] the following proposition does not require the extension of the risk map to the entire space $ca(\mathbb{R})$.

Proposition 82 *(i) Any $\sigma(\mathcal{P}, C_b)-$lsc and quasi-convex functional $\Phi : \mathcal{P} \to \mathbb{R} \cup \{\infty\}$ can be represented as*

$$\Phi(P) = \sup_{f \in C_b} R\left(\int f dP, f \right) \tag{3.8}$$

where $R : \mathbb{R} \times C_b \to \overline{\mathbb{R}}$ is defined by

$$R(t, f) := \inf_{Q \in \mathcal{P}} \left\{ \Phi(Q) \mid \int f dQ \geq t \right\}. \tag{3.9}$$

(ii) If in addition Φ is monotone then (3.8) holds with C_b replaced by C_b^-.

Proof. We will use the fact that $\sigma(\mathcal{P}, C_b)$ is the relativization of $\sigma(ca, C_b)$ to the set $\mathcal{P}$. In particular the lower level sets will be $\sigma(ca, C_b)$-closed.

(i) By definition, for any $f \in C_b(\mathbb{R})$, $R\left(\int f dP, f\right) \leq \Phi(P)$ and therefore

$$\sup_{f \in C_b} R\left(\int f dP, f \right) \leq \Phi(P), \quad P \in \mathcal{P}.$$

Fix any $P \in \mathcal{P}$ and take $\varepsilon \in \mathbb{R}$ such that $\varepsilon > 0$. Then P does not belong to the $\sigma(ca, C_b)$-closed convex set

$$\mathcal{C}_\varepsilon := \{Q \in \mathcal{P} : \Phi(Q) \leq \Phi(P) - \varepsilon\}$$

(if $\Phi(P) = +\infty$, replace the set $\mathcal{C}_\varepsilon$ with $\{Q \in \mathcal{P} : \Phi(Q) \leq M\}$, for any M). By the Hahn Banach theorem there exists a continuous linear functional that strongly separates P and $\mathcal{C}_\varepsilon$, i.e. there exists $\alpha \in \mathbb{R}$ and $f_\varepsilon \in C_b$ such that

$$\int f_\varepsilon dP > \alpha > \int f_\varepsilon dQ \quad \text{for all } Q \in \mathcal{C}_\varepsilon. \tag{3.10}$$

Hence:

$$\left\{Q \in \mathcal{P} : \int f_\varepsilon dP \leq \int f_\varepsilon dQ\right\} \subseteq (\mathcal{C}_\varepsilon)^C = \{Q \in \mathcal{P} : \Phi(Q) > \Phi(P) - \varepsilon\} \tag{3.11}$$

and

$$\begin{aligned}\Phi(P) &\geq \sup_{f \in C_b} R\left(\int f dP, f\right) \geq R\left(\int f_\varepsilon dP, f_\varepsilon\right) \\ &= \inf\left\{\Phi(Q) \mid Q \in \mathcal{P} \text{ such that } \int f_\varepsilon dP \leq \int f_\varepsilon dQ\right\} \\ &\geq \inf\left\{\Phi(Q) \mid Q \in \mathcal{P} \text{ satisfying } \Phi(Q) > \Phi(P) - \varepsilon\right\} \geq \Phi(P) - \varepsilon. \end{aligned} \tag{3.12}$$

(ii) We furthermore assume that Φ is monotone. As shown in (i), for every $\varepsilon > 0$ we find f_ε such that (3.10) holds true. We claim that there exists $g_\varepsilon \in C_b^-$ satisfying:

$$\int g_\varepsilon dP > \alpha > \int g_\varepsilon dQ \quad \text{for all } Q \in \mathcal{C}_\varepsilon. \tag{3.13}$$

and then the above argument (in equations (3.10)-(3.12)) implies the thesis.

We define the decreasing function

$$g_\varepsilon(x) =: \sup_{y \geq x} f_\varepsilon(y) \in C_b^-.$$

First case: suppose that $g_\varepsilon(x) = \sup_{x\in\mathbb{R}} f_\varepsilon(x) =: s$. In this case there exists a sequence of $\{x_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ such that $x_n \to +\infty$ and $f_\varepsilon(x_n) \to s$, as $n \to \infty$. Define

$$g_n(x) = s\mathbf{1}_{(-\infty,x_n]} + f_\varepsilon(x)\mathbf{1}_{(x_n,+\infty)}$$

and notice that $s \geq g_n \geq f_\varepsilon$ and $g_n \uparrow s$. For any $Q \in \mathcal{C}_\varepsilon$ we consider Q_n defined by $F_{Q_n}(x) = F_Q(x)\mathbf{1}_{[x_n,+\infty)}$. Since $Q_n \preccurlyeq Q$, monotonicity of Φ implies $Q_n \in \mathcal{C}_\varepsilon$. Notice that

$$\int g_n dQ - \int f_\varepsilon dQ_n = (s - f_\varepsilon(x_n))Q(-\infty, x_n] \overset{n\to+\infty}{\longrightarrow} 0, \text{ as } n \to \infty. \tag{3.14}$$

From equation (3.10) we have

$$s \geq \int f_\varepsilon dP > \alpha > \int f_\varepsilon dQ_n \quad \text{for all } n \in \mathbb{N}. \tag{3.15}$$

Letting $\delta = s - \alpha > 0$ we obtain $s > \int f_\varepsilon dQ_n + \frac{\delta}{2}$. From (3.14), there exists $\overline{n} \in \mathbb{N}$ such that $0 \leq \int g_n dQ - \int f_\varepsilon dQ_n < \frac{\delta}{4}$ for every $n \geq \overline{n}$. Therefore $\forall\, n \geq \overline{n}$

$$s > \int f_\varepsilon dQ_n + \frac{\delta}{2} > \int g_n dQ - \frac{\delta}{4} + \frac{\delta}{2} = \int g_n dQ + \frac{\delta}{4}$$

and this leads to a contradiction since $g_n \uparrow s$. So the first case is excluded.

Second case: suppose that $g_\varepsilon(x) < s$ for any $x > \overline{x}$. As the function $g_\varepsilon \in C_b^-$ is decreasing, there will exists at most a countable sequence of intervals $\{A_n\}_{n\geq 0}$ on which g_ε is constant. Set $A_0 = (-\infty, b_0)$, $A_n = [a_n, b_n) \subset \mathbb{R}$ for $n \geq 1$. W.l.o.g. we suppose that $A_n \cap A_m = \emptyset$ for all $n \neq m$ (else, we paste together the sets) and $a_n < a_{n+1}$ for every $n \geq 1$. We stress that $f_\varepsilon(x) = g_\varepsilon(x)$ on $D =: \bigcap_{n\geq 0} A_n^C$. For every $Q \in \mathcal{C}_\varepsilon$ we define the probability $\overline{Q}$ by its distribution function as

$$F_{\overline{Q}}(x) = F_Q(x)\mathbf{1}_D + \sum_{n\geq 1} F_Q(a_n)\mathbf{1}_{[a_n,b_n)}.$$

As before, $\overline{Q} \preccurlyeq Q$ and monotonicity of Φ implies $\overline{Q} \in \mathcal{C}_\varepsilon$. Moreover

$$\int g_\varepsilon dQ = \int_D f_\varepsilon dQ + f_\varepsilon(b_0)Q(A_0) + \sum_{n\geq 1} f_\varepsilon(a_n)Q(A_n) = \int f_\varepsilon d\overline{Q}.$$

From $g_\varepsilon \geq f_\varepsilon$ and equation (3.10) we deduce

$$\int g_\varepsilon dP \geq \int f_\varepsilon dP > \alpha > \int f_\varepsilon d\overline{Q} = \int g_\varepsilon dQ \quad \text{for all } Q \in \mathcal{C}_\varepsilon.$$

■

We reformulate the Proposition 82 and provide two dual representations of $\sigma(\mathcal{P}(\mathbb{R}), C_b)$-lsc Risk Measure $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{\infty\}$. The first one is given in terms of the dual function R used by [CMMMa]. The second one is obtained from Proposition 82 considering the left continuous version of R and rewriting it (see Lemma 84) in the formulation proposed by [DK10]. If $R : \mathbb{R} \times C_b(\mathbb{R}) \to \overline{\mathbb{R}}$, the left continuous version of $R(\cdot, f)$ is defined by:

$$R^-(t, f) := \sup \{R(s, f) \mid s < t\}. \tag{3.16}$$

Proposition 83 *Any $\sigma(\mathcal{P}(\mathbb{R}), C_b)$-lsc Risk Measure $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R}\cup\{\infty\}$ can be represented as*

$$\Phi(P) = \sup_{f\in C_b^-} R\left(\int fdP, f\right) = \sup_{f\in C_b^-} R^-\left(\int fdP, f\right). \tag{3.17}$$

The function $R^-(t, f)$ can be written as

$$R^-(t, f) = \inf \{m \in \mathbb{R} \mid \gamma(m, f) \geq t\}, \tag{3.18}$$

where $\gamma : \mathbb{R} \times C_b(\mathbb{R}) \to \overline{\mathbb{R}}$ is given by:

$$\gamma(m, f) := \sup_{Q\in\mathcal{P}} \left\{\int fdQ \mid \Phi(Q) \leq m\right\},\ m \in \mathbb{R}. \tag{3.19}$$

Proof. Notice that $R(\cdot, f)$ is increasing and $R(t,f) \geq R^-(t,f)$. If $f \in C_b^-$ then $Q \preceq P \Rightarrow \int f dQ \leq \int f dP$. Therefore,

$$R^-\left(\int f dP, f\right) := \sup_{s<\int f dP} R(s,f) \geq \lim_{P_n \uparrow P} R(\int f dP_n, f).$$

From Proposition 82 (ii) we obtain:

$$\Phi(P) = \sup_{f \in C_b^-} R\left(\int f dP, f\right) \geq \sup_{f \in C_b^-} R^-\left(\int f dP, f\right) \geq \sup_{f \in C_b^-} \lim_{P_n \uparrow P} R(\int f dP_n, f)$$
$$= \lim_{P_n \uparrow P} \sup_{f \in C_b^-} R(\int f dP_n, f) = \lim_{P_n \uparrow P} \Phi(P_n) = \Phi(P).$$

by (CfB). This proves (3.17). The second statement follows from the Lemma 84. ■

Lemma 84 *Let Φ be any map $\Phi : \mathcal{P}(\mathbb{R}) \to \mathbb{R} \cup \{\infty\}$ and $R : \mathbb{R} \times C_b(\mathbb{R}) \to \overline{\mathbb{R}}$ be defined in (3.9). The left continuous version of $R(\cdot, f)$ can be written as:*

$$R^-(t,f) := \sup\{R(s,f) \mid s < t\} = \inf\{m \in \mathbb{R} \mid \gamma(m,f) \geq t\}, \tag{3.20}$$

where $\gamma : \mathbb{R} \times C_b(\mathbb{R}) \to \overline{\mathbb{R}}$ is given in (3.19).

Proof. Let the RHS of equation (3.20) be denoted by

$$S(t,f) := \inf\{m \in \mathbb{R} \mid \gamma(m,f) \geq t\},\ (t,f) \in \mathbb{R} \times C_b(\mathbb{R}),$$

and note that $S(\cdot, f)$ is the left inverse of the increasing function $\gamma(\cdot, f)$ and therefore $S(\cdot, f)$ is left continuous.

Step I. To prove that $R^-(t,f) \geq S(t,f)$ it is sufficient to show that for all $s < t$ we have:

$$R(s,f) \geq S(s,f), \tag{3.21}$$

Indeed, if (3.21) is true

$$R^-(t,f) = \sup_{s<t} R(s,f) \geq \sup_{s<t} S(s,f) = S(t,f),$$

as both R^- and S are left continuous in the first argument.

Writing explicitly the inequality (3.21)

$$\inf_{Q\in\mathcal{P}}\left\{\Phi(Q)\mid\int fdQ\geq s\right\}\geq\inf\left\{m\in\mathbb{R}\mid\gamma(m,f)\geq s\right\}$$

and letting $Q\in\mathcal{P}$ satisfying $\int fdQ\geq s$, we see that it is sufficient to show the existence of $m\in\mathbb{R}$ such that $\gamma(m,f)\geq s$ and $m\leq\Phi(Q)$. If $\Phi(Q)=-\infty$ then $\gamma(m,f)\geq s$ for any m and therefore $S(s,f)=R(s,f)=-\infty$.

Suppose now that $\infty>\Phi(Q)>-\infty$ and define $m:=\Phi(Q)$. As $\int fdQ\geq s$ we have:

$$\gamma(m,f):=\sup_{Q\in\mathcal{P}}\left\{\int fdQ\mid\Phi(Q)\leq m\right\}\geq s$$

Then $m\in\mathbb{R}$ satisfies the required conditions.

Step II : To obtain $R^-(t,f):=\sup_{s<t}R(s,f)\leq S(t,f)$ it is sufficient to prove that, for all $s<t$, $R(s,f)\leq S(t,f)$, that is

$$\inf_{Q\in\mathcal{P}}\left\{\Phi(Q)\mid\int fdQ\geq s\right\}\leq\inf\left\{m\in\mathbb{R}\mid\gamma(m,f)\geq t\right\}. \tag{3.22}$$

Fix any $s<t$ and consider any $m\in\mathbb{R}$ such that $\gamma(m,f)\geq t$. By the definition of γ, for all $\varepsilon>0$ there exists $Q_\varepsilon\in\mathcal{P}$ such that $\Phi(Q_\varepsilon)\leq m$ and $\int fdQ_\varepsilon>t-\varepsilon$. Take ε such that $0<\varepsilon<t-s$. Then $\int fdQ_\varepsilon\geq s$ and $\Phi(Q_\varepsilon)\leq m$ and (3.22) follows. ■

3.4.3 Computation of the dual function

The following proposition is useful to compute the dual function $R^-(t,f)$ for the examples considered in this paper.

Proposition 85 *Let $\{F_m\}_{m\in\mathbb{R}}$ be a feasible family and suppose in addition that, for every m, $F_m(x)$ is increasing in x and $\lim_{x\to+\infty}F_m(x)=1$. The associated map $\Phi:\mathcal{P}\to$*

$\mathbb{R}\cup\{+\infty\}$ *defined in (3.2) is well defined, (Mon), (Qco) and* $\sigma(\mathcal{P},C_b)-$*l.s.c. and the representation (3.17) holds true with* R^- *given in (3.18) and*

$$\gamma(m,f)=\int f dF_{-m}+F_{-m}(-\infty)f(-\infty). \tag{3.23}$$

Proof. From equations (3.1) and (3.3) we obtain:

$$\mathcal{A}^{-m}=\{Q\in\mathcal{P}(\mathbb{R})\mid F_Q\leq F_{-m}\}=\{Q\in\mathcal{P}\mid\Phi(Q)\leq m\}$$

so that

$$\gamma(m,f):=\sup_{Q\in\mathcal{P}}\left\{\int fdQ\mid\Phi(Q)\leq m\right\}=\sup_{Q\in\mathcal{P}}\left\{\int fdQ\mid F_Q\leq F_{-m}\right\}.$$

Fix $m\in\mathbb{R}$, $f\in C_b^-$ and define the distribution function $F_{Q_n}(x)=F_{-m}(x)\mathbf{1}_{[-n,+\infty)}$ for every $n\in\mathbb{N}$. Obviously $F_{Q_n}\leq F_{-m}$, $Q_n\uparrow$ and, taking into account (3.7), $\int fdQ_n$ is increasing. For any $\varepsilon>0$, let $Q^\varepsilon\in\mathcal{P}$ satisfy $F_{Q^\varepsilon}\leq F_{-m}$ and $\int fdQ^\varepsilon>\gamma(m,f)-\varepsilon$. Then: $F_{Q_n^\varepsilon}(x):=F_{Q^\varepsilon}(x)\mathbf{1}_{[-n,+\infty)}\uparrow F_{Q^\varepsilon}$, $F_{Q_n^\varepsilon}\leq F_{Q_n}$ and

$$\int fdQ_n\geq\int fdQ_n^\varepsilon\uparrow\int fdQ^\varepsilon>\gamma(m,f)-\varepsilon.$$

We deduce that $\int fdQ_n\uparrow\gamma(m,f)$ and, since

$$\int fdQ_n=\int_{-n}^{+\infty}fdF_{-m}+F_{-m}(-n)f(-n),$$

we obtain (3.23). ■

In the following examples $m\in\mathbb{R}$, $f\in C_b^-$ and f^l is the left inverse of f.

Example 86 *Computation of the dual function* R^- *for the* $V@R$ *and the worst case measure.*

The family $\{F_m\}_{m\in\mathbb{R}}$ *is given by (see the Examples 74 and 75)* $F_m = \lambda 1_{(-\infty,m)} + 1_{[m,+\infty)}$, *for* $\lambda \in [0,1)$. *Hence we get from (3.23)*

$$\gamma(m,f) = (1-\lambda)f(-m) + \lambda f(-\infty).$$

If $\lambda > 0$, *from (3.18) and (3.17)*

$$\begin{aligned} R^{-}(t,f) &= -f^{l}\left(\frac{t-\lambda f(-\infty)}{1-\lambda}\right), \\ \Phi_{V@R_\lambda}(P) &= -\inf_{f\in C_b^-} f^{l}\left(\frac{\int f dP - \lambda f(-\infty)}{1-\lambda}\right) \end{aligned}$$

If $\lambda = 0$, $\gamma(m,f) = f(-m)$ *and from (3.18), (3.17)*

$$\begin{aligned} R^{-}(t,f) &= -f^{l}(t), \\ \Phi_w(P) &= -\inf_{f\in C_b^-} f^{l}\left(\int f dP\right) \end{aligned}$$

Example 87 *Computation of* $\gamma(m,f)$ *for the* $\Lambda V@R$.

As $F_m = \Lambda(x)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x)$, *we compute from (3.23):*

$$\gamma(m,f) = \int_{-\infty}^{-m} f d\Lambda + (1-\Lambda(-m))f(-m) + \Lambda(-\infty)f(-\infty).$$

If Λ *is decreasing we may use Remark 79 to derive a simpler formula for* γ. *Indeed,* $\Lambda V@R(P) = \Lambda\widetilde{V}@R(P)$ *where* $\forall m \in \mathbb{R}$

$$\widetilde{F}_m(x) = \Lambda(m)\mathbf{1}_{(-\infty,m)}(x) + \mathbf{1}_{[m,+\infty)}(x)$$

and so

$$\gamma(m,f) = \Lambda(-m)f(-\infty) + (1-\Lambda(-m))f(-m),$$

which is increasing in m.

Chapter 4

Existing Bibliometric Indices

This chapter wants to be an introduction to the issue of the evaluation of the scientific research. The set of quantitative methods used to measure scientific production is known as *bibliometrics*. This science is a branch of the *scientometrics* risen during the 60s and 80s of the 20th century in order to measure and analyze science. The larger availability of the online database, such as Google Scholar, ISI Web of Science by Thomson Reuters, Scopus and MathSciNet has recently brought a big diffusion of citation metrics. Thus *citation analysis* is actually one of the most widely used methods of bibliometrics.

We can distinguish between citation metrics for ranking journals from those for ranking scientists. In such direction a lots of indices were developed, starting with the *h-index* by Hirsh, 2005 [H05], that was the first to consider at same time the productivity in terms of number of publications and the research quality or impact in terms of citations per publications. After its introduction, the h-index has received wide attention from the scientific community and it has been extended by many authors who have proposed other indices in order to overcome some of its drawbacks (a recent overview has been provided by [ACHH].

We introduce in Section 4.1 the issue of the evaluation of the scientific research, with a particular regard to the Italian case. In Section 4.2 we present an historical overview of the bibliometrics indices, pointing out the relative advantages and disadvantages. In Section 4.3 we report some existing results on the axiomatic approach.

4.1 On the valuation of scientific research

In the recent years the evaluation of the scientist's performance has become increasingly important. In fact, most crucial decisions regarding faculty recruitment, accepting research projects, research time, academic positions, travel money, award of grants and promotions depend on great extent upon the scientific merits of the involved researchers.

The scope of the valuation of the scientific research is mainly twofold:

- Provide an updated picture of the existing research activity, in order to allocate financial resources in relation to the scientific quality and scientific production;
- Determine an increase in the quality of the scientific research (of the structures).

The methodologies for the valuation can be divided into two categories, that can be used individually or jointly:

- *qualitative analysis* (or content analysis), based on:
  - internal judgments committee;

  - external reviews of peer panels;

- *quantitative analysis* (or context valuation), based on:
  - bibliometrics (i.e. statistics derived from citation data);
  - characteristics of the Journals associated to the publications.

Both the methodologies have advantages and disadvantages. We are going to see them in the details. In particular, since our aim is to introduce in the next chapter a new measure for the scientific research evaluation we are going to highlight which drawbacks our Scientific Research Measure is able to resolve.

Qualitative (or content) analysis: Pros & Cons

The qualitative analysis is based on expert reviews and it has the advantage that actually evaluates the *quality* of the scientific content. On the other hand it essentially has the following drawbacks; it is:

- **expensive,** in terms of time and people involved; therefore it *can not be used systematically;*

- **subjective,** since the result *depends on the referees.* Hence the qualitative analysis generates new problems regarding the criteria for the choice of the referee, that has to be competent and reliable, and regarding the evaluation of his way of operating, that has to be honest and without any conflict of interests. We have also to consider the *non-uniformity of the judgment*, as each evaluator has a personal scale of preferences leading to different ranking (specially in different areas).

Quantitative (or context) analysis: Pros & Cons

In order to resolve the disadvantages of the qualitative evaluation and thank to the larger availability of the online database several different bibliometric measures have been introduced.

In the 1960s Garfield proposed *citations* like measure of the research quality-impact, introducing the '*impact factor*' that now represents one of the most used citation metrics. The impact factor evaluates a journal by calculating for each article in a particular year the average number of citations to articles of the same journal over a time interval of two years.

We are agree with the UMI report (2010) [UMI]: citation metrics oriented to rank journals, as the impact factor, give some information on the quality of the researcher's work, since the best papers tend to be sent to journals that the scientific community retain to be the best, and, vice versa, the best journals tend to be more selective. On the other hand, one of the most recognized misuses of the impact factor is its use to compare individual papers and scientists. In favour of this opinion the International Mathematical Union has pointed out in the Citation Statistics Report (2008) [CIT] that even if two journals have different impact factor, most of the times a randomly selected paper from one has at least as many citations as a randomly selected article from the other one. Hence, once citations are accepted as measure of quality of the scientific production, the impact factor of the journal in which the paper appears is not a reliable criterion of evaluation.

In the past years many indices were developed to quantify the production of the researchers, such as the total number published papers and the number of papers published in a period of time; or the impact of their publications, such as the total number of citations, the average number of citations per paper, the number and percentage of significant papers (with more than a certain amount of citations).

In 2005 Hirsch [H05] proposed the *h-index,* that is now the most popular and used citation-based metric, the first one to consider at same time the productivity in terms of number of publications and the research quality or impact in terms of citations per publications.

After its introduction, the h-index has received wide attention from the scientific community and it has been extended by many authors who have proposed other indices in order to overcome some of its drawbacks (see Bornmann and Daniel, 2007 [BD07]).

Advantages The use of bibliometric indices is facilitated by the fact that they are :

- **easily accessible,** from the online databases. For this reason, they also give an idea of the *international relevance* of the paper. In this occasion we mention the most important and recognized citation databases:
  - *ISI WoS*: http://portal.isiknowledge.com/portal.cgi?DestApp=WOS&Func=Frame
  - *Google Scholar* (free accessibly): http: //scholar.google.com/
  - *Scopus*: http: //www.scopus.com/
  - *MathSciNet* (for the mathematical sector): http://www.ams.org/mathscinet/
- **not expensive**: can be used systematically, especially if tested - every n years - with peer review.

- “**objective**”.

Drawbacks On the other hand, bibliometric indices based on citations have some *disadvantages* that make them a complex subject during the evaluation procedures. Several critics to the use of the citations as key factor of the scientific quality are underlined by the Citation Statistics Report of the International Mathematical Union (2008) [CIT]. However, many of these critics can be satisfactorily addressed. Hence, we are going to highlight the drawbacks and our point of view or proposal of solution:

- **improper comparison of researchers belonging to different fields**, since the size of the scientific communities is different and also their habits regarding citations.

  Response: Each bibliometric index has to be only used to rank each author inside his scientific community, providing a ranking in **relative** terms (e.g.: top 10% - top 30% - average -median...). In such way we have a comparison among different areas, that in the Italian system corresponds to the scientific field ("Settore Concorsuale" as proposed in [ANVUR]).

- **improper comparison of researchers having different seniorities.**

  Response: Each bibliometric index has to be only used to rank each author having the same seniority. We propose an index calibrated to different researcher’s seniorities as well as different areas.

- **different databases provide different citations.**

  Response: We are agree, but there are studies that have shown that different databases (especially Scopus and ISI WoS) provide different numbers (in terms of citation for each paper) but maintain - more or less - the overall ranking. The study of Bar-Ilan 2008 [B08] brings out that except for few cases the differences in the h-indices (calculated on a group of researchers belonging to different fields) between ISI WoS and Scopus are not significant. Regarding Google Scholar, we are agree with the common opinion that the data collection is often inaccurate, but as pointed out in the American Scientist Open Access Forum (2008) [ASOAF] Google Scholar's accuracy is growing daily.

- **disincentive** for young researchers to study subjects more innovative but less popular.

  Response: True, even though this could be compensated by the consideration that innovative paper (in a new field) typically receive many citations.

- **self citations.**

  Response: On this issue there are several point of view. Many authors propose to exclude self citations from the calculation of each index, because cause changes in the behavior of scientists publishing and the total number of citations can be artificially increased. Other authors argue that there is nothing wrong with self-citations (Katsaros, Akritidis, and Bozanis, 2009 [KAB09]); in many cases, they can effectively describe the "authoritativeness" of an article as in the case the self-cited author is a pioneer in his field.

- **precision problem** due to the difficulty to collect the scientist's complete publication list and hence to distinguish between scientists that have the same name.

  Response: This problem can be easily addressed by the systematic use of Author

Codes (a code that identify the author). In the italian system the issue could be faced and solved by the next constitution of the official web site of the professor (CINECA) and, in future, by the official database of the publication (ANPRePS).

- **a single number is insufficient for the evaluation of a complex feature**, such as the scientific research.

  Response: We agree, infact in our opinion it is necessary to have multiple metrics (including time-based metrics) and we propose one of them. Furthermore, we believe that this argument should not lead to abandon the search of appropriate multiple metrics.

- **quality** of the scientific research **can not be reduced to citations.**

  *Response:*Agree, indeed it is only one component that however should be properly quantified.

- **subjective interpretation of citations**, it can be more subjective than the judgment of experts - see Citation Statistics Report of the International Mathematical Union (2008) ([CIT]).

  Response: It has been pointed out - see the discussion in the American Scientist Open Access Forum, 2008 [ASOAF]- that citation metrics are extremely correlated with peer reviews. Hence the issue could be overpassed validating the new metric against other non metric criterion already validated.

Hence, we conclude that economic considerations strongly depone of using the quantitative analysis on a systematic (yearly) base, while peer review is more plausible on a multiple year base and should also be finalized to check, harmonize, and tune the outcomes based on bibliometric indices. Furthermore, we highlight that the output of the valuation should be the classification of authors (and structures) into few classes of homogeneous research quality: it is not intended to provide a fine ranking.

In the followings we are going to treat an hot problem of the Italian system: the determination of the criteria for the choice of the referees and for the candidate's access to the national scientific qualification test for the position of associate professor and full professor.

4.1.1 ANVUR criteria for the scientific qualification license in the Italian system

The National Agency of Evaluation and Research University System (ANVUR) was established by the art 2 in the DL n.262 in order to streamline the assessment system of the *quality* of the universities and research institutions recipients of public funds, as well as the *efficiency* and *effectiveness* of public programs for the research funding.

We report its contribution (expressed in the recent documents [ANVUR] and [ANVURb]) to the pending debate on the criteria and parameters for the selection of candidates who aim to obtain the national scientific license and the choice of the full professors who apply to be referees.

The principles Because of the significant differences in the scientific fields, ANVUR has stated *general criteria* leading to *different thresholds for different sectors.* The choice of the criteria satisfies the principle of the *progressive improvement of the scientific quality* of the qualified professors*, as measured by indices* of the scientific productivity.

This principle turns out to be necessary on the light of the growing global competition to attract funding and researchers, but also responds to a specific condition of the Italian system, as the retirement of about one-third of the professors between 2015 and 2020.

The ANVUR's objective is to avoid the negative effects of a massive recruitment in a short time period, such as the decay of the scientific quality. This principle also aims to bring the recruitment system of the Italian universities at the level of the advanced countries, in order to attract international researchers and solve the issue of the so-called brain drain, in particular of young researchers.

ANVUR remarks that the criteria are defined as necessary condition to the *access* to the qualification procedure. The evaluation of each candidate is leaved to the referees.

Criteria for the access to the national qualification procedure According to ANVUR, the fundamental criteria that meet the above principles have to be considered as *conditions on indices of scientific production* (literally, parameters denoting the quality of the scientific production), and they are:

1. Candidates must have indices of scientific production *normalized by the academic age* (i.e. years since the first scientific publication, taking into account periods of leave in according to the law of at least 5 months, except for study reason) *above the median* of the *specific sector* and *academic title* (associate/full professor) for which

the candidate is applying;

2. Candidates must have a reasonable *continuity* in the scientific production, measured over the past 5 years for the qualification to associate professor (respectively, 10 years for the position of full professor), even in this case, taking into account periods of leave as specified in 1.

The criterion of the median aims to raise the academic scientific quality over time, since it is flessible and dynamic. There have been two important different critics to this criterion:

- the first one is the possible exclusion of scholars with an high profile, who don't deliberately publish so much. The ANVUR answer has been that it is not appropriate to use outliers (individuals who are placed in extreme tails of the distribution) to define the statistical properties of the distribution. Furthermore, none of the critics have brought an empirical evidence of a wide group of scientists that would have been penalized by the adoption of the median criterion.
- the second one is of opposite sign: the possible selection of scholars who published numerous works of low quality. The ANVUR answer will be discussed later .

The choice of the normalization with respect to the academic age aims not to disadvantage younger applicants. Furthermore, the decision to anchor the computing of the academic age at the date of the first publication is essentially practical: it is easily obtainable by bibliometric databases.

ANVUR suggests a further criterion: the *quality profile* of the figure of associate professor and full professor. These profiles, accurately defined in the document, will serve as a guide for the referees in the evaluation of the candidates.

Criteria similar to these for candidates must be applied to the choice of the referees (in the position of full professor). The only difference consists in the lack of academic age normalization.

The indices of scientific production Regarding the candidates, ANVUR proposes to use the following indices of scientific production:

1. For the CUN areas from 1 to 9 (mathematics, computer science, physics, chemistry, earth, biological, medical and agricultural science, civil and industrial engineering) and for some specific scientific field (in which we find also mathematics for economics and finance) in the area from 10 to 14, the parameters are:

   a) the number of articles in journals and monographs appeared in *ISI WoS* or *Scopus* in the last 10 years . This index has to be normalized for the academic seniority only in the case this age is less than 10 years;

   b) the total number of citations;

   c) the h-index (eventually integrated or substituted with other new validated index in the next future).

   The specified indices must be calculated by using any database that have reached a general international consensus, as the ISI WoS, Scopus and Google Scholar, and must be validated by ANVUR.

The procedure for the application of these indices of scientific production is:

- scholars with the first index (at the point a) less than or equal to the median can not obtain the qualification license;
- scholars with the first index (at the point a) greater than the median, must have at least one of the other two indices (at the points b and c) greater than the median if they want to gain the access to the selection procedure.

The combined use of the h-index and the number of citations has the scope to avoid the h-index distortion of not considering the further citations in the h-core.

2. For the CUN areas from 10 to 14 (antiquity, historical, legal, economic and political science) and except for the specific scientific sector in 1, the index is the number of publications (with the exception of congress acts) over the past 10 years, opportunely weighted taking into account the different commitment in the production of articles and monographs and the different impact of work published in international and national journal.

The indices of scientific production for the full professors that apply for the role of referees are the same of those for candidates, with the only exception that the age normalization is not required.

In order to guarantee the scientific nature of the publications, each candidate and referee must fill the personal profile in the CINECA web site, waiting for the institution of the official database of the publications (ANPRePS).

There have been several critics about the choice of the *bibliometric source*. Regarding the limitation to use Scopus and ISI WoS for the computation of the indices, ANVUR retains that must be used sources that are easily accessible, with wide impact and straight approach. ANVUR is available to the examination of further sectorial source having appropriate and straight criteria (the approval will be anticipated before the publication of the median computation). Regarding Google Scholar, ANVUR is aware of the limits and suggests to use it indirectly, by using the h-index in the *Publish or Perish* web site, which uses the Google Scholar database integrated with an information retrieval system.

Regarding the *different quality* of the journals in ISI WoS and Scopus, the use of the first index (number of publications) could penalize authors that publish fewer articles in journals of higher quality. This critic is strictly linked with the above one risen about the median criterion. It is wide recognized that the average quality of papers in the journals collected by ISI and Scopus is higher, but it's necessary to take into account that anyway the quality of these journals remains different. To this end, ANVUR has proposed a new formulation that aims to turn the choice of the three criteria hierarchically ordered with the choice of *two over the three indices* seen at point 1. So that, the three medians will be distinctly calculated and then will be included to the qualification test any scholar who satisfies two of these three medians, with all possible combinations. In such way it is also possible to find a solution for the second issue underlined for the median criterion, no incentive will be provided for young scholars to produce many low quality works.

4.2 Overview of the bibliometric indices

Originally, many indices were developed to quantify only the production of researchers, e.g. the total number of published papers in a period of time, or only the impact of their publications, e.g. the total number of citations, the average number of citations per paper, the number and percentage of significant papers (with more than a certain amount of citations).

In order to evaluate better the performance of a scientist in 2005 Hirsch [H05] proposed the **h-index.**

Definition 88 (h-index) *A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each.*

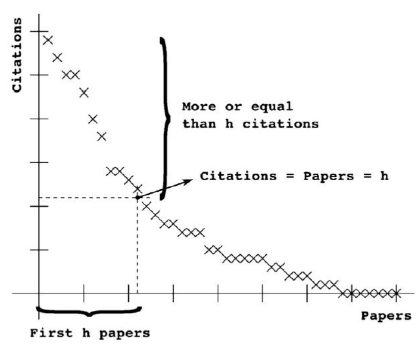


Fig. 4.2. Grafical interpretation of the

h-index

The main **advantage** of the *h*-index is that it measures at the same time the *productivity* in terms of number of publications and the research *quality* in terms of citations per publication. It is also *simple* and *easy to compute*, for example using ISI WoS it is enough to order the scientist's publication list by the field 'Time Cited'.

This new performance measure is *robust* since it doesn't consider the lowly cited papers and it also doesn't count the excess of citations of the papers belonging to the top h (highly cited papers). This feature can be considered also a **disadvantage**, indeed the h-index considers having the same performance the following profiles of scientists: one with 10 publications of 10 citations each, the second one with 100 publications but only 10 of 10 citations and a further scientist with only 10 papers but of 100 citations each.

A further disadvantage is that the h-index is *field and time dependent*, hence it is not possible to compare authors in different disciplines and/or with a different career length. Another aspect is that the h-index can *never decrease*, therefore retired researchers (or simply inactive) may remain with a high h-index even if they have stopped to publish new papers since the citations of previous papers may increase. According to Jin et al. (2007) [JLRE] the h-index lacks sensitivity to performance changes.

As all indices based on citations the h-index suffers of the problems due to the *self-citations*. Even if it is less vulnerable (more robust) to self-citations than the traditional metrics (i.e., total, average, max, min, median citation count) (Katsaros et al. [KAB09]) it is not immune to them (Schreiber, 2007 [S07]). The usual practice is to exclude them, leading to a more complicated computation of the h-index, because the bibliometric databases do not automatically allow to avoid self-citations.

Furthermore, the h-index is negatively influenced by the *"effect Matthew"*, the phenomenon that fame breeds fame, or in other words often-cited papers get cited disproportionately more often than those that are less widely known, and influential authors gain more influence.

Finally, the h-index does not take into account several different variables that are often useful to evaluate the production of researchers, like *time* and the number of *co-authors*.

Due to these disadvantages in the years after several authors proposed corrections and complementary indices.

In order to take into account the *career length*, Hirsch himself [H05] proposed the quotient **m**, obtained dividing the scientist's h value by the number of years y since the first scientific publication, formally

$$m = \frac{h}{y}$$

Later, Egghe [E06] considered an advantage of the h-index not to take into account the papers with a low number of citations and a disadvantage the lack of consideration of the further citations received over the time by the highly cited papers. For this reason he introduced the **g-index**, that he saw "as an improvement of the h-index to measure the global citation performance of a set of articles".

Definition 89 (g-index) *Given a set of articles ranked in decreasing order of the number of citations that they received, the* g-index *is the (unique) largest number such that the top g articles received (together) at least g^2 citations.*

This means that g articles received on average g citations. The g-index is more suitable than the h-index to characterize the overall impact of the publications of a scientist (Schreiber (2008a) [S08a]) and it is more sensitive than the h-index in the evaluation of the "selective scientists", who show in average higher g-index/h-index ratio and a better position in g-index ranking than in the h-index ones (Costas and Bordons (2008) [CB08]).

However, the g-index has some problems. It is more influenced by self-citations than the h-index (Schreiber (2008b) [S08b]) and it is very highly influenced by isolate successful paper: if the articles of a researcher usually receive few citations, but a particular article gets a big success with a large number of citations, the g-index may rise to a large amount compared to those researchers with a lot of average citations per article (Alonso et al. 2010 [ACHHb]). These results led to consider that the g- and h-index can not be substitute each other, but they are complementary.

Following the same idea of the g-index to give more weight to the most cited articles, Kosmulski (2006) [K06] described the h^2**-index**:

Definition 90 (h^2-index) *A scientist's h^2-index is defined as the highest natural number such that his h^2 most-cited papers received each at least $[h^2]^2$ citations and the other papers have $\leq [h^2]^2$ citations each.*

Clearly, the h^2-index is always lower than the h-index. This idea can be still generalized, by the definition of the h^x**-index**. Compared to the h-index, the h^2-index is probably appropriate in the fields where the typical number of citations per article is relatively high, like chemistry and physics, while the h-index is preferred in mathematics or astronomy. When the typical number of citations per article is higher than in chemistry, like in medicine and biology, $x = 2.5$ may be more appropriate. According to Jin et al (2007) [JLRE] the main advantage of the h^2-index is that it reduces the precision problem but it is not enough sensitive.

"Since the scientists do not publish the same number of articles the original h-index is not the fairer metric", so Sidiropoulos et al. (2007) [SKM07] motived the introduction of a normalized version of the h-index, namely the **normalized h-index**:

$$h^n = \frac{h}{N_p}$$

where h is the h-index and N_p is the total number of articles published by the scientist. This particular index has a negative property: it rewards less productive researchers.

The authors also introduced also a pair of generalizations of the h-index. They defined the **contemporary h-index** in order to take into account the *age of the paper*, pointing out senior scientists with an high h-index but now inactive or retired and brilliant young scientists who are expected to contribute with a large number of significant works in the near future, but that nowadays have a small number of important papers due to the time constraint.

Definition 91 (contemporary h-index) *A researcher has contemporary h-index h^c if h^c of its N_p articles get a score of $S_c(j) \geq h^c$ each, and the rest $(N_p - h^c)$ articles get a score of $S_c(j) \leq h^c$, where $S_c(j)$ is a novel score of the $j-th$ article such that:*

$$S_c(j) := \gamma \cdot (Y(now) - Y(j) + 1)^{-\delta} \cdot cit_j$$

where $Y(now)$ is the current year, $Y(j)$ and cit_j are respectively the publication year and the number of citations of the $j-th$ article.

Setting $\delta = 1$, then $S_c(j)$ is the number of citations that article j has received, divided by the "age" of the article. If δ is close to zero, then the impact of the time penalty is reduced, and this variant coincides with the original h-index for $\gamma = 1$. The coefficient γ is used in order to create a meaningful h-index, otherwise dividing the number of citations by the time interval, the quantities $S_c(j)$ would became too small. In their experiments the authors used $\gamma = 4$, so that the citations of an article published during the current year count four times, while the citations of an article published four years ago count only one time and so on. In this way, an old article gradually loses its "value", even if it still cited. In other words, they mainly take into account the newer articles.

For the second case, the authors considered the *"age" of each citation*, namely the year when an article acquired a particular citation. This allow to identify senior researchers whose contributions are still influential even if published long time ago, and also to disclose trendsetters, namely scientists whose work is considered pioneering, such to create a new "trendy" line of research. To handle this, they proceeded assigning to each citation of a paper an exponentially decaying weight, which is as a function of the "age" of the citation. Thus, they defined the **trend h-index,** h^t-index, as in the Definition 91, with the only difference that:

$$S^t(j) := \gamma \cdot \sum_{\forall c \in cit_j} (Y(now) - Y(c) + 1)^{-\delta}$$

and $Y(c)$ is the year of the citation $c \in cit_j$.

In 2006 Rousseau ([R06]) had coined the term **Hirsch-core** (or h-core) as the set consisting of the first h articles. Later Jin et al. (2007) ([JLRE, JLRE]) proposed a following definition:"The **Hirsch core** can be considered as a group of high-performance publications, with respect to the scientist's career". On the basis of this new concept many authors proposed further indices.

Jin (2006) ([J06]) was the first to propose the use of the average of the Hirsch core citations, by the A**-index**:

$$A = \frac{1}{h}\sum_{j=1}^{h} cit_j$$

where h is the h-index and cit_j is the number of citations of the j-th most cited article. Like the g-index it keeps into account all the citations of the articles in the h-core. The A-index can increase even when the h-index remains unchanged.

Jin et al. (2007) ([JLRE, JLRE]) noted also that the A-index penalizes the best scientists with an high h-index as involves a division by h, thus they proposed the R**-index**, defined as

$$R = \sqrt{\sum_{j=1}^{h} cit_j}$$

where h is the h-index and cit_j is the number of citations of the most cited j-th article. We can observe that there is a relationship between the indices R, h, A: $R = \sqrt{h \cdot A}$. However it doesn't have any effect on the precision problem since uses the same publication list of the h-index .

Jin et al. (2007) [JLRE] also introduced a complementary to the R-index called **AR-index** , in order to take into consideration both the citations and *age of the publications* in the Hirsch core.

Definition 92 (AR-index) *The* age-dependent *R-index or AR-index is defined as the square root of the sum of the average citations per year of the articles in the h-core. Formally:*

$$AR = \sqrt{\sum_{j=1}^{h} \frac{cit_j}{a_j}}$$

where h is the h-index, a_j denotes the number of years since the publication of the j-th article in the h-core and cit_j its number of citations.

Contrary to the h-index, the AR-index can actually decrease over time and this allows to take into account the changing in the scientist's performance.

Bornmann et al. (2008) [BMD08] pointed out that both the R-index and the A-index, measuring the citation intensity in the Hirsch core, can be very sensitive to just very few papers with extremely high citations. So in order to reduce the impact of these isolated successful papers, they proposed the **m-index,** defined as the *median* of the citations received by papers in the Hirsch core. The idea of using the median rather than the arithmetic mean, was suggested by the citation distribution, which is often skewed.

Later, Egghe and Rousseau (2008) [ER08] introduced a new variation of the h-index, called the **weighted h-index,** or h_w-index, defined by:

$$h_w = \sqrt{\sum_{j=1}^{r_0} cit_j}$$

where cit_j is the number of citations of the j-th most cited paper and r_0 is the largest row index i such that $r_w(i) \leq cit_i$ and $r_w(i) = \sqrt{\sum_{j=1}^{i} \frac{cit_j}{h}}$. As well as the AR-index, it is has the main advantage to be sensitive to the performance variations.

Eck and Waltman (2008) [EW08] argued the arbitrariness of the definition of the h-index and proposed the **$\mathbf{h}_\alpha$-index.**

Definition 93 *A scientist has h_α-**index** h_α if h_α of his n papers have at least αh_α citations each and the other $n - h_\alpha$ papers have fewer than αh_α citations each. Numerically:*

$$h_\alpha(x) = \max \{u \mid x(u) \geq \alpha u\}.$$

where $\alpha \in (0,1)$.

Hence, $h_1 = h$ and for $\alpha \neq 1$ the h_α-index is not necessarily an integer. This approach can be useful when the h-index is restrictive. For example, if several authors share a high h-index it may be difficult to discriminate them. In that cases an h_α-index with $\alpha < 1$ might show greater granularity among the scientists.

In addition to the indices that take into account time variables, other authors were interested to adapt the h-index in order to consider the number of *co-authors* of the publications. Hirsch (2005) had already suggested that 'it may be useful...to normalize h by a factor that reflects the average number of co-authors'. This proposal, later denoted with h_I, is formalized by:

$$h_I = \frac{h}{N^a}$$

where N^a is the average number of authors of the papers in the h-core.

In 2008 Schreiber [S08b] argued that this normalization is not fair as it penalizes authors with some papers with a large number of co-authors, because the average is sensitive to extreme values. Thus he proposed a new index called the **h$_m$-index** that keeps into account the influence of the number of co-authors for a researcher's publication, counting the papers fractionally according to the number of authors (i.e. only as one third for three authors).

Definition 94 *The h_m-index is defined as follows:*

$$h_m = \max\{r : r_{eff}(r) \leq c(r)\}$$

where r is the rank of the author's paper when the publication list is sorted by the number $c(r)$ of citations and the effective rank r_{eff} is given by

$$r_{eff}(r) = \sum_{j=1}^{r} \frac{1}{a(j)}$$

where $a(j)$ is the number of authors for the paper in the rank j.

The value $\frac{1}{a(j)}$ corresponds to the fraction attributed to the paper j. In case of the h-index each paper is fully counted, thus the h-index turns out to be a special case of the previous definition:

$$h = \max\{r : r \leq c(r)\} \tag{4.1}$$

The authors also observed that the h_I-index can be obtained by the definition of the h_m-index, simply substituting to r in (4.1) the product $r_I(r) \cdot \bar{a}(r)$, where $\bar{a}(r)$ is the average number of authors of the first r papers: $\bar{a}(r) := \frac{1}{r}\sum_{j=1}^{r} a(j)$, and $r_I(r) := \frac{1}{\bar{a}(r)}\sum_{j=1}^{r} 1 = \frac{r}{\bar{a}(r)}$ represents the ranking . So that:

$$h_I := \frac{h}{\bar{a}(h)} = \max\left\{r : r_I(r) \leq \frac{c(r)}{\bar{a}(r)}\right\}$$

Egghe (2008) [E08] also considered the problem of *multiple authors*. This paper studies the h-index (Hirsch index) and the g-index of authors, in case one counts authorship of the cited articles in a fractional way. He studied two different approaches: the first one is called “fractional counting on citations”, and consists in giving to an author of an m-authored paper only a credit of $\frac{c}{m}$ if the paper received c citations. The second one is called “fractional counting on papers” where for each author in a m-authored paper, the paper occupies only a fractional rank of $\frac{1}{m}$ (note that this approach is equivalent to that seen for the h_m-index).

This multiple authors approach might be useful to reward scientists whose papers are entirely produced by themselves from the authors that work in big groups (and which naturally publish a larger amount of papers).

In order to take into account *all citations* not just those in the Hirsch core, Anderson et al. (2008) [AHK08] proposed the **tapered h-index.** Their idea was to represent the citations of the author's papers in a Ferrers graph (as in the Fig.4.2), where each column represents a partition of the citations amongst the articles. The largest completed (filled in) square of points in the upper left hand corner of a Ferrers graph is called the Durfee square. Following this representation, the h-index is defined as the length of the side of the Durfee square (in the case in Fig.4.2, $h = 3$), effectively assigning no credit (zero score) to all points that fall outside.

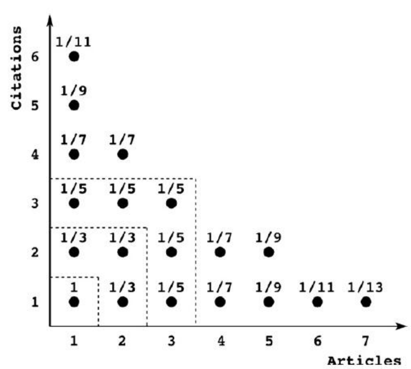


Fig.4.2

The main idea of the tapered h-index is to add to the h-index value all the citations with a weight equal to the inverse of the increment that makes the h-index increase of one unit. So the tapered h-index in Fig.4.2 is calculated as follows: $h_T = 3+\frac{1}{7}\cdot 4+\frac{1}{9}\cdot 3+\frac{1}{11}\cdot 2+\frac{1}{13}\cdot 1 \simeq 4.16$.

Definition 95 *If an author has N papers with associated citations $cit_1, cit_2, cit_3, \ldots, cit_N$ (ranked in descending order as in a Ferrers graph), the total* tapered h-index*, h_T, is calculated by:*

$$h_T = \sum_{j=1}^{N} h_T(j)$$

where $h_T(j)$ is the score for any j-th paper of the list (with cit_j citations), such that:

$$h_T(j) = \begin{cases} \frac{cit_j}{2j-1} & cit_j \leq j \\ \frac{j}{2j-1} + \sum_{i=j+1}^{cit_j} \frac{j}{2i-1} & cit_j > j \end{cases}$$

This approach has the advantage of taking into account the complete researcher's production. The difficulty consists in the implementation, given the difficulty to detect the complete and precise publication list from the bibliometric databases.

By a similar but simpler approach, Ruane and Tol (2008) [RT08] presented the **rational h-index**, h_{rat}-index, which has the advantage to provide more granularity in the evaluation process since it increases in smaller steps than the original h-index.

Definition 96 *The h_{rat}-index is is defined as $(h+1)$ minus the relative number of scores necessary for obtaining a value $h+1$. Formally:*

$$h_{rat} = (h+1) - \frac{n_c}{2h+1}$$

where h is the h-index, n_c is the least number of citations necessary to obtain an h-index of $h+1$ (it corresponds to the blank spaces in the Durfee square of size $h+1$) and $2h+1$ is the maximum amount of cites that could be necessary to increment the h-index in one unit.

It is obvious that $h \leq h_{rat} < h + 1$.

Guns and Rousseau (2009) [GR09] reviewed different real and rational variants of the h- and g- indices.

Definition 97 (real h- and g- index) *Let $P(r)$ denote the number of citations of the $r-th$ publication and let $P(x)$ denote its piecewise linear interpolation, then the **real** h-**index** h_r is the abscissa of the intersection of the function $P(x)$ and the angle bisector $y = x$. Numerically:*

$$h_r = \frac{(h+1) \cdot P(h) - h \cdot P(h+1)}{1 - P(h+1) + P(h)}$$

Let $Q(r)$ denote the cumulative citation count of all publications up to (and including) r, i.e.$Q(r) = \sum_{i=1}^{r} P(i)$, and let $Q(x)$ denote its piecewise linear interpolation, then the **real g-index**, g_r is the abscissa of the intersection of the function $Q(x)$ and the curve $y = x^2$. See the Fig.4.2.

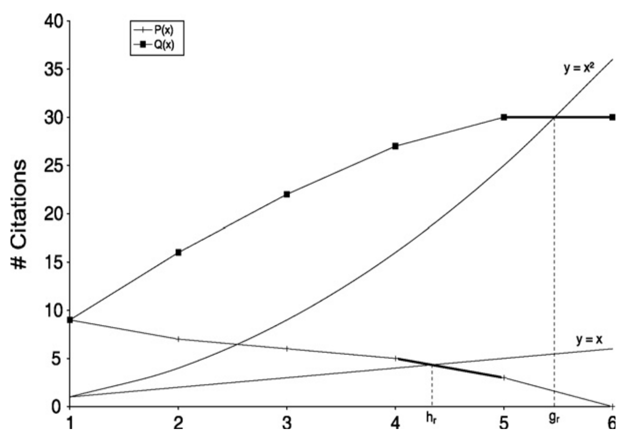


Fig.4.2. Graphical construction for the calculation of h_r and g_r.

The real variants can also be used when citation scores are not natural numbers, for instance when citations are counted fractionally.

A further *time-dependent* h-type index, the so called **dynamic h-type index**, was proposed by Rousseau and Ye (2008) [RY08]. In particular, it depends on the size of the h-core, the actual number of citations received by articles belonging to the h-core, and the *recent increase* in h. It tries to detect situations where two scientists have the same h-index and the same number of citations in the h-core, but that one has no change in his h-index for a long time while the other scientist's h-index is on the rise. So the definition contains three time-dependent elements: the size and contents of the Hirsch core, the number of citations received, and the h-velocity.

Definition 98 *The dynamic h-type index h_d is defined as*

$$h_d = R(T) \cdot v_h(T)$$

where $R(T)$ denotes the R-index defined by Jin et al. (2007) [JLRE] computed at time T and v_h is the h-velocity.

In practice, it is necessary to determine a starting point, $T = 0$ and a way to determine v_h. As suggested by the authors, the starting point should not be the beginning of a scientist's career, but when T is "now", then $T = 0$ can be taken 10 or 5 years ago (or any other appropriate time). Regarding the determination of $v_h(T)$ over this period, should be used a good-fitting continuous model for $h(t)$ (or better for h_{rat}).

Zhang in 2009 [Z09] proposed the **e-index**. The aim of the author was to find a solution of two disadvantages of the h-index: the lack consideration of the citations above the h level and the "low resolution" of this index which doesn't allow a precise evaluation of the scientists. The g-index doesn't manage to resolve completely these problems suffering from the loss of citation informations.

Definition 99 (e-index) *The excess citations received by all papers in the h-core, denoted by* e^2*, are*

$$e^2 = \sum_{j=1}^{h}(cit_j - h) = \sum_{j=1}^{h} cit_j - h^2$$

where cit_j *are the citations received by the* j^{th} *paper. The* e-index *is defined as*

$$e = \sqrt{\sum_{j=1}^{h} cit_j - h^2}$$

Note that $e \geq 0$, and e is a real number. The Fig.4.2 suggests a graphical interpretation.

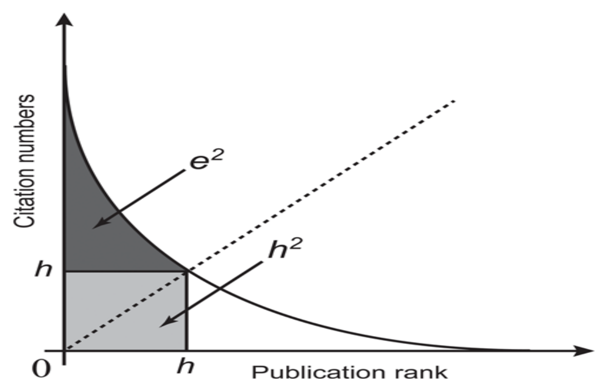


Fig. 4.2. Graphical interpretation of the e-index.

The e-index must be considered a complement of the h-index for the ignored excess citations, especially useful for the comparison of scientists with the same h-index. It can be also used with the h-index to derive the A- and R- indices:

$$A = h + \frac{e^2}{h}$$
$$R = \sqrt{h^2 + e^2}$$

According to this formalization the h- and e- index are fundamental indices, while A and R are derived.

Katsaros et al. (2009) [KAB09] argued that the issue of the citations is much more complex and goes beyond *self-citations*; it involves the essential meaning of a citation. They introduced the concept of *coterminal citations*, as the pattern of citation where some author has (co)authored multiple papers citing another paper, as shown in the following figure:

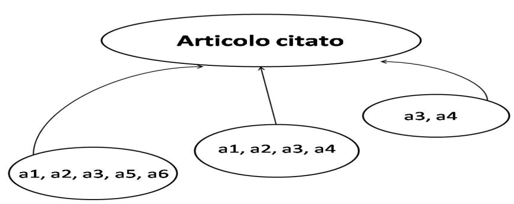


Fig.4.2 Citing articles with author overlap.

Coterminal citations can be considered as a generalization of what is widely known as cocitation, and their introduction attempts to capture the "inflationary" trends in scholarly communication which are reflected by coauthorship and "exaggerate" citing.

To this end, the authors developed the **f-index** that is not affected by the coterminal citations, in the sense that "appropriately" weighs them (**f** stands for "fractional" citation counting scheme).

Definition 100 *Given a set of quantities $N_f^{A_i}$ representing the value of the citations of the author's A_i article, ranked in a decreasing order. The point where the rank becomes larger than the respective $N_f^{A_i}$ in the sorted sequence defines the author's f index. The quantity $N_f^{A_i}$ is computed for each article A_i by:*

$$N_f^{A_i} := f^{A_i} \cdot s$$

where s is a "weighting" vector and $f^{A_i} = \left\{ f_1^{A_i},\ f_{2,}^{A_i},\ f_x^{A_i}, ...f_{nca^{A_i}}^{A_i} \right\}$ is a $nca^{A_i}-$dimensional vector of probability mass (nca^{A_i} is the number of articles citing the A_i article), representing the penetration of A_i's ideas to the scientific community. Each component of f^{A_i} is defined by:

$$f_x^{A_i} := \frac{\mid F_x^{A_i} \mid}{N_i}$$

where$\left\{ F_x^{A_i} \right\}_x$ is a family of sets such that

$$F_x^{A_i} := \{a_j : author\ a_j\ appears\ in\ exactly\ x\ articles\ citing\ A_i\}$$

with $x = 1, 2, ...nca^{A_i}$ and N_i is the total number of distinct authors citing A_i.

Consider the citing example shown in Fig.4.2, we call A the article cited by three other articles. The number of articles citing A is 3, then $F_1^A = \{a_5, a_6\}$, $F_2^A = \{a_1, a_2, a_4\}$ and $F_3^A = \{a_3\}$. Thus $f^A = \left\{ \frac{2}{7}, \frac{3}{7}, \frac{1}{7} \right\}$. Naturally, for successful scholars, we would prefer the probability mass to be concentrated to the first coordinates. The simplest choice of the weighting vector s is $s = \{nca, nca - 1, ..., 1\}$. Then $N_f^A = f^A \cdot s = \frac{2}{7}.3 + \frac{3}{7}.2 + \frac{1}{7}.1 = \frac{13}{7} \approx 1.86$.

The authors were aware that the detection and weighting of coterminal citations is a difficult procedure, so they suggested that the scientific community should set some rules about citing, for example qualifying each reference in context, self or oppose citations.

In 2010 Alonso et al. [ACHHb] presented the **hg-index**, which is based both on h-index on the g-index, trying to maintain a balance between their benefits and minimize the disadvantages:

Definition 101 *The hg-index of a scientist is calculated as the geometric mean of his indices h and g, i.e.*

$$hg = \sqrt{h \cdot g}$$

It was shown that $h \leq hg \leq g$ and $hg - h \leq g - hg$, that is, the hg-index corresponds to a value nearer to h than to g. In this way it is avoid the g-index problem consisting in the influence of an isolated big successful paper. The authors noted that this new index has several advantages when compared with h- and g-indices individually: it is very simple to compute given the h- and g- index, but it provides more granularity, making easier to compare scientists with similar h- or g indices; as well as the g-index, it takes into account the citations of the author's most cited papers, but softens the influence of very successful paper.

A further index, introduced by Vinkler (2009) [V09, V09], is the **π-index** . Unlike other indicators (h-index included), which generally use data from all the published articles, the π-index prefers a selection of the *most influential publications*, P_π, defined as the square root of total papers $P_\pi = \sqrt{P}$.

Definition 102 *The π-index is equal to one hundredth of the number $C(P_\pi)$ of citations obtained to the top square root P_π of the total number of journal papers P ('elite set of* *papers') ranked by a decreasing number of citations. Formally:*

$$\pi = 0.01 \cdot C(P_\pi)$$

According to the author, selecting influential scientists and publications seem to be the essential problem for scientometricians. The π-index prefers authors with outstandingly cited papers, which may represent greatly influential publications in the field.

One of the most recent impact indices is the **q^2-index**, proposed by Cabrezizo et al. (2010) [CAHH]. It is based on two indices which stand for very different dimensions of the scientist's research output:

- the h-index, that informs about the number of papers in the h-core (quantitative dimension),
- the m-index by [BMD08], which depicts the impact of the papers in the h-core (qualitative dimension).

Definition 103 *The q^2-index of a scientist is calculated as the geometric mean of his $h-$ and $m-$ index, i.e.:*

$$q^2 = \sqrt{h \cdot m}$$

Note that $h \leq q^2 \leq m$ and $q^2 - h \leq q^2 - m$, the q^2-index is nearer to h than to m. It is simple to compute, given the h- and the m- index and it provides more granularity than the h-index, since it takes also real values. It takes into account both the quantitative and qualitative dimensions of the researcher's productive core and, therefore, it obtains a more global and balanced view of the scientific production of researchers than if we use the h- and m-indices separately.

4.3 Axiomatic approach

In the previous section we have seen an historical overview of the many variants of bibliometric indices. The next step will be to report the contribute of the literature to the axiomatic approach.

In 2008, Woeginger [W08a] proposed an axiomatic characterization of the h-index in terms of three natural axioms. According to him, a researcher is represented by a n-dimensional vector $X = (X_1, ..., X_k, ..., X_n)$, where n is the number of the author's publications and X_k is the number of the citations to the his k-th most important (in terms of impact) publication. The vector components are non-negative and decreasingly ordered, i.e. $X_1 \leq X_2 \leq ... \leq X_k \leq ... \leq X_n$ in the sense that a vector $X = (X_1, ..., X_n)$ is dominated by a vector $Y = (Y_1, ..., Y_m)$, $X \leq Y$, if $n \leq m$ and $X_k \leq Y_k$ for $1 \leq k \leq n$.

Definition 104 *Let $\mathcal{X}$ be the set of the vectors of the researchers. A* scientific impact index *is a function $f : \mathcal{X} \rightarrow \mathbb{N}$ that satisfies two conditions:*

- *if $X = (0, 0, ..., 0)$ or if X is the empty vector (if $n = 0$), then $f(X) = 0$;*
- *monotonicity: if $X \leq Y$ then $f(X) \leq f(Y)$.*

Both properties seem to characterize well a measure of scientific production (impact): a researcher whose scientific research fails to generate citations (all-zero vector X) as well as a researcher without publications (empty vector X) has no impact. If the citations to the scientific production of researcher Y dominate the citations to the scientific production of researcher X publication by publication, then Y has more impact than X.

On the basis of the previous definition, Woeginger provided a new formal definition of the h-index.

Definition 105 *The* h-index *is the scientific impact index* $h : \mathcal{X} \rightarrow \mathbb{N}$ *that assigns to the vector* X *the value*

$$\begin{aligned} h(X) &:= \max\{k : X_m \geq k \quad \text{for all } m \leq k\} \\ &= \max\{k : X_k \geq k\} \end{aligned}$$

Woeginger also introduced a new index, called *w-index*: "A w-index of at least k means that there are k distinct publications that have at least $1, 2, 3, 4, ..., k$ citations respectively" that is a scientific impact index as well.

Definition 106 *The* w-index *is the scientific impact index* $w : \mathcal{X} \rightarrow \mathbb{N}$ *that assigns to the vector* X *the value*

$$w(X) := \max\{k : X_m \geq k - m + 1 \quad \text{for all } m \leq k\}$$

The main contribute of Woeginger (2008) is the formulation of the following five axioms and the characterization of the h-index and the w-index in terms of these axioms.

Axiom 107 (A1) *If the* $(n+1)-$*dimensional vector* Y *results from the* $n-$*dimensional vector* X *by adding a new article with* $f(X)$ *citations, then* $f(Y) \leq f(X)$.

Axiom 108 (A2) *If the* $(n+1)-$*dimensional vector* Y *results from the* $n-$*dimensional vector* X *by adding a new article with* $f(X)+1$ *citations, then* $f(Y) > f(X)$.

Axiom 109 (B) *If the* $n-$*dimensional vector* Y *results from the* $n-$*dimensional vector* X *by increasing the number of citations of a single article, then* $f(Y) \leq f(X) + 1$.

Axiom 110 (C) *If the $n-$dimensional vector Y results from the $n-$dimensional vector X by increasing the number of citations of every article by at most one, then $f(Y) \leq f(X)+1$.*

Axiom 111 (D) *If the $(n+1)-$dimensional vector Y results from the $n-$dimensional vector X by first adding an article with $f(X)$ citations and afterwards increasing the number of citations of every article by at least one, then $f(Y) > f(X)$.*

Theorem 112 (Th.4.1. [W08a]) *A scientific impact index $f : \mathcal{X} \to \mathbb{N}$ satisfies the three Axioms $A1$, B, and D, if and only if it is the h-index.*

Theorem 113 (Th.4.2. [W08a]) *A scientific impact index $f : \mathcal{X} \to \mathbb{N}$ satisfies the three Axioms $A2$, B, and C, if and only if it is the w-index.*

In a sequent work [W08b] Woeginger provided a new axiomatic characterization of the h-index, based on a simple *symmetry axiom* which essentially imposes that the number of citations and the number of publications should be treated in the same way and should be measured in the same scale.

To this end, he defined for any citation vector $X = (X_1, ..., X_k, ..., X_n)$ the corresponding reflected publication vector $R(X) := (X'_1, ..., X'_m)$ where

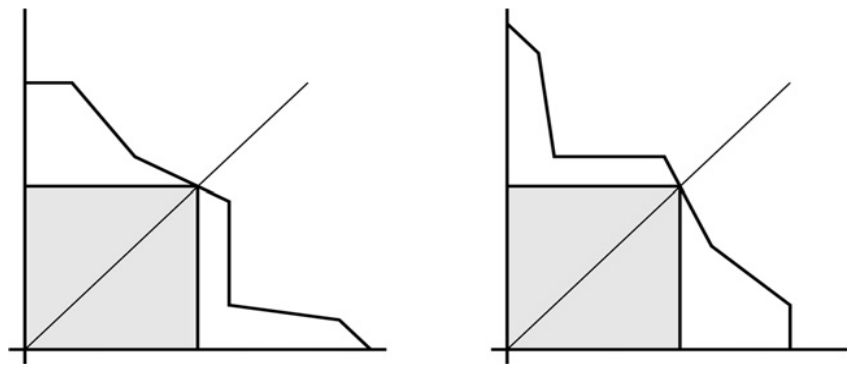
$$X'_l = |\{k : X_k \geq l\}| \, .$$

If the vector X doesn't contain any null components, then $R(R(X)) = X$.

Axiom 114 (S) $f(X) = f(R(X))$ *for any* $X \in \mathcal{X}$.

Proposition 115 (Prop.3.1 [W08b]) *The h-index and the w-index both satisfy Axiom S.*

In the following figure is given a graphical interpretation of the symmetry axiom for the h-index. Publications are on the horizontal axis, and the numbers of citations per publication are on the vertical axis. The side length of the shaded square yields the h-index.



The reflection leaves the value of the h-index .

In [DW09] and [W09] we found further generalizations to the h^2-index and the g-index, respectively.

Chapter 5

Scientific Research Measures

The aim of this chapter is to present an interesting application of the quasi-concave analysis to the bibliometrics. We propose a new scientific performance measures. Differently from any existing approach, our formulation is clearly germinated from the theory of risk measures. The axiomatic approach developed in the seminal paper by Artzner et al. [ADEH99] turned out to be, in this last decade, very influential for the theory of risk measures: instead of focusing on some particular measurement of the risk carried by financial positions (the variance, the $V@R$, etc. etc.), [ADEH99] proposed a class of measures satisfying some reasonable properties (the "coherent" axioms, see Definition 1). Ideally, each institution could select its own risk measure, provided it obeyed the structural coherent properties. This approach added flexibility in the selection of the risk measure and, at the same time, established a unified framework. We propose the same approach in order to determine a good class of scientific performance measures, that we call Scientific Research Measures (SRMs).

We have seen in the Chapter 2 that the theory of coherent risk measures was later extended to the class of convex risk measures (Föllmer and Schied [FS02], Frittelli and Rosazza [FR02]). The origin of our proposal can be traced in the more recent development of this theory, leading to the notion of quasi-convex risk measures introduced by Cerreia-Vioglio et al. [CMMMa] and further developed in the dynamic frame work by Frittelli and Maggis [FM11]. Additional papers in this area include: Cherny and Madan [CM09], that introduced the concept of an Acceptability Index (see Definition 13) having the property of quasi-concavity; Drapeau and Kupper [DK10], where the correspondence between a quasi-convex risk measure and the associated family of acceptance sets - already present in [CM09] - is fully analyzed. The representation of quasi-convex monotone maps in terms of family of acceptance sets, as well as their dual formulations, are the key ideas underlying our definition of SRM.

We propose a family of SRMs that are:

- *flexible* in order to fit peculiarities of different areas and ages;
- *inclusive*, as they comprehends several popular indices;
- *calibrated* to the particular scientific community;
- *coherent*, as they share the same structural properties - based on an axiomatic approach;
- *granular,* as they allow a more precise comparison between scientists and are based on *the whole citation curve* of a scientist.

We propose a general class of scientific performance measures (SRMs), in order to provide an unifying framework for most of the popular indices. The definition of the SRM, the relative properties and some examples are given in Section 5.1. A new interesting approach to the whole area of bibliometric indices is provided by the dual representation of a SRM, discussed in Section 5.2. We also show in Section 5.3 the method to compute a particular SRM, called ϕ-index, and we report some empirical results obtained by *calibrating* the performance curves to a specific data set (built using Google Scholar).

5.1 On a class of Scientific Research Measures

We represent each author by a vector X of citations, where the i-th component of X represents the number of citations of the i-th publication and the components of X are ranked in decreasing order. We consider the whole *citation curve* of an author as a decreasing bounded step functions X (see Fig.5.1) in the convex cone:

$$\mathcal{X}^+ = \left\{ \begin{array}{c} X : \mathbb{R} \to \mathbb{R}_+ \mid X \text{ is bounded, with only a finite numbers of values,} \\ \text{decreasing on } \mathbb{R}_+ \text{ and such that } X(x) = 0 \text{ for } x < 0. \end{array} \right\}$$

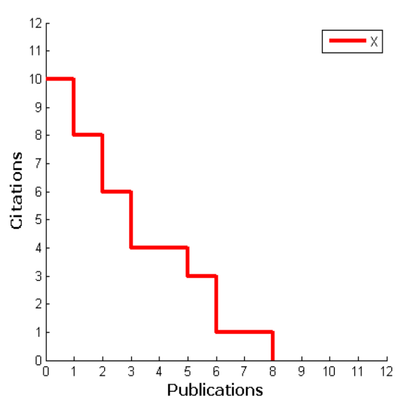


Fig. 5.1. Author's Citation Curve

We compare the citation curve X of an author with a theoretical citation curve f_q representing the desiderata citations at a fix performance level q. For this purpose we introduce the following class of curves. Let $\mathcal{I} \subseteq \mathbb{R}$ be the index set of the *performance level*. For any $q \in \mathcal{I}$ we define the theoretical *performance curve of level* q as a function $f_q : \mathbb{R} \to \mathbb{R}_+$ that associates to each publication $x \in \mathbb{R}$ the corresponding number of citations $f_q(x) \in \mathbb{R}_+$.

Definition 116 (Performance curves) *Given a index set $\mathcal{I} \subseteq \mathbb{R}$ of performance levels $q \in \mathcal{I}$, a class $\mathbb{F} := \{f_q\}_{q\in\mathcal{I}}$ of functions $f_q : \mathbb{R} \to \mathbb{R}_+$ is a* family of performance curves *if*

i) $\{f_q\}$ is increasing in q, i.e. if $q \geq p$ then $f_q(x) \geq f_p(x)$ for all x;

ii) for each q, $f_q(x)$ is left continuous in x;

iii) $f_q(x) = 0$ for all $x < 0$ and all q.

The main feature of these curves is that a higher performance level implies a higher number of citations. This family of curves is crucial for our objective to build a SRM able to comprehend many of the popular indices and calibrated to the scientific area and the seniority of the authors.

Definition 117 (Performance sets and SRM) *Given a family of performance curves $\mathbb{F} = \{f_q\}_q$, we define the* family of performance sets *$\mathcal{A}_{\mathbb{F}} := \{\mathcal{A}_q\}_q$ by*

$$\mathcal{A}_q := \left\{X \in \mathcal{X}^+ \mid X(x) \geq f_q(x) \text{ for all } x \in \mathbb{R}\right\}.$$

The Scientific Research Measure (SRM) *is the map $\phi_{\mathbb{F}} : \mathcal{X}^+ \to [0, \infty]$ associated to $\mathbb{F}$ and $\mathcal{A}_{\mathbb{F}}$ given by*

$$\begin{aligned} \phi_{\mathbb{F}}(X) \quad &: \quad = \sup\{q \in \mathcal{I} \mid X \in \mathcal{A}_q\} \\ &= \quad \sup\{q \in \mathcal{I} \mid X(x) \geq f_q(x) \text{ for all } x \in \mathbb{R}\}. \end{aligned} \tag{5.1}$$

The SRM $\phi_{\mathbb{F}}$ is obtained by the comparison between the real citation curve of an author X (the red line in Fig.5.1) and the family $\mathbb{F}$ of performance curves (the blue line in Fig.5.1): the $\phi_{\mathbb{F}}(X)$ is the greatest level q of the performance curve f_q below the author's citation curve X.

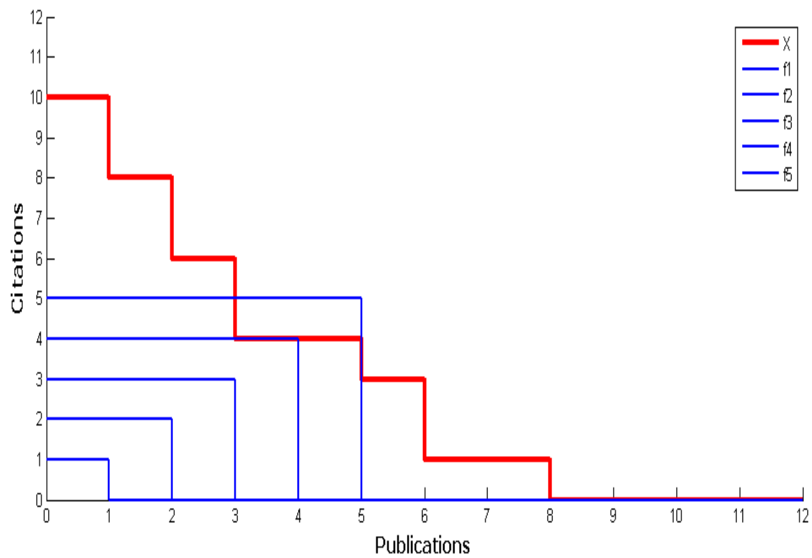


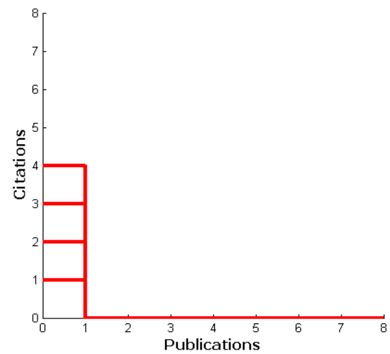
Fig. 5.1. Determination of a particular SRM, the h-index (that in this example is equal to 4).

5.1.1 Some examples of existing SRMs

The previous definition points out the importance of the family of theoretical performance curves for the determination of the SRM. It is clear that different choices of $\mathbb{F} := \{f_q\}_q$ lead to different SRM $\phi_{\mathbb{F}}$. The following examples show that some well known indices of scientific performance are particular cases of our SRM. In the following examples, if X has $p \geq 1$ publications that received at least one citation, we set: $X = \sum_{i=1}^{p} x_i 1_{(i-1,i]}$, with $x_i \geq x_{i+1}$ for all i.

Example 118 (max # of citations) *The* maximum number of citations *of the most cited author's paper is the SRM* $\phi_{\mathbb{F}_{c_{\max}}}$ *defined by (5.1), where the family* $\mathbb{F}_{c_{\max}}$ *of performance curves is:*

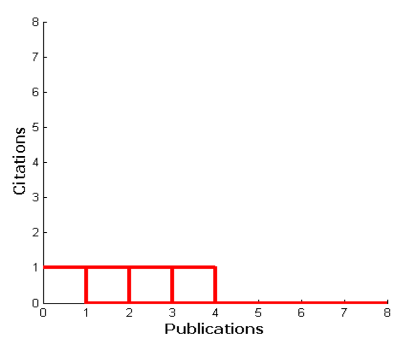
$$f_q(x) = \begin{cases} q & 0 < x \le 1 \\ 0 & x > 1 \end{cases} \quad \text{for all } x \in \mathbb{R}_+. \tag{5.2}$$



(5.3)

Example 119 (total number of publications) *The* total number of publications *with at least one citation is the SRM* $\phi_{\mathbb{F}_p}$ *defined by (5.1), where the family* $\mathbb{F}_p$ *of performance curves is:*

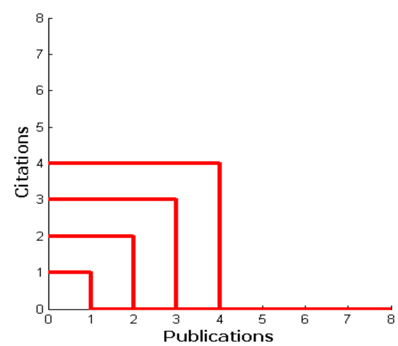
$$f_q(x) = \begin{cases} 1 & 0 < x \le q \\ 0 & x > q \end{cases} \quad \text{for all } x \in \mathbb{R}_+. \tag{5.4}$$



(5.5)

Example 120 (h-index) *According to the definition given by Hirsch, 2005 [H05]: "A scientist has index h if h of his or her N_p papers have at least h citations each and the other $(N_p - h)$ papers have $\leq h$ citations each". The* h-index *is the SRM $\phi_{\mathbb{F}_h}$ defined by (5.1), where the family $\mathbb{F}_h$ of performance curves is:*

$$f_q(x) = \begin{cases} q & 0 < x \leq q \\ 0 & x > q \end{cases} . \quad \text{for } x \in \mathbb{R}_+. \tag{5.6}$$



(5.7)

Example 121 (h^2-index) *Kosmulski, 2006 [K06] defined a scientist's* h^2-index *"as the highest natural number such that his h^2 most cited papers received each at least $[h^2]^2$ citations". This index is the SRM $\phi_{\mathbb{F}_{h^2}}$ defined by (5.1), where the family $\mathbb{F}_{h^2}$ of performance curves is:*

$$f_q(x) = \begin{cases} q^2 & 0 < x \leq q \\ 0 & x > q \end{cases} \quad \text{for } x \in \mathbb{R}_+.$$

Example 122 (h$_\alpha$-index) *Eck and Waltman, 2008 [EW08] proposed the* h$_\alpha$-index *as a generalization of the h-index so defined: "a scientist has* h$_\alpha$*-index* h$_\alpha$ *if* h$_\alpha$ *of his n papers have at least $\alpha \cdot$h$_\alpha$ citations each and the other $n-$h$_\alpha$ papers have $\leq \alpha\cdot$h$_\alpha$ citations each". Hence,* h$_\alpha$-index *is the SRM $\phi_{\mathbb{F}_{h_\alpha}}$ defined by (5.1),where the family $\mathbb{F}_{h_\alpha}$ of performance* *curves is:*

$$f_q(x) = \begin{cases} \alpha q & 0 < x \le q \\ 0 & x > q \end{cases} \qquad \textit{for } x \in \mathbb{R}_+ \textit{ and } \alpha \in (0, \infty).$$

Example 123 (w-index) *Woeginger, 2008 [W08a] introduced the* w-index *defined as: "a* w-index *of at least k means that there are k distinct publications that have at least* 1, 2, 3, 4,..., *k citations, respectively". It is the SRM* $\phi_{\mathbb{F}_w}$ *defined by (5.1), where the family* $\mathbb{F}_w$ *of performance curves is:*

$$f_q(x) = \begin{cases} q - x + 1 & 0 < x \le q \\ 0 & x > q \end{cases} \qquad \textit{for all } x \in \mathbb{R}_+. \tag{5.8}$$

Example 124 ($\mathbf{h}_{rat}$-index & $\mathbf{h}_r$-index) *The rational and the real h-index,* h_{rat}-index *and* h_r-index, *introduced respectively by Ruane and Tol, 2008 [RT08] and Guns and Rousseau, 2009 [GR09] are SRMs, indeed they could be defined as the h-index but taking respectively* $q \in \mathcal{I} \subseteq \mathbb{Q}$ *and* $q \in \mathcal{I} \subseteq \mathbb{R}$.

Example 125 ($\mathbf{h}_m$-index) *Schreiber, 2008 [S08b] proposed a new index called* h_m-index *that keeps into account the influence of the number of co-authors for a researcher's publication, counting the papers fractionally according to the number of authors. The* h_m-index *is the SRM* $\phi_{\mathbb{F}_{h_m}}$ *defined by (5.1),where the family* $\mathbb{F}_{h_m}$ *of performance curves is:*

$$f_q(x) = \begin{cases} \sum_{j=1}^{q} \frac{1}{a(j)} & 0 < x \le q \\ 0 & x > q \end{cases} \qquad \textit{for } x \in \mathbb{R}_+,$$

where $a(j)$ *is the number of authors for the paper* j.

5.1.2 Key properties of the SRMs

Now we point out some relevant properties of the family $\mathcal{A}_{\mathbb{F}} = \{\mathcal{A}_q\}_q$ of performance sets and of the SRM $\phi_{\mathbb{F}}$.

Proposition 126 *Let $X_1, X_2 \in \mathcal{X}^+$.*

1. *If $\mathcal{A}_{\mathbb{F}} = \{\mathcal{A}_q\}_q$ is a family of* performance sets *then:*

   i) *$\{\mathcal{A}_q\}$ is decreasing monotone: $\mathcal{A}_q \subseteq \mathcal{A}_p$ for any level $q \geq p$;*

   ii) *$\mathcal{A}_q$ is monotone for any q: $X_1 \in \mathcal{A}_q$ and $X_2 \geq X_1$ implies $X_2 \in \mathcal{A}_q$;*

   iii) *$\mathcal{A}_q$ is convex for any q: if $X_1, X_2 \in \mathcal{A}_q$ then $\lambda X_1 + (1-\lambda)X_2 \in \mathcal{A}_q$ for $\lambda \in [0,1]$;*

2. *If $\phi_{\mathbb{F}}$ is a* SRM *then it is:*

   i) *increasing monotone: if $X_1 \leq X_2 \Rightarrow \phi_{\mathbb{F}}(X_1) \leq \phi_{\mathbb{F}}(X_2)$;*

   ii) *quasi-concave: $\phi_{\mathbb{F}}(\lambda X_1 + (1-\lambda)X_2) \geq \min(\phi_{\mathbb{F}}(X_1),\ \phi_{\mathbb{F}}(X_2))$ for all $\lambda \in [0,1]$.*

Proof.

1) The proof of the monotonicity and convexity of $\mathcal{A}_{\mathbb{F}}$ follows from the definition.

2.i) It is sufficient to show that

$$\{q \in \mathcal{I} \mid X_1 \geq f_q\} \subseteq \{q \in \mathcal{I} \mid X_2 \geq f_q\}.$$

As $X_1 \leq X_2$, $X_1 \geq f_{\overline{q}}$ implies $X_2 \geq f_{\overline{q}}$.

2.ii) Let $\phi_{\mathbb{F}}(X_1) \geq m$ and $\phi_{\mathbb{F}}(X_2) \geq m$. By definition of $\phi_{\mathbb{F}}$, $\forall \varepsilon > 0$ $\exists q_i$ s.t. $X_i \geq f_{q_i}$ and $q_i > \phi_{\mathbb{F}}(X_i) - \varepsilon \geq m - \varepsilon$. Then $X_i \geq f_{q_i} \geq f_{m-\varepsilon}$, as $\{f_q\}_q$ is an increasing family, and therefore $\lambda X_1 + (1-\lambda)X_2 \geq f_{m-\varepsilon}$. As this holds for any $\varepsilon > 0$, we conclude that $\phi_{\mathbb{F}}(\lambda X_1 + (1-\lambda)X_2) \geq m$ and $\phi_{\mathbb{F}}$ is quasi-concave.

■

It is obviously reasonable that a SRM should be *increasing:* if the citations of a researcher X_2 dominate the citations of another researcher X_1 publication by publication, then X_2 has a performance greater than X_1.

Now, we introduce a counterexample in order to show that a SRM is not in general quasi-convex, that is $\phi_{\mathbb{F}}(\lambda X_1 + (1-\lambda)X_2) \leq \max(\phi_{\mathbb{F}}(X_1),\ \phi_{\mathbb{F}}(X_2))$ for all $\lambda \in [0,1]$. We consider two researchers, $X_1 = [8 \quad 6 \quad 4 \quad 2]$ and $X_2 = [4 \quad 2 \quad 2 \quad 2 \quad 2]$, where X_2 has more publications than X_1 but less cited. If we compute for example the w-index we obtain that $\phi_{\mathbb{F}_w}(X_1) = 4$ and $\phi_{\mathbb{F}_w}(X_2) = 3$, while taking $\lambda = \frac{1}{2}$ the SRM $\phi_{\mathbb{F}_w}$ of the combined citation curve $X = \frac{1}{2}X_1 + \frac{1}{2}X_2 = [6 \quad 4 \quad 3 \quad 2 \quad 1]$ is $\phi_{\mathbb{F}_w}(X) = 5$.

5.1.3 Additional properties of SRMs

We have seen that all the SRMs $\phi_{\mathbb{F}}$ share the same structural properties of monotonicity and quasiconcavity. We start this section classifying the SRMs on the basis of the addition of citations to the old papers.

Definition 127 (Additional citation properties) *A SRM* $\phi_{\mathbb{F}} : \mathcal{X}^+ \to [0, \infty]$ *is*

a) C-*superadditive if* $\phi_{\mathbb{F}}(X+m) \geq \phi_{\mathbb{F}}(X) + m$ *for all* $m \in \mathbb{R}_+$*;*

b) C-*subadditive if* $\phi_{\mathbb{F}}(X+m) \leq \phi_{\mathbb{F}}(X) + m$ *for all* $m \in \mathbb{R}_+$*;*

c) C-*additive if* $\phi_{\mathbb{F}}(X+m) = \phi_{\mathbb{F}}(X) + m$ *for all* $m \in \mathbb{R}_+$*.*

A SRM is C-superadditive (C-subadditive) if the additional citations to the old papers lead an increase of the measure more (less) than linear. In other terms, a C-superadditive SRM gives more weight than the C-subadditive SRM to the additional citations to the oldest papers.

We have seen that the SRM $\phi_{\mathbb{F}}$ depends on the family of performance curves $\mathbb{F} := \{f_q\}_q$ under consideration. The main feature of this family of curves is that is increasing monotone over q. We provide a characterization of this family in terms of the speed in the increase of the performance curves.

Definition 128 *Let $\mathbb{F}$ a family of performance curve. We say that:*

a) $\mathbb{F}$ *is* slowly increasing in q *if $f_{q+m} - f_q \leq m$ for all $m \in \mathbb{R}_+$;*

b) $\mathbb{F}$ *is* fast increasing in q *if $f_{q+m} - f_q \geq m$ for all $m \in \mathbb{R}_+$;*

c) $\mathbb{F}$ *is* linear increasing in q *if $f_{q+m} - f_q = m$ for all $m \in \mathbb{R}_+$.*

These properties of the family of performance curves can be express in terms of corresponding properties of the family $\mathcal{A}_{\mathbb{F}}$ of performance sets.

Lemma 129 *Let $\mathbb{F}$ a family of performance curve.*

1. *$\mathbb{F}$ is slowly increasing in q, if and only if*

$$\mathcal{A}_q + m \subseteq \mathcal{A}_{q+m} \tag{5.9}$$

for all $m \in \mathbb{R}_+$ and $q \in \mathcal{I}$;

2. $\mathbb{F}$ *is fast increasing in* q*, if and only if*

$$\mathcal{A}_{q+m} \subseteq \mathcal{A}_q + m \tag{5.10}$$

for all $m \in \mathbb{R}_+$ *and* $q \in \mathcal{I}$;

3. $\mathbb{F}$ *is linear increasing in* q*, if and only if*

$$\mathcal{A}_{q+m} = \mathcal{A}_q + m$$

for all $m \in \mathbb{R}_+$ *and* $q \in \mathcal{I}$;

Proof. (1) In order to show that $\mathcal{A}_q + m \subseteq \mathcal{A}_{q+m}$ we observe that:

$$\mathcal{A}_{q+m} := \left\{X \in \mathcal{X}^+ \mid X \geq f_{q+m}\right\}$$

$$\mathcal{A}_q + m = \left\{X \in \mathcal{X}^+ \mid X \geq f_q\right\} + m$$

$$= \{X \mid X \geq f_q + m\}$$

As $f_q + m \geq f_{q+m}$, if $\overline{X}$ is such that $\overline{X} \geq f_q + m$ then $\overline{X} \geq f_{q+m}$. This means that $\overline{X} \in \mathcal{A}_q + m$ implies that $\overline{X} \in \mathcal{A}_{q+m}$.

Regarding the other implication, we know that if $X \in \mathcal{A}_q + m$ then $X \in \mathcal{A}_{q+m}$, that is $X \geq f_q + m$ implies $X \geq f_{q+m}$. This implies that $f_q + m \geq f_{q+m}$.

(2) By hypothesis we know that $f_{q+m} \geq f_q + m$. Hence, if $\overline{X}$ is such that $\overline{X} \geq f_{q+m}$ then $\overline{X} \geq f_q + m$. This means that $\overline{X} \in \mathcal{A}_{q+m}$ implies that $\overline{X} \in \mathcal{A}_q + m$.

Regarding the other side, we know that if $X \in \mathcal{A}_{q+m}$ then $X \in \mathcal{A}_q + m$, that is $X \geq f_{q+m}$ implies $X \geq f_q + m$. This implies that $f_{q+m} \geq f_q + m$.

(3) The proof of this point follows directly from the previous ones, observing that $\mathbb{F}$ is linear increasing in q if and only if $\mathbb{F}$ is slowly and fast increasing in q and $\mathcal{A}_{q+m} = \mathcal{A}_q + m$ if both of the inclusions (5.9) and (5.10) hold. ■

The following lemma shows that the additional citation properties of the SRM $\phi_{\mathbb{F}}$ can be *built in* from the corresponding properties of the family $\mathbb{F}$ of the performance curves or $\mathcal{A}_{\mathbb{F}}$ of performance sets.

Lemma 130 *Let $\mathbb{F}$ a family of performance curves.*

1. *If $\mathbb{F}$ is slowly increasing in q, then $\phi_{\mathbb{F}}$ is* C-*superadditive;*

2. *If $\mathbb{F}$ is fast increasing in q, then $\phi_{\mathbb{F}}$ is* C-*subadditive;*

3. *If $\mathbb{F}$ is linear increasing in q, then $\phi_{\mathbb{F}}$ is* C-*additive.*

Proof. (1) In order to show that $\phi_{\mathbb{F}}(X+m)-m \geq \phi_{\mathbb{F}}(X)$ for all $m \in \mathbb{R}_+$ we use the definition in (5.1) and we observe that

$$\begin{aligned} \phi_{\mathbb{F}}(X+m)-m &= \sup\{q \in \mathcal{I} \mid X+m \geq f_q\} - m \\ &= \sup\{q-m \in \mathcal{I} \mid X \geq f_q - m\} \\ &= \sup\{q \in \mathcal{I} \mid X \geq f_{q+m} - m\} \end{aligned} \tag{5.11}$$

Hence it's sufficient to show that $\{q \mid \ X \geq f_q\} \subseteq \{q \mid X \geq f_{q+m} - m\}$ and this is true since $f_q \geq f_{q+m} - m$;

(2) In order to show that $\phi_{\mathbb{F}}(X+m)-m \leq \phi_{\mathbb{F}}(X)$ for all $m \in \mathbb{R}^+$ we use the definition in (5.1) and the relation (5.11). Hence it's sufficient to show that $\{q \mid \ X \geq f_{q+m} - m\} \subseteq \{q \mid X \geq f_q\}$ and this is true since $f_{q+m} - m \geq f_q$;

(3) It follows directly from the previous points observing that $\mathbb{F}$ is linear increasing in q if and only if it is slowly and fast increasing in q and that $\phi_{\mathbb{F}}$ is C-additive if and only if it is C-superadditive and C-subadditive. ■

Now we give some examples using some popular SRMs.

Example 131 *The h-index in the example (120) is a* C*-subadditive SRM, but the associated family* $\mathbb{F}$ *of performance curves defined in (5.6) is fast increasing in* q*. Indeed the property is true only on* $[0, q+m]$ *for any* $m \in \mathbb{R}_+$ *since the performance curves are equal to zero outside. Hence, the performance curves of the h-index are fast increasing only in the Hirsch core.*

The same considerations hold for the h^2*- and* h_α*- index (see examples (121) and (122)).*

Example 132 *The family* $\mathbb{F}$ *defined in 5.8 of the w-index (see example 123) is slowly increasing in* q*. This condition is sufficient to say that the w-index is a* C*-superadditive SRM.*

Example 133 *The maximum number of citations of an article (see example 118) is a* C*-additive SRM, even if the family* $\mathbb{F}$ *of performance curves defined in 5.2 is not linear increasing in* q*. This property holds only on* $[0, 1]$*, since the performance curves are equal to zero outside.*

Example 134 *The total number of publications (see example 119) is a* C*-superadditive SRM since the family* $\mathbb{F}$ *of performance curves defined in 5.4 is slowly increasing in* q*.*

We now define further properties linked to the addition of a single publication to the author's citation curve.

Definition 135 (Additional paper properties) *Let* $p := \max\{x : X(x) > 0\}$ *the maximum number of publications with at least one citation of the author* X*. A SRM* $\phi_{\mathbb{F}} : \mathcal{X}^+ \to \mathbb{R}_+$ *is*

a) P-*superadditive if* $\phi_{\mathbb{F}}(X + 1_{\{p+1\}}) \geq \phi_{\mathbb{F}}(X) + 1$*;*

b) P-*subadditive if* $\phi_{\mathbb{F}}(X + 1_{\{p+1\}}) \leq \phi_{\mathbb{F}}(X) + 1$*;*

c) P-*additive if* $\phi_{\mathbb{F}}(X + 1_{\{p+1\}}) = \phi_{\mathbb{F}}(X) + 1$*;*

c) P-*invariance if* $\phi_{\mathbb{F}}(X + 1_{\{p+1\}}) = \phi_{\mathbb{F}}(X)$*.*

A SRM is P-superadditive if the addition of one citation to a new publication leads to an increase of the measure more than linear. Someway if we use a P-superadditive SRM in our evaluation we are giving more weight to the additional publication than in case of P-subadditive SRM. Many known SRMs are P-invariance (i.e. the $c_{\max}$, h-, h^2- and h_α-index in the examples (118) (120), (121) and (122)) as the addition of one citation to a new publication leaves the SRM invariant. The w-index (in the example (123)) is P-subadditive as the addition of one citation to a new publication makes it greater at most of 1 unit. While the total number of publications p with at least one citation (in the example (119)) is clearly P-additive.

5.2 On the Dual Representation of the SRMs

The goal of this section is to provide a dual representation of the SRM. To this scope, we need some topological structure. Let $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ be a probability space, where $\mathcal{B}(\mathbb{R})$ is the σ-algebra of the Borel sets, μ is a probability measure on $\mathcal{B}(\mathbb{R})$. Since the citation curve of an author X is a bounded function, it appears natural to take $X \in L^{\infty}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$, where $L^{\infty}(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu)$ is the space of $\mathcal{B}(\mathbb{R})$-measurable functions that are μ almost surely bounded. If we endow L^{∞} with the weak topology $\sigma(L^{\infty}, L^1)$ then $L^1 = (L^{\infty}, \sigma(L^{\infty}, L^1))'$ is its topological dual. In the dual pairing $(L^{\infty}, L^1, \langle\cdot,\cdot\rangle)$ the bilinear form $\langle\cdot,\cdot\rangle : L^{\infty} \times L^1 \to \mathbb{R}$ is given by $\langle X, Z\rangle = E[ZX]$, the linear function $X \mapsto E[ZX]$, with $Z \in L^1$, is $\sigma(L^{\infty}, L^1)$ continuous and $(L^{\infty}, \sigma(L^{\infty}, L^1))$ is a locally convex topological vector space.

We have seen in the Section 5.1 that the SRM is a quasi-concave and monotone map. Under appropriate continuity assumptions, the dual representation of these type of maps can be found in [PV90],[Vo98], [CMMMa].

Definition 136 *A map $\phi : L^{\infty}(\mathbb{R}) \to \mathbb{R}$ is $\sigma(L^{\infty}, L^1)-$upper-semicontinuous if the upper level sets*

$$\{X \in L^{\infty}(\mathbb{R}) \mid \phi(X) \geq q\}$$

are $\sigma(L^{\infty}, L^1)-$closed for all $q \in \mathbb{R}$.

Lemma 137 *If $\mathcal{A}_{\mathbb{F}} = \{\mathcal{A}_q\}_q$ is a family of performance sets then $\mathcal{A}_q$ is $\sigma(L^{\infty}, L^1)$-closed for any q.*

Proof. To prove that $\mathcal{A}_q$ is $\sigma(L^{\infty}, L^1)$-closed let $Y_n \in \mathcal{A}_q := \{X \in L^{\infty} \mid X \geq f_q\}$ satisfy $Y_n \overset{\sigma(L^{\infty},L^1)}{\to} Y$. By contradiction, suppose that $\mu(A) > 0$ where $A := \{x \in \mathbb{R} \mid Y(x) < f_q(x)\} \in \mathcal{B}(\mathbb{R})$. Taking as a continuous linear functional $Z = 1_A \in L^1$, from $Y_n \overset{\sigma(L^{\infty},L^1)}{\to} Y$ we deduce: $E[1_A f_q] \leq E[1_A Y_n] \to E[1_A Y] < E[1_A f_q]$. ■

The following lemma shows the relation between the continuity property of the family $\mathbb{F}$ of performance curves, those of the family $\mathcal{A}_{\mathbb{F}}$ of performance sets and those of the

SRM $\phi_{\mathbb{F}}$.

Lemma 138 *Let $\mathbb{F}$ be a family of performance curves. If $\mathbb{F}$ is* left continuous in q, *that is*

$$f_{q-\varepsilon}(x) \uparrow f_q(x) \text{ for } \varepsilon \downarrow 0, \text{ for all } x,$$

then:

1. *$\mathcal{A}_{\mathbb{F}}$ is left-continuous in q, that is*

$$\mathcal{A}_q = \bigcap_{\epsilon>0} \mathcal{A}_{q-\varepsilon},$$

2.

$$\mathcal{A}_q = \{X \in L^\infty \mid \phi_{\mathbb{F}}(X) \geq q\}, \text{ for all } q \in \mathcal{I}. \tag{5.12}$$

3. *$\phi_{\mathbb{F}}$ is $\sigma(L^\infty, L^1)$-upper-semicontinuous.*

Proof.

1. By assumption we have that $f_{q-\varepsilon}(x) \uparrow f_q(x)$ for $\varepsilon \to 0$, for all $x \in \mathbb{R}$. We have proved in Proposition (126) that $\{\mathcal{A}_q\}$ is decreasing monotone, hence we know that $\mathcal{A}_q \subseteq \bigcap_{\varepsilon>0} \mathcal{A}_{q-\varepsilon}$. By contradiction we suppose that

$$\bigcap_{\varepsilon>0} \mathcal{A}_{q-\varepsilon} \supsetneqq \mathcal{A}_q,$$

so that there will exist $X \in L^\infty$ such that $X \geq f_{q-\varepsilon}$ for every $\varepsilon > 0$ but $X(A) < f_q(A)$ for some $A \in \mathcal{B}(\mathbb{R})$ such that $\mu(A) > 0$. Then there exists $\delta > 0$ such that $f_q(x) - X(x) \geq \delta$ for any x in $B \subseteq A$ such that $\mu(B) > 0$. Then $f_q(x) - X(x) > \frac{\delta}{2}$ for any $x \in B$. Since $f_{q-\varepsilon} \uparrow f_q$ we may find $\varepsilon > 0$ such that $f_q(x) - f_{q-\varepsilon}(x) < \frac{\delta}{2}$ for $x \in B$. Thus $X(x) \geq f_{q-\varepsilon}(x) > f_q(x) - \frac{\delta}{2}$ for $x \in B$ and this is a contradiction.

2. Now let

$$B_q := \{X \in L^\infty \mid \phi_{\mathbb{F}}(X) \geq q\}.$$

$\mathcal{A}_q \subseteq B_q$ follows directly from the definition of $\phi_{\mathbb{F}}$. We have to show that $B_q \subseteq \mathcal{A}_q$. Let $\overline{X} \in B_q$. Hence $\phi_{\mathbb{F}}(\overline{X}) \geq q$ and for all $\varepsilon > 0$ there exists $\overline{q}$ such that $\overline{q} + \varepsilon \geq q$ and $\overline{X}(x) \geq f_{\overline{q}}(x)$ for all $x \in \mathbb{R}$. Since f_q are increasing in q we have that $\overline{X}(x) \geq f_{q-\varepsilon}(x)$ for all $x \in \mathbb{R}$ and $\varepsilon > 0$, therefore $\overline{X} \in \mathcal{A}_{q-\varepsilon}$. By the left continuity in q of the family $\mathbb{F}$ we have know that $\{\mathcal{A}_q\}$ is left-continuous in q for the previous item and so: $\overline{X} \in \bigcap_{\epsilon > 0} \mathcal{A}_{q-\varepsilon} = \mathcal{A}_q$.

3. By Lemma (137) we know that $\mathcal{A}_q$ is $\sigma(L^\infty, L^1)-$closed for any q and therefore the upper level sets $B_q = \mathcal{A}_q$ are $\sigma(L^\infty, L^1)-$closed and $\phi_{\mathbb{F}}$ is $\sigma(L^\infty, L^1)$ upper semicontinuous.

■

Notice that $\sigma(L^\infty, L^1)$-upper semicontinuity is equal to the continuity from above of a SRM. This fact can be proved in a way similar to the convex case (see for example [FS04]).

Lemma 139 *Let $\phi_{\mathbb{F}} : L^\infty \to \mathbb{R}_+$ be a SRM. Then the following are equivalent:*

$\phi_{\mathbb{F}}$ is $\sigma(L^\infty, L^1)$-upper semicontinuous;

$\phi_{\mathbb{F}}$ is continuous from above: $X_n, X \in L^\infty$ and $X_n \downarrow X$ imply $\phi_{\mathbb{F}}(X_n) \downarrow \phi_{\mathbb{F}}(X)$

Proof. Let $\phi_{\mathbb{F}}$ be $\sigma(L^\infty, L^1)$-upper semicontinuous and suppose that $X_n \downarrow X$. As the elements in L^1 are order continuous, we also have: $X_n \overset{\sigma(L^\infty, L^1)}{\longrightarrow} X$. The monotonicity of $\phi_{\mathbb{F}}$ implies $\phi_{\mathbb{F}}(X_n) \downarrow$ and $q := \lim_n \phi_{\mathbb{F}}(X_n) \geq \phi_{\mathbb{F}}(X)$. Hence $\phi_{\mathbb{F}}(X_n) \geq q$ and $X_n \in B_q :=$

$\{Y \in L^\infty \mid \phi_{\mathbb{F}}(Y) \geq q\}$ which is closed by assumption. Hence $X \in B_q$, which implies that $\phi_{\mathbb{F}}(X) = q$ and that $\phi_{\mathbb{F}}$ is continuous from above.

Conversely, suppose that $\phi_{\mathbb{F}}$ is continuous from above. We have to show that the convex set B_q is $\sigma(L^\infty, L^1)$-closed for any q. By the Krein Smulian Theorem it is sufficient to prove that $C := B_q \cap \{X \in L^\infty \mid \; \| X \|_\infty < r\}$ is $\sigma(L^\infty, L^1)$-closed for any fixed $r > 0$ and $q \in \mathbb{R}$. As $C \subseteq L^\infty \subseteq L^1$ and as the embedding

$$(L^\infty, \sigma(L^\infty, L^1)) \hookrightarrow (L^1, \sigma(L^1, L^\infty))$$

is continuous it is sufficient to show that C is $\sigma(L^1, L^\infty)$-closed. Since the $\sigma(L^1, L^\infty)$ topology and the L^1 norm topology are compatible, and C is convex, it is sufficient to prove that C is closed in L^1. Take $X_n \in C$ such that $X_n \to X$ in L^1. Then there exists a subsequence $\{Y_n\}_n \subseteq \{X_n\}_n$ such that $Y_n \to X$ a.s. and $\phi_{\mathbb{F}}(Y_n) \geq q$ for all n. Set $Z_m := \sup_{n \geq m} Y_n \vee X$. Then $Z_m \in L^\infty$, since $\{Y_n\}_n$ is uniformly bounded, and $Z_m \geq Y_m$, $\phi_{\mathbb{F}}(Z_m) \geq \phi_{\mathbb{F}}(Y_m)$ and $Z_m \downarrow X$. From the continuity from above we conclude: $\phi_{\mathbb{F}}(X) = \lim_m \phi_{\mathbb{F}}(Z_m) \geq \limsup_m \phi_{\mathbb{F}}(Y_m) \geq q$. Thus $X \in B_q$ and consequently $X \in C$. ■

When the family of performance curves $\mathbb{F}$ is left continuous, Lemma (138) shows that the SRM is $\sigma(L^\infty, L^1)$-upper semicontinuous. Hence we can provide a dual representation for the SRM in the same spirit of [Vo98] and [DK10].

Denote

$$\mathcal{P} := \{Q \ll P\} \text{ and } \mathcal{Z} := \left\{Z = \frac{dQ}{dP} \mid Q \in \mathcal{P}\right\} = \left\{Z \in L^1_+ \mid E[Z] = 1\right\}$$

Theorem 140 *Suppose that the family of performance curves $\mathbb{F}$ is left continuous. Each*

SRM $\phi_{\mathbb{F}} : L^\infty(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mu) \to \mathbb{R}$ *defined in (5.1) can be represented as*

$$\begin{aligned} \phi_{\mathbb{F}}(X) &= \inf_{Z \in \mathcal{Z}} H(Z, E[ZX]) = \inf_{Z \in \mathcal{Z}} H^+(Z, E[ZX]) \qquad (5.13) \\ &= \inf_{Q \in \mathcal{P}} H^+(Q, E_Q[X]) \quad \textit{for all } X \in L^\infty \end{aligned}$$

where $H : L^1 \times \mathbb{R} \to \overline{\mathbb{R}}$ *is defined by*

$$H(Z,t) := \sup_{\xi \in L^\infty} \{\phi_{\mathbb{F}}(\xi) \mid E[\xi Z] \le t\},$$

$H^+(Z, \cdot)$ *is its right continuous version:*

$$H^+(Z,t) := \inf_{s>t} H(Z,s) = \sup \{q \in \mathbb{R} \mid t \ge \gamma(Z,q)\}, \qquad (5.14)$$

and $\gamma : L^1 \times \mathbb{R} \to \overline{\mathbb{R}}$ *is defined by:*

$$\gamma(Z,q) := \inf_{X \in L^\infty} \{E[ZX] \mid \phi_{\mathbb{F}}(X) \ge q\}. \qquad (5.15)$$

Proof. Step 1: $\phi_{\mathbb{F}}(X) = \inf_{Z \in \mathcal{Z}} H(Z, E[ZX])$.

Fix $X \in L^\infty$. As $X \in \{\xi \in L^\infty \mid E[Z\xi] \le E[ZX]\}$, by the definition of $H(Z,t)$ we deduce that, for all $Z \in L^1$,

$$H(Z, E[ZX]) \ge \phi_{\mathbb{F}}(X)$$

hence

$$\inf_{Z \in L^1} H(Z, E[ZX]) \ge \phi_{\mathbb{F}}(X).$$

We prove the opposite inequality. Let $\varepsilon > 0$ and define the set

$$C_\varepsilon := \{\xi \in L^\infty \mid \phi_{\mathbb{F}}(\xi) \ge \phi_{\mathbb{F}}(X) + \varepsilon\}$$

As $\phi_{\mathbb{F}}$ is quasi-concave and $\sigma(L^\infty, L^1)$-upper semicontinuous, C is convex and $\sigma(L^\infty, L^1)-$closed. Since $X \notin C_\varepsilon$, the Hahn Banach theorem implies the existence of a continuous linear functional that strongly separates X and C_ε, that is there exist $k \in \mathbb{R}$ and $Z_\varepsilon \in L^1$ such that

$$E[\xi Z_\varepsilon] > k > E[XZ_\varepsilon] \text{ for all } \xi \in C_\varepsilon.$$

Hence

$$\{\xi \in L^\infty \mid E[\xi Z_\varepsilon] \le E[XZ_\varepsilon]\} \subseteq C_\varepsilon^c := \{\xi \in L^\infty \mid \phi_{\mathbb{F}}(\xi) < \phi_{\mathbb{F}}(X) + \varepsilon\}$$

and

$$\begin{aligned}\phi_{\mathbb{F}}(X) &\le \inf_{Z \in L^1} H(Z, E[ZX]) \le H(Z_\varepsilon, E[XZ_\varepsilon]) \\ &= \sup\{\phi_{\mathbb{F}}(\xi) \mid \xi \in L^\infty \text{ and } E[\xi Z_\varepsilon] \le E[XZ_\varepsilon]\} \\ &\le \sup\{\phi_{\mathbb{F}}(\xi) \mid \xi \in L^\infty \text{ and } \phi_{\mathbb{F}}(\xi) < \phi_{\mathbb{F}}(X) + \varepsilon\} \le \phi_{\mathbb{F}}(X) + \varepsilon.\end{aligned}$$

Therefore, $\phi_{\mathbb{F}}(X) = \inf_{Z \in L^1} H(Z, E[ZX])$. To show that the *inf* can be taken over the positive cone L^1_+, it is sufficient to prove that $Z_\varepsilon \subseteq L^1_+$. Let $Y \in L^\infty_+$ and $\xi \in C_\varepsilon$. Given that $\phi_{\mathbb{F}}$ is monotone increasing, $\xi + nY \in C_\varepsilon$ for every $n \in \mathbb{N}$ and we have:

$$E[(\xi + nY)Z_\varepsilon] > k > E[XZ_\varepsilon] \Rightarrow E[YZ_\varepsilon] > \frac{E[Z_\varepsilon(X - \xi)]}{n} \to 0, \text{ as } n \to \infty.$$

As this holds for any $Y \in L^\infty_+$ we deduce that $Z_\varepsilon \subseteq L^1_+$. Therefore, $\phi_{\mathbb{F}}(X) = \inf_{Z \in L^1_+} H(Z, E[ZX])$.

By definition of $H(Z, t)$,

$$H(Z, E[ZX]) = H(\lambda Z, E[X(\lambda Z)]) \quad \forall Z \in L^1 \text{ and } Z \neq 0.$$

Hence we deduce

$$\phi_{\mathbb{F}}(X) = \inf_{Z \in L^1_+(\mathbb{R})} H(Z, E[ZX]) = \inf_{Z \in \mathcal{Z}} H(Z, E[ZX]) = \inf_{Q \in \mathcal{P}} H(Q, E_Q[X]).$$

Step 2: $\phi_{\mathbb{F}}(X) = \inf_{Z\in\mathcal{Z}} H^{+}(Z, E[ZX])$.

Since $H(Z,\cdot)$ is increasing and $Z \in L^1_+$ we obtain

$$H^{+}(Z, E[ZX]) := \inf_{s>E[ZX]} H(Z,s) \leq \lim_{X_m \downarrow X} H(Z, E[X_m Z]),$$

$$\phi_{\mathbb{F}}(X) = \inf_{Z\in L^1_+} H(Z, E[ZX]) \leq \inf_{Z\in L^1_+} H^{+}(Z, E[ZX]) \leq \inf_{Z\in L^1_+} \lim_{X_m\downarrow X} H(Z, E[X_m Z])$$
$$= \lim_{X_m\downarrow X} \inf_{Z\in L^1_+} H(Z, E[X_m Z]) = \lim_{X_m\downarrow X} \phi_{\mathbb{F}}(X_m) \overset{(CFA)}{=} \phi_{\mathbb{F}}(X).$$

Step 3: $H^{+}(Z,t) := \inf_{s>t} H(Z,s) = \sup\{q \in \mathbb{R} \mid t \geq \gamma(Z,q)\}$.

Now let the RHS of equation (5.14) be denoted by

$$S(Z,t) := \sup\{q \in \mathbb{R} \mid \gamma(Z,q) \leq t\}, \ (Z,t) \in L^1 \times \mathbb{R},$$

and note that $S(Z,\cdot)$ is the right inverse of the increasing function $\gamma(Z,\cdot)$ and therefore $S(Z,\cdot)$ is right continuous.

To prove that $H^{+}(Z,t) \leq S(Z,t)$ it is sufficient to show that for all $p > t$ we have:

$$H(Z,p) \leq S(Z,p), \tag{5.16}$$

Indeed, if (5.16) is true

$$H^{+}(Z,t) = \inf_{p>t} H(Z,p) \leq \inf_{p>t} S(Z,p) = S(Z,t),$$

as both H^{+} and S are right continuous in the second argument.

Writing explicitly the inequality (5.16)

$$\sup_{\xi\in L^\infty}\{\phi_{\mathbb{F}}(\xi) \mid E[\xi Z] \leq p\} \leq \sup\{q \in \mathbb{R} \mid \gamma(Z,q) \leq p\}$$

and letting $\xi \in L^\infty$ satisfying $E[\xi Z] \leq p$, we see that it is sufficient to show the existence of $q \in \mathbb{R}$ such that $\gamma(Z,q) \leq p$ and $q \geq \phi_{\mathbb{F}}(\xi)$. If $\phi_{\mathbb{F}}(\xi) = \infty$ then $\gamma(Z,q) \leq p$ for any q and therefore $S(Z,p) = H(Z,p) = \infty$.

Suppose now that $\infty > \phi_{\mathbb{F}}(\xi) > -\infty$ and define $q := \phi_{\mathbb{F}}(\xi)$. As $E[\xi Z] \leq p$ we have:

$$\gamma(Z,q) := \inf \{E[\xi Z] \mid \phi_{\mathbb{F}}(\xi) \geq q\} \leq p.$$

Then $q \in \mathbb{R}$ satisfies the required conditions.

To obtain $H^+(Z,t) := \inf_{p>t} H(Z,p) \geq S(Z,t)$ it is sufficient to prove that, for all $p > t$, $H(Z,p) \geq S(Z,t)$, that is :

$$\sup_{\xi \in L^\infty} \{\phi_{\mathbb{F}}(\xi) \mid E[\xi Z] \leq p\} \geq \sup \{q \in \mathbb{R} \mid \gamma(Z,q) \leq t\}. \tag{5.17}$$

Fix any $p > t$ and consider any $q \in \mathbb{R}$ such that $\gamma(Z,q) \leq t$. By the definition of γ, for all $\varepsilon > 0$ there exists $\xi_\varepsilon \in L^\infty$ such that $\phi_{\mathbb{F}}(\xi_\varepsilon) \geq q$ and $E[\xi_\varepsilon Z] \leq t + \varepsilon$. Take ε such that $0 < \varepsilon < p - t$. Then $E[\xi_\varepsilon Z] \leq p$ and $\phi_{\mathbb{F}}(\xi_\varepsilon) \geq q$ and (5.17) follows. ■

Remark 141 *This dual representation provides an interesting interpretation of the SRM. Let Q be the 'weight' that we can assign to the author's publications (for example, the impact factor of the Journal where the article is published). For a fixed Q, the term $\gamma(Q,q) := \inf \{E_Q[\xi] \mid \phi_{\mathbb{F}}(\xi) \geq q\}$ represents the smallest Q-average of citations that a generic author needs in order to have the SRM at least of q. We observe that this term is independent from the citations of the author X.*

On the light of these considerations we can interpret the term $H^+(Q, E_Q[X]) := \sup \{q \in \mathbb{R} \mid E_Q[X] \geq \gamma(Q,q)\}$ as the greatest performance level that the author X can reach, *in the case that we attribute the weight Q to the publications. Namely, we compare the Q-average of the author X, $E_Q[X]$, with the minimum Q-average necessary to reach each level q, that is $\gamma(Q,q)$.*

Finally, the SRM of the author X, $\phi_{\mathbb{F}}(X) = \inf_{Q\in\mathcal{P}} H^+(Q, E_Q[X])$, corresponds to the smallest performance level obtained changing the weight attributed to the journals.

The theorem exhibits the relationship between the performance curve approach and this average approach.

In the following examples we find *the dual representation of some existing indices.* In all these examples the family $\mathbb{F}$ of performance curves is left continuous hence, by Lemma (138), the associated SRM $\phi_{\mathbb{F}}$ is $\sigma(L^\infty, L^1)$-upper semicontinuous and X satisfies: $\phi_{\mathbb{F}}(X) \geq q$ iff $X \in \mathcal{A}_q$ iff $X \geq f_q$. Therefore, we find the dual representation computing γ, H^+ and $\phi_{\mathbb{F}}$ applying the formulas: (5.15),(5.14) and (5.13). Recall that $X = \sum_{i=1}^p x_i 1_{(i-1,i]}$, with $x_i \geq x_{i+1}$ for all i.

Example 142 (max # of citations) *Consider the example (118). For $Z \in L^1_+$, we compute $\gamma(Z,q)$*

$$\gamma(Z,q) := \inf_{\phi_{\mathbb{F}_{c\max}}(X)\geq q} E[ZX] = \inf_{X(x)\geq q 1_{(0,1]}(x)} E[ZX] = qE[1_{(0,1]}Z]$$

where the first equality is due to (5.12). We obtain

$$H^+(Z, E[ZX]) := \sup\left\{q \in \mathbb{R} \mid E[ZX] \geq qE[1_{(0,1]}Z]\right\} = \frac{E[ZX]}{E[1_{(0,1]}Z]}.$$

In our application, any non zero citation vector X always satisfies $X \geq x_1 1_{(0,1]}$ and, since $E[X1_{(0,1]}] = x_1 E[1_{(0,1]}]$, we also have: $\frac{1_{(0,1]}}{E[1_{(0,1]}]} \leq \frac{X}{E[X1_{(0,1]}]}$. Therefore,

$$E\left[Z\frac{1_{(0,1]}}{E[1_{(0,1]}]}\right] \leq E\left[Z\frac{X}{E[X1_{(0,1]}]}\right] \quad \forall Z \in L^1_+(\mathbb{R})$$

and

$$\frac{E\left[ZX\right]}{E\left[Z1_{(0,1]}\right]} \geq \frac{E[1_{(0,1]}X]}{E[1_{(0,1]}]} \quad \forall Z \in L^1_+(\mathbb{R}).$$

Hence:

$$\begin{aligned}\phi_{\mathbb{F}_{c_{\max}}}(X) &= \inf_{Z \in L^1_+(\mathbb{R})} H^+(Z, E[ZX]) = \inf_{Z \in L^1_+(\mathbb{R})} \frac{E[ZX]}{E[Z1_{(0,1]}]} \\ &= \frac{E[1_{(0,1]}X]}{E[1_{(0,1]}1_{(0,1]}]},\end{aligned}$$

i.e. the infimum is attained at $Z = 1_{(0,1]} \in L^1_+$, which is of course natural as this SRM weights only to the first publication.

Example 143 (total # of publications) *Consider the example (119). For $Z \in L^1_+$, we compute $\gamma(Z, q)$ as in the previous example:*

$$\gamma(Z, q) = \inf_{X \geq 1_{(0,q]}} E[ZX] = E[1_{(0,q]}Z]$$

We obtain

$$H^+(Z, E[ZX]) := \sup\left\{q \in \mathbb{R} \mid E[ZX] \geq E[1_{(0,q]}Z]\right\}$$

Hence the dual representation of the total number of publications p *with at least one citation is*

$$\phi_{\mathbb{F}_p}(X) = \inf_{Z \in L^1_+(\mathbb{R})} \sup_{E[ZX] \geq E[1_{[0,q]}Z]} q$$

We show indeed that $\phi_{\mathbb{F}_p}(X) = p$, where p is such that $X = X1_{(0,p]} \in L^\infty_+$. First we check that $\phi_{\mathbb{F}_p}(X) \geq p$. For all $Z \in L^1_+$, and $q \leq p$ we have

$$E[ZX] = E[ZX1_{(0,p]}] \geq E[1_{(0,q]}Z]$$

and therefore

$$\sup_{E[ZX] \geq E[1_{(0,q]}Z]} q \geq p \qquad \forall Z \in L^1_+,$$

and $\phi_{\mathbb{F}_p}(X) \geq p$. *Regarding the* $\leq$ *inequality, it is enough to take* $Z = 1_{(p,p+\delta]}$, *with* $\delta > 0$, *for* $X = X1_{(0,p]}$. *In this case, the condition* $E[ZX] \geq E[1_{(0,q]}Z]$ *becomes*

$$0 = E[1_{(p,p+\delta]}X] \geq E[1_{(0,q]}1_{(p,p+\delta]}]$$

that holds only for $q \leq p$, *hence*

$$\sup_{E[X1_{(p,p+\delta]}] \geq E[1_{(0,q]}1_{(p,p+\delta]}]} q = p$$

and $\phi_{\mathbb{F}_p}(X) \leq p$.

Example 144 (h-index) *Consider the example (120). For* $Z \in L^1_+$,

$$\gamma(Z, q) = \inf_{X(x) \geq q1_{(0,q]}(x)} E[ZX] = E[Zq1_{(0,q]}]$$

We obtain

$$H^+(Z, E[ZX]) := \sup\left\{q \in \mathbb{R} \mid E[ZX] \geq E[Zq1_{(0,q]}]\right\}$$

Hence the dual representation of the h-index *is*

$$\phi_{\mathbb{F}_h}(X) = \inf_{Z \in L^1_+(\mathbb{R}^+)} \sup_{E[ZX] \geq E[Zq1_{(0,q]}]} q$$

We indeed show that $\phi_{\mathbb{F}_h}(X) = h$, *where* h *is such that* $X1_{(0,h]} \geq h1_{(0,h]}$ *and* $X1_{(h,+\infty)} \leq h1_{(h,+\infty)}$. *First we check that* $\phi_{\mathbb{F}_h}(X) \geq h$. *For all* $Z \in L^1_+$, *and* $q \leq h$ *we have*

$$E[ZX] \geq E[ZX1_{(0,h]}] \geq E[Zq1_{(0,q]}],$$

hence

$$\sup_{E[ZX] \geq E[q1_{(0,q]}Z]} q \geq h \qquad \forall Z \in L^1_+$$

and $\phi_{\mathbb{F}_h}(X) \geq h$.

Regarding the $\leq$ side, take $Z = 1_{(h,h+\delta]}$ with $\delta > 0$. For any $q \leq h$ the condition

$$E[X1_{(h,h+\delta]}] \geq E[q1_{(0,q]}1_{(h,h+\delta]}] = 0$$

holds. Instead, $\forall q > h$ there exists $\delta > 0$ such that $h + \delta < q$ and then

$$E[X1_{(h,h+\delta]}] \leq E[h1_{(h,h+\delta]}] < E[q1_{(0,q]}1_{(h,h+\delta]}]$$

hence

$$\sup_{E[X1_{(h,h+\delta]}] \geq E[q1_{(0,q]}1_{(h,h+\delta]}]} q \leq h$$

and $\phi_{\mathbb{F}_h}(X) \leq h$.

On an alternative approach to SRMs

The dual representation suggests us another approach for the definition of a generic class of SRMs. This approach is based on the assumption that we can represent the author's citation as a function $X(w)$ defined on the events $w \in \Omega$, where each event now corresponds to the journal in which the paper appeared.

We start fixing a plausible family $\mathcal{P} \subseteq \{Q \ll P\}$ where each $Q(w)$ represents the *'value'* attributed to the journal $w \in \Omega$. It is clear that the valuation criterion for journals (i.e. the selection of the family $\mathcal{P}$) has to be determined a priori and could be based on the *'impact factor'* or other criterion. A specific Q could attribute more importance to the journals with a large number of citations (a large impact factor); another particular Q to the journals having a high quality.

As suggested from the dual representation results and in particular from the equations (5.13) and (5.14) we consider, independently to the particular scientist X, a family

$\{\gamma_\beta\}_{\beta\in\mathbb{R}}$ of functions $\gamma_\beta : \mathcal{P} \to \overline{\mathbb{R}}$ that associate to each Q the value $\gamma_\beta(Q)$, that should represent the smallest Q-average of citations in order to reach a quality index at least of β.

So given a particular value $Q(w_i)$ for each i^{th}-journal and the average citations $\gamma_\beta(Q)$ necessary to have an index level greater than β, we build the SRM in the following way. We define the function $H^+ : \mathcal{P} \times \mathbb{R} \to \overline{\mathbb{R}}$ that associates to each pair $(Q, E_Q(X))$ the number

$$H^+(Q, E_Q(X)) := \sup\left\{\beta \in \mathbb{R} \mid E_Q(X) \geq \gamma_\beta(Q)\right\},$$

which represents the greatest quality index that the author X can reach when Q is fixed, and we build the SRM as follows:

$$\phi(X) := \inf_{Q\in\mathcal{P}} H^+(Q, E_Q(X))$$

which represents a prudential and robust approach with respect to $\mathcal{P}$, the plausible different selections of the evaluation of the Journals. This SRM is by construction *quasi-concave* and *monotone increasing*.

5.3 Empirical results

The SRM $\phi_{\mathbb{F}}(X)$ depends on the family of performance curves $\mathbb{F} := \{f_q\}_{q\in\mathcal{I}}$, which depends on the **scientific area** and **seniority** under consideration.

We suggest two kinds of calibration:

- *Calibration to scientific areas*: each scientific area should determine their own family $\mathbb{F} := \{f_q\}_{q\in\mathcal{I}}$ by using existing data from a sample of "well known established scientific expert" researchers in the same area. Then this family will reflect the characteristics of the citation records of that particular scientific community.

- *Calibration to age:* similarly, in each area, it could be possible to determine n (two/three) families of performance curves $\mathbb{F}^i := \left\{f_q^i\right\}_{q\in\mathcal{I}}$, $i = 1, ...n$, that correspond to different ages (seniority in research), each determining n indices $\phi_{\mathbb{F}}^i$ that could be used to compute the scientific research quality of authors of the same seniority. A more advanced aspect is to calibrate the *time evolution* of the performance curves, in order to capture also the *scientific productivity*.

The SRM should be used only in *relative* terms (to compare the author quality with respect to the other researchers in the same area) in order to classify the authors (and structures) into few classes of homogeneous research quality.

In this section we provide a procedure to *calibrate* the family $\mathbb{F}$ from the historic data available for a group of "well known established scientific expert" researchers in the same area.

Sample setting

The first step consists in the selection of a representative sample of M authors belonging to the same scientific area and with the same seniority.

If p is the total number of the author's publications with at least one citation, then $X = \sum_{i=1}^{p} x_i 1_{(i-1,i]}$, with $x_i \geq x_{i+1}$ for all i, where the first component x_1 corresponds to the number of citations received by the most cited article and similarly for $x_1 \geq x_2 \geq ... \geq x_p$.

The citation data of each author are downloaded from Google Scholar by a procedure implemented in Pyton. This procedure performs a filter on the name of the author and on the scientific area we are analyzing.

Determination of the family $\{f_q\}_q$ and of the SRM

First of all we need to determine the family of curves $\{f_q\}_q$ that better fit the citation curve of the sample of the selected scientists. By the analysis of the data we found that the theoretical model is the following hyperbole-type equation:

$$y = f_q(x) = \frac{q}{x^{\beta}} \tag{5.18}$$

with $q, \beta \in \mathbb{R}_+$. Setting $\ln y = Y$, $\ln(q) = \hat{q}$, $\ln x = X$, $\beta = \hat{\beta}$ we obtain the linearized model

$$Y = \hat{q} - \hat{\beta} X. \tag{5.19}$$

For each i-th author of the sample we determine $\hat{\beta}_i$ that minimizes the sum of the square distances of the points from the line (5.19). Therefore, we determine $\bar{\beta}$ as the average of the $\hat{\beta}_i$:

$$\bar{\beta} = \frac{1}{M} \sum_{i=1}^{M} \hat{\beta}_i.$$

Fixing the parameter $\bar{\beta}$, common to every author's citation curve, we obtain the ϕ-index of each author X as

$$\phi(X) = \sup \left\{ q \in \mathbb{R} \mid X(x) \geq \frac{q}{x^{\bar{\beta}}} \qquad \forall x \right\} \tag{5.20}$$

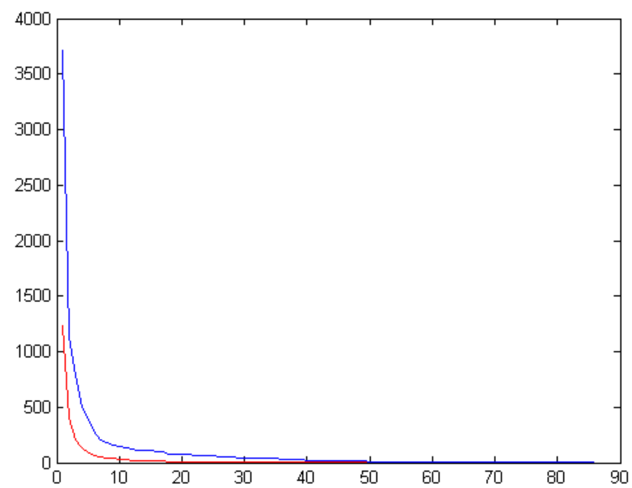
using the following simple computation:

$$\phi(X) = \min_{x \in \{1,2,\dots p\}} q_x \qquad \text{where } q_x = X(x) \cdot x^{\bar{\beta}}$$

Namely, the procedure consists in two steps:

1. we determine p parameters q_x, each one obtained forcing the curve to pass in each one of the p points $(x, X(x))$ of the scientist citation curve. In this way we guarantee that the curve we are going to find is the greatest;

2. we take $q = \min q_x$, in this way we are sure that the scientist citation vector (blue line in the following figure) is greater than the hyperbole-type curve (red line in the following figure).



Afterwards, we suggest the use of these results in order to classify the authors into few classes of homogeneous research quality, in order to facilitate a comparison between scientific areas and seniority.

Our Results

We have chosen a group of 20 well established researchers in the mathematical finance area. The analysis of the citation vectors of each author (see Fig.5.3.a) brings out that the theoretical model is that in the formula (5.18). We have computed the $\hat{\beta}_i$ for each author, that we report in the Fig.5.3.b, finding that $\bar{\beta} = 1,62$.

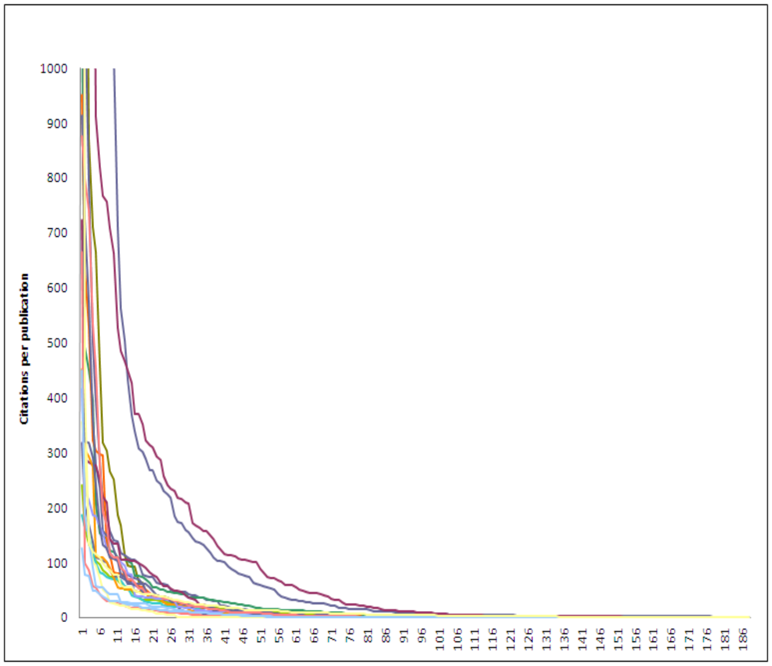


Fig.5.3.a. Citation curves of 20 senior authors in Math Finance area.

Author	beta
Author A	2,175691523
Author B	2,056970317
Author C	2,055140708
Author D	1,798296016
Author E	1,440536232
Author F	1,661309419
Author G	1,927231658
Author H	1,674063201
Author I	1,309767545
Author J	1,377028962
Author K	1,223621198
Author L	1,554745643
Author M	1,586690298
Author N	1,529379342
Author O	1,605636041
Author P	1,524849198
Author Q	1,358412433
Author R	1,839198418
Author S	1,297305299
Author T	1,477010726

Fig. 5.3.b

In the following table (Fig.5.3) we show the results and the ranking obtained calculating the ϕ-index as in (5.20) and the h-index for each author.

Author	ø-index	rank ø-index	h-index	rank h-index
Author A	4423	1	53	2
Author B	2985	2	60	1
Author D	1235	3	35	3
Author E	1136	4	35	4
Author F	950	5	25	14
Author C	908	6	28	7
Author R	875	7	28	8
Author T	800	8	29	6
Author P	780	9	28	9
Author H	723	10	33	5
Author G	511	11	26	11
Author L	451	12	24	15
Author Q	449	13	20	17
Author M	417	14	27	10
Author J	318	15	26	12
Author N	304	16	17	19
Author I	240	17	26	13
Author O	221	18	15	20
Author K	186	19	23	16
Author S	127	20	18	18

Fig. 5.3.

We note that the ϕ-index is more granular, allowing a more precise comparison between scientists. For example, the author F increases his index, moving from the position 14 of h-index to 5 of ϕ-index. If we compare this author with the author I, we note that they have almost the same h-index but the F's ϕ-index is definitely greater than the I's ϕ-index. Analyzing their citation curves we observed that they have the same number of publications, but F has in general many more citations for any publication than I, especially those in the Hirsh-core. The same reasons can be provide for the comparison between the author H and the author D, in this case we noticed also that D has also more publications.

Bibliography

[CIT] Adler, R., Ewing, J., Taylor, P. (2008) “Citation Statistics”, http://www.mathunion.org/fileadmin/IMU/Report/CitationStatistics.pdf.

[Ali] Aliprantis, C.D. and Border K.C. (2005), "Infinite dimensional analysis", *Springer, Berlin*, 3rd edition.

[ACHH] Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E. and Herrera, F. (2009a), "h-index: A review focused in its variants, computation and standardization for different scientific fields", *Journal of Informetrics*, 3, 273-289

[ACHHb] Alonso, S., Cabrerizo, F.J., Herrera-Viedma, E. and Herrera, F. (2010). "hg-index: A new index to characterize the scientific output of researchers based on the h-and g- indices", *Scientometrics*, 82:391–400.

[ASOAF] American Scientist Open Access Forum (2008) "Comments to the Citation Statistics Report", http://openaccess.eprints.org/index.php?/archives/417-Citation-Statistics-International-Mathematical-Union-Report.html.

[AHK08] Anderson, T., Hankin, K. and Killworth, P. (2008). "Beyond the durfee square: Enhancing the h-index to score total publication output", *Scientometrics*, 76(3), 577–588.

[ANVUR] ANVUR (2011a) "Criteri e parametri di valutazione dei candidati e dei commissari dell'abilitazione scientifica nazionale".

[ANVURb] ANVUR (2011b) "Commenti, osservazioni critiche e proposte di soluzione sul documento ANVUR relativo ai criteri di abilitazione scientifica nazionale".

[ADEH99] Artzner, P., Delbaen, F., Eber, J.-M., and Heath, D. (1999) "Coherent Measure of Risk", *Mathematical Finance*, 9, 203-228.

[ADHEK] Artzner, P., Delbaen, F., Eber, J.-M., Heath, D., Ku, H. (2004) "Coherent multiperiod risk adjusted values and Bellman's principle" Preprint, ETH-Zürich.

[B08] Bar-Ilan, J. (2008) "Which h-index? - A comparison of WoS, Scopus and Google Scholar", *Scientometrics*, 74(2), 257-271.

[BL00] Bernardo, A. and Ledoit, O. (2000), "Gain, Loss, and Asset Pricing", *Journal of Political Economy*, 108, 144-172.

[BF09] Biagini, S. and Frittelli, M. (2009) "On the extension of the Namioka-Klee theorem and on the Fatou property for risk measures", In: *Optimality and risk: modern trends in mathematical finance*, Eds: F. Delbaen, M. Rasonyi, Ch. Stricker, Springer, Berlin, 1–29.

[BCZ11] Bielecki, T.R, Cialenco, I. and Zhang, Z. (2011) "Dynamic Coherent Acceptability Indices and their Applications to Finance", preprint

[BD07] Bornmann, L., Daniel, H. (2007) "What do we know about the h-index?", *Journal of the American Society for Information Science and Technology*, 58(9), 1381–1385.

[BMD08] Bornmann, L., Mutz, R., and Daniel, H. (2008). "Are there better indices for evaluation purposes than the h-index? A comparison of nine different variants of the h-index using data from biomedicine", *Journal of the American Society for Information Science and Technology*, 59(5), 830–837.

[CAHH] Cabrerizo, F.J., Alonso, S., Herrera-Viedma, E. and Herrera, F. (2010) "q^2-Index: Quantitative and qualitative evaluation based on the number and impact of papers in the Hirsch core", *Journal of Informetrics*, 4 (2010) 23–28.

[CMMMa] Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011a) "Risk measures: rationality and diversification", *Mathematical Finance*, 21, 743-774.

[CMMMb] Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M. and Montrucchio, L. (2011b) “Complete Monotone Quasiconcave Duality ”, *Mathematics of Operations Research*, 36, 321-339.

[CM09] Cherny, A. and Madan, D. (2009) "New measures for performance evaluation", *Review of Financial Studies*, 22, 2571-2606.

[CB08] Costas, R. and Bordons, M. (2008), "Is g-index better than h-index? An exploratory study at the individual level", *Scientometrics*, 77(2), 267–288.

[DeFin] De Finetti, B. (1949) Sulle stratificazioni convesse, *Annali di Matematica pura e applicata* **30**, 173-183.

[DW09] Deineko, V. and Woeginger G.J. (2009) "A new family of scientific impact measures: The generalized Kosmulski-indices", *Scientometrics*, Vol. 80, n.3 (2009) 821–828

[DS05] Detlefsen, K. and Scandolo, G. (2005) "Conditional and dynamic convex risk measures", *Fin. Stoc.* 9, 539-561.

[DK10] Drapeau, S. and Kupper, M. (2010) "Risk preferences and their robust representation", *preprint*.

[DKR10] Drapeau, S., Kupper, M. and Reda, R. (2010) "A Note on Robust Representations of Law-Invariant Quasiconvex Functions", *preprint*.

[EW08] Eck, N.V., Waltman, L. (2008) "Generalizing the h- and g-indices" *Journal of Informetrics*, 2(4), 263–271.

[E06] Egghe, L. (2006) "Theory and practise of the g-index", *Scientometrics*, 69 (1), 131–152.

[E08] Egghe, L. (2008) "Mathematical theory of the h- and g-index in case of fractional counting of authorship", *Journal of the American Society for Information Science and Technology*, 59(10), 1608–1616.

[ER08] EGGHE, L. AND ROUSSEAU, R. (2008) "An h-index weighted by citation impact", *Information Processing & Management*, 44(2), 770–780.

[ER09] EL KAROUI, N. AND RAVANELLI, C. (2009) "Cash sub-additive risk measures and interest rate ambiguity", *Mathematical Finance*, 19 (4), 561-590.

[Fe] FENCHEL, W. (1952) "A remark on convex sets and polarity", *Meddelanden Lunds Universitets Matematiska Seminarium*, 82-89.

[FS02] FOLLMER, H., AND SCHIED, A. (2002) "Convex measures of risk and trading constraints", *Finance and Stochastics*, 6, 429-447.

[FS04] FOLLMER, H. AND SCHIED, A (2004) "Stochastic Finance. An Introduction in Discrete Time", 2nd ed., *de Gruyter Studies in Mathematics,* 27.

[FM11] FRITTELLI, M. AND MAGGIS, M. (2011) "Dual representation of quasiconvex conditional maps", *SIAM J. Financial Math*, Vol 2, pp 357-382.

[FR02] FRITTELLI, M., AND ROSAZZA GIANIN, E. (2002) "Putting order in risk measures", *Journal of Banking and Finance*, 26, 1473–1486.

[FR04] FRITTELLI, M., AND ROSAZZA GIANIN, E. (2004) "Dynamic convex risk measures", Risk Measures for the 21st Century, G. Szego ed., J. Wiley, 227-248.

[FR05] FRITTELLI, M., AND ROSAZZA GIANIN, E. (2005), "Law-invariant convex risk measures", *Advances in Mathematical Economics* 7, 33-46.

[GR09] GUNS, R., ROUSSEAU, R. (2009) "Real and rational variants of the h-index and the g-index", *Journal of Informetrics*, 3, 64–71.

[H05] HIRSCH, J.E. (2005) "An index to quantify an individual's scientific research output", *Proceedings of the National Academy of Sciences of the United States of America*, 102 (46), 16569–16572.

[J06] JIN, B. (2006) "h-Index: An evaluation indicator proposed by scientist", *Science Focus*, 1(1), 8–9.

[JLRE] JIN, B., LIANG, L., ROUSSEAU, R., EGGHE, L. (2007) "The R- and AR-indices: Complementing the h-index", *Chinese Science Bulletin*, 52 (6), 855-863.

[JST06] JOUINI, E.,SCHACHERMAYER, W. AND TOUZI, N. (2006), "Law invariant risk measures have the Fatou property", *Advances in Mathem. Economics*, 9, 49-71.

[KAB09] KATSAROS, D., AKRITIDIS, L. AND BOZANIS, P. (2009) "The f Index: Quantifying the Impact of Coterminal Citations on Scientists' Ranking", *Journal of the American Society for Information Science and Technology*, 60(5), 1051–1056.

[K06] KOSMULSKI, M. (2006) "A new Hirsch-type index saves time and works equally well as the original h-index", *ISSI Newsletter*, 2(3), 4–6.

[K01] KUSUOKA, S. (2001) "On law invariant coherent risk measures", *Advances in Mathematical Economics*, Vol. 3, Springer, Tokyo, 83-95.

[PV90] PENOT, J.P. AND VOLLE, M. (1990) "On Quasi-Convex Duality", *Mathematics of Operations Research*, Vol.**15**(4), 597-625 .

[ROC] Rockafellar, R.T (1974) "Conjugate duality and optimization", *Regional series in Applied Mathematics*, vol. 16, SIAM.

[R06] Rousseau, R. (2006) "New developments related to the Hirsch index", *Science Focus*, 1(4), p. 23–25.

[RY08] Rousseau, R. and Ye, F. (2008) "A proposal for a dynamic h-type index", *Journal of the American Society for Information Science and Technology*, 59(11), 1853–1855.

[RT08] Ruane, F., and Tol, R. (2008) "Rational (successive) h-indices: An application to economics in the Republic of Ireland", *Scientometrics*, 75(2), 395–405.

[S07] Schreiber, M. (2007) "Self-citation corrections for the Hirsch index", *Europhysics Letters*, 78(3), n. 30002.1-30002.6.

[S08a] Schreiber, M. (2008a) "An empirical investigation of the g-index for 26 physicists in comparison with the h-index, the a-index, and the r-index", *Journal of the American Society for Information Science and Technology*, 59(9), 1513–1522.

[S08b] Schreiber, M. (2008b), "To share the fame in a fair way, h_m for modifies h for multi-authored manuscripts", *New Journal of Physics*, 10 (040201), 1–9.

[Sh64] Sharpe, W. (1964), "Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 19, 425-442.

[Shi] Shiryaev, A.N. (1995), "Probability", *Graduate Texts in Mathematics*, Springer 2nd Edition.

[SKM07] Sidiropoulos, A., Katsaros, D. and Manolopoulos, Y. (2007) "Generalized hirsch h-index for disclosing latent facts in citation networks", *Scientometrics*, 72(2), 253–280.

[S10] Svindland, G. (2010) "Continuity Properties of Law-Invariant (Quasi-)Convex Risk Functions", *Math. and Financial Economics*.

[UMI] UMI (2010), *Notiziario dell'Unione Matematica Italiana*, Edizioni dell'UMI, Giugno-Luglio 2010, Anno XXXVII n. 6-7.

[V09] Vinkler, P. (2009) "The π-index: a new indicator for assessing scientific impact", *Journal of Information Science*, 35 (5) 2009, 602–612.

[Vo98] Volle, M. (1998) "Duality for the level sum of quasiconvex functions and applications", *Control, Optimisation and Calculus of variations*, Vol.**3**, 329-343.

[W99] Wang, T. (1999) "A class of dynamic risk measures", Preprint, University of British Columbia.

[W06] Weber, S. (2006), "Distribution-Invariant Risk Measures, Information, and Dynamic Consinstency", *Math. Finance*, Vol.**16**(2) , 419–441.

[W07] Wendl, M. (2007). "h-index: However ranked, citations need context", *Nature*, 449(7161), 403–1403.

[W08a] Woeginger, G. J. (2008a) "An axiomatic characterization of the Hirsch-index", *Mathematical Social Sciences,* 56, 224–232.

[W08b] Woeginger, G. J. (2008b) "A symmetry axiom for scientific impact indices", *Journal of Infometrics,* 2 (2008), 298–303.

[W09] Woeginger, G. J. (2009) "Generalizations of Egghe's g-Index", *Journal of the American Society for Information Science and Technology*, 60(6):1267–1273, 2009

[Z09] Zhang, C.-T. (2009) "The e-Index, Complementing the h-Index for Excess Citations", *PLoS ONE*, 4 (5), e5429.