Nendel, Max; Streicher, Jan Maximilian

Working Paper

An axiomatic approach to default risk and model uncertainty in rating systems

Center for Mathematical Economics Working Papers, No. 725

Provided in Cooperation with:
Center for Mathematical Economics (IMW), Bielefeld University

Center for Mathematical Economics
Working Papers

September 2023

725

An axiomatic approach to default risk and model uncertainty in rating systems

Max Nendel and Jan Streicher

Center for Mathematical Economics (IMW)
Bielefeld University
Universitätsstraße 25
D-33615 Bielefeld · Germany

e-mail: imw@uni-bielefeld.de
uni-bielefeld.de/zwe/imw/research/working-papers
ISSN: 0931-6558

AN AXIOMATIC APPROACH TO DEFAULT RISK AND MODEL UNCERTAINTY IN RATING SYSTEMS

MAX NENDEL AND JAN STREICHER

Abstract. In this paper, we deal with an axiomatic approach to default risk. We introduce the notion of a default risk measure, which generalizes the classical probability of default (PD), and allows to incorporate model risk in various forms. We discuss different properties and representations of default risk measures via monetary risk measures, families of related tail risk measures, and Choquet capacities. In a second step, we turn our focus on default risk measures, which are given as worst-case PDs and distorted PDs. The latter are frequently used in order to take into account model risk for the computation of capital requirements through risk-weighted assets (RWAs), as demanded by the Capital Requirement Regulation (CRR). In this context, we discuss the impact of different default risk measures and margins of conservatism on the amount of risk-weighted assets.

Key words: default risk measure, model uncertainty, probability of default, Choquet capacity, margin of conservatism, monetary risk measure, value at risk, risk-weighted assets

JEL Classification: G21; G28; G32

AMS 2020 Subject Classification: 91G40; 91G70; 28A12

1. Introduction

Financial institutions and corporate firms have a variety of different risk types, such as credit risk, market price risk, operational risk, or liquidity risk. These are ubiquitous in business operations and affect different business areas. In the context of corporate risk management, credit risk is of fundamental importance for a correct assessment of existing and potential future risks and an adequate management of business operations, in general. For most financial institutions, such as banks, credit risk is even the most important type of risk and therefore also in the focus of supervisory regulations such as the Capital Requirements Regulation (CRR) [17], which is part of Basel III and also known as Regulation (EU) 575/2013.

From the institutes' point of view, credit risk generally involves the risk that borrowers cannot repay loans granted to them in accordance with the predetermined contracts. In the context of counterparty credit risk, ratings serve the purpose of dividing borrowers or business partners, in general, into risk classes, which are specified via a probability of default over a one-year horizon. Since default risk represents a complex issue, there are different rating systems depending on the type of customers (individuals, corporate customers, banks, insurance companies, etc.). For banks, ratings represent the basis of credit risk controlling and reporting, and lead to more differentiated pricing, greater

Date: September 21, 2023.

The authors thank Felix-Benedikt Liebrich, Markus Klein, Ruodu Wang, as well as two anonymous referees for helpful comments and discussions related to this work. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – SFB 1283/2 2021 – 317210226. The second author gratefully acknowledges the support of the Landesbank Baden-Württemberg related to this work.

competitiveness, and an improvement in credit decisions. For a general introduction to credit risk modeling, we refer to Bluhm et al. [8] and Lando [30]. We also refer to Guo et al. [26] for an axiomatic study of credit rating criteria.

The focus of a rating is on the probability of default (PD), which refers to the creditworthiness of the borrower, and *not* to credit-specific terms, such as exposure at default (EaD) or loss given default (LGD). Nevertheless, all three terms play an important role in the context of default risk, are included in the calculation of expected losses (EL) and risk-weighted assets (RWAs), and are therefore part of supervisory requirements for risk-differentiated capital backing. Since rating systems are mathematical statistical models that transform a borrower's default-relevant characteristics into a statement of creditworthiness, they are subject to model risks and model uncertainties, which can lead to major discrepancies in credit risk management if neglected. Additionally, the European Banking Authority's (EBA) guidelines on PD and LGD estimation (EBA-GL-2017-16) [20] serve to reduce fluctuations in risk parameters, and focus on modeling techniques used in the estimation of risk parameters. In particular, the PD estimation in low-default portfolios is a subtle issue, cf. Pluto and Tasche [37] and Tasche [41].

In the aftermath of the subprime mortgage crisis, the topic of model uncertainty or Knightian uncertainty has become increasingly significant for financial institutions and found its way into regulatory requirements in various forms. As a consequence, this classical and already very prominent topic in economic theory has received even more attention in the literature on theoretical economics, mathematical finance, and actuarial sciences. Model uncertainty appears in the economic literature, for example, in the context of preference relations, cf. Gilboa and Schmeidler [25] and Maccheroni et al. [33], general equilibrium theory, cf. Beissner and Riedel [4], insurance pricing, cf. Castagnoli et al. [13], Nendel et al. [35], and Wang et al. [44], as well as hedging and no-arbitrage conditions, see, for instance, Bouchard and Nutz [9] and Burzoni et al. [12]. However, to the best of our knowledge, a detailed study of model uncertainty in credit risk management is not present in the literature.

This paper therefore aims to provide a decision-theoretic foundation for the treatment of default risk and model uncertainty in rating systems. While the work of Guo et al. [26] focuses on an axiomatic study of credit rating criteria, including, among others, generalized PD criteria, our axiomatic approach introduces the notion of a default risk measure, which aims to provide a more general perspective on PDs, and allows to include, for example, model uncertainty in the form of worst-case PDs and distorted PDs, warning signals, and default risk arising from regulatory risk measures. In particular, we explore the use of generalized versions of PDs in the context of credit risk, but do not aim to identify relevant criteria for credit ratings. In contrast to monetary risk measures, cf. Artzner et al. [3], Frittelli and Rosazza Gianin [23], and Föllmer and Schied [22], and nonlinear expectations, cf. Coquet et al. [16] and Peng [36], default risk measures do not behave linearly along constants but only take the values zero (no default) or one (default) for constant functions.

Throughout, we consider a set C of customers, i.e., a set of bounded measurable functions on a given measurable space $(\Omega, \mathcal{F})$ containing all constant functions. A default risk measure ϱ is a monotone functional that assigns to each customer $X \in C$ a default risk $\varrho(X) \in [0, 1]$. Here, positive values of X represent a negative total cash flow or, loosely speaking, a default. In a first step, we show that every default risk measure can be extended from the set C to the space $\mathrm{B_b}$ of all bounded measurable functions, cf. Theorem 10. The extension procedure is constructive, and shows how default risk can be assigned consistently to new customers based on a financial institution's stock of existing clients.

In a second step, we consider tail risk measures related to default risk measures. As noted by Liu and Wang [32], the consideration of tail risk, i.e., the risk beyond a given threshold, is crucial in today's financial regulation. We also refer to Bignozzi et al. [7] for a generalization of the value at risk that depends on the size of potential losses in the form of quantile-based risk measures, to Fadina et al. [21] for an axiomatic study of quantiles, and to Burzoni et al. [11] for a study of adjusted expected shortfall. In Section 3, we show that each default risk measure induces its own notion of a value at risk, and establish a ono-to-one relation between default risk measures and so-called generalized quantile functions. For particular choices of default risk measures, e.g., distorted PDs and worst-case PDs, we provide explicit representations of the related value at risk. We point out that our notion of a generalized quantile function follows a different philosophy than the notion of a tail risk measure introduced by Liu and Wang [32].

A key property of the PD is that it is specified only by the states of the world where a negative total cash flow is realized, independent of the amount of capital given liquidity or illiquidity. In mathematical terms, this means that the PD of X is the same as the PD of $\mathbb{1}_{\{X>0\}}$. In Section 4, we characterize default risk measures that have this property, and connect them to Choquet capacities, cf. Dellacherie and Meyer [18]. In this context, the notions of default scaling invariance, liquidity invariance, and illiquidity invariance play a fundamental role. Using continuity properties of Choquet integrals, we derive sufficient and necessary conditions for default risk measures to admit a representation via probability measures or, in other words, as worst-case PDs.

For the calculation of RWAs, the concept of a margin of conservatism (MoC) is used, in practice, to quantify the amount of model uncertainty. The regulatory need to consider model uncertainty regarding default risks by calculating a MoC that reflects the expected range of estimation errors can be found in Article 179 (f) or, PD-specific, in Article 180 (e) of the CRR [17], among others. In Section 6, we characterize default risk measures that are given in terms of a MoC or, equivalently, as distorted PDs. The characterization generalizes the fact that law-invariant capacities on an atomless probability space can be represented as distorted probabilities, cf. Wang et al. [44], where this result is established in the context of insurance premia that are given as Choquet integrals and Amarante and Liebrich [1] for a detailed study of distortion risk measures, i.e., law-invariant and comonotonically additive risk measures.

Moreover, we establish a connection between distorted PDs and worst-case PDs or, equivalently, the margin of conservatism and a suitable set of probability measures, based on the Kusuoka representation of law-invariant risk measures, cf. Kusuoka [29], and the well-known Fréchet-Hoeffding bounds for joint distributions, cf. Burgert and Rüschendorf [10], which also play a fundamental role for the collapse to the mean of law-invariant risk functionals, cf. Bellini et al. [5] and Liebrich and Munari [31].

In Section 7, we use the results on distorted PDs in a case study on capital requirements as demanded by current regulations. There, we discuss the impact of model uncertainty in rating systems on financial institutions' RWAs. Since, from a regulatory perspective, model uncertainty only has to be considered for unexpected losses and not for expected losses (EL), ironically, a high degree of model uncertainty can actually reduce the amount of capital requirement for badly rated customers, since it transforms unexpected losses into expected ones. We refer to Example 53 for the details.

The rest of the paper is organized as follows. In Section 2, we define the notion of a default risk measure, illustrate the definition in several examples, and state our extension result for default risk measures (Theorem 10). The link between default risk measures and generalized quantile functions is discussed in Section 3. Section 4 is devoted to default risk measures that are given only in terms of default scenarios, see Theorem 31. Section 5 contains several results on robust representations as worst-case PDs. In Section 6, we focus on law-invariant risk measures and distorted PDs. There, we connect distortion functions with certain properties to sets of absolutely continuous probability measures based on the value at risk and expected shortfall of probability densities. In Section 7, we discuss capital requirements for rating systems, and illustrate the impact of different default risk measures on the amount of financial institutions' risk weighted assets. In Appendix A, we provide a short proof for a characterization of exact capacities and distorted probabilities, cf. Aouani and Chateauneuf [2] and Kadane and Wassermann [28]. The proofs of Section 2 are contained in Appendix B. The proofs of Section 3 can be found in Appendix C. The proofs of Section 4 are collected in Appendix D. The proofs of Section 5 are given in Appendix E and the proofs of Section 6 in Appendix F.

2. Default Risk Measures: Definition and Examples

In this section, we introduce the concept of a default risk measure, which is strongly motivated by the probability of default (PD) as a prime example. Like monetary risk measures, default risk measures are monotone functionals defined on suitable sets of measurable functions, cf. [22]. However, they exhibit a completely different behaviour along constants.

Throughout, let $(\Omega, \mathcal{F})$ be a measurable space and $\mathrm{B_b} = \mathrm{B_b}(\Omega, \mathcal{F})$ denote the space of all bounded measurable functions $\Omega \to \mathbb{R}$. We consider a set $C \subset \mathrm{B_b}$, containing the set of all constant functions. A function $X \in C$ can be interpreted as a customer of a financial institution with $-X(\omega)$ being the sum of all financial flows (earnings, spendings, and maturities combined) at the end of the respective observation period if a scenario $\omega \in \Omega$ is realized. Thus, positive values of X resemble a negative sum of all financial flows, which we will, loosely speaking, refer to as a *default*. Choosing this, in comparison to the literature on monetary risk measures, inverted sign convention leads to an easier exposition since it avoids confusion arising from repeated sign changes on several occasions. As in the theory of monetary risk measures, we do not differentiate between a real constant $m \in \mathbb{R}$ and the constant function $X \colon \Omega \to \mathbb{R}$ with $X(\omega) = m$ for all $\omega \in \Omega$, and write $X = m$, thinking of it as *cash*. For $X, Y \in \mathrm{B_b}$, we write $X \leq Y$ if $X(\omega) \leq Y(\omega)$ for all $\omega \in \Omega$. Moreover, we define

$$\inf X := \inf_{\omega\in\Omega} X(\omega) \quad \text{and} \quad \sup X := \sup_{\omega\in\Omega} X(\omega) \quad \text{for all } X \in \mathrm{B_b}.$$

Additionally, for $X \in \mathrm{B_b}$, we consider the standard decomposition $X = X^+ - X^-$ with

$$X^+ := X\mathbb{1}_{\{X>0\}} \quad \text{and} \quad X^- := -X\mathbb{1}_{\{X<0\}}.$$

For any two real numbers $x, y \in \mathbb{R}$, we use the notation $x \vee y := \max\{x, y\}$ and $x \wedge y := \min\{x, y\}$. In a similar fashion, we write $X \vee Y$ and $X \wedge Y$ for the pointwise maximum and minimum of $X, Y \in \mathrm{B_b}$, respectively. Throughout, we use the following slightly modified notion of a (monetary) risk measure, and refer to [22] for a detailed discussion on this topic.

Definition 1. We say that a map $R\colon \mathrm{B_b} \to \mathbb{R}$ is a *(monetary) risk measure* if

(i) $R(X) \leq R(Y)$ for all $X, Y \in \mathrm{B_b}$ with $X \leq Y$,
(ii) $R(0) = 0$ and $R(X + m) = R(X) + m$ for all $X \in \mathrm{B_b}$ and $m \in \mathbb{R}$.

We now introduce the central object of our study.

Definition 2. A map $\varrho\colon C \to [0, 1]$ is called a *default risk measure* if

(i) $\varrho(X) \leq \varrho(Y)$ for all $X, Y \in C$ with $X \leq Y$,
(ii) $\varrho(0) = 0$ and $\varrho(m) = 1$ for all $m \in \mathbb{R}$ with $m > 0$.

Thinking of PDs, the respective properties seem to be very canonical. For instance, comparing two customers it is obvious that the one with the higher total cash flow in all scenarios exhibits a lower risk of default (Property (i)). Moreover, for all $X \in C$ with $X \leq 0$,

$$0 \leq \varrho(X) \leq \varrho(0) = 0,$$

i.e., if all obligations can be payed in any scenario the customer's default risk will be zero, and a constant negative total cash flow ($m > 0$) leads at least to an unlikely repayment, and hence to a sure default (Property (ii)). Although Property (i) in the definition of a default risk measure is analogous to the monotonicity of monetary risk measures, Property (ii) is substantially different from the standard *cash additivity* or *translation invariance*. To that end, consider a default risk measure $\varrho\colon \mathrm{B_b} \to [0, 1]$ and observe that

$$\varrho(X - \sup X) = 0 \quad \text{for all } X \in \mathrm{B_b}.$$

By definition, neither convexity nor positive homogeneity (of degree 1) are meaningful properties for default risk measures, since

$$\varrho(\lambda) = 1 > \lambda = \lambda\varrho(1) \quad \text{for all } \lambda \in (0, 1).$$

We thus observe that the properties of monetary risk measures differ substantially from those of default risk measures despite the similarity of their very general definitions. Nevertheless, there is the possibility to construct default risk measures from monetary risk measures as the following example illustrates.

Example 3 (Default risk measure defined by a monetary risk measure)**.** Given a monetary risk measure $R\colon \mathrm{B_b} \to \mathbb{R}$ as in Definition 1, we are able to construct a default risk measure via

$$\varrho(X) := R\big(\mathbb{1}_{\{X>0\}}\big) \quad \text{for all } X \in \mathrm{B_b}.$$

For $X, Y \in \mathrm{B_b}$ with $X \leq Y$, we have $\mathbb{1}_{\{X>0\}} \leq \mathbb{1}_{\{Y>0\}}$, so that

$$0 \leq R\big(\mathbb{1}_{\{X>0\}}\big) \leq R\big(\mathbb{1}_{\{Y>0\}}\big) \leq 1.$$

Moreover, $\varrho(0) = R\big(\mathbb{1}_\emptyset\big) = R(0) = 0$ and $\varrho(m) = R(1) = 1$ for all $m \in \mathbb{R}$ with $m > 0$.

We continue with several examples of default risk measures, and begin with the most prominent one.

Example 4 (Probability of default)**.** We fix a reference probability measure $\mathbb{P}$ on $\mathcal{F}$, and consider the *probability of default* (PD), given by

$$\mathrm{PD}_\mathbb{P}(X) := \mathbb{P}(X > 0) \quad \text{for all } X \in \mathrm{B_b}.$$

If the probability measure $\mathbb{P}$ is discrete, this default risk measure could be interpreted as a mapping to common rating classes. Clearly, $\mathrm{PD}_{\mathbb{P}}$ satisfies all properties of a default risk measure. For $X, Y \in \mathrm{B_b}$ with $X \leq Y$, $\{X > 0\} \subset \{Y > 0\}$, so that

$$0 \leq \mathbb{P}(X > 0) \leq \mathbb{P}(Y > 0) \leq 1.$$

Furthermore, $\mathrm{PD}_{\mathbb{P}}(0) = \mathbb{P}(\emptyset) = 0$ and $\mathrm{PD}_{\mathbb{P}}(m) = \mathbb{P}(\Omega) = 1$ for all $m \in \mathbb{R}$ with $m > 0$.

Building on this example, we can also consider the case, where model uncertainty is taken into account via a distortion function.

Example 5 (Distorted PD)**.** Again, we fix a reference probability measure $\mathbb{P}$ on $\mathcal{F}$. Due to a lack of data, bad data quality, or changing economic environments, the consideration of uncertainties in form of a *margin of conservatism* (MoC) becomes more and more important for financial institutions. Since such model uncertainties are part of any model, including rating models, it is possible that the reference probability measure $\mathbb{P}$ is not the 'precise' probability measure that represents the default risk of costumers over a one-year time horizon. We therefore consider a nondecreasing distortion function $T\colon [0,1] \to [0,1]$ with $T(0) = 0$ and $T(1) = 1$. We define

$$\varrho(X) := T\big(\mathbb{P}(X > 0)\big) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \text{for all } X \in \mathrm{B_b}.$$

In this case, the distortion function T can be regarded as a benchmark for model uncertainty, and the margin of conservatism is given by

$$\mathrm{MoC}(p) := \frac{T(p)}{p} - 1 \quad \text{for all } p \in (0,1].$$

Clearly, the two properties of a default risk measure carry over from the classical PD, cf. Example 4, to the distorted PD for any probability measure $\mathbb{P}$ and any nondecreasing distortion function $T\colon [0,1] \to [0,1]$ with $T(0) = 0$ and $T(1) = 1$.

Apart from distorting a reference probability measure as in the previous example, there is also the possibility of incorporating model uncertainty via worst-case considerations among sets of probability measures.

Example 6 (Worst-case PD)**.** In order to properly account for uncertainties w.r.t. model specifications, it is often necessary to consider various models at the same time. This becomes particularly relevant, if the models have different sets of measure zero, since then one model neglects certain events that occur with positive probability under a different model. We therefore consider the following generalization of Example 4. Let $\mathcal{P}$ be a nonempty set of probability measures and

$$\varrho(X) := \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{Q}(X > 0) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathrm{PD}_{\mathbb{Q}}(X) \quad \text{for all } X \in \mathrm{B_b}.$$

As before, the properties (i) and (ii) have been shown for classical PDs in Example 4, and remain valid when taking the supremum over PDs. Following [2], we will see that, in many cases, distorted PDs allow for a representation as worst-case PDs and vice versa, see Section 6.

Up to now, all examples for default risk measures have been of the form

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) \quad \text{for all } X \in \mathrm{B_b}, \tag{1}$$

i.e., $\varrho(X)$ only depends on the set where $X \in \mathrm{B_b}$ is larger than zero, completely independent of its values. Thinking of PDs from rating systems, this is a very desirable property, since it implies that the financial institution is only interested in the customers' ability to pay their dues. In Section 4, we derive sufficient and necessary conditions for default risk measures in order to satisfy (1). The following two examples show that, however, not every default risk measure needs to allow for such a representation as it is also possible to define default risk measures that depend on specific values of X, for instance, using a warning signal that leads to a more conservative risk assessment.

Example 7 (Warning signal)**.** Consider two default risk measures ρ_0 and ρ_c with $\rho_0 \le \rho_c$, i.e., ρ_0 is less conservative then ρ_c. The idea is that ρ_c acts as a warning signal if scenarios are possible where the loss exceeds a given maximum level. For $X \in \mathrm{B_b}$ and $\gamma > 0$, we define

$$\varrho_\gamma(X) := \begin{cases} \varrho_0(X), & \varrho_c\big(X - \frac{1}{\gamma}\big) = 0, \\ \varrho_c(X), & \varrho_c\big(X - \frac{1}{\gamma}\big) > 0. \end{cases}$$

Thus, for $\gamma > 0$, we change from ϱ_0 to the more conservative default risk measure ϱ_c when the potential loss exceeds the level $\frac{1}{\gamma}$ under ϱ_c. We observe that

$$\lim_{\gamma \to 0} \varrho_\gamma(X) = \varrho_0(X) \quad \text{for all } X \in \mathrm{B_b}.$$

In fact, $\varrho_\gamma(X) = 0 = \varrho_0(X)$ for all $X \in \mathrm{B_b}$ with $X \le 0$, and $\varrho_c(X - \sup X) = 0$ for all $X \in \mathrm{B_b}$ with $\sup X > 0$. A concrete choice for ϱ_c and ϱ_0 are, for example, $\varrho_0 = \mathrm{PD}_\mathbb{P}$ and $\varrho_c(x) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathrm{PD}_\mathbb{Q}(X)$ (worst-case PD) for all $X \in \mathrm{B_b}$, where $\mathcal{P}$ is a nonempty set of probability measures containing $\mathbb{P}$. For example, it is conceivable that customers, for whom the possible loss exceeds the limit $\frac{1}{\gamma}$, might exhibit additional risk factors that increase their probability of default.

At first glance, the following example is reminiscent of Example 6, in which the worst-case PD was considered. As an additional criterion, the risk of $X \in \mathrm{B_b}$, described by a monetary risk measure, determines how many models are considered in the calculation of the supremum.

Example 8 (Increasing conservatism)**.** Let $\mathcal{P}$ be a set of probability measures on $\mathcal{F}$, $R\colon \mathrm{B_b} \to \mathbb{R}$ be a monetary risk measure, and $\alpha\colon \mathcal{P} \to [0, \infty)$ with $\inf_{\mathbb{Q} \in \mathcal{P}} \alpha(\mathbb{Q}) = 0$. For $X \in \mathrm{B_b}$, let $\varrho(X) := 0$ if $R(X) \le 0$ and

$$\varrho(X) := \sup \big\{ \mathrm{PD}_\mathbb{Q}(X) \,\big|\, \mathbb{Q} \in \mathcal{P},\, \alpha(\mathbb{Q}) \le R(X) \big\} \quad \text{if } R(X) > 0,$$

i.e., the larger the risk associated to $X \in \mathrm{B_b}$, the more models are taken into account, when assessing the default risk. In this case, $\alpha(\mathbb{Q})$ measures the degree of confidence that the model $\mathbb{Q} \in \mathcal{P}$ is the 'correct' model, where $\alpha(\mathbb{Q}) = 0$ corresponds to maximal confidence. The (monetary) risk measure R can be interpreted as the outcome of some internal risk assessment. Since, by assumption, $\inf_{\mathbb{Q} \in \mathcal{P}} \alpha(\mathbb{Q}) = 0$, it follows that

$$\big\{ \mathrm{PD}_\mathbb{Q}(X) \,\big|\, \mathbb{Q} \in \mathcal{P},\, \alpha(\mathbb{Q}) \le R(X) \big\} \ne \emptyset \quad \text{for all } X \in \mathrm{B_b} \text{ with } R(X) > 0.$$

In particular, $\varrho(X) \in [0, 1]$ for all $X \in \mathrm{B_b}$ and $\varrho(m) = 1$ for all $m \in \mathbb{R}$ with $m > 0$. For $X, Y \in \mathrm{B_b}$ with $X \le Y$, it follows that $R(X) \le R(Y)$ and $\mathrm{PD}_\mathbb{Q}(X) \le \mathrm{PD}_\mathbb{Q}(Y)$ for all $\mathbb{Q} \in \mathcal{P}$, so that $\varrho(X) \le \varrho(Y)$.

After some examples have been discussed, we now turn our focus on extensions of a given default risk measure $\varrho\colon C \to [0, 1]$ from the set of existing customers C to the space $\mathrm{B_b}$ of all bounded measurable functions. We start with following definition.

Definition 9. Let $\varrho\colon C \to [0, 1]$ be a default risk measure and $F\colon \mathrm{B_b} \to \mathbb{R}$.

a) We say that F is *monotone* if $F(X) \leq F(Y)$ for all $X, Y \in \mathrm{B_b}$ with $X \leq Y$.
b) We say that F is *compatible* with ϱ if $F(0) = 0$ and, for $X, Y \in C$, $F(X-Y) \leq 0$ implies that $\varrho(X) \leq \varrho(Y)$.

Theorem 10. *Let $\varrho\colon C \to [0,1]$ be a default risk measure and $F\colon \mathrm{B_b} \to \mathbb{R}$ be monotone and compatible with ϱ. Then, $\varrho_F\colon \mathrm{B_b} \to \mathbb{R}$, given by*

$$\varrho_F(X) := \inf\big\{\varrho(X_0)\,\big|\, X_0 \in C,\ F(X - X_0) \leq 0\big\} \quad \textit{for all } X \in \mathrm{B_b},$$

defines a default risk measure on $\mathrm{B_b}$ with $\varrho_F(X) = \varrho(X)$ for all $X \in C$.

The previous theorem shows that any default risk measure ϱ on C can be extended to a default risk measure on $\mathrm{B_b}$. Assume that a financial institution has a set C of existing customers and a default risk measure ϱ, which represents a rating system, assigning to each customer a rating class in the form of a probability of default. If the financial institution now aims to allocate new customers, which are characterized by different or additional risk factors, within the existing rating classes, this is possible with the extension procedure described in Theorem 10. Due to the construction of ϱ_F, no new rating classes are introduced and the monotone function $F\colon \mathrm{B_b} \to \mathbb{R}$ might resemble an internal risk measurement procedure. Here, the compatibility condition, which may not seem very intuitive at first, entails a comparison in the risk assessment between new and existing customers.

A particularly interesting choice for a monotone map $F\colon \mathrm{B_b} \to \mathbb{R}$, as in Theorem 10, is given by the choice $F(X) := \sup X$ for all $X \in \mathrm{B_b}$. This leads to the default risk measure

$$\varrho_{\sup}(X) := \inf\big\{\varrho(X_0)\,\big|\, X_0 \in C,\ X \leq X_0\big\} \quad \text{for all } X \in \mathrm{B_b},$$

which is akin to the idea of superhedging. Note that this choice of F is compatible with *every* default risk measure $\varrho\colon C \to [0,1]$. Furthermore, this choice of an extension leads to the most conservative default risk measure which is consistent with ϱ as the following corollary indicates.

Corollary 11. *Let $\varrho\colon C \to [0,1]$ be a default risk measure. Then, for every default risk measure $\overline{\varrho}\colon \mathrm{B_b} \to [0,1]$ with $\overline{\varrho}(X_0) = \varrho(X_0)$ for all $X_0 \in C$, it holds*

$$\overline{\varrho}(X) \leq \varrho_{\sup}(X) \quad \textit{for all } X \in \mathrm{B_b}.$$

Another interesting example is given by the case, where C consists of only constant functions and F is a monetary risk measure, e.g., the expected value with respect to a probability measure $\mathbb{P}$ on $\mathcal{F}$. This case is discussed in the following example.

Example 12 (Binary default risk measure)**.** Let $R\colon \mathrm{B_b} \to \mathbb{R}$ be a monetary risk measure. For $X \in \mathrm{B_b}$, let

$$\varrho(X) := \inf\big\{\mathbb{1}_{(0,\infty)}(m)\,\big|\, m \in \mathbb{R},\ R(X - m) \leq 0\big\}.$$

Then, for all $X \in \mathrm{B_b}$,

$$\varrho(X) = \begin{cases} 0, & \text{if } R(X) \leq 0, \\ 1, & \text{otherwise.} \end{cases}$$

Since ϱ only takes two values, it is an almost trivial example of a default risk measure, which, however, has some surprising properties in terms of acceptance sets and its related value at risk, cf. Section 3, below.

3. Additional properties and value at risk

Recall that a monetary risk measure $R\colon \mathrm{B_b} \to \mathbb{R}$ is uniquely determined by its acceptance set

$$\mathscr{A}_R := \big\{X \in \mathrm{B_b} \,\big|\, R(X) \leq 0\big\}$$

and, given a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$, the *value at risk* (VaR) at level $\alpha \in (0,1)$ is given by

$$\mathrm{VaR}_{\mathbb{P}}^{\alpha}(X) := \inf\big\{m \in \mathbb{R} \,\big|\, \mathbb{P}(X - m > 0) \leq \alpha\big\}.$$

For a default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$, the condition $\varrho(X) \leq 0$ implies that $X \in \mathrm{B_b}$ exhibits no default risk whatsoever, so that an acceptance set similar to the one of a monetary risk measure is not meaningful for default risk measures. In this context, it is worth noting that default risk is not expressed in monetary units. In particular, default risk measures are not cash additive. Having the PD with respect to a probability measure $\mathbb{P}$ in mind as a prime example for a default risk measure, and looking at its connection to the VaR, a different approach seems more natural. The aim of this section is to formalize this relation between PD and VaR, and transfer it to general default risk measures. In view of Theorem 10, we focus on the case $C = \mathrm{B_b}$ throughout this section, and start with the following definition, which is central for the subsequent discussion.

Definition 13. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure and $\alpha \in (0,1)$. For $X \in \mathrm{B_b}$, let

$$\mathrm{VaR}_{\varrho}^{\alpha}(X) := \inf\big\{m \in \mathbb{R} \,\big|\, \varrho(X - m) \leq \alpha\big\}.$$

Then, $\mathrm{VaR}_{\varrho}^{\alpha}$ is called the *ϱ-value at risk* at level α.

For default risk measures, we consider the following additional properties.

Definition 14. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure.

a) We say that ϱ is *quasi-convex* if

$$\varrho(\lambda X + (1-\lambda)Y) \leq \varrho(X) \vee \varrho(Y) \quad \text{for all } \lambda \in [0,1] \text{ and } X, Y \in \mathrm{B_b}.$$

b) We say that ϱ is *scaling invariant* if

$$\varrho(X) = \varrho(\lambda X) \quad \text{for all } \lambda > 0 \text{ and } X \in \mathrm{B_b}. \tag{2}$$

We point out that quasi-convexity and scaling invariance, i.e., positive homogeneity of degree zero, are well-known properties in theoretical economics and mathematical finance. In the context of credit risk, scaling invariance has appeared in [26] in a different setting under the name *nominal-invariance*. In the economic literature, positive homogeneity of degree zero or, simply, *homogeneity* is also known under the name *scale independence*, see for example [38, 39].

Proposition 15. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure.*

a) For all $\alpha \in (0,1)$, $\mathrm{VaR}_{\varrho}^{\alpha}$ is a monetary risk measure, cf. Definition 1.

b) ϱ can be recovered from the family $(\mathrm{VaR}_{\varrho}^{\alpha})_{\alpha \in (0,1)}$ via

$$\varrho(X) = \inf\Big(\big\{\alpha \in (0,1) \,\big|\, \mathrm{VaR}_{\varrho}^{\alpha}(X) \leq 0\big\} \cup \{1\}\Big). \tag{3}$$

c) ϱ is scaling invariant if and only if $\mathrm{VaR}_{\varrho}^{\alpha}$ is positively homogeneous for all $\alpha \in (0,1)$.

d) ϱ is quasi-convex if and only if $\mathrm{VaR}_{\varrho}^{\alpha}$ is convex for all $\alpha \in (0,1)$.

Remark 16. Although a default risk measure $\varrho\colon \mathrm{B_b} \to \mathbb{R}$ is never convex, there are default risk measures that are quasi-convex. An example for such a default risk measure is given by

$$\varrho(X) := \begin{cases} 0, & \text{if } R(X) \leq 0, \\ 1, & \text{else,} \end{cases} \quad \text{for all } X \in \mathrm{B_b},$$

where $R\colon \mathrm{B_b} \to \mathbb{R}$ is a monetary risk measure in the sense of Definition 1. Then, for $X \in \mathrm{B_b}$ and $m \in \mathbb{R}$, $\varrho(X - m) \leq 0$ if and only if $R(X) \leq m$. Hence,

$$\mathrm{VaR}^\alpha_\varrho(X) = R(X) \quad \text{for all } \alpha \in (0,1).$$

By Proposition 15 d), ϱ is quasi-convex if and only if R is convex. Moreover, by Proposition 15 c), ϱ is scaling invariant if and only if R is positively homogeneous.

Thinking of the family $(\mathrm{VaR}^\alpha_\varrho)_{\alpha\in(0,1)}$ for a default risk measure ϱ, leads to the following definition.

Definition 17. A *generalized quantile function* is a family $(R^\alpha)_{\alpha\in(0,1)}$, where

(i) $R^\alpha\colon \mathrm{B_b} \to \mathbb{R}$ is a monetary risk measure for each $\alpha \in (0,1)$, cf. Definition 1,
(ii) $R^\beta(X) \leq R^\alpha(X)$ for all $X \in \mathrm{B_b}$ and $\alpha, \beta \in (0,1)$ with $\alpha \leq \beta$,
(iii) for all $\alpha \in (0,1)$ and $X \in \mathrm{B_b}$,

$$R^\alpha(X) = \sup_{\beta\in(\alpha,1)} R^\beta(X).$$

Remark 18. For each default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$, the family $\big(\mathrm{VaR}^\alpha_\varrho\big)_{\alpha\in(0,1)}$ is a generalized quantile function. The properties (i) and (ii) are immediate consequences of Proposition 15 a) and the definition of the family $\big(\mathrm{VaR}^\alpha_\varrho\big)_{\alpha\in(0,1)}$, respectively. We prove (iii) by contradiction. To that end, let $\alpha \in (0,1)$ and $X \in \mathrm{B_b}$, and suppose that

$$\mathrm{VaR}^\alpha_\varrho(X) > \sup_{\beta\in(\alpha,1)} \mathrm{VaR}^\beta_\varrho(X).$$

Then, there exists some $m \in \mathbb{R}$ with

$$\mathrm{VaR}^\alpha_\varrho(X) > m > \mathrm{VaR}^\beta_\varrho(X) \quad \text{for all } \beta \in (\alpha,1),$$

which implies that $\alpha < \varrho(X-m) \leq \beta$ for all $\beta \in (\alpha,1)$, and thus leads to contradiction.

The following theorem together with Proposition 15 b) and the previous remark shows that there is a one-to-one relation between default risk measures and generalized quantile functions.

Theorem 19. *Let $(R^\alpha)_{\alpha\in(0,1)}$ be a generalized quantile function and*

$$\varrho(X) := \inf\Big(\big\{\alpha \in (0,1) \,\big|\, R^\alpha(X) \leq 0\big\} \cup \{1\}\Big).$$

Then, $\varrho\colon \mathrm{B_b} \to [0,1]$ is a default risk measure and $R^\alpha = \mathrm{VaR}^\alpha_\varrho$ for all $\alpha \in (0,1)$.

Proposition 20. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure and $\mathcal{P}$ be a nonempty set of probability measures. Then, the following are equivalent:*

(i) For all $X \in \mathrm{B_b}$,

$$\varrho(X) = \sup_{\mathbb{P}\in\mathcal{P}} \mathbb{P}(X > 0).$$

(ii) For all $X \in \mathrm{B_b}$ and all $\alpha \in (0,1)$,

$$\mathrm{VaR}^\alpha_\varrho(X) = \sup_{\mathbb{P}\in\mathcal{P}} \mathrm{VaR}^\alpha_\mathbb{P}(X). \tag{4}$$

Remark 21.

a) In [43, Proposition 4.6], it is shown that the worst-case value at risk, given by (4), is *comonotonically additive*. More generally, [14, Corollary 1] implies that $\mathrm{VaR}_\varrho^\alpha$ is comonotonically additive if $\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$ for all $X \in \mathrm{B_b}$, i.e., ϱ is given by a capacity $c\colon \mathcal{F} \to [0,1]$, cf. Section 5 below. In fact, [14, Corollary 1] states that, for every $\alpha \in (0,1)$, there exists a capacity $c_\alpha\colon \mathcal{F} \to [0,1]$ such that $\mathrm{VaR}_\varrho^\alpha$ is given by the Choquet integral w.r.t. c_α, i.e.,

$$\mathrm{VaR}_\varrho^\alpha(X) = \int X \,\mathrm{d}c_\alpha \quad \text{for all } X \in \mathrm{B_b}, \tag{5}$$

see Section 5 below for a definition of the Choquet integral. Clearly, if $\mathrm{VaR}_\varrho^\alpha$ satisfies (5), it is positively homogeneous for all $\alpha \in (0,1)$, so that ϱ is scaling invariant by Proposition 15 c). The natural question arises if a representation of $\mathrm{VaR}_\varrho^\alpha$ in terms of (5) for all $\alpha \in (0,1)$ already implies that ϱ is given by a capacity as well. We provide a negative answer to this question in part b).

b) Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$, and assume that there exists some event $A \in \mathcal{F}$ with $p := \mathbb{P}(A) \in (0,1)$. Choosing, in the situation of Remark 16, $R = \mathbb{E}_\mathbb{P}$ as the expected value under $\mathbb{P}$, it follows that $\mathrm{VaR}_\varrho^\alpha = \mathbb{E}_\mathbb{P}$, so that $\mathrm{VaR}_\varrho^\alpha$ satisfies (5) for all $\alpha \in (0,1)$. However, for $X := \frac{1}{p}\mathbb{1}_A - \frac{1}{1-p}\mathbb{1}_{A^c}$, it follows that

$$\mathbb{E}_\mathbb{P}(X) = 0 \neq 1 = \mathbb{E}_\mathbb{P}(X^+),$$

so that $\varrho(X) = 0 \neq 1 = \varrho(X^+)$. Hence, by Theorem 31 below, ϱ is *not* given by a capacity.

Proposition 22. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure, $\mathbb{P}$ a probability measure, and $T\colon [0,1] \to [0,1]$ a nondecreasing and lower semicontinuous function with $T(0) = 0$ and $T(1) = 1$. Then, the following are equivalent:*

(i) For all $X \in \mathrm{B_b}$,

$$\varrho(X) = T\big(\mathbb{P}(X > 0)\big).$$

(ii) For all $X \in \mathrm{B_b}$ and all $\alpha \in (0,1)$,

$$\mathrm{VaR}_\varrho^\alpha(X) = \mathrm{VaR}_\mathbb{P}^{T^{-1}(\alpha)}(X), \tag{6}$$

where $T^{-1}(a) := \inf\{b \in (0,1) \,|\, T(b) > a\}$ for all $a \in (0,1)$.

We conclude this section with two examples for generalized quantile functions arising in the context of regulatory risk measures and their resulting default risk measures.

Example 23. Let $\mathbb{P}\colon \mathcal{F} \to [0,1]$ be a probability measure.

a) We consider the *expected shortfall*

$$\mathrm{ES}_\mathbb{P}^\alpha(X) := \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_\mathbb{P}^\lambda(X)\,\mathrm{d}\lambda \quad \text{of } X \in \mathrm{B_b} \text{ at level } \alpha \in (0,1).$$

Then, $\mathrm{ES}_\mathbb{P}^\alpha\colon \mathrm{B_b} \to \mathbb{R}$ defines a monetary risk measure for each $\alpha \in (0,1)$, and $\mathrm{ES}_\mathbb{P}^\beta(X) \leq \mathrm{ES}_\mathbb{P}^\alpha(X)$ for all $X \in \mathrm{B_b}$ and $\alpha, \beta \in (0,1)$ with $\alpha \leq \beta$. Moreover, by the monotone convergence theorem, Property (iii) in the definition of a generalized quantile function is satisfied. Then, the default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$ related to the generalized quantile function $(\mathrm{ES}_\mathbb{P}^\alpha)_{\alpha\in(0,1)}$ is given by

$$\varrho(X) := \inf\Big(\big\{\alpha \in (0,1) \,\big|\, \mathrm{ES}_\mathbb{P}^\alpha(X) \leq 0\big\} \cup \{1\}\Big) \quad \text{for all } X \in \mathrm{B_b}.$$

Let $X \in \mathrm{B_b}$ with $X \neq 0$. First, observe that

$$\varrho(X) = 1 \quad \text{if and only if} \quad \mathbb{E}_{\mathbb{P}}(X) = \int_0^1 \mathrm{VaR}_{\mathbb{P}}^{\lambda}(X)\,\mathrm{d}\lambda \geq 0.$$

If $\mathbb{E}(X) < 0$, then the continuity of the map $[0,1] \to \mathbb{R},\ \alpha \mapsto \int_0^\alpha \mathrm{VaR}_{\mathbb{P}}^{\lambda}(X)\,\mathrm{d}\lambda$ implies that

$$\begin{aligned}\varrho(X) &= \inf\left\{\alpha \in (0,1)\,\middle|\, \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_{\mathbb{P}}^{\lambda}(X)\,\mathrm{d}\lambda \leq 0\right\}\\ &= \inf\left\{\alpha \in [0,1)\,\middle|\, \int_0^\alpha \mathrm{VaR}_{\mathbb{P}}^{\lambda}(X)\,\mathrm{d}\lambda = 0\right\}.\end{aligned}$$

In particular, by Theorem 19 and Proposition 20 with $\mathcal{P} = \{\mathbb{P}\}$,

$$\varrho(X) \geq \mathbb{P}(X > 0) = \mathrm{PD}_{\mathbb{P}}(X) \quad \text{for all } X \in \mathrm{B_b},$$

i.e., ϱ is a more conservative default risk measure than $\mathrm{PD}_{\mathbb{P}}$.

b) For $X \in \mathrm{B_b}$ and $\alpha \in (0,1)$, we consider the α-*expectile*, given by

$$\mathrm{Exp}_{\mathbb{P}}^{\alpha}(X) := \inf\big\{m \in \mathbb{R}\,\big|\, (1-\alpha)\mathbb{E}_{\mathbb{P}}[(X-m)^-] - \alpha\mathbb{E}_{\mathbb{P}}[(X-m)^+] \leq 0\big\}.$$

Then, $\mathrm{Exp}_{\mathbb{P}}^{\alpha}\colon \mathrm{B_b} \to \mathbb{R}$ is a monetary risk measure for all $\alpha \in (0,1)$. By the monotone convergence theorem, for all $X \in \mathrm{B_b}$,

$$\alpha\mathbb{E}_{\mathbb{P}}\big[\big(X - \mathrm{Exp}_{\mathbb{P}}^{\alpha}(X)\big)^+\big] = (1-\alpha)\mathbb{E}_{\mathbb{P}}\big[\big(X - \mathrm{Exp}_{\mathbb{P}}^{\alpha}(X)\big)^-\big].$$

In particular, $\mathrm{Exp}_{\mathbb{P}}^{1/2}(X) = \mathbb{E}_{\mathbb{P}}(X)$. Since the map $(0,1) \to (0,\infty),\ \alpha \mapsto \frac{1-\alpha}{\alpha}$ is decreasing and continuous, the properties (ii) and (iii) of a generalized quantile function follow. Let $X \in \mathrm{B_b}$ and

$$\varrho(X) := \inf\Big(\big\{\alpha \in (0,1)\,\big|\, \mathrm{Exp}_{\mathbb{P}}^{\alpha}(X) \leq 0\big\} \cup \{1\}\Big).$$

Then, $\varrho(X) = 1$ if and only if $X \geq 0$ $\mathbb{P}$-a.s and $\mathbb{P}(X > 0) > 0$. Moreover, $\varrho(X) = 0$ if and only if $X \leq 0$ $\mathbb{P}$-a.s. In all other cases, i.e., if $\varrho(X) \in (0,1)$, then $\varrho(X)$ is given by the unique solution $\varrho(X) := \alpha^* \in (0,1)$ to the equation

$$\alpha^*\mathbb{E}_{\mathbb{P}}(X^+) = (1-\alpha^*)\mathbb{E}_{\mathbb{P}}(X^-).$$

4. Representation via capacities

After discussing first characteristics, generalizations of the VaR, and different examples for default risk measures, we now aim for equivalent conditions for default risk measures to admit a representation of the form

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) \quad \text{for all } X \in \mathrm{B_b}. \tag{7}$$

This representation allows to describe the default risk measure ϱ via its related capacity, given by $c(A) := \varrho(\mathbb{1}_A)$ for $A \in \mathcal{F}$. This description allows to derive explicit representations of ϱ as worst-case PDs and distorted PDs in Section 5 and Section 6, respectively.

In the following, we discuss a slightly weaker notion of scaling invariance for default risk measures.

Definition 24. We say that a default risk measure $\varrho\colon \mathrm{B_b} \to \mathbb{R}$ is *default scaling invariant* if

$$\varrho(X) = \varrho(\lambda X) \quad \text{for all } \lambda > 0 \text{ and } X \in \mathrm{B_b} \text{ with } X \geq 0.$$

Default risk measures that are not default scaling invariant can depend, for example, on the maximum possible default. This applies, for example to the default risk measure in Example 7, where a warning signal γ must be exceeded by $\sup X$ in order to switch to a more conservative default risk measure. Another example for a default risk measure that is not default scaling invariant is given in Example 8.

From an economic point of view, default scaling invariance means that the default risk does not change if all potential losses, i.e., $X(\omega) > 0$ are multiplied with a positive constant. Hence, for a given customer $X \geq 0$, we are not interested how high this customer defaults, which would be the case in Example 7, but rather if they default or not. All scenarios in which the customer cannot pay a due at the end of the observation period are equivalent to the scenarios in which the due multiplied by $\lambda > 0$ cannot be paid. Certainly, therefore, it is a necessary condition in order to obtain the desired representation (7).

Remark 25. Before we provide sufficient and necessary conditions for ϱ to have the form (7), we first discuss various properties of ϱ that will appear at a later stage in this section.

a) Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default scaling invariant default risk measure. Then, for all $X \in \mathrm{B_b}$,
$$\varrho(X) \leq \varrho\big(\mathbb{1}_{\{X>0\}}\big), \tag{8}$$
Indeed, for $X \in \mathrm{B_b}$ with $X \leq 0$, $\varrho(X) \leq \varrho(0) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$. On the other hand, if $X \in \mathrm{B_b}$ with $\sup X > 0$, the default scaling invariance of ϱ yields that
$$\varrho(X) \leq \varrho(X^+) \leq \varrho\big((\sup X)\mathbb{1}_{\{X>0\}}\big) = \varrho\big(\mathbb{1}_{\{X>0\}}\big).$$
Condition (8) states that, under the mild assumption of default scaling invariance, the most conservative choice of a default risk measure ϱ is a representation via a capacity.

b) To see that scaling invariance, in particular, default scaling invariance is not enough to come up with the representation $\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$ for all $X \in \mathrm{B_b}$, we again pick up Example 12, where a default risk measure was defined by means of a monetary risk measure R. Let $\mathbb{P}$ be a probability measure on $\mathcal{F}$, and consider the case $R(X) = \mathbb{E}_{\mathbb{P}}(X)$ for all $X \in \mathrm{B_b}$, where $\mathbb{E}_{\mathbb{P}}(\,\cdot\,)$ denotes the mean value under $\mathbb{P}$. Then, the default risk measure in Example 12 is scaling invariant by Remark 16. However, the default risk measure very much depends on both the positive and negative values of $X \in \mathrm{B_b}$ as soon as there exists an event $A \in \mathcal{F}$ with $\mathbb{P}(A) \in (0,1)$. In fact, let $A \in \mathcal{F}$ with $p := \mathbb{P}(A) \in (0,1)$ and define
$$X := \mathbb{1}_A - \frac{2p}{1-p}\mathbb{1}_{A^c}.$$
Then,
$$\mathbb{E}_{\mathbb{P}}(X) = \mathbb{P}(A) - \frac{1+p}{1-p} \cdot \mathbb{P}(A^c) = -1 < 0,$$
so that $\varrho(X) = 0$. On the other hand, $\mathbb{P}(X > 0) = \mathbb{P}(A) = p > 0$, which implies that $\varrho\big(\mathbb{1}_{\{X>0\}}\big) = 1$. Nevertheless, in Proposition 26 below, we will see that, up to certain continuity requirements, default scaling invariance implies the equality $\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$ for all $X \in \mathrm{B_b}$ with $X \geq 0$. Later we seek for additional properties of ϱ such as $\varrho(X) = \varrho(X^+)$ for all $X \in \mathrm{B_b}$.

c) In order to explain why additional continuity assumptions are needed in Proposition 26 below, we regard, for a given probability measure $\mathbb{P}$ on $\mathcal{F}$, the default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$, given by

$$\varrho(X) := \begin{cases} 1, & \text{if } \mathbb{P}(\{X > \varepsilon\}) = 1 \text{ for some } \varepsilon > 0, \\ \frac{\mathbb{P}(\{X>0\})}{2}, & \text{otherwise.} \end{cases}$$

As such ϱ is default scaling invariant and does not depend on the part, where $X \in \mathrm{B_b}$ is less or equal to 0. However, if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and $X \in \mathrm{B_b}$ is uniformly distributed on $[0,1]$ under $\mathbb{P}$, then

$$\mathbb{P}\big(X \in (0,\varepsilon)\big) = \varepsilon > 0 \quad \text{for all } \varepsilon > 0,$$

so that

$$\varrho(X) = \frac{1}{2} \neq 1 = \varrho\big(\mathbb{1}_{\{X>0\}}\big).$$

The following proposition clarifies this elementary example.

Proposition 26. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default scaling invariant default risk measure.*

a) For all $X \in \mathrm{B_b}$ with $X \geq 0$,

$$\varrho(X) \leq \varrho\big(\mathbb{1}_{\{X>0\}}\big) \leq \inf_{\varepsilon>0} \varrho\big(X + \varepsilon \mathbb{1}_{\{X>0\}}\big). \tag{9}$$

b) For all $X \in \mathrm{B_b}$ with $X \geq 0$,

$$\sup_{\varepsilon>0} \varrho\big(\mathbb{1}_{\{X>\varepsilon\}}\big) \leq \varrho(X) \leq \varrho\big(\mathbb{1}_{\{X>0\}}\big). \tag{10}$$

The previous proposition shows that default scaling invariance together with $\varrho(X) = \varrho(X^+)$, for all $X \in \mathrm{B_b}$, implies that the default risk measure can be represented via a capacity except for some additional continuity properties. The inequality in part a) is quite mild from a theoretical perspective, because reducing a customer's cash flows by an amount of money $\varepsilon > 0$ in scenarios, where the customer is defaulting anyways, has no impact on the default scenarios. In contrast to this, the continuity assumption

$$\sup_{\varepsilon>0} \varrho\big(\mathbb{1}_{\{X>\varepsilon\}}\big) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) \quad \text{for all } X \in \mathrm{B_b}$$

is not for free since, in mathematical terms, it requires some mild form of continuity from below for the related capacity. From an economic point of view, however, it is quite intuitive since $\varepsilon > 0$ can be chosen arbitrarily small, e.g., in such a way that it falls below the smallest possible amount of money. A theoretical increase in capital by 10^{-10} euros should not change the default risk of the respective customer. This motivates to introduce the following definition.

Definition 27. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. We say that ϱ is *illiquidity invariant*, if, for all $m > 0$ and $X \in \mathrm{B_b}$ with $X \geq 0$,

$$\varrho(X) = \varrho\big(X + m\mathbb{1}_{\{X>0\}}\big).$$

In other words, *illiquidity invariance* means that customers that are in default or, loosely speaking, illiquid remain illiquid if their debt due increases or their capital decreases in the sense that their default risk does not change. The following proposition establishes a first connection between illiquidity invariance, default scaling invariance, and a representation via indicator functions.

Proposition 28. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. Then, the following statements are equivalent.*

(i) ϱ *is default scaling invariant and*

$$\varrho(X) = \inf_{\varepsilon>0} \varrho(X + \varepsilon \mathbb{1}_{\{X>0\}}) \quad \textit{for all } X \in \mathrm{B_b} \textit{ with } X \geq 0.$$

(ii) ϱ *is illiquidity invariant.*

(iii) *For all* $X \in \mathrm{B_b}$ *with* $X \geq 0$,

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big).$$

In analogy to Definition 27, we speak of *liquidity invariance*, when the default risk remains the same if capital is added in scenarios where the customer is able to make all payments due.

Definition 29. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. We say that ϱ is *liquidity invariant*, if, for all $m > 0$ and $X \in \mathrm{B_b}$ with $X \geq 0$,

$$\varrho(X) = \varrho\big(X - m\mathbb{1}_{\{X=0\}}\big).$$

Lemma 30. *Let* $\varrho\colon \mathrm{B_b} \to [0,1]$ *be a default risk measure. Then, the following conditions are equivalent.*

(i) $\varrho(X) = \varrho(X^+)$ *for all* $X \in \mathrm{B_b}$.

(ii) $\varrho(X) = \varrho\big(X - m\mathbb{1}_{\{X\leq 0\}}\big)$ *for all* $m > 0$ *and* $X \in \mathrm{B_b}$.

(iii) ϱ *is liquidity invariant.*

In the context of risk functionals, Property (i) in Lemma 30 (with a different sign convention) is known under the names *surplus invariance*, cf. [24], *loss dependence*, cf. [15], and *excess invariance*, cf. [40]. A combination of Proposition 28 and Lemma 30 leads to the following theorem, which is the main result of this section.

Theorem 31. *Let* $\varrho\colon \mathrm{B_b} \to [0,1]$ *be a default risk measure. Then, the following conditions are equivalent.*

(i) *For all* $X \in \mathrm{B_b}$,

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big).$$

(ii) ϱ *is default scaling invariant with*

$$\varrho(X) = \varrho(X^+) = \inf_{\varepsilon>0} \varrho\big(X + \varepsilon \mathbb{1}_{\{X>0\}}\big) \quad \textit{for all } X \in \mathrm{B_b}.$$

(iii) ϱ *is liquidity invariant and illiquidity invariant.*

To get a link to Section 5, where the focus lies on the related capacities and their Choquet integrals, we introduce the notion of *submodularity* for default risk measures. In theoretical economics, submodularity is a classical property, which is closely related to substitute goods, cf. [42].

Definition 32. Let C be a sublattice of $\mathrm{B_b}$, which contains all constant functions. A default risk measure $\varrho\colon C \to [0,1]$ is called *submodular* if

$$\varrho(X \wedge Y) + \varrho(X \vee Y) \leq \varrho(X) + \varrho(Y) \quad \text{for all } X, Y \in C.$$

The natural question arises, whether submodularity is a sensible notion for default risk measures. For classical PDs, we have equality, i.e.,

$$\mathrm{PD}_{\mathbb{P}}(X \wedge Y) + \mathrm{PD}_{\mathbb{P}}(X \vee Y) = \mathrm{PD}_{\mathbb{P}}(X) + \mathrm{PD}_{\mathbb{P}}(Y) \quad \text{for all } X, Y \in \mathrm{B_b}.$$

Moreover, submodularity is given for certain classes of distorted PDs, for instance, if the distortion function is concave, see Section 6.

Remark 33. Let C be a sublattice of $\mathrm{B_b}$, which contains all constant functions, and $\varrho\colon C \to [0,1]$ be a submodular default risk measure. Then,

$$\varrho(X) = \varrho(X^+) \quad \text{for all } X \in C.$$

In fact, let $X \in C$. The submodularity of ϱ together with $0 \in C$ implies that

$$\varrho(X^+) + \varrho(-X^-) \leq \varrho(X).$$

Since $-X^- \leq 0$, it follows that $\varrho(-X^-) = 0$ by the defining properties of a default risk measure. Therefore,

$$\varrho(X) \leq \varrho(X^+) = \varrho(X^+) + \varrho(-X^-) \leq \varrho(X).$$

As a consequence of Theorem 31 and Remark 33, we obtain the following corollary.

Corollary 34. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a submodular default risk measure. Then, the following statements are equivalent.*

(i) ϱ is illiquidity invariant.
(ii) For all $X \in \mathrm{B_b}$,

$$\varrho(X) = \varrho\left(\mathbb{1}_{\{X>0\}}\right).$$

5. Choquet integrals and robust representations

In Section 4, we discussed equivalent conditions for a default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$ in order to be the form (7). Building on this representation, the aim of this section is to connect default risk measures with Choquet integrals and monetary risk measures and use this connection to attain a representation for ϱ via probability measures. Recall that a *capacity* is a map $c\colon \mathcal{F} \to [0,1]$ with $c(\emptyset) = 0$, $c(\Omega) = 1$, and $c(A) \leq c(B)$ for all $A, B \in \mathcal{F}$ with $A \subset B$. We start with the following observation.

Remark 35. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. Then, we can define a capacity $c\colon \mathcal{F} \to [0,1]$ by

$$c(A) := \varrho\left(\mathbb{1}_A\right) \quad \text{for all } A \in \mathcal{F}. \tag{11}$$

By definition of a default risk measure,

$$c(\emptyset) = c(\mathbb{1}_\emptyset) = \varrho(0) = 0 \quad \text{and} \quad c(\Omega) = c(\mathbb{1}_\Omega) = \varrho(1) = 1.$$

Moreover, for all $A, B \in \mathcal{F}$ with $A \subset B$,

$$c(A) = \varrho\left(\mathbb{1}_A\right) \leq \varrho\left(\mathbb{1}_B\right) = c(B)$$

due to the monotonicity of ϱ. For $X \in \mathrm{B_b}$, the *Choquet integral* with respect to c is defined as

$$\int X \,\mathrm{d}c := \int_{-\infty}^0 \big(c(\{X > s\}) - 1\big)\,\mathrm{d}s + \int_0^\infty c(\{X > s\})\,\mathrm{d}s.$$

Although the Choquet integral is, in general, not a linear functional, it defines a monetary risk measure $R\colon \mathrm{B_b} \to \mathbb{R}$ via

$$R(X) := \int X\,\mathrm{d}c \quad \text{for all } X \in \mathrm{B_b}.$$

By definition of the Choquet integral, the monetary risk measure R is positively homogeneous, i.e., $R(\lambda X) = \lambda R(X)$ for all $X \in \mathrm{B_b}$ and $\lambda > 0$. If ϱ satisfies (7), then

$$\varrho(X) = R\big(\mathbb{1}_{\{X>0\}}\big) \quad \text{for all } X \in \mathrm{B_b},$$

which leads to the following proposition.

Proposition 36. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. Then, the following two statements are equivalent.*

(i) For all $X \in \mathrm{B_b}$,

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big).$$

(ii) There exists a positively homogeneous monetary risk measure $R\colon \mathrm{B_b} \to \mathbb{R}$, cf. Definition 1, with

$$\varrho(X) = R\big(\mathbb{1}_{\{X>0\}}\big) \quad \textit{for all } X \in \mathrm{B_b}.$$

Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure, which satisfies (7). Then, ϱ is submodular if and only if the related capacity $c\colon \mathcal{F} \to [0,1]$, given by (11), is *2-alternating*, i.e.,

$$c(A \cup B) + c(A \cap B) \le c(A) + c(B) \quad \text{for all } A, B \in \mathcal{F}.$$

A well-known fact is that a capacity is 2-alternating if and only if the related Choquet integral defines a coherent risk measure. We recall that a monetary risk measure $R\colon \mathrm{B_b} \to \mathbb{R}$ is coherent if and only if there exists a nonempty set $\mathcal{P}$ of finitely additive probability measures such that

$$R(X) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}(X) \quad \text{for all } X \in \mathrm{B_b},$$

cf. [22]. In this case, we obtain the robust representation

$$\varrho(X) = \int \mathbb{1}_{\{X>0\}}\,\mathrm{d}c = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{E}_{\mathbb{Q}}(\mathbb{1}_{\{X>0\}}) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(X>0) \quad \text{for all } X \in \mathrm{B_b}. \tag{12}$$

We now investigate additional continuity properties for ϱ that guarantee a representation via countably additive probability measures. In the sequel, for a sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ and $X \in \mathrm{B_b}$, we write $X_n \nearrow X$ or $X_n \searrow X$ as $n \to \infty$ if $X_n \le X_{n+1}$ or $X_n \ge X_{n+1}$ for all $n \in \mathbb{N}$ and $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ for all $\omega \in \Omega$, respectively.

Proposition 37. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure with $\varrho(X) = \varrho\left(\mathbb{1}_{\{X>0\}}\right)$ for all $X \in \mathrm{B_b}$ and $c(A) := \varrho(\mathbb{1}_A)$ for all $A \in \mathcal{F}$. Then, the following statements are equivalent.*

(i) ϱ is continuous from below, i.e., for every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \nearrow X \in \mathrm{B_b}$ as $n \to \infty$,

$$\varrho(X) = \lim_{n\to\infty} \varrho\left(X_n\right).$$

(ii) For every sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ with $A_n \subset A_{n+1}$ for all $n \in \mathbb{N}$

$$c\bigg(\bigcup_{n\in\mathbb{N}} A_n\bigg) = \lim_{n\to\infty} c(A_n).$$

(iii) For every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \nearrow X \in \mathrm{B_b}$ as $n \to \infty$,

$$\int X\,\mathrm{d}c = \lim_{n\to\infty} \int X_n\,\mathrm{d}c.$$

Unfortunately, even if ϱ is submodular, continuity from below of ϱ is, in general, not a sufficient condition in order to guarantee a representation via countably additive probability measures on $\mathrm{B_b}$ as the following example shows.

Example 38. Let $\mathbb{P}\colon \mathcal{B} \to [0,1]$ be the Lebesgue measure defined on the Borel σ-algebra $\mathcal{B}$ of the closed interval $\Omega := [0,1]$ and $\mathcal{F}$ be the power set of Ω. Let $\mathcal{L}$ denote the space of all bounded $\mathcal{B}$-measurable functions $\Omega \to \mathbb{R}$ and

$$R(X) := \inf\{\mathbb{E}_{\mathbb{P}}(X_0) |\, X_0 \in \mathcal{L},\, X_0 \geq X\}.$$

Then, by [19, Proposition 2.2 and Lemma 3.6], R is a coherent risk measure, which is continuous from below and the maximal extension of $\mathbb{E}_{\mathbb{P}}$ to $\mathrm{B_b}$. Consider the capacity $c\colon \mathcal{F} \to [0,1]$, given by

$$c(A) := R(\mathbb{1}_A) \quad \text{for all } A \in \mathcal{F},$$

and $\varrho(X) := c(\{X > 0\})$ for all $X \in \mathrm{B_b}$. Let $\mathbb{Q}\colon \mathcal{F} \to [0,1]$ be a finitely additive probability with

$$\mathbb{E}_{\mathbb{Q}}(X) \leq R(X) \quad \text{for all } X \in \mathrm{B_b}. \tag{13}$$

Then, for all $X \in \mathrm{B_b}$ with $X \geq 0$,

$$\int X \,\mathrm{d}c = \int_0^\infty c(\{X > s\})\,\mathrm{d}s \geq \int_0^\infty \mathbb{Q}(\{X > s\})\,\mathrm{d}s = \mathbb{E}_{\mathbb{Q}}(X).$$

Moreover, for $X \in \mathcal{L}$ with $X \geq 0$,

$$\int X \,\mathrm{d}c = \int_0^\infty c(\{X > s\})\,\mathrm{d}s = \int_0^\infty \mathbb{P}(\{X > s\})\,\mathrm{d}s = \mathbb{E}_{\mathbb{P}}(X).$$

Hence, by the maximality of R, it follows that $\int X\,\mathrm{d}c \leq R(X)$ for all $X \in \mathrm{B_b}$. We have therefore shown that $R(X) = \int X\,\mathrm{d}c$ for all $X \in \mathrm{B_b}$. In particular, ϱ is submodular. On the other hand, assuming the continuum hypothesis, by [6, Satz 1C], there exists not a single countably additive probability measure $\mathbb{Q}\colon \mathcal{F} \to [0,1]$ with (13). In fact, with a similar construction, using a probability measure which is different from the Lebesgue measure, one can avoid invoking the continuum hypothesis, see [19, Example 3.7] for further details.

If, however, ϱ is a distorted PD, which is submodular, continuity from below is a sufficient condition in order to allow for a representation in terms of countably additive probability measures as the following remark discusses.

Remark 39. Let $\mathbb{P}$ be a probability measure of $\mathcal{F}$ and $T\colon [0,1] \to [0,1]$ be a nondecreasing function with $T(0) = 0$ and $T(1) = 1$. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be given by

$$\varrho(X) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \text{for all } X \in \mathrm{B_b},$$

and assume that ϱ is submodular, so that $c\colon \mathcal{F} \to [0,1],\ A \mapsto \varrho(\mathbb{1}_A)$ is a 2-alternating capacity.

a) Assume that T is left-continuous or, equivalently, lower semicontinuous, i.e.,

$$T(p) = \sup_{q \in (0,p)} T(q) \quad \text{for all } p \in (0,1].$$

Then, by Proposition 37, the Choquet integral w.r.t. c is a coherent and law-invariant risk measure, which is continuous from below. Hence, by [22, Theorem 4.33], there exists a set of countably additive and (w.r.t. $\mathbb{P}$) absolutely continuous probability measures $\mathcal{P}$ on $\mathcal{F}$ with

$$\varrho(X) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathrm{PD}_{\mathbb{Q}}(X) \quad \text{for all } X \in \mathrm{B_b}.$$

b) Now, assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless. Then, the Choquet integral w.r.t. c is a coherent and law-invariant risk measure on an atomless probability space. By [27, Theorem 2.1], it follows that the Choquet integral w.r.t. c is continuous from below. Hence, by [22, Theorem 4.33], there exists a set of countably additive and (w.r.t. $\mathbb{P}$)

absolutely continuous probability measures $\mathcal{P}$ on $\mathcal{F}$ with

$$\varrho(X) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathrm{PD}_{\mathbb{Q}}(X) \quad \text{for all } X \in \mathrm{B_b}.$$

If Ω is a Polish space and $\mathcal{F}$ is the Borel σ-algebra, we have the following characterization for general submodular default risk measures on the space $\mathrm{L_b}$ of all bounded lower semicontinuous functions $\Omega \to \mathbb{R}$.

Proposition 40. *Let Ω be a Polish space and $\varrho\colon \mathrm{L_b} \to [0,1]$ be a submodular default risk measure with*

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) \quad \textit{for all } X \in \mathrm{L_b} \textit{ with } X \geq 0. \tag{14}$$

Then, the following statements are equivalent.

(i) ϱ is continuous from below, i.e.,

$$\varrho(X) = \lim_{n\to\infty} \varrho(X_n)$$

for every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{L_b}$ with $X_n \nearrow X \in \mathrm{L_b}$ as $n \to \infty$.

(ii) There exists a nonempty set $\mathcal{P}$ of probability measures on $(\Omega, \mathcal{F})$ such that

$$\varrho(X) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(X > 0) \quad \textit{for all } X \in \mathrm{L_b}.$$

Remark 41. Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure. Then, *continuity from above* in the sense that

$$\varrho(X) = \lim_{n\to\infty} \varrho(X_n)$$

for every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \searrow X \in \mathrm{B_b}$ as $n \to \infty$, is *not* a sensible property. Considering a sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ defined by $X_n = \frac{1}{n}$, it follows that $\varrho(X_n) = 1$, while $\varrho(0) = 0$. However, by (12), it is enough to express the Choquet integral via a set of (countably additive) probability measures. We can therefore weaken the requirement of continuity from above and obtain the following proposition.

Proposition 42. *Let $\varrho\colon \mathrm{B_b} \to [0,1]$ be a default risk measure with $\varrho(X) = \varrho\left(\mathbb{1}_{\{X>0\}}\right)$ for all $X \in \mathrm{B_b}$ and $c(A) := \varrho(\mathbb{1}_A)$ for all $A \in \mathcal{F}$. Then, the following statements are equivalent.*

(i) For every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \searrow 0$ as $n \to \infty$ and all $\varepsilon > 0$,

$$\lim_{n\to\infty} \varrho\left(X_n - \varepsilon\right) = 0.$$

(ii) For every sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ with $A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$ and $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$,

$$\lim_{n\to\infty} \varrho\left(\mathbb{1}_{A_n}\right) = 0.$$

(iii) For every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \searrow 0$ as $n \to \infty$,

$$\lim_{n\to\infty} \int X_n \,\mathrm{d}c = 0.$$

If ϱ is additionally submodular, either of these conditions implies that there exists a nonempty set $\mathcal{P}$ of probability measures on $(\Omega, \mathcal{F})$ with

$$\varrho(X) = \max_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(X > 0) \quad \textit{for all } X \in \mathrm{B_b}.$$

Example 43. Let $\mathbb{P}$ be a probability measure on $\mathcal{F}$ and $T\colon [0,1] \to [0,1]$ be nondecreasing with $T(0) = 0$ and $T(1) = 1$. Then, the default risk measure $\varrho\colon \mathrm{B_b} \to [0,1]$, given by

$$\varrho(X) := T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \text{for all } X \in \mathrm{B_b},$$

satisfies Property (ii) in Proposition 42 if $\inf_{p>0} T(p) = 0$. In case there exists a sequence $(A_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ with $\emptyset \neq A_{n+1} \subset A_n$ for all $n \in \mathbb{N}$ and $\bigcap_{n\in\mathbb{N}} A_n = \emptyset$, this is also a necessary condition for ϱ to satisfy Property (ii) in Proposition 42 using the continuity from above of $\mathbb{P}$.

6. Law-invariant default risk measures and distorted PDs

Let C be a set of customers $C \subset \mathrm{B_b}$, which contains the set of all constant functions, and satisfies $\mathbb{1}_{\{X>0\}} \in C$ for all $X \in C$. In this section, we fix a reference probability measure $\mathbb{P}$ on $\mathcal{F}$, and specialize on *law-invariant* default risk measures $\varrho\colon C \to [0,1]$, i.e., $\varrho(X) = \varrho(Y)$ whenever $X \in C$ and $Y \in C$ have the same distribution under $\mathbb{P}$. In rating systems, customers in the same rating class are considered to be statistically identical in terms of their default behaviour. Hence, when choosing a default risk measure in order to quantify the probability of default including uncertainty, it makes sense to require law-invariance. Since customers are usually divided into rating classes and not every default probability is realized, we also consider the case where $(\Omega, \mathcal{F}, \mathbb{P})$ is *not* atomless. We start with several characterizations of law-invariance, which do not hinge on the standard assumption of an atomless probability space, before we switch to an atomless setting in order to derive finer properties and representations of law-invariant default risk measures.

In the sequel, we say that a function $T\colon [0,1] \to [0,1]$ is a *distortion function* if $T(0) = 0$ and $T(1) = 1$. The following theorem adopts an argument from Wang et al. [44, Proof of Theorem 2], and provides a characterization of distorted PDs.

Theorem 44. *Let $\varrho\colon C \to [0,1]$ be a default risk measure. Then, the following statements are equivalent.*

(i) ϱ is law-invariant and $\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$ for all $X \in C$.
(ii) There exists a distortion function $T\colon [0,1] \to [0,1]$ with

$$\varrho(X) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \textit{for all } X \in C.$$

A natural question that arises is whether the distortion function T in Theorem 44 is nondecreasing or, in other words, if the default risk measure $\varrho\colon C \to [0,1]$ is *consistent* with $\mathbb{P}$, i.e.,

$$\varrho(X) \le \varrho(Y) \quad \text{for all } X, Y \in C \text{ with } \mathrm{PD}_{\mathbb{P}}(X) \le \mathrm{PD}_{\mathbb{P}}(Y).$$

This property is very natural since one would expect the default risk of X to be smaller than the default risk of Y if the PD of X is smaller than the PD of Y. In the situation of Theorem 44, it is, however, possible that the distortion function T is *not* monotone as the following simple example shows.

Example 45. Let $\Omega = \{0,1\}$, $\mathcal{F}$ be the power set, and C consist of all constants and the two functions $X := \mathbb{1}_{\{0\}}$ and $Y := \mathbb{1}_{\{1\}}$. Assume that

$$0 < p := \mathbb{P}(X > 0) < \mathbb{P}(Y > 0) =: q < 1 \text{ and } 0 < \varrho(Y) < \varrho(X) < 1.$$

Then ϱ is a default risk measure, which satisfies Property (i) in Theorem 44 but, for any distortion function $T\colon [0,1] \to [0,1]$ with

$$\varrho(Z) = T\big(\mathrm{PD}_{\mathbb{P}}(Z)\big) \quad \text{for all } Z \in C,$$

it follows that $T(p) = \varrho(X) > \varrho(Y) = T(q)$, so that T is not monotone.

We now aim towards a characterization in terms of a nondecreasing distortion function. The proof of Theorem 10 indicates that the set

$$P := \big\{p \in [0,1] \,\big|\, \exists X \in C\colon \mathbb{P}(X > 0) = p\big\}. \tag{15}$$

plays a fundamental role for the monotonicity of the distortion function.

Definition 46. We say C contains an *ordered subset* if there exists a family $\big(X_p\big)_{p\in P}$ with $\mathrm{PD}_{\mathbb{P}}(X_p) = p$ for all $p \in P$ and $\{X_p > 0\} \subset \{X_q > 0\}$ for all $p, q \in P$ with $p \leq q$.

Clearly, if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and $C = \mathrm{B_b}$, then C contains an ordered subset.

Theorem 47. *Let $\varrho\colon C \to [0,1]$ be a default risk measure, and assume that C contains an ordered subset. Then, the following statements are equivalent.*

(i) ϱ is law-invariant and $\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$ for all $X \in C$.
(ii) There exists a nondecreasing distortion function $T\colon [0,1] \to [0,1]$ with

$$\varrho(X) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \textit{for all } X \in C.$$

Remark 48. Under the assumptions of Theorem 47, including one of the equivalences, the default risk measure ϱ can be extended to a law-invariant default risk measure $\overline{\varrho}\colon \mathrm{B_b} \to [0,1]$ by means of the distortion function T. The extension $\overline{\varrho}$ is given by

$$\overline{\varrho}(X) := T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \text{for all } X \in \mathrm{B_b}.$$

Obviously, $\overline{\varrho}(X) = \varrho(X)$ for all $X \in C$ and $\overline{\varrho}$ is a default risk measure, cf. Example 5. Note that a default risk measure can, in general, not be extended using a nonmonotone distortion function as in Example 45. This can be seen by considering, for example, the set $\Omega = \{0, 1, 2\}$ together with the power set $\mathcal{F}$ and C consisting of $X := \mathbb{1}_{\{0\}}$, $Y := \mathbb{1}_{\{1\}}$ and all constant functions. Let $\mathbb{P}(X > 0) = 0.5$, $\mathbb{P}(Y > 0) = 0.3$, $\varrho(X) = 0.5$, $\varrho(Y) = 0.7$, and $T\colon [0,1] \to [0,1]$ be a distortion function with

$$\varrho(Z) = T\big(\mathrm{PD}_{\mathbb{P}}(Z)\big) \quad \text{for all } Z \in C.$$

Then, for $U := \mathbb{1}_{\{1,2\}}$, it follows that $\mathbb{P}(U > 0) = 0.5 = \mathbb{P}(X > 0)$. Hence,

$$T\big(\mathrm{PD}_{\mathbb{P}}(U)\big) = T(0.5) = \varrho(X) = 0.5 < 0.7 = \varrho(Y)$$

despite the fact that $Y \leq U$.

Remark 49. We recall some well-known facts about distorted probabilities, cf. [22]. For the reader's convenience, we provide short proofs of some of the statements collected in this remark in the Appendix A. In the following, let $T\colon [0,1] \to [0,1]$ be a distortion function.

a) A well-known fact is that the capacity $c\colon \mathcal{F} \to [0,1]$, given by

$$c(A) := T\big(\mathbb{P}(A)\big) \quad \text{for all } A \in \mathcal{F}. \tag{16}$$

is 2-alternating if T is concave. If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, also the converse statement holds, cf. [22, Section 4.6].

b) A capacity $c\colon \mathcal{F} \to [0,1]$ is called *exact* if there exists a set of countably additive probability measures $\mathcal{P}$ on $\mathcal{F}$ with

$$c(A) = \sup_{\mathbb{Q}\in\mathcal{P}} \mathbb{Q}(A) \quad \text{for all } A \in \mathcal{F}.$$

It is well-known that the distorted probability $c\colon \mathcal{F} \to [0,1]$, given by (16), is an exact capacity if, for all $p, q \in (0,1)$ with $p < q$,

$$\frac{T(p)}{p} \geq \frac{T(q) - T(p)}{q - p} \geq \frac{1 - T(q)}{1 - q}. \tag{17}$$

If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, also the converse holds true, cf. [2]. In this case,

$$T(p) = \sup_{\mathbb{Q} \in \mathcal{P}} \int_{1-p}^{1} q_{\mathbb{Q}}(s)\,\mathrm{d}s \quad \text{for all } p \in [0,1], \tag{18}$$

where $q_{\mathbb{Q}}$ denotes the quantile function of the density $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$, see, e.g., [22, Lemma 4.60]. In particular,

$$\mathrm{MoC}(p) = \frac{T(p)}{p} - 1 = \sup_{\mathbb{Q} \in \mathcal{P}} \mathrm{ES}_{\mathbb{P}}^{1-p}\left(\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) - 1 \quad \text{for all } p \in (0,1],$$

where, for $\alpha \in (0,1]$ and $\mathbb{Q} \in \mathcal{P}$, $\mathrm{ES}_{\mathbb{P}}^{\alpha}\left(\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\right) := \frac{1}{1-\alpha}\int_\alpha^1 q_{\mathbb{Q}}(s)\,\mathrm{d}s$ denotes the *expected shortfall* of the density $\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}$ with confidence level α. Note that (17) is, for example, satisfied if T is concave, and observe that (17) implies that the function T is nondecreasing and absolutely continuous as soon as

$$\inf_{p \in (0,1]} T(p) = 0. \tag{19}$$

Recall that a monotone function is a.e. differentiable and that absolute continuity of T implies that

$$T(p) = \int_0^p T'(s)\,\mathrm{d}s \quad \text{for all } p \in [0,1],$$

where T' denotes the weak derivative of T. Moreover, (17) implies that

$$\mathrm{MoC}(q) \leq \mathrm{MoC}(p) \quad \text{for all } p, q \in (0,1] \text{ with } p \leq q.$$

In fact, assume that (17) is satisfied, and let $p, q \in (0,1]$. If $p = q$, the statement is trivial, and if $p < q$,

$$\frac{T(p)\frac{q}{p} - T(p)}{q - p} = \frac{T(p)}{p} \geq \frac{T(q) - T(p)}{q - p},$$

which yields that $T(p)\frac{q}{p} \geq T(q)$.

We conclude this section with the following characterization of the minorants of distorted PDs, which can be found in a similar yet different form in [22, Theorem 4.79].

Proposition 50. *Let $T\colon [0,1] \to [0,1]$ be a distortion function, which satisfies (17) and (19), and $\mathbb{Q}$ be a probability measure on $\mathcal{F}$. If*

$$\int_0^p q_{\mathbb{Q}}(1-s)\,\mathrm{d}s \leq T(p) \quad \textit{for all } p \in [0,1], \tag{20}$$

then

$$\mathrm{PD}_{\mathbb{Q}}(X) \leq T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \quad \textit{for all } X \in \mathrm{B_b}. \tag{21}$$

If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, also the converse holds true.

Remark 51. Consider the situation of Proposition 50. In view of Remark 49, a sufficient condition for (20) and thus (21) to be satisfied is that $q_{\mathbb{Q}}(1-p) \leq T'(p)$ for a.a. $p \in (0,1)$. That this is, however, not a necessary condition, can easily be seen by considering $T(p)=\sqrt{p}$ for all $p\in[0,1]$ and $\mathbb{Q}=\mathbb{P}$. Then, $q_{\mathbb{P}}(1-p)=1$ for all $p\in(0,1)$ and $T'(p)=\frac{1}{2\sqrt{p}}<1$ for all $p\in\left(\frac{1}{4},1\right)$. However,

$$T(p)=\sqrt{p}\geq p=\int_0^p q_{\mathbb{P}}(1-s)\,\mathrm{d}s.$$

7. A Case Study on Capital Requirements for Financial Institutions

In this section, we link our axiomatic study of default risk measures to financial institutions' capital requirements accounting for model uncertainty. According to the guidelines of the European Banking Authority [20], the PDs of a rating system are calibrated to a 'best estimate' level without using systematically conservative input values for the calculation. Those PDs are then used in the risk-oriented group management, among others. In order to establish the connection between model uncertainty and capital requirements, we are guided by Article 179 (f) and Article 180 (e) of the CRR [17] that determine that an appropriate margin of conservatism, reflecting the expected range of estimation errors, must be formed for the 'best estimate' PD of the rating system.

Usually, one differentiates between the *expected loss* (EL) and unexpected losses, which are covered by *risk-weighted assets* (RWAs). According to CRR Article 153 [17], the dependence of RWAs on the probability of default including model uncertainty is described by the function $\mathrm{RWA}\colon[0,1]\to\mathbb{R}_{\geq 0}$, via $\mathrm{RWA}(0)=0$, $\mathrm{RWA}(1)=1$, and

$$\mathrm{RWA}(p):=1.06\cdot 12.5\cdot \mathrm{EaD}\cdot \mathrm{LGD}\left(N\left(\frac{G(p)-\sqrt{R(p)}\cdot G(0.999)}{\sqrt{1-R(p)}}\right)-p\right)\tag{22}$$

with

$$R(p):=0.12\cdot\frac{1-e^{-50p}}{1-e^{-50}}+0.24\cdot\left(1-\frac{1-e^{-50p}}{1-e^{-50}}\right)\quad\text{for all } p\in(0,1),$$

where N is the cumulative distribution function of the standard normal distribution, G denotes the inverse distribution function of the standard normal distribution, and R can be seen as a correlation factor.[1] On a practical level, a frequent choice for the PD including model uncertainty is the 'best-estimate' PD of the rating system multiplied with suitable a margin of conservatism ($1+\mathrm{MoC}$ as a multiplier). In order to get a better understanding for the RWA formula, we briefly explain the terms EaD and LGD. The *exposure at default* (EaD) can be seen as the amount of credits of a borrower at the time of its default, for instance 1 million. The *loss given default* (LGD) on the other hand is the height of the loss in relation to the amount of exposure at the time of default, i.e., it is a number between 0 and 1.

Risk-weighted assets are essential for financial institutions' capital requirements, since they must hold at least 8% of RWAs as equity capital. Hence, they play an important role for financial institutions' risk provisions. In comparison to the expected loss (EL), which is calculated as the product of 'best-estimate' PD, LGD, and EaD. RWAs focus on unexpected losses from exposures, which show different characteristics compared to expected losses.

In Figure 1, we see the dependence of RWAs on p, where we assume the EaD to be 1 and the LGD to be 0.4. The standardization of setting the EaD as 1 is also referred

[1] For simplicity, we assume the effective maturity M in Article 153 of the CRR [17] to be equal to 1.

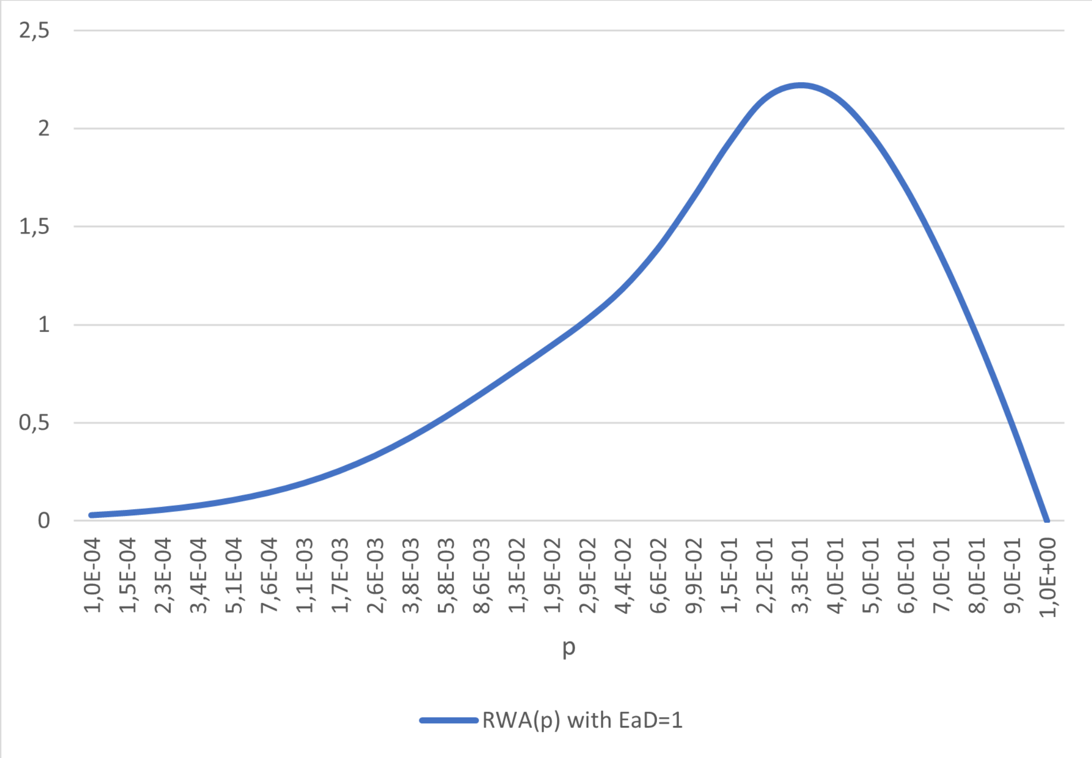

Figure 1. RWA dependence on PD

to as risk weight (RW) in Article 153 of the CRR [17]. We observe that the function is not monotonically increasing with respect to the probability of default $p \in [0,1]$ – a characteristic that applies, e.g., to expected losses. This property is not surprising since for very high PDs the loss is expected and thus the unexpected loss becomes smaller (for the same exposure). Nonetheless, this could potentially lead to problems concerning risk provisions as we will see in Example 53 below.

In the following examples, we now focus on a concrete rating system consisting of 22 rating classes and regard two distinct methods of quantifying model uncertainties. Our aim is to analyze the effect of different default risk measures on capital requirements.

For the calculation of capital requirements corresponding to RWAs, the PD including the MoC (hence including model uncertainty) is used. In general, a common problem occurring in the course of a model estimation is the lack of sufficient data. As a consequence, the model is sometimes estimated on pooled data, using combined information of many financial institutions in order to attain a higher validity.

There are of course several ways to specify a margin of conservatism. On the one hand, it is conceivable to calculate a MoC that is constant for all rating classes, since uncertainties cannot always be precisely quantified and the total data provides a more stable result for the entire rating system. However, on the other hand, the highly different amount of data per rating class strongly suggests that the uncertainty of the model should depend on the rating class as well. In the following examples, we compare the influence of these different approaches on capital requirements.

To display the 22 rating classes, we fix a discrete probability measure $\mathbb{P}$ and the reference default risk measure $\mathrm{PD}_{\mathbb{P}}(X) := \mathbb{P}(X > 0)$. To express the variable model uncertainty, which, loosely speaking, will be referred to as the 'true' model uncertainty, we use a distorted PD (which is a default risk measure according to 5), where the distortion function $T\colon [0,1] \to [0,1]$ is assumed to be nondecreasing with $T(0) = 0$ and $T(1) = 1$, and the margin of conservatism is given by

$$\mathrm{MoC}(p) = \frac{T(p)}{p} - 1 \quad \text{for all } p \in (0,1].$$

If we allowed T to be decreasing on a subinterval of $[0,1]$, then $\varrho(X) := T\big(\mathrm{PD}_{\mathbb{P}}(X)\big)$ is, in general, no default risk measure, cf. Example 45.

From a practical point of view, the exclusion of decreasing functions T in not a big restriction, since it would imply that there are at least two neighboring rating classes k and $k+1$ with corresponding PDs denoted by p_k and p_{k+1} such that $T(p_k)$ is larger than $T(p_{k+1})$. Assuming, for example, that the PD of two neighboring rating classes grows by the factor 1.5, i.e., $p_{k+1} = 1.5 \cdot p_k$, we would attain the inequality

$$\mathrm{MoC}(p_k) = \frac{T(p_k)}{p_k} - 1 > \frac{T(p_{k+1})}{p_k} - 1 = 1.5 \cdot \frac{T(p_{k+1})}{p_{k+1}} - 1 = 1.5 \cdot \mathrm{MoC}(p_{k+1}) + 0.5,$$

where the uncertainty of the rating classes k and $k+1$ is expressed by $\mathrm{MoC}(p_k)$ and $\mathrm{MoC}(p_{k+1})$, respectively. If, for instance, $\mathrm{MoC}(p_{k+1}) = 0.2$, then $\mathrm{MoC}(p_k) > 0.8$, which implies a significantly larger uncertainty in rating class k both in relative and absolute numbers. Such a difference between two rating classes neither seems to be realistic nor is observed in practice.

Example 52 (Monotonically decreasing uncertainty with respect to the rating class)**.** In this example, we focus on a model where the 'true' uncertainty decreases with increasing rating classes, i.e., for two arbitrary rating classes k and $k+1$ with PDs p_k and p_{k+1}, respectively, we have

$$\mathrm{MoC}(p_k) \geq \mathrm{MoC}(p_{k+1}).$$

Note that this is consistent with the choice of a distortion function T that satisfies (17), which we will assume throughout the remainder of this example. As a consequence, the distorted PD has a representation over probability measures, which is not unusual since the lack of defaults in the 'good' rating classes typically makes it very difficult to adequately estimate a PD – thus the uncertainty here is very high. In the *intermediate* rating classes, there are both a lot of customers and defaults. The set of very 'bad' customers is often smaller but most of the defaults occur there implying a higher validity.

In Figure 2, we depict three different functions. The green function is just the identity function, the dashed blue line is the model uncertainty where a constant MoC is used for all rating classes, and the orange line describes the uncertainty through the function $T(p) = p + \frac{\sqrt{p}-p}{40}$ for $p \in [0,1]$, which represents a 'small' perturbation of the identity expressing the 'true' uncertainty of the model per rating class. We observe that

$$T(p_k) > p_k\big(1 + \mathrm{MoC}_{\text{constant}}\big) \text{ for } k \leq 8$$

and

$$T(p_k) < p_k\big(1 + \mathrm{MoC}_{\text{constant}}\big) \text{ for } k > 8.$$

To illustrate the impact of different uncertainties on RWAs, we consider the quotient between RWAs with model uncertainty and RWAs without model uncertainty once with variable uncertainty and then with a constant uncertainty per rating class. In other

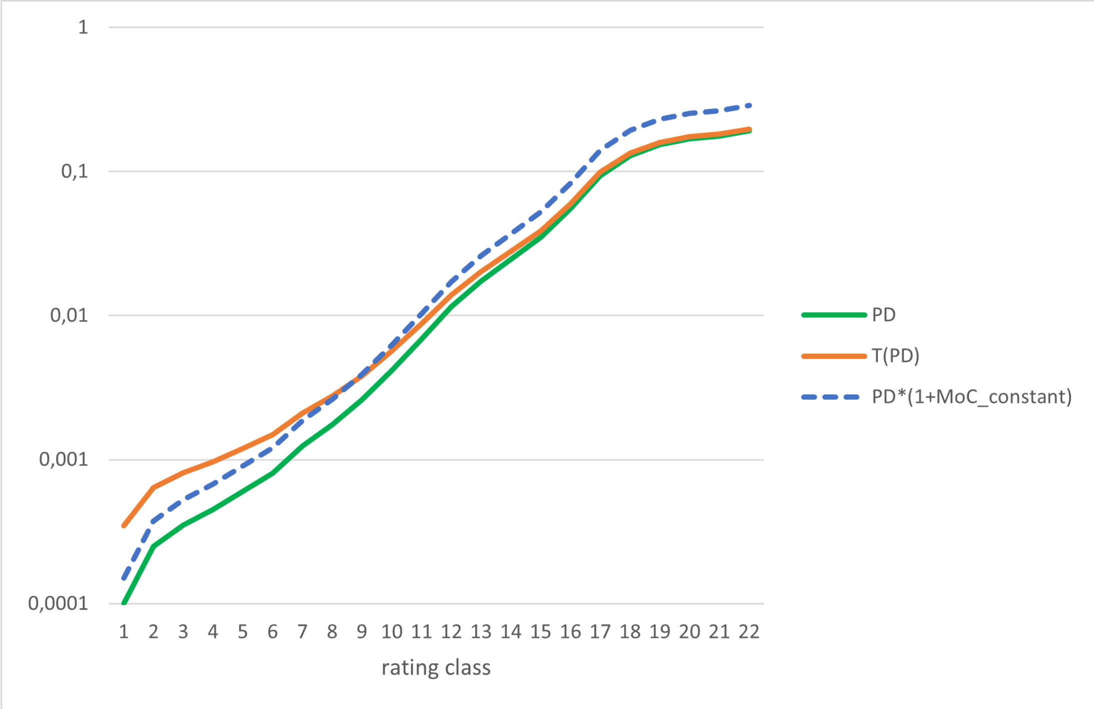

FIGURE 2. PDs per rating class including and excluding uncertainty

words, the quotient tells us with which factor the capital requirement grows per rating class due to uncertainty.

Unsurprisingly, the described factor in Figure 3 is larger for the variable uncertainty in the rating classes 1 to 8. For the other rating classes, it is the other way around. In this case, we can conclude that a higher model uncertainty leads to more RWAs and thus a higher amount of capital requirement. As a result, we could have a scenario where a financial institution uses the model uncertainty with a constant MoC to calculate its risk provisions. If the institution has customers primary in the good rating classes then it would not have enough capital requirement, again assuming that the 'true' model uncertainty is expressed by T. On the other hand, an institution with customers in bad rating classes would have a larger amount of capital requirement than necessary by regulation, thus leading to a possible competitive disadvantage.

Example 53 (Nonmonotonic uncertainty per rating class)**.** In a second example, we focus on a rating model that is estimated only on customers that are in good or middle rating classes. Thus, the uncertainty for bad rating classes is high since the calibration function has to be extrapolated on it without direct information from backtesting or benchmarking data. Such constellations can appear in low default portfolios where, historically, no 'bad' customers have been observed yet. Nonetheless, it could happen that some customers will be classified in bad rating classes in the future and, in that case, the accuracy of the extrapolation and its uncertainty cannot be neglected.

Again, the distortion function T expresses the 'true' degree of uncertainty, and is assumed to be nondecreasing. Hence, the 'true' uncertainty of the model again depends on the rating classes and peaks on both the very good and the very bad classes. The distortion function T can explicitly be constructed in such a way that the orange line lies above the dashed blue line for the rating classes 1 to 4 and 18 to 22.

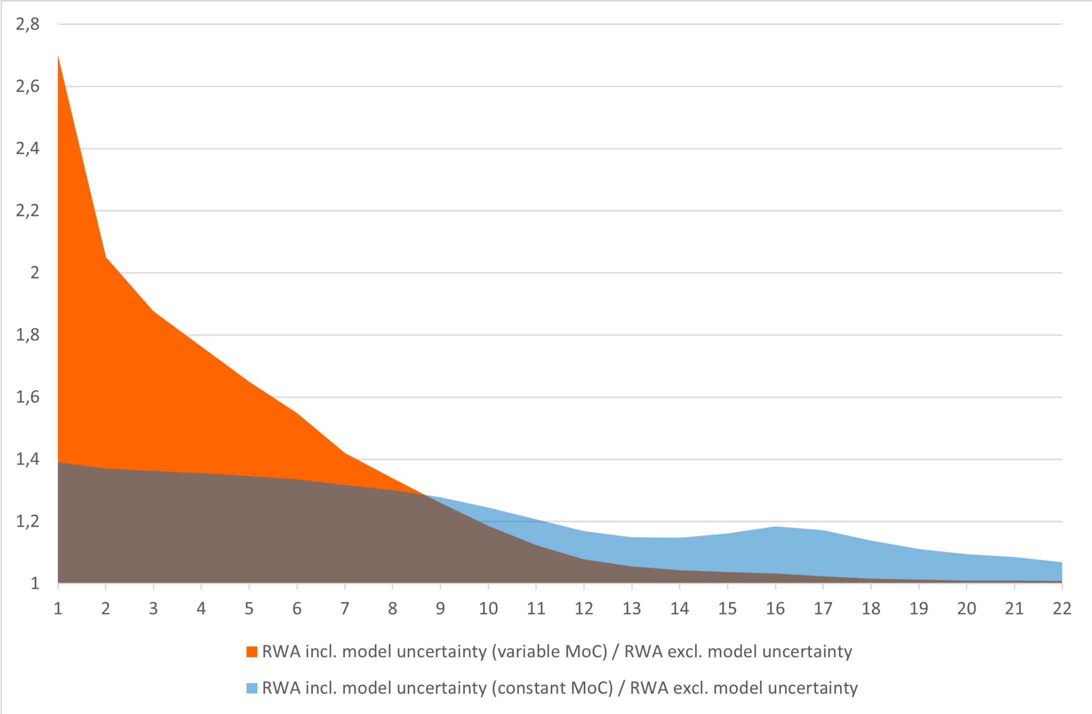

FIGURE 3. RWA growth due to uncertainty

Eye-catching is the fact that, for the rating classes from 20 to 22, the growth of RWAs caused by uncertainty is smaller when, instead of the constant MoC, the distortion function T is used although the uncertainty here is much higher. This can be explained by recalling the property of RWAs of not being monotonically increasing with respect to the PD (including uncertainty) since the unexpected loss gets smaller for very high PDs. However, regarding capital requirements, model uncertainty is only considered for RWAs and not for ELs. So, in contrast to the regulatory intention to increase equity capital by considering model uncertainty, for those rating classes (20 to 22) the capital requirement is de facto smaller when higher model uncertainty is taken as a basis.

APPENDIX A. DISTORTED PROBABILITIES

Throughout this section, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $T\colon [0,1] \to [0,1]$ with $T(0) = 0$ and $T(1) = 1$. Moreover, let

$$c(A) := T\big(\mathbb{P}(A)\big) \quad \text{for all } A \in \mathcal{F}.$$

The following lemma can be found, e.g., in [22]. For the reader's convenience, we provide a short proof.

Lemma 54. *The capacity c is 2-alternating if T is concave. If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, also the converse is true.*

Proof. First assume that T is concave, and let $A, B \in \mathcal{F}$. Since $\mathbb{P}(A \cap B) \leq \mathbb{P}(A) \leq \mathbb{P}(A \cup B)$, there exists some $\lambda \in [0,1]$ such that

$$\mathbb{P}(A) = \lambda \mathbb{P}(A \cup B) + (1 - \lambda)\mathbb{P}(A \cap B).$$

Since $\mathbb{P}(A) + \mathbb{P}(B) = \mathbb{P}(A \cup B) + \mathbb{P}(A \cap B)$, it follows that

$$\mathbb{P}(B) = (1 - \lambda)\mathbb{P}(A \cup B) + \lambda \mathbb{P}(A \cap B).$$

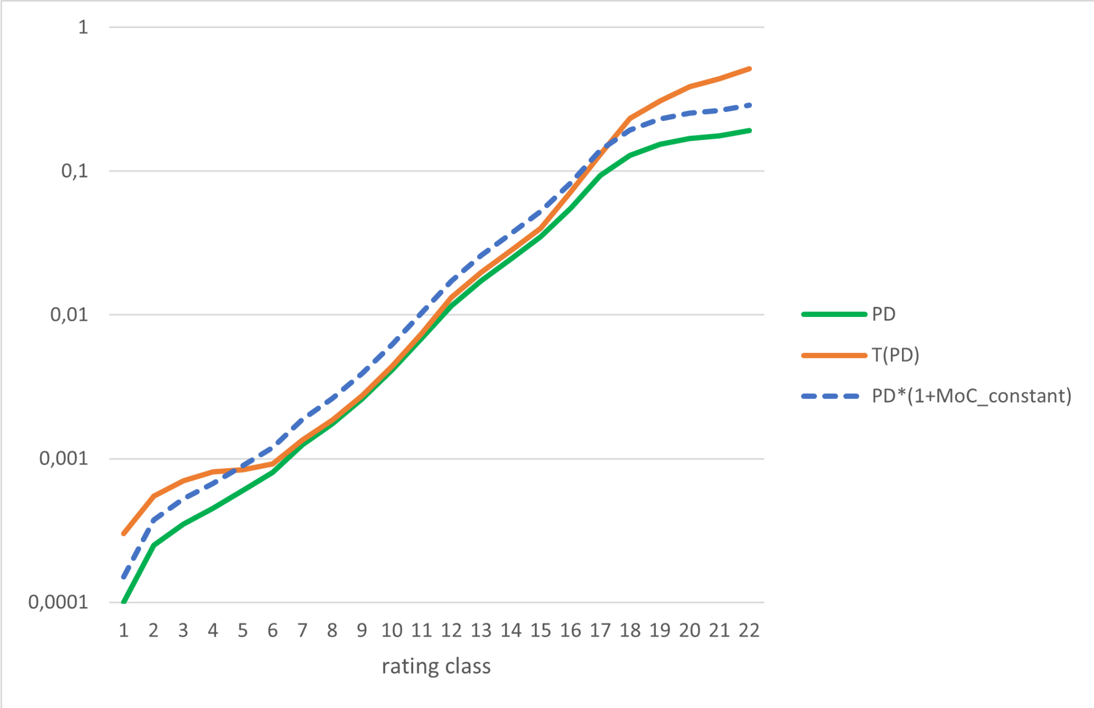

FIGURE 4. PDs per rating class including and excluding uncertainty

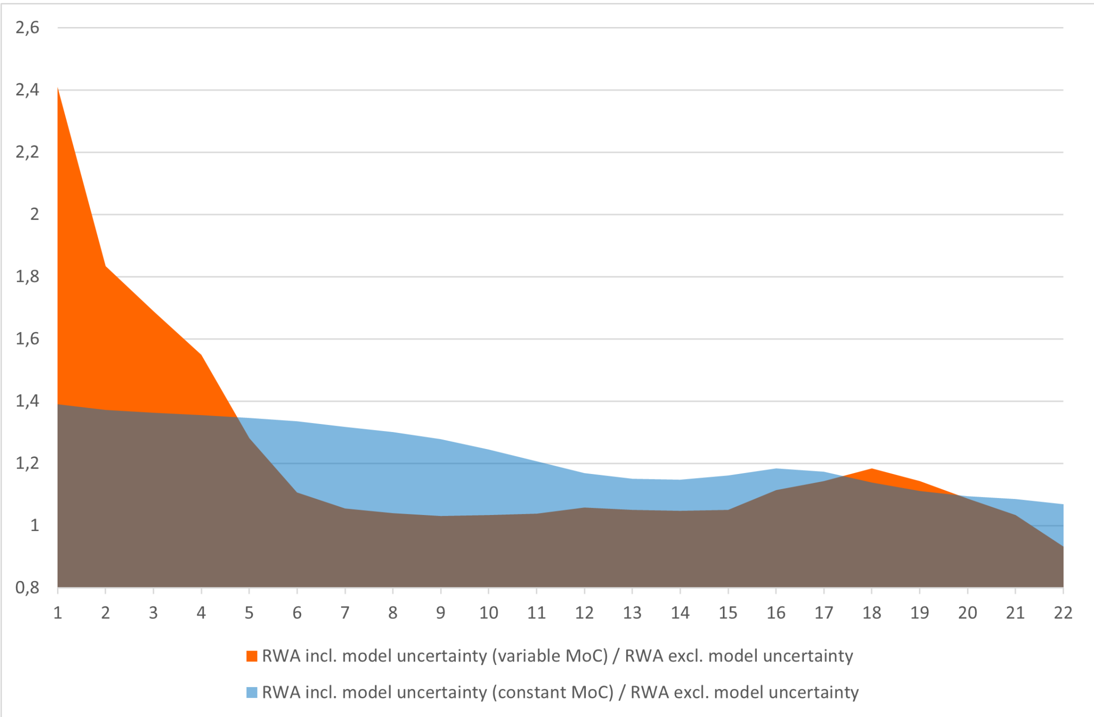

FIGURE 5. RWA growth due to uncertainty

Hence, the concavity of T implies that

$$T\big(\mathbb{P}(A)\big) + T\big(\mathbb{P}(B)\big) \geq T\big(\mathbb{P}(A \cup B)\big) + T\big(\mathbb{P}(A \cap B)\big).$$

Now, assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and that c is 2-alternating. By Remark 39 b), there exists a set $\mathcal{P}$ of countably additive and (w.r.t. $\mathbb{P}$) absolutely continuous probability measures with

$$\int X \, \mathrm{d}c = \sup_{\mathbb{P} \in \mathcal{P}} \mathbb{E}_{\mathbb{P}}(X) \quad \text{for all } X \in \mathrm{B_b}.$$

Let $p, q \in [0,1]$ with $p \leq q$. Since $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, there exist $A_1, A_2 \in \mathcal{F}$ with $\mathbb{P}(A_1) = p$, $\mathbb{P}(A_2) = q$, and $A_1 \subset A_2$. Let $\varepsilon > 0$ and $X := \mathbb{1}_{A_1} + \mathbb{1}_{A_2}$. Then, there exists some $\mathbb{Q} \in \mathcal{P}$ with

$$c(A_1) + c(A_2) = \int X \, \mathrm{d}c \leq \mathbb{E}_{\mathbb{Q}}(X) + \varepsilon = \mathbb{Q}(A_1) + \mathbb{Q}(A_2) + \varepsilon.$$

Since $\mathbb{Q}(A_1) \leq c(A_1)$ and $\mathbb{Q}(A_2) \leq c(A_2)$, it follows that $c(A_1) \leq \mathbb{Q}(A_1) + \varepsilon$ and $c(A_2) \leq \mathbb{Q}(A_2) + \varepsilon$. By [22, Lemma 4.60], for all $\lambda \in [0,1]$,

$$\begin{aligned} T\big((1-\lambda)p + \lambda q\big) &\geq \int_{1-(1-\lambda)p-\lambda q}^{1} q_{\mathbb{Q}}(s) \, \mathrm{d}s \geq (1-\lambda) \int_{1-p}^{1} q_{\mathbb{Q}}(s) \, \mathrm{d}s + \lambda \int_{1-q}^{1} q_{\mathbb{Q}}(s) \, \mathrm{d}s \\ &\geq (1-\lambda) c(A_1) + \lambda c(A_2) - \varepsilon = (1-\lambda) T(p) + \lambda T(q) - \varepsilon, \end{aligned}$$

where, in the second step, we used the fact that the map $[0,1] \to \mathbb{R}$, $t \mapsto \int_{1-t}^{1} q_{\mathbb{Q}}(s) \, \mathrm{d}s$ is concave. $\square$

Remark 55. By [2, Lemma 3.1], the function T satisfies (17) if and only if

$$T(p) = \sup_{i \in I} T_i(p) \quad \text{for all } p \in [0,1], \tag{23}$$

where I is a nonempty set and $T_i \colon [0,1] \to [0,1]$ is concave with $T_i(0) = 0$ and $T_i(1) = 1$ for all $i \in I$. In this case, the supremum in (23) is in fact a maximum.

Using Lemma 54, we can prove the following result from [2], see also [28].

Lemma 56. *The capacity c is exact if T satisfies (17). If $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless, also the converse is true.*

Proof. First assume that T satisfies (17). Then, by Remark 55, $T(p) = \sup_{i \in I} T_i(p)$ for all $p \in [0,1]$, where I is a nonempty set and $T_i \colon [0,1] \to [0,1]$ is concave with $T_i(0) = 0$ and $T_i(1) = 1$ for all $i \in I$. By Lemma 54, for each $i \in I$, there exists a nonempty set $\mathcal{P}_i$ of countably additive probability measures with

$$T_i\big(\mathbb{P}(A)\big) = \sup_{\mathbb{Q} \in \mathcal{P}_i} \mathbb{Q}(A) \quad \text{for all } A \in \mathcal{F}.$$

Let $\mathcal{P} := \bigcup_{i \in I} \mathcal{P}_i$. Then,

$$T\big(\mathbb{P}(A)\big) = \sup_{i \in I} T_i\big(\mathbb{P}(A)\big) = \sup_{i \in I} \sup_{\mathbb{Q} \in \mathcal{P}_i} \mathbb{Q}(A) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{Q}(A) \quad \text{for all } A \in \mathcal{F}.$$

Now, assume that $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and that c is exact. Let $\mathcal{P}$ be a nonempty set of countably additive probability measures with

$$T\big(\mathbb{P}(A)\big) = \sup_{\mathbb{Q} \in \mathcal{P}} \mathbb{Q}(A) \quad \text{for all } A \in \mathcal{F}.$$

Then, By Remark 39 and [22, Lemma 4.60],

$$T(p) = \sup_{\mathbb{Q} \in \mathcal{P}} \int_{1-p}^{1} q_{\mathbb{Q}}(s) \, \mathrm{d}s \quad \text{for all } p \in [0,1].$$

Since the map $[0,1]\to\mathbb{R}$, $t\mapsto\int_{1-t}^1 q_{\mathbb{Q}}(s)\,\mathrm{d}s$ is concave for all $\mathbb{Q}\in\mathcal{P}$, it follows that T satisfies (17) by Remark 55. $\square$

Remark 57. Alternatively, Lemma 54 and Lemma 56 can also be proved without invoking the Kusuoka representation of law-invariant risk measures. In this case, one uses the fact that, for every probability measure $\mathbb{Q}$, which is absolutely continuous w.r.t. $\mathbb{P}$, every set $C\in\mathcal{F}$ with $\mathbb{P}(C)>0$, and every $\lambda\in(0,1)$, there exists a set $A\in\mathcal{F}$ with $A\subset C$, $\mathbb{P}(A)=\lambda\mathbb{P}(C)$, and $\mathbb{Q}(A)\geq\lambda\mathbb{Q}(C)$. In fact, let

$$t:=\inf\left\{s>0\,\middle|\,\mathbb{P}\Big(\big\{\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}>s\big\}\cap C\Big)\leq\lambda\mathbb{P}(C)\right\}.$$

Since $(\Omega,\mathcal{F},\mathbb{P})$ is atomless, there exists a set $B\in\mathcal{F}$ with $B\subset\big\{\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}=t\big\}$ and

$$\mathbb{P}\bigg(\Big(B\cup\Big\{\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}>t\Big\}\Big)\cap C\bigg)=\lambda\mathbb{P}(C)>0.$$

Then, for $A:=\big(B\cup\big\{\frac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}>t\big\}\big)\cap C$,

$$\begin{aligned}\mathbb{Q}(C)&\leq t\mathbb{P}(C\setminus A)+\mathbb{Q}(A)\leq\frac{\mathbb{P}(C\setminus A)}{\mathbb{P}(A)}\mathbb{E}_{\mathbb{P}}\Big(\mathbb{1}_A\tfrac{\mathrm{d}\mathbb{Q}}{\mathrm{d}\mathbb{P}}\Big)+\mathbb{Q}(A)\\&\leq\frac{\mathbb{P}(C)}{\mathbb{P}(A)}\mathbb{Q}(A)=\frac{\mathbb{Q}(A)}{\lambda}.\end{aligned}$$

Note that $\mathbb{P}(C\setminus A)=(1-\lambda)\mathbb{P}(C)$ and $\mathbb{Q}(C\setminus A)\leq(1-\lambda)\mathbb{Q}(C)$.

Appendix B. Proofs of Section 2

Proof of Theorem 10. We first show that $\varrho_F(X)=\varrho(X)$ for all $X\in C$. In order to do so, let $X\in C$. Since $F(X-X)=F(0)=0$, it follows that $\varrho_F(X)\leq\varrho(X)$. On the other hand, $\varrho(X)\leq\varrho(X_0)$ for all $X_0\in C$ with $F(X-X_0)\leq 0$, and we obtain that $\varrho(X)\leq\varrho_F(X)$. Since C contains all constants, we have already verified Property (ii) in Definition 2 for ϱ_F. In order to prove Property (i), let $X,Y\in\mathrm{B_b}$ with $X\leq Y$. Since F is monotone,

$$\big\{X_0\in C\,\big|\,F(X-X_0)\leq 0\big\}\subset\big\{X_0\in C\,\big|\,F(Y-X_0)\leq 0\big\},$$

which implies that $\varrho_F(X)\leq\varrho_F(Y)$. $\square$

Proof of Corollary 11. Let $X\in\mathrm{B_b}$. Since

$$\overline{\varrho}(X)\leq\overline{\varrho}(X_0)=\varrho(X_0)\quad\text{for all }X_0\in C\text{ with }X\leq X_0,$$

it follows $\overline{\varrho}(X)\leq\varrho_{\sup}(X)$. $\square$

Appendix C. Proofs of Section 3

Proof of Proposition 15. In order to prove part a), let $\alpha\in(0,1)$. Then,

$$\varrho(X-m)\leq\varrho(Y-m)\quad\text{for all }m\in\mathbb{R}\text{ and }X,Y\in\mathrm{B_b}\text{ with }X\leq Y.$$

Hence, $\mathrm{VaR}_\varrho^\alpha(X)\leq\mathrm{VaR}_\varrho^\alpha(Y)$. Since $\varrho(0)=0$, it follows that $\mathrm{VaR}_\varrho^\alpha(0)=0$. Moreover, by definition of $\mathrm{VaR}_\varrho^\alpha$, $\mathrm{VaR}_\varrho^\alpha(X+m)=\mathrm{VaR}_\varrho^\alpha(X)+m$ for all $X\in\mathrm{B_b}$ and $m\in\mathbb{R}$. We have therefore shown that $\mathrm{VaR}_\varrho^\alpha$ is a monetary risk measure.

Next, we prove part b). To that end, let $X\in\mathrm{B_b}$. First assume that $\varrho(X)=1$. Then, $\mathrm{VaR}_\varrho^\alpha(X)>0$ for all $\alpha\in(0,1)$, so that

$$\inf\Big(\big\{\alpha\in(0,1)\,\big|\,\mathrm{VaR}_\varrho^\alpha(X)\leq 0\big\}\cup\{1\}\Big)=1=\varrho(X).$$

Next, assume that $\varrho(X) < 1$. Then, for all $\alpha \in (0,1)$, $\mathrm{VaR}_\varrho^\alpha(X) \le 0$ if and only if $\varrho(X) \le \alpha$. Hence,

$$\inf\Big(\big\{\alpha \in (0,1)\,\big|\,\mathrm{VaR}_\varrho^\alpha(X) \le 0\big\} \cup \{1\}\Big) = \varrho(X).$$

We proceed with the proof of part c). If $\mathrm{VaR}_\varrho^\alpha$ is positively homogeneous for all $\alpha \in (0,1)$, then ϱ is scaling invariant by part b). On the other hand, if ϱ is scaling invariant, then, for all $\alpha \in (0,1)$, $X \in \mathrm{B_b}$, and $\lambda > 0$,

$$\big\{m \in \mathbb{R}\,\big|\,\varrho(\lambda X - m) \le \alpha\big\} = \big\{\lambda m \in \mathbb{R}\,\big|\,\varrho(X - m) \le \alpha\big\}.$$

Hence, for all $\alpha \in (0,1)$, $X \in \mathrm{B_b}$, and $\lambda > 0$, it follows that

$$\mathrm{VaR}_\varrho^\alpha(\lambda X) = \lambda\,\mathrm{VaR}_\varrho^\alpha(X).$$

It remains to prove part d). If ϱ is quasi-convex, it follows that the level set

$$\mathcal{A}_\alpha := \{X \in \mathrm{B_b}\,|\,\varrho(X) \le \alpha\}$$

is convex for all $\alpha \in (0,1)$. Hence, by [22, Proposition 4.7],

$$\mathrm{VaR}_\varrho^\alpha(X) := \inf\{m \in \mathbb{R}\,|\,X - m \in \mathcal{A}_\alpha\}, \quad \text{for } X \in \mathrm{B_b},$$

defines a convex risk measure for all $\alpha \in (0,1)$. Next, assume that $\mathrm{VaR}_\varrho^\alpha$ is convex for all $\alpha \in (0,1)$. Then, for all $\alpha \in (0,1)$, $\lambda \in (0,1)$, and $X, Y \in \mathrm{B_b}$ with $\varrho(X) \le \alpha$ and $\varrho(Y) \le \alpha$, it follows that

$$\mathrm{VaR}_\varrho^\alpha\big(\lambda X + (1-\lambda)Y\big) \le \lambda\,\mathrm{VaR}_\varrho^\alpha(X) + (1-\lambda)\,\mathrm{VaR}_\varrho^\alpha(Y) \le 0,$$

so that, by part b),

$$\varrho\big(\lambda X + (1-\lambda)Y\big) \le \alpha.$$

This shows that ϱ is quasi-convex. □

Proof of Theorem 19. Let $X, Y \in \mathrm{B_b}$ with $X \le Y$. Since $R^\alpha(X) \le R^\alpha(Y)$ for all $\alpha \in (0,1)$, it follows that $\varrho(X) \le \varrho(Y)$. Moreover, $R^\alpha(0) = 0$ for all $\alpha \in (0,1)$, so that $\varrho(0) = 0$. On the other hand, for all $m \in \mathbb{R}$ with $m > 0$, $R^\alpha(m) = m > 0$ for all $\alpha \in (0,1)$. Therefore,

$$\big\{\alpha \in (0,1)\,\big|\,R^\alpha(X) \le 0\big\} \cup \{1\} = \{1\}$$

and it follows that $\varrho(X) = 1$. We have therefore shown that ϱ is a default risk measure.

It remains to show the equality $R^\alpha = \mathrm{VaR}_\varrho^\alpha$ for all $\alpha \in (0,1)$. To that end, let $\alpha \in (0,1)$ and $X \in \mathrm{B_b}$. First, observe that

$$\varrho\big(X - R^\alpha(X)\big) \le \alpha.$$

Hence, $\mathrm{VaR}_\varrho^\alpha(X) \le R^\alpha(X)$. Now, let $\beta \in (\alpha, 1)$ and $m \in \mathbb{R}$ with $\varrho(X - m) \le \alpha$. Then, $R^\beta(X - m) \le 0$, which implies that $R^\beta(X) \le m$. Taking the supremum over all $\beta \in (\alpha, 1)$ and the infimum over all $m \in \mathbb{R}$ with $\varrho(X - m) \le \alpha$, it follows that $R^\alpha(X) \le \mathrm{VaR}_\varrho^\alpha(X)$. □

Proof of Proposition 20. The implication (ii) ⇒ (i) follows from Proposition 15 b), once we have shown that (i) implies (ii). To that end, let $X \in \mathrm{B_b}$ and $\alpha \in (0,1)$. Then, for all $m \in \mathbb{R}$ with

$$\sup_{\mathbb{P} \in \mathcal{P}} \mathbb{P}(X > m) = \varrho(X - m) \le \alpha,$$

it follows that $m \geq \mathrm{VaR}_{\mathbb{P}}^{\alpha}(X)$ for all $\mathbb{P} \in \mathcal{P}$. This implies that

$$\sup_{\mathbb{P}\in\mathcal{P}} \mathrm{VaR}_{\mathbb{P}}^{\alpha}(X) \leq \mathrm{VaR}_{\varrho}^{\alpha}(X).$$

Now, let $\varepsilon > 0$ and

$$m := \sup_{\mathbb{P}\in\mathcal{P}} \mathrm{VaR}_{\mathbb{P}}^{\alpha}(X) + \varepsilon.$$

Then, for all $\mathbb{P} \in \mathcal{P}$, $\mathbb{P}(X > m) \leq \alpha$. This implies that

$$\mathrm{VaR}_{\varrho}^{\alpha}(X) \leq m = \sup_{\mathbb{P}\in\mathcal{P}} \mathrm{VaR}_{\mathbb{P}}^{\alpha}(X) + \varepsilon.$$

Taking the limit $\varepsilon \downarrow 0$, the claim follows. □

Proof of Proposition 22. Again, the implication (ii) ⇒ (i) follows from Proposition 15 b), once we have shown that (i) implies (ii). To that end, let $X \in \mathrm{B_b}$ and $\alpha \in (0,1)$. Then, for all $m \in \mathbb{R}$ with

$$T\big(\mathbb{P}(X > m)\big) = \varrho(X - m) \leq \alpha,$$

it follows that $\mathbb{P}(X > m) \leq T^{-1}(\alpha)$. Now, let $m \in \mathbb{R}$ with $\mathbb{P}(X > m) \leq T^{-1}(\alpha)$. Since T is lower semicontinuous, it follows that

$$T\big(\mathbb{P}(X > m)\big) \leq \alpha.$$

Hence,

$$\big\{m \in \mathbb{R}\,\big|\,\varrho(X - m) \leq \alpha\big\} = \big\{m \in \mathbb{R}\,\big|\,\mathbb{P}(X > m) \leq T^{-1}(\alpha)\big\}.$$

Taking the infimum, both, on the left and the right-hand side, the claim follows. □

Appendix D. Proofs of Section 4

Proof of Proposition 26. Let $X \in \mathrm{B_b}$ with $X \geq 0$. In Remark 25 a), we have already seen that $\varrho(X) \leq \varrho(\mathbb{1}_{\{X>0\}})$. On the other hand, for all $\varepsilon > 0$,

$$\varrho\big(X + \varepsilon\mathbb{1}_{\{X>0\}}\big) \geq \varrho\big(\varepsilon\mathbb{1}_{\{X>0\}}\big) = \varrho\big(\mathbb{1}_{\{X>0\}}\big)$$

and, using again Remark 25 a),

$$\varrho(X) \geq \varrho\big(X\mathbb{1}_{\{X>\varepsilon\}}\big) \geq \varrho\big(\varepsilon\mathbb{1}_{\{X>\varepsilon\}}\big) = \varrho\big(\mathbb{1}_{\{X>\varepsilon\}}\big).$$

The proof is complete. □

Proof of Proposition 28. We start with the implication (i) ⇒ (ii). Let $m > 0$ and $X \in \mathrm{B_b}$ with $X \geq 0$. Then, $\varrho(X) \leq \varrho\big(X + m\mathbb{1}_{\{X>0\}}\big)$. Moreover, for every $\lambda \in (0,1)$,

$$\varrho\big(X + m\mathbb{1}_{\{X>0\}}\big) = \varrho\big(\lambda X + \lambda m\mathbb{1}_{\{X>0\}}\big) \leq \varrho\big(X + \lambda m\mathbb{1}_{\{X>0\}}\big).$$

Therefore, by assumption,

$$\varrho\big(X + m\mathbb{1}_{\{X>0\}}\big) \leq \inf_{\lambda\in(0,1)} \varrho\big(X + \lambda m\mathbb{1}_{\{X>0\}}\big) = \varrho(X).$$

In order to prove the implication (ii) ⇒ (i), let $X \in \mathrm{B_b}$ with $X \geq 0$ and $\lambda > 0$. If $X = 0$ or $\lambda = 1$, it follows that $\varrho(\lambda X) = \varrho(X)$. Therefore, assume that $\sup X > 0$ and $\lambda \neq 1$. First, we consider the case, where $\lambda > 1$. Then, $\lambda X \geq X$, so that $\varrho(\lambda X) \geq \varrho(X)$. Moreover, $\lambda X \leq X + (\lambda - 1)\sup X\mathbb{1}_{\{X>0\}}$. Defining $m := (\lambda - 1)\sup X$, it follows that

$$\varrho(\lambda X) \leq \varrho\big(X + m\mathbb{1}_{\{X>0\}}\big) = \varrho(X).$$

Now, let $\lambda < 1$. Then, $\varrho(\lambda X) \leq \varrho(X)$. On the other hand,

$$X \leq \lambda X + (1 - \lambda)(\sup X)\mathbb{1}_{\{X>0\}} = \lambda X + (1 - \lambda)(\sup X)\mathbb{1}_{\{\lambda X>0\}}.$$

Therefore, defining $m := (1-\lambda)\sup X$, we find that

$$\varrho(X) \le \varrho\big(\lambda X + m\mathbb{1}_{\{\lambda X>0\}}\big) = \varrho(\lambda X).$$

By Proposition 26, (i) implies (iii) and, trivially, (iii) implies (i). □

Proof of Lemma 30. Trivially, (i) ⇒ (ii) and (ii) ⇒ (iii). In order to prove the remaining implication (iii) ⇒ (i), let $X \in \mathrm{B_b}$. Then, $X \le X^+$ and therefore $\varrho(X) \le \varrho(X^+)$. If $X \ge 0$, then $X = X^+$ and it follows that $\varrho(X) = \varrho(X^+)$. Therefore, assume that $\inf X < 0$ and define $m := -\inf X$. Then,

$$X^+ - m\mathbb{1}_{\{X^+=0\}} = X^+ - m\mathbb{1}_{\{X\le 0\}} \le X.$$

By assumption, we obtain that

$$\varrho(X^+) = \varrho\big(X^+ - m\mathbb{1}_{\{X^+=0\}}\big) \le \varrho(X).$$

□

Appendix E. Proofs of Section 5

Proof of Proposition 36. The implication (i) ⇒ (ii) follows from Remark 35 and the implication (ii) ⇒ (i) is trivial. □

Proof of Proposition 37. Clearly, (iii) implies (ii). The implication (ii) ⇒ (i) follows directly from the fact that, for any sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \nearrow X \in \mathrm{B_b}$ as $n \to \infty$, $\{X_n > 0\} \subset \{X_{n+1} > 0\}$ for all $n \in \mathbb{N}$ and

$$\bigcup_{n\in\mathbb{N}} \{X_n > 0\} = \{X > 0\}.$$

It remains to prove the implication (i) ⇒ (iii). Let $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ and $X \in \mathrm{B_b}$ with $X_n \nearrow X$ as $n \to \infty$. By potentially adding $\|X_1\|_\infty$ to X and X_n for all $n \in \mathbb{N}$ and using the fact that the Choquet integral is a monetary risk measure, we may w.l.o.g. assume that $X_1 \ge 0$. Then, using the monotone convergence theorem,

$$\lim_{n\to\infty} \int X_n \,\mathrm{d}c = \lim_{n\to\infty} \int_0^\infty \varrho(X_n - s)\,\mathrm{d}s = \int_0^\infty \varrho(X - s)\,\mathrm{d}s = \int X \,\mathrm{d}c.$$

The proof is complete. □

Proof of Proposition 40. Clearly (ii) implies (i). We prove the nontrivial implication (i) ⇒ (ii). Let $\mathcal{O}$ denote the set of all open subsets of Ω and $c(B) := \varrho(\mathbb{1}_B)$ for all $B \in \mathcal{O}$. Since ϱ is submodular with (14), it follows that

$$\varrho(X) = c\big(\{X > 0\}\big) \quad \text{for all } X \in \mathrm{L_b}.$$

The continuity from below of ϱ implies that $c\big(\bigcup_{n\in\mathbb{N}} B_n\big) = \lim_{n\to\infty} c(B_n)$ for all sequences $(B_n)_{n\in\mathbb{N}} \subset \mathcal{O}$ with $B_n \subset B_{n+1}$ for all $n \in \mathbb{N}$. The statement now follows from [34, Corollary 2.6]. □

Proof of Proposition 42. We first prove the implication (i) ⇒ (iii). Let $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \searrow 0$ as $n \to \infty$. Using the monotone convergence theorem,

$$\lim_{n\to\infty} \int X_n \,\mathrm{d}c = \lim_{n\to\infty} \int_0^\infty \varrho(X_n - s)\,\mathrm{d}s = 0.$$

Clearly, (iii) implies (ii), and it remains to prove that (ii) implies (i). To that end, observe that, for every sequence $(X_n)_{n\in\mathbb{N}} \subset \mathrm{B_b}$ with $X_n \searrow 0$ as $n \to \infty$ and $\varepsilon > 0$,

$\{X_{n+1} > \varepsilon\} \subset \{X_n > \varepsilon\}$ for all $n \in \mathbb{N}$ and $\bigcap_{n\in\mathbb{N}}\{X_n > \varepsilon\} = \emptyset$. Hence,

$$\lim_{n\to\infty} \varrho\,(X_n - \varepsilon) = \lim_{n\to\infty} \varrho\big(\mathbb{1}_{\{X_n>\varepsilon\}}\big) = 0.$$

The representation via (countably additive) probability measures now follows from the standard theory on coherent risk measures, cf. [22]. □

Appendix F. Proofs of Section 6

Proof of Theorem 44. Clearly, (ii) implies (i). For the other implication, let

$$P := \big\{p \in [0,1] \,\big|\, \exists X \in C\colon \mathbb{P}(X > 0) = p\big\}$$

Thus, for any $p \in P$, there exists some $X_p \in C$ with $\mathbb{P}\,(X_p > 0) = p$ and, by our global assumption on C, $I_p := \mathbb{1}_{\{X_p>0\}} \in C$. For $p \in [0,1]$, we define

$$T(p) := \varrho(I_{q_p}) \quad \text{with} \quad q_p := \sup\big([0,p] \cap P\big).$$

As a result $T(p) = \varrho(I_p)$ for all $p \in P$. In particular, $T(0) = 0$ and $T(1) = 1$. Let $X \in C$ and $p_X := \mathrm{PD}_{\mathbb{P}}(X) = \mathbb{P}(X > 0)$. Then, by assumption,

$$\varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) = \varrho(I_{p_X}) = T(p_X) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big).$$

□

Proof of Theorem 47. In view of Theorem 44, we only have to prove that the distortion function in (ii) is nondecreasing. To that end, assume that $\varrho(X) = T\big(\mathrm{PD}_{\mathbb{P}}(X)\big)$ for all $X \in C$ with a distortion function $T\colon [0,1] \to [0,1]$. Let $p, q \in P$. Since C contains an ordered subset, there exist $X, Y \in C$ with $\mathrm{PD}_{\mathbb{P}}(X) = p$, $\mathrm{PD}_{\mathbb{P}}(Y) = q$, and $\{X > 0\} \subset \{Y > 0\}$. Due to the monotonicity of ϱ, it follows that

$$T(p) = \varrho(X) = \varrho\big(\mathbb{1}_{\{X>0\}}\big) \le \varrho\big(\mathbb{1}_{\{Y>0\}}\big) = \varrho(Y) = T(q).$$

The proof is complete. □

Proof of Proposition 50. First assume that $\mathbb{Q}$ satisfies (20). Then,

$$T(p) \ge \int_0^p q_{\mathbb{Q}}(1-s)\,\mathrm{d}s = \int_{1-p}^1 q_{\mathbb{Q}}(s)\,\mathrm{d}s \quad \text{for all } p \in [0,1].$$

Therefore, using [22, Lemma 4.60],

$$T\big(\mathrm{PD}_{\mathbb{P}}(X)\big) \ge \int_{1-\mathrm{PD}_{\mathbb{P}}(X)}^1 q_{\mathbb{Q}}(s)\,\mathrm{d}s \ge \mathrm{PD}_{\mathbb{Q}}(X) \quad \text{for all } X \in \mathrm{B_b}.$$

On the other hand, if $(\Omega, \mathcal{F}, \mathbb{P})$ is atomless and $\mathbb{Q}$ satisfies (21), it follows from (18) that

$$T(p) \ge \int_{1-p}^1 q_{\mathbb{Q}}(s)\,\mathrm{d}s = \int_0^p q_{\mathbb{Q}}(1-s)\,\mathrm{d}s \quad \text{for all } p \in [0,1].$$

□

References

[1] M. Amarante and F.-B. Liebrich. Distortion Risk Measures: Prudence, Coherence, and the Expected Shortfall. *Preprint*, 2023.

[2] Z. Aouani and A. Chateauneuf. Exact capacities and star-shaped distorted probabilities. *Math. Social Sci.*, 56(2):185–194, 2008.

[3] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[4] P. Beissner and F. Riedel. Equilibria under Knightian price uncertainty. *Econometrica*, 87(1):37–64, 2019.

[5] F. Bellini, P. Koch-Medina, C. Munari, and G. Svindland. Law-invariant functionals that collapse to the mean. *Insurance Math. Econom.*, 98:83–91, 2021.

[6] D. Bierlein. Über die Fortsetzung von Wahrscheinlichkeitsfeldern. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 1:28–46, 1962/63.

[7] V. Bignozzi, M. Burzoni, and C. Munari. Risk measures based on benchmark loss distributions. *Journal of Risk and Insurance*, 87(2):437–475, 2020.

[8] C. Bluhm, L. Overbeck, and C. Wagner. *Introduction to credit risk modeling*. Chapman & Hall/CRC, 2016.

[9] B. Bouchard and M. Nutz. Arbitrage and duality in nondominated discrete-time models. *Ann. Appl. Probab.*, 25(2):823–859, 2015.

[10] C. Burgert and L. Rüschendorf. On the optimal risk allocation problem. *Statist. Decisions*, 24(1):153–171, 2006.

[11] M. Burzoni, C. Munari, and R. Wang. Adjusted expected shortfall. *Journal of Banking & Finance*, 134:106297, 2022.

[12] M. Burzoni, F. Riedel, and H. M. Soner. Viability and arbitrage under Knightian uncertainty. *Econometrica*, 89(3):1207–1234, 2021.

[13] E. Castagnoli, F. Maccheroni, and M. Marinacci. Choquet insurance pricing: a caveat. *Math. Finance*, 14(3):481–485, 2004.

[14] C. P. Chambers. Ordinal aggregation and quantiles. *J. Econom. Theory*, 137(1):416–431, 2007.

[15] R. Cont, R. Deguest, and X. D. He. Loss-based risk measures. *Stat. Risk Model.*, 30(2):133–167, 2013.

[16] F. Coquet, Y. Hu, J. Mémin, and S. Peng. Filtration-consistent nonlinear expectations and related g-expectations. *Probab. Theory Related Fields*, 123(1):1–27, 2002.

[17] Council of European Union. Regulation (EU) No 575/2013 of the European Parliament and of the Council of 26 June 2013 on prudential requirements for credit institutions and investment firms and amending regulation (EU) No 648/2012. *Official Journal of the European Union*, L 176:1, 2013. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32013R0575.

[18] C. Dellacherie and P.-A. Meyer. *Probabilities and potential*, volume 29 of *North-Holland Mathematics Studies*. North-Holland Publishing Co., Amsterdam-New York, 1978.

[19] R. Denk, M. Kupper, and M. Nendel. Kolmogorov-type and general extension results for nonlinear expectations. *Banach J. Math. Anal.*, 12(3):515–540, 2018.

[20] European Banking Authority (EBA). *Guidelines on PD estimation, LGD estimation and treatment of defaulted assets* https://www.eba.europa.eu/regulation-and-policy/model-validation/guidelines-on-pd-lgd-estimation-and-treatment-of-defaulted-assets.

[21] T. Fadina, P. Liu, and R. Wang. One axiom to rule them all: A minimalist axiomatization of quantiles. *SIAM J. Financial Math.*, 14(2):644–662, 2023.

[22] H. Föllmer and A. Schied. *Stochastic finance*. De Gruyter Graduate. De Gruyter, Berlin, 2016. An introduction in discrete time.

[23] M. Frittelli and E. Rosazza Gianin. Law invariant convex risk measures. In *Advances in mathematical economics*. Volume 7, pages 33–46. Springer, Tokyo, 2005.

[24] N. Gao and C. Munari. Surplus-invariant risk measures. *Math. Oper. Res.*, 45(4):1342–1370, 2020.

[25] I. Gilboa and D. Schmeidler. Maxmin expected utility with nonunique prior. *J. Math. Econom.*, 18(2):141–153, 1989.

[26] N. Guo, S. Kou, B. Wang, and R. Wang. A theory of credit rating criteria. *SSRN Preprint 3504065*, 2020.

[27] E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the Fatou property. In *Advances in mathematical economics*. Volume 9, pages 49–71. Springer, Tokyo, 2006.

[28] J. B. Kadane and L. Wasserman. Symmetric, coherent, Choquet capacities. *Ann. Statist.*, 24(3):1250–1264, 1996.

[29] S. Kusuoka. On law invariant coherent risk measures. In *Advances in mathematical economics*. Volume 3, pages 83–95. Springer, Tokyo, 2001.

[30] D. Lando. *Credit Risk Modeling*. Princeton Series in Finance. Princeton University Press, 2009. Theory and Applications.

[31] F.-B. Liebrich and C. Munari. Law-invariant functionals that collapse to the mean: beyond convexity. *Math. Financ. Econ.*, 16(3):447–480, 2022.

[32] F. Liu and R. Wang. A theory for measures of tail risk. *Math. Oper. Res.*, 46(3):1109–1128, 2021.

[33] F. Maccheroni, M. Marinacci, and A. Rustichini. Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, 74(6):1447–1498, 2006.

[34] M. Nendel. Lower semicontinuity of monotone functionals in the mixed topology on C_b. *Preprint*, arXiv:2210.09133, 2022.

[35] M. Nendel, F. Riedel, and M. D. Schmeck. A decomposition of general premium principles into risk and deviation. *Insurance Math. Econom.*, 100:193–209, 2021.

[36] S. Peng. *Nonlinear expectations and stochastic calculus under uncertainty*, volume 95 of *Probability Theory and Stochastic Modelling*. Springer, Berlin, 2019. With robust CLT and G-Brownian motion.

[37] K. Pluto and D. Tasche. Estimating probabilities of default for low default portfolios. *The Basel II Risk Parameters: Estimation, Validation, Stress Testing-with Applications to Loan Risk Management*, pages 75–101, 2011.

[38] A. F. Shorrocks. The class of additively decomposable inequality measures. *Econometrica*, 48(3):613–625, 1980.

[39] A. F. Shorrocks. Inequality decomposition by population subgroups. *Econometrica*, 52(6):1369–1385, 1984.

[40] J. Staum. Excess invariance and shortfall risk measures. *Oper. Res. Lett.*, 41(1):47–53, 2013.

[41] D. Tasche. Bayesian estimation of probabilities of default for low default portfolios. *Journal of Risk Management in Financial Institutions*, 6(3):302–326, 2013.

[42] D. M. Topkis. *Supermodularity and complementarity*. Frontiers of Economic Research. Princeton University Press, Princeton, NJ, 1998.

[43] R. Wang and J. F. Ziegel. Scenario-based risk evaluation. *Finance Stoch.*, 25(4):725–756, 2021.

[44] S. S. Wang, V. R. Young, and H. H. Panjer. Axiomatic characterization of insurance prices. *Insurance Math. Econom.*, 21(2):173–183, 1997.

Center for Mathematical Economics, Bielefeld University, Germany
Email address: `max.nendel@uni-bielefeld.de`

Center for Mathematical Economics, Bielefeld University, Germany and Landesbank Baden-Württemberg, Stuttgart, Germany
Email address: `jan.streicher@uni-bielefeld.de`