Received: 1 July 2016 | Accepted: 1 May 2017

DOI: 10.1111/mafi.12160

WILEY MATHEMATICAL FINANCE An International Journal of Mathematics, Statistics and Financial Economics

ORIGINAL ARTICLE

Risk management with weighted VaR

Pengyu Wei

Mathematical Institute and Oxford-Man Institute of Quantitative Finance, The University of Oxford, Oxford, UK

Correspondence
Pengyu Wei, Mathematical Institute and Oxford-Man Institute of Quantitative Finance, The University of Oxford, Oxford OX2 6GG, UK.
Email: weip@maths.ox.ac.uk

Abstract

This article studies the optimal portfolio selection of expected utility-maximizing investors who must also manage their market-risk exposures. The risk is measured by a so-called weighted value-at-risk (WVaR) risk measure, which is a generalization of both value-at-risk (VaR) and expected shortfall (ES). The feasibility, well-posedness, and existence of the optimal solution are examined. We obtain the optimal solution (when it exists) and show how risk measures change asset allocation patterns. In particular, we characterize three classes of risk measures: the first class will lead to models that do not admit an optimal solution, the second class can give rise to endogenous portfolio insurance, and the third class, which includes VaR and ES, two popular regulatory risk measures, will allow economic agents to engage in "regulatory capital arbitrage," incurring larger losses when losses occur.

KEYWORDS

expected shortfall, portfolio insurance, portfolio selection, regulatory capital arbitrage, risk measure, value-at-risk, weighted value-at-risk

1 | INTRODUCTION

Value-at-risk (VaR) has long been an industry standard, whether by choice or by regulation (BCBS, 2011; Dowd, 1998; Jorion, 1997, 2002; Saunders, 2000; SEC, 1997). However, since its introduction in approximately 1994, VaR has been criticized in both academia and industry, for its weaknesses as a benchmark. VaR fails to capture "tail risk" and it is not subadditive, defying the notion of diversification. Recognizing the shortcomings of VaR, Artzner, Delbaen, Eber, and Heath (1999) argue that a good risk measure should satisfy a set of reasonable axioms, leading to the so-called coherent risk measures. Recently there has been a movement in both academia (Acerbi & Tasche, 2002; Alexander & Baptista, 2006; Artzner et al., 1999; Embrechts, Puccetti, Rüschendorf, Wang, & Beleraj, 2014;

Rockafellar & Uryasev, 2000, 2002) and industry (BCBS, 2016) to replace VaR with expected shortfall (ES).[1] Simultaneously, alternative risk measures, such as spectral risk measures and distortion risk measures, have arisen in the portfolio selection literature (Acerbi & Simonetti, 2002; Adam, Houkari, & Laurent, 2008; Sereda et al., 2010). However, despite rich research on risk measurement, little is known about the effects of these risk measures on portfolio selection, especially in the expected utility maximization framework. This paper contributes to filling that gap.

We study the optimal portfolio selection of expected utility-maximizing investors who must also manage their market-risk exposures in a continuous-time, complete market. The risk is quantified by the weighted value-at-risk (WVaR) proposed in He, Jin, and Zhou (2015), which is a generalization of both VaR and ES and covers spectral risk measures, distortion risk measures, and many law-invariant coherent risk measures.

We first solve the problem completely, with the help of the so-called quantile formulation, which was developed recently mainly in the context of behavioral portfolio selection. Feasibility, well-posedness, and attainability are examined in greater detail. These issues have been more or less overlooked in the related literature. We propose the notion of *risk reduction per cost* (RRPC), which depends on only the risk measure and the market, to measure the trade-off between reducing the risk and incurring the cost. We find that if a risk measure's RRPC for extreme gains is infinite, then the model is unattainable whenever the constraint is nonredundant, i.e., the optimal value is finite but is not achievable by any portfolios, indicating that the model is misformulated.

We then study how different risk measures can change optimal trading patterns when the optimal solution exists. In particular, we characterize two classes of risk measures. On the one hand, if a risk measure's RRPC for extreme losses is infinite, then the agent will follow a trading strategy that creates endogenous portfolio insurance, shielding himself from large losses as it is the most efficient way to meet the requirement. This could be of particular interest to regulators. In general, portfolio insurance is costly, and it is highly unlikely that the agent will utilize such a strategy. Regulators can encourage economic agents to use the portfolio insurance strategy by imposing a constraint within this type, such as the Wang (2000) risk measure and the beta family of distortion risk measures. On the other hand, if a risk measure's RRPC for extreme losses is 0, then the agent will ignore losses in the bad states and leave himself completely uninsured because it is costly and inefficient to insure against such losses. Moreover, a further inspection reveals that these risk measures allow economic agents to engage in the "regulatory capital arbitrage" defined by Jones (2000), in the sense that if a large loss occurs, it is likely to be even larger than it would have been in the absence of the risk constraint. This could be a source of concern for both regulators and real-world risk managers. Risk measures are viewed by many as a tool to protect economic agents from large losses, which could cause credit and solvency problems. However, risk measures within this type, such as VaR and ES, actually defeat the purpose of such regulations. Basak and Shapiro (2001) obtain similar results for VaR, but to the best of our knowledge, we are the first to characterize risk measures that can create endogenous portfolio insurance or lead to "regulatory capital arbitrage" within a relatively general class of risk measures.

Our paper contributes to the literature on utility maximization with risk constraints. Basak and Shapiro (2001) consider utility maximization in a continuous-time complete market, assuming that agents must limit their risks as measured by VaR. Basak, Shapiro, and Teplá (2006) study the portfolio choice problem in which VaR is evaluated relative to a benchmark. Gabih, Grecksch, and Wunderlich (2005) and Gabih, Sass, and Wunderlich (2009) consider the problem when the risk is measured by either VaR or the expected loss. Gundel and Weber (2008) generalize VaR to a class of convex risk measures. Rogers (2009) considers law-invariant coherent risk measures. Cahuich and Hernández-Hernández (2013) study rank-dependent utility maximization with a risk constraint. However, all of the above papers, except for those by Basak and Shapiro (2001) and Basak et al. (2006), are more of a mathematical treatment of the problem with little analysis of its economic implications. In particular, less attention has been paid to how various risk measures can alter portfolio choice patterns.

A different approach in portfolio selection, pioneered by Markowitz (1952), is to use the mean and variance to measure the return and risk, respectively, and the economic agent chooses among the portfolios that yield a prespecified level of expected return while minimizing the variance of the portfolio's return. In addition to variance, researchers have also considered alternative risk measures (Acerbi & Simonetti, 2002; Adam et al., 2008; Alexander & Baptista, 2002, 2004; Campbell, Huisman, & Koedijk, 2001; Kast, Luciano, & Peccati, 1999; Rockafellar & Uryasev, 2000, 2002). These researchers all study single-period mean-risk portfolio selection problems. There have also been extensions of the mean-risk model from the single-period setting to the dynamic, continuous-time setting, including but not limited to Bielecki, Jin, Pliska, and Zhou (2005), Jin, Yan, and Zhou (2005), Basak and Chabakauri (2010), etc. In particular, He et al. (2015) recently considered continuous-time mean-risk portfolio choice problems with WVaR. They found that the model is prone to being ill-posed, especially when bankruptcy is allowed, leading to extreme risk-taking behaviors.

Our paper makes two main contributions. First, by completely solving the corresponding constrained continuous-time utility maximization problem, we can characterize the optimal terminal wealth when the risk is measured by various risk measures. Our model offers a variety of attractive features compared to its mean-WVaR counterpart (He et al., 2015). In particular, we provide a critique of the current risk management practices and especially the new Basel Accord (BCBS, 2016).

Second, the technical analysis performed in our paper contributes to the mathematical aspects of the portfolio selection literature. Typically, the continuous-time utility maximization problem is solved by the Lagrange dual method. We find that with an additional constraint on the risk, the dual method can fail under some circumstances; in other words, the optimal solution to the original problem can exist but might not be given by the Lagrange dual problem. After establishing the relationship between the original problem and the Lagrange dual problem, we solve the dual problem completely. Although the main techniques for solving the dual problem, i.e., quantile formulation and the concave envelope relaxation, have been employed in the literature (Rogers, 2009; Xia & Zhou, 2016; Xu, 2016), the existence of optimal solutions requires a more nuanced analysis. For example, in finding the Lagrange multipliers, we characterize the monotonicity and continuity of a function's concave envelope's right derivative with respect to the original function's parameter.

Finally, let us comment on the settings of our model. Following Basak and Shapiro (2001) and Basak et al. (2006), we measure the investment risk by applying a risk measure to the terminal wealth. The risk is evaluated at the beginning of the investment and the agent must commit himself to comply with the constraint in all future dates. Some papers, e.g., Yiu (2004), Cuoco and Liu (2006), Leippold, Trojani, and Vanini (2006), and Cuoco, He, and Isaenko (2008), apply VaR dynamically. The agent then maximizes the expected utility of her terminal wealth or consumption while the portfolio risk at any time is controlled at a certain level. Generally, closed-form solutions are unavailable and one must resort to either numerical solutions or asymptotic solutions. Alternatively, one can apply dynamic time-consistent risk measures. However, Kupper and Schachermayer (2009) show that the only dynamic risk measure that is both law invariant and time consistent is the entropic risk measure, which is too restrictive to accommodate VaR and ES, two popular regulatory risk measures. To maintain analytical tractability and a relatively general framework, we choose our current setting.

The remainder of this paper is organized as follows: In Section 2, we introduce WVaR. In Section 3, we formulate the risk management with weighted VaR (WVaR-RM) problem, which is then solved completely in Section 4. Impacts on portfolio choice are analyzed in Section 5. Section 6 concludes the paper. Appendix A gives an overview of the quantile formulation. All of the remaining proofs are placed in Appendix B.

2 | RISK MEASURES

In this section, we introduce the WVaR risk measure. Consider a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and denote by $L^\infty := L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ the set of all $\mathbb{P}$-essentially bounded random variables; in addition, $LB := LB(\Omega, \mathcal{F}, \mathbb{P})$, the set of all lower-bounded finite-valued random variables (but different random variables could have different lower bounds). Let $X \in LB$ represent the profit and loss (P&L) of an investment.

Let $\mathbb{F}$ be the set of cumulative distribution functions (CDFs) of all lower bounded random variables that take values in $\mathbb{R}$, in other words,

$$\mathbb{F} = \{F(\cdot) : \mathbb{R} \to [0,1], \text{ nondecreasing, right continuous,}$$
$$F(a) = 0 \text{ for some } a \in \mathbb{R} \text{ and } F(+\infty) = 1\}.$$

The lower boundedness above corresponds to the required tameness of the portfolios (to be discussed in the next section). For any $F(\cdot) \in \mathbb{F}$, denote by $F^{-1}(\cdot)$ its right-inverse, in other words,

$$F^{-1}(t) = \inf\{x \in \mathbb{R} : F(x) > t\} = \sup\{x \in \mathbb{R} : F(x) \le t\},\ t \in [0,1].$$

Let $\mathbb{G} := \{F^{-1}(\cdot) : F(\cdot) \in \mathbb{F}\}$ be the corresponding set of quantile functions, or

$$\mathbb{G} = \{G(\cdot) : [0,1) \to \mathbb{R}, \text{ nondecreasing, right continuous with left limits (RCLL)}\},$$

where $G(1) := G(1-)$. Throughout the paper, we will use $F_X(\cdot)$ and $G_X(\cdot)$ to denote a random variable X's CDF and quantile function, respectively.

Consider a functional over the set of random P&Ls, $\rho : X \to \rho(X) \in \mathbb{R}$. It may fulfill some of the following axioms:

A1 *Monotonicity*: $\rho(X) \ge \rho(Y)$ for any $X, Y \in LB$ such that $X \le Y$;
A2 *Translation invariance*: $\rho(X + a) = \rho(X) - a$ for any $X \in LB$ and $a \in \mathbb{R}$;
A3 *Truncation continuity*: $\rho(X) = \lim_{n\to+\infty} \rho(X \wedge n)$ for any $X \in LB$;
A4 *Positive homogeneity*: $\rho(\lambda X) = \lambda\rho(X)$ for any $X \in LB$ and $\lambda > 0$;
A5 *Subadditivity*: $\rho(X + Y) \le \rho(X) + \rho(Y)$ for any $X, Y \in LB$;
A6 *Convexity*:[2] $\rho(\alpha X + (1-\alpha)Y) \le \alpha\rho(X) + (1-\alpha)\rho(Y)$ for any $X, Y \in LB$ and $\alpha \in (0,1)$;
A7 *Law invariance*: $\rho(X) = \rho(Y)$ for any $X, Y \in LB$ with the same distribution function;
A8 *Comonotonic additivity*: $\rho(X + Y) = \rho(X) + \rho(Y)$ for any $X, Y \in LB$ such that X any Y are comonotonic, i.e.,

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \ge 0 \text{ for all } (\omega, \omega') \in \Omega \times \Omega.$$

ρ is called a risk measure if it satisfies A1–A3[3] and it is a coherent risk measure if it satisfies A1–A5, as introduced in Artzner et al. (1999).

Based on the above axioms, we introduce WVaR as a generalization of VaR and ES, two popular regulatory risk measures.

2.1 | VaR, ES, and WVaR

VaR is one of the most important measures (if not the most used measure) of risk in finance. VaR describes the loss that can occur over a given period, at a given confidence level. It has long been an industry standard, whether by choice or by regulation (BCBS, 2011; Dowd, 1998; Jorion, 1997; Saunders, 2000; SEC, 1997). The VaR at a specified threshold $\alpha \in (0,1)$ is defined as

$$\mathrm{VaR}_\alpha(X) = -G_X(\alpha).$$

Since its introduction in approximately 1994, VaR has been criticized in both academia and industry, for its weaknesses as a benchmark. VaR fails to capture "tail risk" and it is not subadditive, which means that the risk of a portfolio can be larger than the sum of the stand-alone risks of its components when measured by VaR (Artzner et al., 1999).

Recognizing the shortcomings of VaR, there has been a movement to replace VaR with ES, also known as average value-at-risk (AVaR) and conditional value-at-risk (CVaR), in both academia (Acerbi & Tasche, 2002; Alexander & Baptista, 2006; Artzner et al., 1999; Embrechts et al., 2014; Rockafellar & Uryasev, 2000, 2002) and industry (BCBS, 2016).[4] ES measures the riskiness of a position by considering both the size and the likelihood of losses above a certain confidence level. The ES at level $\alpha \in (0,1)$ is defined as

$$\mathrm{ES}_\alpha(X) = -\frac{1}{\alpha}\int_0^\alpha G_X(z)dz.$$

ES satisfies A1–A8, and it is the smallest coherent, comonotonic additive and law-invariant risk measure that dominates VaR, according to Dhaene et al. (2004). Moreover, ES is in accordance with second stochastic dominance, according to Bertsimas, Lauprete, and Samarov (2004) and Leitner (2005).

In this paper, we focus on the weighted VaR (WVaR) risk measures[5] introduced in He et al. (2015), which take the following form:

$$\rho_\Phi(X) = -\int_{[0,1]} G_X(z)\Phi(dz), \tag{2.1}$$

where $\Phi \in P[0,1]$, and $P[0,1]$ is the set of all probability measures on $[0,1]$. This is a generalization of VaR and ES: when Φ is a Dirac measure, it becomes VaR; when $\Phi([0,z]) = \frac{z}{\alpha} \wedge 1, z \in [0,1]$, it becomes ES_α. WVaR is a law-invariant comonotonic additive risk measure, and it covers a large class of law-invariant coherent risk measures and all law-invariant risk measures that are both convex and comonotonic additive. See He et al. (2015) for a more detailed discussion.

We now use examples to show how WVaR generalizes many well-known risk measures that are widely used in finance and actuarial sciences. The generality of WVaR allows us to solve the risk-constrained utility maximization problem in a unified framework. Readers who are already familiar with risk measures can skip the remainder of this section.

2.2 | Spectral risk measures

Spectral risk measures, proposed in Acerbi (2002), cover a large class of coherent risk measures. One distinctive feature of such risk measures is that they map any rational investor's subjective risk aversion onto a coherent measure and vice versa.

Definition 2.1. *An element $\phi \in L^1([0,1])$ is said to be an admissible risk spectrum if*

1. *ϕ is nonnegative;*
2. *ϕ is nonincreasing;*
3. $\| \phi \| := \int_0^1 \phi(z)dz = 1$.

The spectral risk measures are defined as

$$M_\phi(X) = -\int_0^1 F_X^{-1}(z)\phi(z)dz = -\int_0^1 G_X(z)\phi(z)dz$$

for each admissible risk spectrum ϕ, and ϕ is also called a risk-aversion function. It is easy to see that spectral risk measures are special examples of WVaR. Spectral risk measures are law invariant, convex, and comonotonic additive.

Examples of spectral risk measures include:

1. The ES_α has the spectrum given by

$$\phi(z) = \begin{cases} \dfrac{1}{\alpha}, & z \in [0, \alpha]; \\ 0, & z \in (\alpha, 1]. \end{cases}$$

2. The exponential spectral risk measures, proposed in Cotter and Dowd (2006) and Dowd, Cotter, and Sorwar (2008), have spectrums given by

$$\phi(z) = \frac{Re^{-Rz}}{1 - e^{-R}},$$

where $R > 0$ is the investor's absolute risk aversion coefficient.

3. The power spectral risk measures, proposed in Dowd et al. (2008), have spectrums given by

$$\phi(z) = \begin{cases} \gamma z^{\gamma-1}, & 0 < \gamma < 1; \\ \gamma(1-z)^{\gamma-1}, & \gamma > 1, \end{cases}$$

where γ is the investor's relative risk aversion coefficient.

2.3 | Distortion risk measures

Distortion risk measures originated from Yarri's dual theory of choice under risk (Yaari, 1987). Yaari's idea consists of measuring risk by applying a distortion function g on F_X,

$$\rho_g(X) = -\int_{-\infty}^{\infty} xd(g(F_X(x))) = -\int_0^1 G_X(z)dg(z),$$

where the distortion function g is continuous, nondecreasing, and satisfies $g(0) = 0$ and $g(1) = 1$. It is a concave distortion risk measure if g is further concave. A distortion risk measure is coherent if and only if it is a concave distortion risk measure, as in Sereda et al. (2010). Distortion risk measures are also special cases of WVaR.

Distortion risk measures are widely used in insurance and actuarial sciences. For a comprehensive review, please refer to Wirch and Hardy (1999) and Sereda et al. (2010). Some well-known examples are presented below:

1. The (negative) expectation risk measure uses $g(z) = z$.
2. The Wang (2000) risk measure uses

$$g(z) = \Phi_N\left(\Phi_N^{-1}(z) - \Phi_N^{-1}(q)\right),$$

where $0 < q \leq 0.5$ is a given parameter, and Φ_N is the standard normal distribution function. The Wang (2000) risk measure is a concave distortion risk measure.

3. The beta family of distortion risk measures, proposed in Wirch and Hardy (1999), uses the distribution function of the beta distribution

$$g(z) = \int_0^z \frac{1}{\beta(a,b)} t^{a-1}(1-t)^{b-1}\,dt,$$

where $\beta(a, b)$ is the beta function with parameters $a > 0$ and $b > 0$. It is concave if and only if $a \leq 1$ and $b \geq 1$; it is strictly concave if $a < 1$ and $b > 1$.

4. The proportional hazard (PH) risk measure, discussed in Wang (1995, 1996) and Wang, Young, and Panjer (1997), uses

$$g(z) = z^{\frac{1}{\gamma}}, \quad \gamma \geq 1.$$

This is an example of the beta family of distortion risk measures, with $a = \frac{1}{\gamma}, b = 1$.

5. The dual power risk measure uses

$$g(z) = 1 - (1-z)^{\kappa}, \quad \kappa \geq 1.$$

It is also an example of the beta family of distortion risk measures, with $a = 1, b = \kappa$.

We see that power spectral risk measures, although they arise from a different context, are in fact proportional hazard/dual power risk measures.

3 | MODEL

In this section, we formulate our risk management with weighted VaR (WVaR-RM) problem.

3.1 | Market

Let $T > 0$ be a given terminal time, and let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P})$ be a filtered probability space on which we define a standard $(\mathcal{F}_t)_{t\in[0,T]}$-adapted n-dimensional Brownian motion $W(t) \equiv (W^1(t), \cdots, W^n(t))^\top$ with $W(0) = 0$, and hence, the probability space is atomless. It is assumed that $\mathcal{F}_t = \sigma\{W(s) : 0 \leq s \leq t\}$, augmented by all the $\mathbb{P}$-null sets in $\mathcal{F}$. Here and henceforth, $A^\top$ denotes the transpose of a matrix A. We define a continuous-time financial market following Karatzas and Shreve (1998). In the market, $n + 1$ assets are being traded continuously. One of the assets is a bank account whose price process $S_0(t)$ is subject to the following equation

$$dS_0(t) = r(t)S_0(t)dt,\ t \in [0, T];\ S_0(0) = s_0 > 0,$$

where the interest rate $r(\cdot)$ is a uniformly bounded, $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable, scalar-valued stochastic process. The other n assets are risky securities whose price processes $S_i(t),\ i = 1, \cdots, n$, satisfy the following stochastic differential equation (SDE)

$$dS_i(t) = S_i(t)\left[b_i(t)dt + \sum_{j=1}^{n} \sigma_{ij}(t)dW^j(t)\right], \quad t \in [0,T];\ S_i(0) = s_i > 0,$$

where $b_i(\cdot)$ and $\sigma_{ij}(\cdot)$, the appreciation and volatility rates respectively, are scalar-valued, $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable stochastic processes with

$$\int_0^T \left[\sum_{i=1}^{n} |b_i(t)| + \sum_{i=1}^{n}\sum_{j=1}^{n} |\sigma_{ij}(t)|^2\right] dt < \infty,\ a.s.$$

Set the excess rate of return process as

$$B(t) := (b_1(t) - r(t), \dots, b_n(t) - r(t))^\top,$$

and define the volatility matrix process $\sigma(t) := (\sigma_{ij}(t))_{n\times n}$. The basic assumptions imposed on the market parameters throughout this paper are summarized as follows.

Assumption 3.1. *There is a unique, $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable, $\mathbb{R}^n$-valued process $\theta(t)$ with $\mathbb{E}e^{\frac{1}{2}\int_0^T |\theta(t)|^2 dt} < \infty$ such that*

$$\sigma(t)\theta(t) = B(t),\ a.s.,\ a.e.\ t \in [0,T].$$

Consequently, we have a complete model of a securities market.

Consider an economic agent, with an initial endowment $x > 0$ and an investment horizon $[0,T]$, whose total wealth at time $t \geq 0$ is denoted by $X(t)$. Assume that the trading of shares occurs continuously in a self-financing fashion and there are no transaction costs. Then, $X(\cdot)$ satisfies

$$dX(t) = [r(t)X(t) + B(t)^\top \pi(t)]dt + \pi(t)^\top \sigma(t)dW(t),\ t \in [0,T];\ X(0) = x,$$

where $\pi_i(t)$ denotes the total market value of the agent's wealth in stock i at time t. The process $\pi(\cdot) := (\pi_1(\cdot), \cdots, \pi_n(\cdot))^\top$ is called an admissible portfolio if it is $(\mathcal{F}_t)_{t\in[0,T]}$-progressively measurable with

$$\int_0^T |\sigma(t)^\top \pi(t)|^2 dt < \infty \text{ and } \int_0^T |B(t)^\top \pi(t)| dt < \infty,\ a.s.$$

and is tame, i.e., the corresponding wealth process $X(\cdot)$ is almost surely bounded from below, although the bound could depend on $\pi(\cdot)$. It is standard in the continuous-time portfolio choice literature for a portfolio to be required to be tame, enabling, among other things, the exclusion of the doubling strategy.

With the complete market assumption, we can define the pricing kernel or state price density process as

$$\xi(t) := \exp\left\{-\int_0^t \left[r(s) + \frac{1}{2}|\theta(s)|^2\right] ds - \int_0^t \theta(s)^\top dW(s)\right\}.$$

Let $\xi := \xi(T)$. It is clear that under Assumption 3.1 and the uniform boundedness of $r(\cdot)$, $0 < \xi < +\infty$ a.s. and $0 < \mathbb{E}\xi < +\infty$. Then, in view of the standard martingale approach (Cox & Huang, 1989;

Karatzas & Shreve, 1998; Pliska, 1986), finding the optimal portfolio in this economy is equivalent to finding the optimal terminal wealth.

3.2 | Benchmark agent

Let $X := X(T)$. The benchmark model is the standard utility maximization problem,

$$\begin{aligned} \max_X \quad & \mathbb{E}[u(X)] \\ \text{subject to} \quad & \mathbb{E}[\xi X] \le x, \end{aligned} \tag{3.1}$$

where u is a utility function with the following assumption as in most of the literature.

Assumption 3.2. *$u(\cdot)$ is twice continuously differentiable, strictly increasing and strictly concave. Furthermore, $u^{'}(\cdot)$ satisfies the Inada condition, i.e., $u^{'}(0+) = +\infty$ and $u^{'}(+\infty) = 0$.*

By convention, we set $u(x) = -\infty$ for $x < 0$.

We also impose the following integrability assumption throughout the paper.

Assumption 3.3. $\mathbb{E}[(u^{'})^{-1}(\lambda_0\xi)\xi] < \infty$ *for all* $\lambda_0 > 0$.

This integrability assumption is standard in expected utility maximization problems under which we have,

Proposition 3.4. *The optimal solution to (3.1) is given by*

$$X_0^* = (u^{'})^{-1}\left(\lambda_0^*\xi\right), \tag{3.2}$$

where λ_0^* solves $\mathbb{E}[\xi X_0^*] = x$.

To avoid unnecessary technical details, we assume throughout the paper that (3.1) is well-posed, in other words,

Assumption 3.5. *The optimal value of (3.1) is finite for all $x > 0$.*

This assumption can be weakened, such as in Kramkov and Schachermayer (1999) and Jin, Xu, and Zhou (2008).

3.3 | VaR-based risk management

We motivate our new portfolio choice model by the VaR-based risk management (VaR-RM) proposed by Basak and Shapiro (2001). Recognizing that risk management is typically not an economic agent's primary objective, those scholars focus on portfolio choice within the familiar (continuous-time) complete markets setting, with the assumption that agents must limit their risks, as measured by VaR, while maximizing the expected utility.

The VaR-RM model is given by

$$\begin{aligned} \max_X \quad & \mathbb{E}[u(X)] \\ \text{subject to} \quad & \mathbb{E}[\xi X] \le x, \\ & \mathbb{P}(X \ge \underline{x}) \ge 1 - \alpha, \end{aligned} \tag{3.3}$$

where $\mathbb{P}(X \ge \underline{x}) \ge 1 - \alpha$ is the VaR constraint and $\underline{x}$ is the "floor" terminal wealth specified exogenously.

3.4 | Risk management with weighted VaR

In this paper, we follow Basak and Shapiro (2001) to embed the risk management with weighted VaR (WVaR-RM) into the standard utility maximization. We assume that an economic agent uses the expected utility model when managing his trading portfolio and that as a consequence of either the internal risk management or the external regulation such as by the SEC (1997) and BCBS (2011, 2016), he has decided to manage the portfolio's risk by imposing a constraint on the portfolio's WVaR. Consequently, the WVaR-RM model is

$$\begin{aligned} \max_{X} \quad & \mathbb{E}[u(X)] \\ \text{subject to} \quad & \mathbb{E}[\xi X] \leq x, \\ & \rho_{\Phi}(X) \leq -\underline{x}, \end{aligned} \tag{3.4}$$

where ρ_{Φ} is the weighted VaR risk measure given by (2.1).

Note that $\rho_{\Phi}(X) \leq -\underline{x}$ is equivalent to $\rho_{\Phi}(X - x) \leq x - \underline{x}$ and therefore is consistent with the literature. Rogers (2009) considers a similar problem with law-invariant coherent risk measures, which is a proper subset of the WVaR risk measure. Moreover, both the feasibility and existence of the optimal solution and the impact on the portfolio choice are absent from his paper.

4 | SOLUTION

In this section, we present the solution to the WVaR-RM (3.4). The second constraint in (3.4) is based on the quantile function instead of the terminal wealth itself; thus, the standard convex duality method that is employed to solve the expected utility maximization problem cannot be applied directly. To overcome this difficulty, we employ the so-called quantile formulation. To this end, we impose the following assumption on ξ,

Assumption 4.1. *ξ is atomless.*

This assumption is satisfied when the investment opportunity set, i.e., the triplet $(r(\cdot), b(\cdot), \sigma(\cdot))$, is deterministic, $\int_0^T |\theta(t)|^2 dt \neq 0$, and ξ is lognormally distributed (which is the case with a Black–Scholes market).

Remark 4.2. *The problem with an atomic pricing kernel can be solved by following the method in Xu (2014).*

The basic idea of the quantile formulation technique is to change the decision variable of a portfolio choice problem from terminal payoffs to their quantile functions. The details of the quantile formulation are provided in Appendix A.

Let $F_{\xi}(\cdot)$ denote the CDF of ξ, and let $F_{\xi}^{-1}(\cdot)$ denote the quantile function of ξ, which is strictly increasing as ξ is atomless. Let us introduce the following assumption,

Assumption 4.3. $\operatorname{ess\,inf} \xi = 0$, $\operatorname{ess\,sup} \xi = +\infty$, *i.e.,* $F_{\xi}^{-1}(0) = 0$ *and* $F_{\xi}^{-1}(1) = +\infty$.

This assumption stipulates that for any given positive value, there is a state of nature in which the market offers a return that is greater (less) than that value. In particular, this assumption is valid when the investment opportunity set, i.e., the triplet $(r(\cdot), b(\cdot), \sigma(\cdot))$, is deterministic, $\int_0^T |\theta(t)|^2 dt \neq 0$, and ξ is lognormally distributed (which is the case with a Black–Scholes market).

We will first present the feasibility and well-posedness of the problem. Next, we will characterize the optimal solution when it exists.

Before we proceed, let us introduce the following set function:

$$\kappa_\Phi(A) = \frac{\Phi(A)}{\int_A F_\xi^{-1}(1-z)dz}, \quad \forall A \in \mathcal{B}[0,1],$$

where $\mathcal{B}[0,1]$ denotes all Borel-measurable sets in $[0,1]$. κ_Φ turns out to be closely related to the feasibility, existence, and properties of the optimal solution.

κ_Φ measures the *risk reduction per cost* (RRPC) across different states of the market at time T. Consider an economic agent who optimally chooses a terminal wealth X starting from an initial wealth x. It is illustrated in Appendix A that if the objective function and constraints (except for the initial budget constraint) in a portfolio choice problem are law invariant and the objective function is improved with a higher level of terminal wealth, which is indeed our case, then we can consider only the terminal wealth of the form $X = G_X(1 - F_\xi(\xi))$, where G_X is the quantile function of X. The risk of X is $\rho_\Phi(X) = -\int_{[0,1]} G_X(z)\Phi(dz)$. Consider a contingent claim whose payoff at time T is $\varepsilon \mathbf{1}_{1-F_\xi(\xi)\in[a,b]}, \varepsilon > 0, 0 \le a < b \le 1$. Its cost at time 0 is given by $\mathbb{E}[\xi\varepsilon\mathbf{1}_{1-F_\xi(\xi)\in[a,b]}] = \varepsilon \int_{[a,b]} F_\xi^{-1}(1-z)dz$. If the agent purchases this contingent claim with additional initial wealth $\varepsilon \int_{[a,b]} F_\xi^{-1}(1-z)dz$ at time 0, then his terminal wealth becomes $X_{\varepsilon,a,b} := X + \varepsilon\mathbf{1}_{1-F_\xi(\xi)\in[a,b]}$. A rough approximation for the quantile function of $X_{\varepsilon,a,b}$ is $G_X(z) + \varepsilon\mathbf{1}_{z\in[a,b]}$, when ε and $b-a$ are sufficiently small. We then have $\rho_\Phi(X_{\varepsilon,a,b}) \approx -\int_{[0,1]}(G_X(z) + \varepsilon\mathbf{1}_{z\in[a,b]})\Phi(dz) = -\int_{[0,1]} G_X(z)\Phi(dz) - \varepsilon\Phi([a,b])$. Thus, the agent can reduce the risk of his terminal wealth (approximately) by $\varepsilon\Phi([a,b])$ at an extra cost of $\varepsilon\int_{[a,b]} F_\xi^{-1}(1-z)dz$. $\kappa_\Phi([a,b])$ is the ratio between the claim's risk reduction and cost and thus measures the trade-off between reducing the risk and incurring the cost by investing in the future state $\{1 - F_\xi(\xi) \in [a,b]\}$. High a and b correspond to good states of the market, because these states are associated with low ξ.

4.1 | Feasibility and well-posedness

An optimization problem is feasible if it admits at least one feasible solution (i.e., a solution that satisfies all the constraints involved), and it is well-posed if it has a finite optimal value. A feasible solution is optimal if it achieves the finite optimal value. See Jin et al. (2008) for a detailed discussion of these terminologies in the context of portfolio selection.

The feasibility of (3.4) depends on the solution to the following problem

$$\begin{aligned} C_\Phi(x) := \max_X \quad & -\rho_\Phi(X) \\ \text{subject to } & \mathbb{E}[\xi X] \le x. \end{aligned} \tag{4.1}$$

$-C_\Phi(x)$ is the minimal risk of the terminal wealth X that an investor can achieve with an initial wealth x.

Proposition 4.4. *We have the following assertions:*

1. *If* $\sup_{0<c<1} \kappa_\Phi((c,1]) > \frac{1}{\mathbb{E}\xi}$, *then* $C_\Phi(x) = +\infty$.

2. *If* $\sup_{0<c<1} \kappa_\Phi((c,1]) \le \frac{1}{\mathbb{E}\xi}$, *then* $C_\Phi(x) = \frac{x}{\mathbb{E}\xi}$.

Define

$$\Delta_1 := \{(x, \underline{x}) : x > 0,\ 0 < \underline{x} < C_\Phi(x)\}. \tag{4.2}$$

Because of Proposition 4.4, (3.4) is feasible if $(x, \underline{x}) \in \Delta_1$, and it is infeasible if $\underline{x} > C_\Phi(x)$. Accordingly, from now on, we focus only on models with $(x, \underline{x}) \in \Delta_1$. We do not include the boundary $\underline{x} = C_\Phi(x) = \frac{x}{\mathbb{E}\xi}$, because the Lagrange method could fail on the boundary; see Remark B.5.

Finally, as the optimal value of the unconstrained problem (3.1) is always larger than or equal to that of the constrained problem (3.4), we know that the optimal value of (3.4) is finite, provided it is feasible and (3.1) is well-posed.

4.2 | Optimal solution

Before we present the optimal solution, let us discuss when the risk constraint is nonredundant. Recall that X_0^* is the optimal solution to (3.1). If $\rho_\Phi(X_0^*) \le -\underline{x}$, then X_0^* is already optimal to (3.4). In other words, the risk constraint is nonredundant if and only if $\rho_\Phi(X_0^*) > -\underline{x}$. Define

$$R_\Phi(x) := -\rho_\Phi((u')^{-1}\left(\lambda_0^*\xi\right)) \text{ where } \lambda_0^* \text{ solves } \mathbb{E}[\xi(u')^{-1}\left(\lambda_0^*\xi\right)] = x, \tag{4.3}$$

and

$$\Delta_2 := \{(x, \underline{x}) : \underline{x} > R_\Phi(x)\} \cap \Delta_1. \tag{4.4}$$

$-R_\Phi(x)$ is the risk of the optimal terminal wealth X_0^* of an investor with an initial wealth x, in the absence of the risk constraint. To exclude trivial cases, we further make the following assumption on Φ,

Assumption 4.5. $\Phi(\{1\}) = 0$ *and* $\kappa_\Phi(A)$ *is not constant for all* $A \in \mathcal{B}[0,1]$.

Remark 4.6.

1. *If* $\Phi(\{1\}) > 0$, *then* $R_\Phi(x) = +\infty$, *and the risk constraint is always satisfied by* X_0^*. *Thus, the risk constraint is redundant.*
2. *If* $\kappa_\Phi(A) = c$ *for all* $A \in \mathcal{B}[0,1]$ *where* c *is a positive constant, then* $\Phi([0,z)) = c\int_{[0,z)} F_\xi^{-1}(1-s)ds,\ z \in [0,1]$. *It is illustrated in Appendix A that if the objective function and constraints (except for the initial budget constraint) in a portfolio choice problem are law invariant and the objective function is improved with a higher level of the terminal wealth, which is indeed our case, then we can consider only the terminal wealth of the form* $X = G_X(1 - F_\xi(\xi))$, *where* G_X *is the quantile function of* X. *We then have* $\rho_\Phi(X) = -\int_{[0,1]} G_X(z)\Phi(dz) = -c\int_{[0,1]} G_X(z)F_\xi^{-1}(1-z)dz = -c\mathbb{E}[\xi X]$, *and the risk constraint becomes a multiple of the budget constraint. In this case, one of the constraints is redundant.*

Let us introduce several notations.

Define

$$\varphi_{\lambda_2}(z) := -\int_{[z,1)} F_\xi^{-1}(1-s)ds + \lambda_2\Phi((z,1]),\ z \in [0,1),$$

with $\varphi_{\lambda_2}(1) := 0$, and its left-continuous version

$$\varphi_{\lambda_2}(z-) := -\int_{[z,1)} F_\xi^{-1}(1-s)ds + \lambda_2\Phi([z,1]),\ z \in [0,1),$$

with $\varphi_{\lambda_2}(1-) := 0$, where we have used $\Phi(\{1\}) = 0$.

Denote the concave envelope of $\varphi_{\lambda_2}(\cdot-)$ by $\delta_{\lambda_2}(\cdot)$, in other words

$$\delta_{\lambda_2}(z) := \sup_{0\le a\le z\le b\le 1} \frac{(b-z)\varphi_{\lambda_2}(a-) + (z-a)\varphi_{\lambda_2}(b-)}{b-a},\ z \in [0,1]. \tag{4.5}$$

Let

$$\begin{aligned}\mathcal{A}_\Phi &:= \{\lambda : \lambda > 0, \delta_\lambda'(z+) > 0,\ \forall z \in [0,1)\}\\ &= \{\lambda : \lambda > 0, \varphi_\lambda(z-) < 0,\ \forall z \in [0,1)\},\end{aligned}$$

where $\delta_\lambda'(\cdot+)$ is the right derivative of δ_λ.

For any given $\lambda_2 \in \{0\} \cup \mathcal{A}_\Phi$, define

$$f_{\lambda_2}(\lambda_1) := \int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta_{\lambda_2}'(z+)\right) F_\xi^{-1}(1-z)dz,\ \lambda_1 > 0,$$

and

$$g(\lambda_2, x) := f_{\lambda_2}^{-1}(x),\ x > 0.$$

For any $\lambda_2 \in \{0\} \cup \mathcal{A}_\Phi,\ x > 0$, define

$$h(\lambda_2, x) := \int_{[0,1)} (u')^{-1}\left(g(\lambda_2, x)\delta_{\lambda_2}'(z+)\right) \Phi(dz),$$

and

$$S_\Phi(x) = \sup_{\lambda_2 \in \mathcal{A}_\Phi} h(\lambda_2, x). \tag{4.6}$$

We can now characterize the solutions to (3.4).

Theorem 4.7. *With C_Φ, R_Φ, and S_Φ defined by (4.3), (4.1), and (4.6), respectively, we have the following:*

1. *When* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$*:*
   - **(a)** *If $\underline{x} \le R_\Phi(x)$, then the optimal solution to (3.4) is given by (3.2).*
   - **(b)** *If $\underline{x} > R_\Phi(x)$, then there is no optimal solution.*
2. *When* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) < \infty$*:*
   - **(a)** *If $\underline{x} \le R_\Phi(x)$, the optimal solution is given by (3.2).*

   If, in addition,

$$\mathbb{E}\left[(u')^{-1}\left(\lambda_1 \delta_{\lambda_2}'((1-F_\xi(\xi))+)\right)\xi\right] < \infty, \tag{4.7}$$

   for all $\lambda_1 > 0$ and $\lambda_2 \in \mathcal{A}_\Phi$, we have the following:
   - **(b)** *If $R_\Phi(x) < \underline{x} < S_\Phi(x)$, the optimal solution is given by*

$$X^* = (u')^{-1}\left(\lambda_1^* \delta_{\lambda_2^*}'((1-F_\xi(\xi))+)\right), \tag{4.8}$$

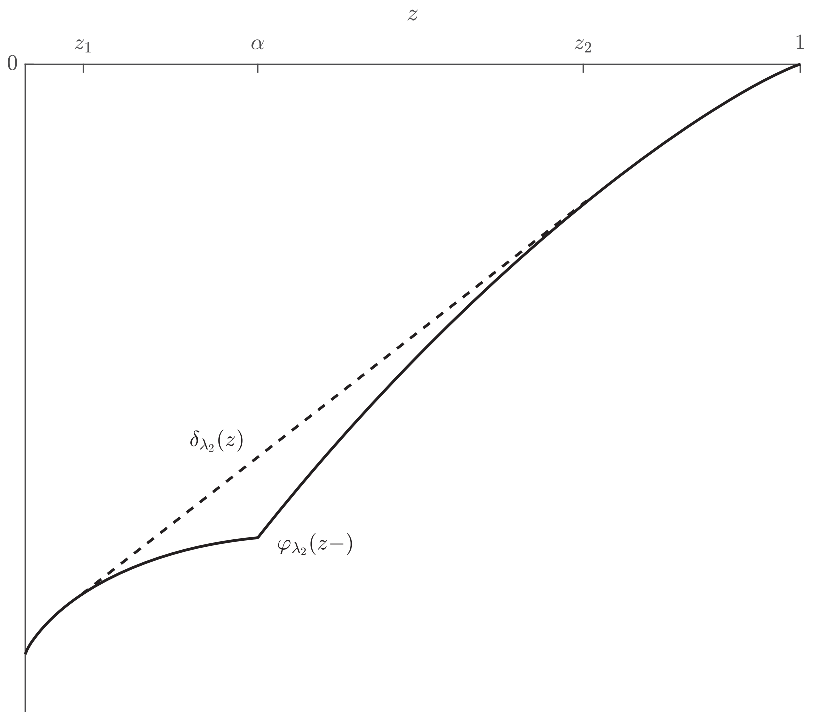

FIGURE 4.1 $\varphi_{\lambda_2}(\cdot-)$ and its concave envelope $\delta_{\lambda_2}(\cdot)$

where $\delta'_{\lambda_2}(\cdot+)$ is the right derivative of $\delta_{\lambda_2}(\cdot)$, and λ_1^* and λ_2^* solve

$$\mathbb{E}[\xi X^*] = x,$$
$$\rho_\Phi(X^*) = -\underline{x}. \tag{4.9}$$

(c) *If $S_\Phi(x) < \underline{x} < C_\Phi(x)$, then there is no optimal solution.*

Remark 4.8.

1. *If $F_\xi^{-1}(\cdot)$ is differentiable, $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$, and there exists $z_\Phi \in [0,1)$ such that $\Phi([z,1))$ (as a function of z) is twice differentiable on $(z_\Phi, 1)$, then (4.7) holds. The differentiability of $F_\xi^{-1}(\cdot)$ is satisfied with a lognormal ξ. The twice differentiability of $\Phi([z,1))$ for z near 1 is satisfied with all of the aforementioned risk measures. Moreover, as seen later, $0 < \liminf_{z\uparrow 1} \kappa_\Phi([z,1)) < \infty$ holds only on rare occasions.*
2. *$S_\Phi(x)$ is the supremum of $\underline{x}$ that can ensure the existence of the optimal solution to (3.4). For $\underline{x} = S_\Phi(x)$, the existence of the optimal solution depends on the attainability of $S_\Phi(x)$; see Remark B.15. In general, $S_\Phi(x)$ and $C_\Phi(x)$ can be different. $C_\Phi(x)$ is related to the feasibility of (3.4) and does not depend on the agent's preference, whereas $S_\Phi(x)$ is related to the existence of the optimal solution and does depend on the agent's preference.*
3. *The Lagrange multipliers, i.e., the solutions to (4.9), might not be unique. However, they must result in the same X^*, as the optimal solution (if it exists) must be unique due to the strict concavity of u. We can take any λ_1^* and λ_2^* that solve (4.9), and they will always result in the same X^*.*

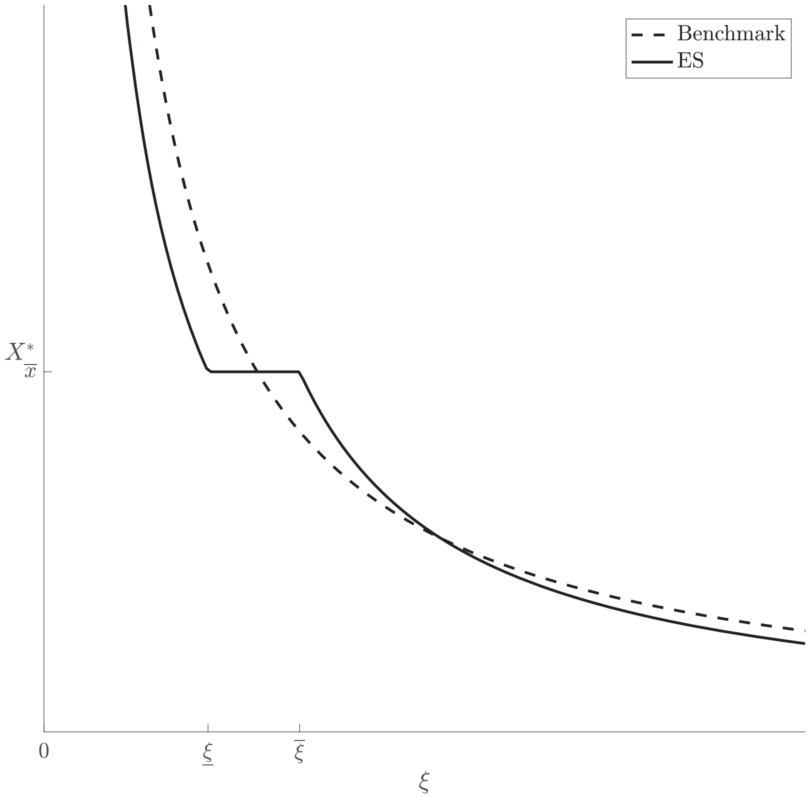

FIGURE 4.2 Optimal terminal wealth under ES. The figure plots the optimal terminal wealth of a benchmark agent (without the risk constraint, dashed line) and an ES agent (with an ES constraint, solid line), with the same initial wealth x, as functions of the horizon state price density ξ. Here, $\underline{\xi} := F_\xi^{-1}(1 - z_2^*)$, $\overline{\xi} := F_\xi^{-1}(1 - z_1^*)$, and $\overline{x} := (u')^{-1}(\lambda_1^* F_\xi^{-1}(1 - z_2^*)) = (u')^{-1}(\lambda_1^* \underline{\xi})$. When ξ is sufficiently large, the terminal wealth of the ES agent is lower than that of the benchmark agent

Before we close this section, we use an example to illustrate Theorem 4.7. We solve (3.4) when the risk constraint is given by $\mathrm{ES}_\alpha(X) \le -\underline{x}$. We assume that the ES constraint is binding and graphically show how the optimal wealth can be obtained.

Example 4.9. *For the constraint $ES_\alpha(X) \le -\underline{x}$, we have $\Phi([z, 1]) = (1 - \frac{z}{\alpha}) \vee 0$ and*

$$\varphi_{\lambda_2}(z-) = \begin{cases} -\int_{[z,1)} F_\xi^{-1}(1-s)ds + \lambda_2(1 - \frac{z}{\alpha}), & z \in [0, \alpha]; \\ -\int_{[z,1)} F_\xi^{-1}(1-s)ds, & z \in (\alpha, 1]. \end{cases}$$

The solid line in Figure 4.1 plots $\varphi_{\lambda_2}(\cdot-)$. Its concave envelope $\delta_{\lambda_2}(\cdot)$ replaces part of $\varphi_{\lambda_2}(\cdot-)$ with a chord (dashed line) between z_1 and z_2, at which the slopes of $\varphi_{\lambda_2}(\cdot-)$ equal the slope of the chord, in other words,

$$\delta_{\lambda_2}(z) = \begin{cases} \varphi_{\lambda_2}(z-), & z \in [0, z_1]; \\ F_\xi^{-1}(1 - z_2)(z - z_1) + \varphi_{\lambda_2}(z_1-), & z \in (z_1, z_2]; \\ \varphi_{\lambda_2}(z-), & z \in (z_2, 1], \end{cases}$$

where $z_1 < \alpha$ *and* $z_2 > \alpha$ *satisfy the following condition*

$$\begin{cases} \delta_{\lambda_2}(z_2) = \varphi_{\lambda_2}(z_2-), \\ F_\xi^{-1}(1-z_1) - \frac{\lambda_2}{\alpha} = F_\xi^{-1}(1-z_2). \end{cases}$$

The optimal terminal wealth is then given by

$$X^* = \begin{cases} (u')^{-1}\left(\lambda_1^*\left(\xi - \frac{\lambda_2^*}{\alpha}\right)\right), & \xi \in \left(F_\xi^{-1}\left(1-z_1^*\right), +\infty\right); \\ (u')^{-1}\left(\lambda_1^* F_\xi^{-1}\left(1-z_2^*\right)\right), & \xi \in \left(F_\xi^{-1}\left(1-z_2^*\right), F_\xi^{-1}\left(1-z_1^*\right)\right]; \\ (u')^{-1}\left(\lambda_1^*\xi\right), & \xi \in \left(0, F_\xi^{-1}\left(1-z_2^*\right)\right], \end{cases}$$

where $z_1^* < \alpha, z_2^* > \alpha, \lambda_1^* > 0$*, and* $\lambda_2^* > 0$ *satisfy the following condition*

$$\begin{cases} \delta_{\lambda_2^*}\left(z_2^*\right) = \varphi_{\lambda_2^*}\left(z_2^*-\right), \\ F_\xi^{-1}\left(1-z_1^*\right) - \frac{\lambda_2^*}{\alpha} = F_\xi^{-1}\left(1-z_2^*\right), \\ \mathbb{E}[\xi X^*] = x, \\ ES_\alpha(X^*) = -\underline{x}. \end{cases}$$

The solid line in Figure 4.2 plots the optimal terminal wealth under the ES constraint. The dashed line represents the optimal terminal wealth of a benchmark agent who solves (3.1).

5 | PROPERTIES

In this section, we perform a detailed analysis of trading behaviors under different risk measures. WVaR-RM exhibits a much richer variety of investment behaviors than its mean-risk counterpart (He et al., 2015). He et al. (2015) find that the mean-WVaR is likely to be ill-posed, and the asymptotically optimal strategy is binary, which is to bank most of the money and invest the remainder in an extremely risky but highly rewarding lottery. In the WVaR-RM, the presence of the utility offers a variety of interesting features. We first study the existence of optimal solutions and then characterize risk-taking behaviors.

5.1 | Existence of optimal solutions

Because the optimal value of (3.4) is always finite, we claim that if the optimal solution does not exist, the model is unattainable, i.e., the optimality of (3.4) cannot be achievable by any admissible portfolio.

When the risk constraint is active, Theorem 4.7 characterizes a class of risk measures that will lead to unattainability. When $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$, i.e., the RRPC is infinity in the extremely good states as the risk measure places too many weights, the optimal solution does not exist for all non-trivial levels of risk constraints. When the risk is measured by this type of risk measure, investments in the extremely good states are not only the cheapest but also the most efficient way to reduce the risk. The agent is thus incentivized to meet the risk constraint by assuming the greatest possible risk exposure. However, the risk aversion (implied by the strictly concave utility) prevents the agent from taking extremely risky positions. Such a conflict will result in an unattainable model.

We now examine this circumstance for the risk measures discussed in Section 2.

Proposition 5.1. *We have the following assertions:*

1. *For the negative expectation,* $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$.
2. *For the VaR*$_\alpha$, $0 < \alpha < 1$, $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$.
3. *For the ES*$_\alpha$, $0 < \alpha < 1$, $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$.
4. *For the exponential spectral risk measures,* $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$.
5. *For the power spectral risk measures with* $0 < \gamma < 1$, $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$.

If we further assume that ξ *is log-normally distributed, i.e.,* $F_\xi(x) = \Phi_N(\frac{\ln x - \mu_\xi}{\sigma_\xi})$, *for some* μ_ξ *and* $\sigma_\xi > 0$, *we have the following:*

6. *For the power spectral risk measures with* $\gamma > 1$, $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$.
7. *For the Wang (2000) risk measure,*

$$\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \begin{cases} 0, & q < \Phi_N(-\sigma_\xi); \\ e^{-\frac{1}{2}(\Phi^{-1}(q))^2 - \mu_\xi}, & q = \Phi_N(-\sigma_\xi); \\ +\infty, & q > \Phi_N(-\sigma_\xi). \end{cases}$$

8. *For the beta family of distortion risk measures,*

$$\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \begin{cases} 0, & b > 1; \\ +\infty, & 0 < b \leq 1. \end{cases}$$

Coherent risk measures are believed to be better alternatives to traditional risk measures. However, the above analysis reveals that even law-invariant, coherent, comonotonic additive risk measures can still be inappropriate in the context of portfolio selection. In contrast, distortion risk measures that are widely used in the actuarial sciences, depending on the parameters, might or might not be appropriate for the purpose of risk management. Economic agents should use caution when adopting these risk measures.

Overall, we suggest that a "good" risk measure, for the purpose of risk management in asset allocation, should only focus on the downside risk.

5.2 | Impacts on asset allocation

We now give a detailed analysis of how risk measures affect asset allocation, assuming that the risk constraint is active and the optimal solution exists. Throughout this section, we assume that $F_\xi^{-1}(\cdot)$ is differentiable. We focus on risk measures such that $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$ and $\Phi([z,1))$ (as a function of z) is twice differentiable on $(z_\Phi, 1)$ for some $z_\Phi \in [0,1)$, and we assume that $R_\Phi(x) < \underline{x} < S_\Phi(x)$. Consequently, X^*, the optimal solution to (3.4), is given by (4.8).

First, we focus on the potential gains from the stock market.

Proposition 5.2. $\operatorname{ess\,sup} X^* = +\infty$.

In the mean-WVaR (He et al., 2015), although the investors can benefit from the stock market, the reward is capped and fixed. The investors receive the same amount of reward when the market is sufficiently good, regardless of how good it is, which makes such strategies less appealing. In contrast, under the WVaR-RM, the potential gains from the stock market are unbounded. This could be attractive to some investors: Although risk management is costly, they can still participate in the potential unlimited gains.

Gains and losses are always associated. Although the gains under WVaR-RM are unbounded irrespective of the risk measures, the losses are in a more complex situation. We first characterize a class of risk measures that can give rise to endogenous portfolio insurance. A portfolio insurance trading strategy is defined as one that guarantees a minimum level of wealth at some specified horizon, yet also participates in the potential gains of some reference portfolio (Grossman & Vila, 1989; Luskin, 1988). Under portfolio insurance, the agent's downside risk is significantly reduced because all losses are capped at a prescribed level.

Proposition 5.3. *If* $\liminf_{z\downarrow 0} \kappa_\Phi([0,z)) = +\infty$, *then* $\operatorname{ess\,inf} X^* > 0$.

This proposition says that if a risk measure's RRPC in the extremely bad states (such as a catastrophic loss) is infinity, then the agent will insure against these states endogenously. Even though these states are the most expensive states to insure against, the risk constraint incentivizes the agent to follow the portfolio insurance strategy because it is the most efficient way to satisfy the requirement. This could be of interest to regulators, because portfolio insurance is, in general, costly. As noted in Leland (1980) and Benninga and Blume (1985), it is highly unlikely that an investor would utilize such a strategy in a complete market. Regulators can encourage economic agents to do so by imposing a risk constraint of this type.

We also provide additional evidence of why people buy portfolio insurance. Although much work has been conducted on the effects on prices of the presence of portfolio insurance in the economy (Basak, 1995, 2002; Grossman & Vila, 1989; Grossman & Zhou, 1996), in which investors are assumed to be portfolio insurers, justifications for the necessity of portfolio insurance are limited, especially in complete markets. To the best of our knowledge, the only work on why investors use portfolio insurance in complete markets is He and Zhou (2016), who find in a rank-dependent portfolio choice model that a sufficiently high level of fear endogenously necessitates portfolio insurance. We suggest that people can use portfolio insurance strategies as a means to manage their market-risk exposure.

However, not all risk measures within WVaR can lead to portfolio insurance. The following theorem characterizes another class of risk measures that can increase risk exposure because they result in larger losses when losses occur.

Proposition 5.4. *If* $\liminf_{z\downarrow 0} \kappa_\Phi([0,z)) = 0$, *then*

1. $\operatorname{ess\,inf} X^* = 0$.
2. *If, in addition,* $\limsup_{z\downarrow 0} \kappa_\Phi([0,z)) = 0$ *and there exists* $\underline{z}_\Phi \in (0,1]$ *such that* $\Phi([0,z))$ *(as a function of z) is twice differentiable on* $(0,\underline{z}_\Phi)$, *then there exists* $\bar{\xi}$ *such that* $X^* < X_0^*$ *when* $\xi > \bar{\xi}$, *where* X_0^* *is the benchmark agent's optimal wealth given by (3.2).*

If a risk measure's RRPC for catastrophic losses is 0, then the agent will ignore these losses and leave himself completely uninsured, incurring all losses, because it is costly and inefficient to insure against these losses. Moreover, in the bad states ($\xi > \bar{\xi}$), the terminal wealth is typically lower than it would have been in the absence of the risk constraint. In other words, under such regulations, if a large loss occurs, then it is likely to be an even larger loss compared to the benchmark agent and consequently, the probability of extreme losses is higher. The economic agent exploits differences between a portfolio's true economic risks and the measurements of risk. This perverse consequence is often referred to as "regulatory capital arbitrage" (Jones, 2000) that banks can reduce substantially their regulatory measures of risk, with little or no corresponding reduction in their overall economic risks.

This could be a source of concern for regulators and real-world risk managers. Risk measures are viewed by many as a tool to shield economic agents from large losses that could drive them out of business. However, many risk measures, although they have some desirable properties such as law-invariance, coherence, and comonotonic-additivity, could backfire, and thus would be more likely to lead to credit and solvency problems, defeating the purpose of such regulations. Such an undesirable property has been observed in Basak and Shapiro (2001) for VaR, but to the best of our knowledge, we are the first to characterize "regulatory capital arbitrage" for general risk measures.

Based on the above characterizations, we now examine the risk measures discussed in Section 2. It is easy to see that there exists $\underline{z}_\Phi \in (0, 1]$ such that $\Phi([0, z))$ (as a function of z) is twice differentiable on $(0, \underline{z}_\Phi)$ for all of the aforementioned risk measures.

Proposition 5.5. *We have the following assertions:*

1. *For the* VaR_α, $0 < \alpha < 1$, $\liminf_{z\downarrow 0} \kappa_\Phi([0, z)) = 0$.
2. *For the* ES_α, $0 < \alpha < 1$, $\liminf_{z\downarrow 0} \kappa_\Phi([0, z)) = 0$.
3. *For the power spectral risk measures with* $\gamma > 1$, $\liminf_{z\downarrow 0} \kappa_\Phi([0, z)) = 0$.

If we further assume that ξ *is log-normally distributed, i.e.,* $F_\xi(x) = \Phi_N(\frac{\ln x - \mu_\xi}{\sigma_\xi})$, *for some* μ_ξ *and* $\sigma_\xi > 0$, *then we have the following:*

4. *For the Wang (2000) risk measure with* $q < \Phi_N(-\sigma_\xi)$, $\liminf_{z\downarrow 0} \kappa_\Phi([0, z)) = +\infty$.
5. *For the beta family of distortion risk measures with* $b > 1$,

$$\liminf_{z\downarrow 0} \kappa_\Phi([0, z)) = \begin{cases} 0, & a \geq 1; \\ +\infty, & 0 < a < 1. \end{cases}$$

To our surprise, distortion risk measures such as the Wang (2000) risk measure and the beta family of distortion risk measures, although originally designed as premium principles and capital adequacy principles, can give rise to endogenous portfolio insurance and thus reduce the magnitude of losses (with a suitable choice of parameters) because they account for the severity of extreme losses. Figure 5.1 plots the optimal terminal wealth under the Wang (2000) risk measure. In the bad states of the market, the agent behaves like a portfolio insurer and his terminal wealth is higher than that of a benchmark agent when ξ is sufficiently large.

However, VaR, ES, and some well-known spectral risk measures actually increase risk exposure in the bad states. The case of VaR is consistent with Basak and Shapiro (2001) and the empirical evidence provided by Berkowitz and Obrien (2002), who document that when a bank suffers losses, such losses are often substantially larger than the bank's reported VaR. ES is proposed as an effective alternative to VaR in financial risk management (Acerbi, Nordio, & Sirtori, 2001; Artzner et al., 1999). It is believed that ES will help to ensure a more prudent capture of "tail risk" and capital adequacy during periods of significant financial market stress (BCBS, 2016). However, we show that this objective might not be achieved because ES actually increases the magnitude of the loss when a loss occurs instead of reducing it. Figure 4.2 plots the optimal terminal wealth under the ES constraint (solid line). When the market is sufficiently bad, the ES agent suffers from larger losses than a benchmark agent. One common criticism of VaR is that it fails to account for the magnitude of losses. Our results reveal that even though ES accounts for the sizes of the losses, it is far from adequate. ES places equal weights for all levels of losses that exceed a certain threshold. However, the costs to insure against these losses differ, and it is costlier to insure against a larger loss. Thus, the agent finds it inefficient to insure against catastrophic losses when the risk is measured by ES.

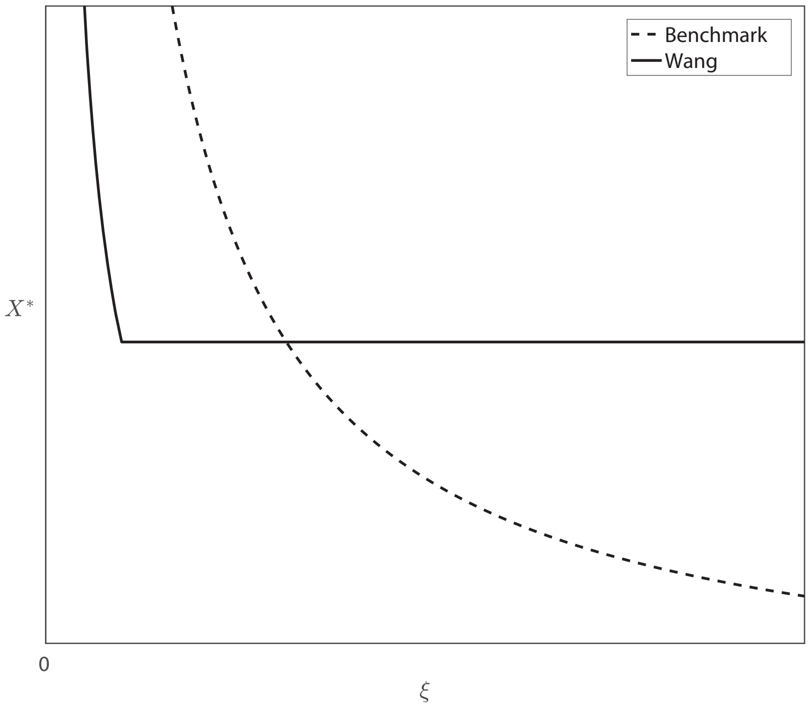

FIGURE 5.1 Optimal terminal wealth under the Wang (2000) risk measure. The figure plots the optimal terminal wealth of a benchmark agent (without the risk constraint, dashed line) and a Wang agent (with a risk constraint and the risk is measured by the Wang (2000) risk measure, solid line), with the same initial wealth x, as functions of the horizon state price density ξ. We assume that $F_\xi(x) = \Phi_N(\frac{\ln x - \mu_\xi}{\sigma_\xi})$ and $q < \Phi_N(-\sigma_\xi)$. When ξ is sufficiently large, the terminal wealth of the Wang agent is higher than that of the benchmark agent.

Overall, to mitigate or even preclude "regulatory capital arbitrage" in asset management, we suggest that the regulatory risk measure's sensitivity to losses should be relevant to the severity of the losses. Distortion risk measures, such as the Wang (2000) risk measure and the beta family of distortion risk measures, should be preferred over current regulatory risk measures such as VaR and ES.

6 | CONCLUSIONS

We study the effects of the WVaR-based risk management on the portfolio choice of expected utility maximizers, who derive utility from wealth at some horizon and must comply with a WVaR constraint imposed at that horizon. The feasibility, well-posedness, and existence of optimal solutions are discussed. When the optimal solution exists, we reveal several interesting effects. We characterize a class of risk measures that allows economic agents to engage in "regulatory capital arbitrage." In particular, VaR and ES, two popular regulatory risk measures, often incur even larger losses in the most adverse states. This provides a critique of the current risk management practices and the Basel Committee's plan to replace VaR with ES for calculating market risk capital requirements. On the other hand, we provide conditions on risk measures that can lead to endogenous portfolio insurance and thus mitigate "regulatory capital arbitrage." These findings may be of potential interest to regulators.

Although we show how various risk measures can alter asset allocation patterns, our analysis is in a partial equilibrium setting. It is equally or even more important to study how they can affect the market price dynamics. In a related paper (Wei, 2017), we have analyzed both the partial equilibrium and the general equilibrium of an economy that features agents who must manage their ES-measured risks.

ACKNOWLEDGMENTS

I am grateful to helpful comments from Suleyman Basak, Hanqing Jin, Michael Monoyios, Xunyu Zhou, seminar participants at Oxford, as well as conference participants at the 10th Actuarial and Financial Mathematics Conference in Brussels. I am also indebted to Jerome Detemple (the editor) and two anonymous referees for their suggestions and comments. All errors are my own. I acknowledge financial support from the Oxford-Man Institute of Quantitative Finance and the Oxford-Nie Financial Big Data Lab.

ENDNOTES

[1] For example, as stated in BCBS (2016), one of the key enhancements in the revised market risk framework is "a shift from Value-at-Risk (VaR) to an Expected Shortfall (ES) measure of risk under stress. Use of ES will help to ensure a more prudent capture of 'tail risk' and capital adequacy during periods of significant financial market stress."

[2] A4 and A5 imply A6.

[3] In the literature, risk measures are often defined on L^∞ and are assumed to satisfy only A1–A2. In this paper, I consider risk measures on LB as in He et al. (2015). The truncation continuity is thus imposed to guarantee that the risk of any unbounded P&L can be computed through its truncations. To verify that ρ satisfies some of the aforementioned axioms, it suffices to show that these axioms are satisfied when ρ is restricted on L^∞ and the truncation continuity is satisfied.

[4] If the underlying distribution of X is a continuous distribution, then ES is equivalent to tail conditional expectation, which is defined as $\mathrm{TCE}_\alpha(X) = -\mathbb{E}[X|X \le -\mathrm{VaR}_\alpha(X)]$.

[5] Cherny (2006) also proposed a version of weighted VaR risk measures, which is a special case of He et al. (2015).

REFERENCES

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, *26*(7), 1505–1518.

Acerbi, C., Nordio, C., & Sirtori, C. (2001). *Expected shortfall as a tool for financial risk management*. Unpublished manuscript.

Acerbi, C., & Simonetti, P. (2002). *Portfolio optimization with spectral measures of risk*. Unpublished manuscript.

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking & Finance*, *26*(7), 1487–1503.

Adam, A., Houkari, M., & Laurent, J.-P. (2008). Spectral risk measures and portfolio selection. *Journal of Banking & Finance*, *32*(9), 1870–1882.

Alexander, G. J., & Baptista, A. M. (2002). Economic implications of using a mean-VaR model for portfolio selection: A comparison with mean-variance analysis. *Journal of Economic Dynamics and Control*, *26*(7), 1159–1193.

Alexander, G. J., & Baptista, A. M. (2004). A comparison of VaR and CVaR constraints on portfolio selection with the mean-variance model. *Management Science*, *50*(9), 1261–1273.

Alexander, G. J., & Baptista, A. M. (2006). Does the Basle Capital Accord reduce bank fragility? An assessment of the value-at-risk approach. *Journal of Monetary Economics*, *53*(7), 1631–1660.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, *9*, 203–228.

Basak, S. (1995). A general equilibrium model of portfolio insurance. *Review of Financial Studies*, *8*(4), 1059–1090.

Basak, S. (2002). A comparative study of portfolio insurance. *Journal of Economic Dynamics & Control*, *26*(7), 1217–1241.

Basak, S., & Chabakauri, G. (2010). Dynamic mean-variance asset allocation. *Review of Financial Studies*, *23*(8), 2970–3016.

Basak, S., & Shapiro, A. (2001). Value-at-risk-based risk management: Optimal policies and asset prices. *Review of Financial Studies*, *14*(2), 371–405.

Basak, S., Shapiro, A., & Teplá, L. (2006). Risk management with benchmarking. *Management Science*, *52*(4), 542–557.

BCBS. (2011). *Revisions to the Basel II market risk framework*. Basel Committee on Banking Supervision, Bank for International Settlements, Basel, Switzerland.

BCBS. (2016). *Standards January 2016. Minimum capital requirements for market risk*. Basel Committee on Banking Supervision, Bank for International Settlements, Basel, Switzerland.

Benninga, S., & Blume, M. (1985). On the optimality of portfolio insurance. *Journal of Finance*, *40*(5), 1341–1352.

Berkowitz, J., & Obrien, J. (2002). How accurate are value-at-risk models at commercial banks? *Journal of Finance*, *57*(3), 1093–1111.

Bertsimas, D., Lauprete, G. J., & Samarov, A. (2004). Shortfall as a risk measure: Properties, optimization and applications. *Journal of Economic Dynamics and Control*, *28*(7), 1353–1381.

Bielecki, T. R., Jin, H., Pliska, S. R., & Zhou, X. Y. (2005). Continuous-time mean–variance portfolio selection with bankruptcy prohibition. *Mathematical Finance*, *15*(2), 213–244.

Cahuich, L. D., & Hernández-Hernández, D. (2013). Quantile portfolio optimization under risk measure constraints. *Applied Mathematics & Optimization*, *68*(2), 157–179.

Campbell, R., Huisman, R., & Koedijk, K. (2001). Optimal portfolio selection in a value-at-risk framework. *Journal of Banking & Finance*, *25*(9), 1789–1804.

Carlier, G., & Dana, R.-A. (2006). Law invariant concave utility functions and optimization problems with monotonicity and comonotonicity constraints. *Statistics & Decisions*, *24*(1/2006), 127–152.

Carlier, G., & Dana, R.-A. (2011). Optimal demand for contingent claims when agents have law invariant utilities. *Mathematical Finance*, *21*(2), 169–201.

Cherny, A. S. (2006). Weighted V@R and its properties. *Finance and Stochastics*, *10*(3), 367–393.

Cotter, J., & Dowd, K. (2006). Extreme spectral risk measures: An application to futures clearinghouse margin requirements. *Journal of Banking & Finance*, *30*(12), 3469–3485.

Cox, J. C., & Huang, C.-F. (1989). Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, *49*(1), 33–83.

Cuoco, D., He, H., & Isaenko, S. (2008). Optimal dynamic trading strategies with risk limits. *Operations Research*, *56*(2), 358–368.

Cuoco, D., & Liu, H. (2006). An analysis of VaR-based capital requirements. *Journal of Financial Intermediation*, *15*(3), 362–394.

Dhaene, J., Vanduffel, S., Tang, Q., Goovaerts, M. J., Kaas, R., & Vyncke, D. (2004). Capital requirements, risk measures and comonotonicity. *Belgian Actuarial Bulletin*, *4*(1), 53–61.

Dowd, K. (1998). *Beyond value at risk: The new science of risk management*. Chichester, UK: Wiley.

Dowd, K., Cotter, J., & Sorwar, G. (2008). Spectral risk measures: Properties and limitations. *Journal of Financial Services Research*, *34*(1), 61–75.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R., & Beleraj, A. (2014). An academic response to Basel 3.5. *Risks*, *2*(1), 25–48.

Gabih, A., Grecksch, W., & Wunderlich, R. (2005). Dynamic portfolio optimization with bounded shortfall risks. *Stochastic Analysis and Applications*, *23*(3), 579–594.

Gabih, A., Sass, J., & Wunderlich, R. (2009). Utility maximization under bounded expected loss. *Stochastic Models*, *25*(3), 375–407.

Grossman, S. J., & Vila, J.-L. (1989). Portfolio insurance in complete markets: A note. *Journal of Business*, *62*(4), 473–476.

Grossman, S. J., & Zhou, Z. (1996). Equilibrium analysis of portfolio insurance. *Journal of Finance*, *51*(4), 1379–1403.

Gundel, A., & Weber, S. (2008). Utility maximization under a shortfall risk constraint. *Journal of Mathematical Economics*, *44*(11), 1126–1151.

He, X. D., Jin, H., & Zhou, X. Y. (2015). Dynamic portfolio choice when risk is measured by weighted VaR. *Mathematics of Operations Research*, *40*(3), 773–796.

He, X. D., & Zhou, X. Y. (2011). Portfolio choice via quantiles. *Mathematical Finance*, *21*(2), 203–231.

He, X. D., & Zhou, X. Y. (2016). Hope, fear, and aspirations. *Mathematical Finance*, *26*(1), 3–50.

Jin, H., Xu, Z. Q., & Zhou, X. Y. (2008). A convex stochastic optimization problem arising from portfolio selection. *Mathematical Finance*, *18*(1), 171–183.

Jin, H., Yan, J.-A., & Zhou, X. Y. (2005). Continuous-time mean–risk portfolio selection. *Annales de l'IHP Probabilités et statistiques*, *41*(3), 559–580.

Jin, H., & Zhou, X. Y. (2008). Behavioral portfolio selection in continuous time. *Mathematical Finance*, *18*(3), 385–426.

Jones, D. (2000). Emerging problems with the Basel Capital Accord: Regulatory capital arbitrage and related issues. *Journal of Banking & Finance*, *24*(1), 35–58.

Jorion, P. (1997). *Value at risk: The new benchmark for controlling market risk*. Chicago: Irwin Professional Publishing.

Jorion, P. (2002). How informative are value-at-risk disclosures? *The Accounting Review*, *77*(4), 911–931.

Karatzas, I., & Shreve, S. E. (1998). *Methods of mathematical finance* (Vol. 39). New York: Springer Science & Business Media.

Kast, R., Luciano, E., & Peccati, L. (1999, May). *Value-at-risk as a decision criterion*. Presented at the EURO Working Group on Financial Modeling Meeting, Paris.

Kramkov, D., & Schachermayer, W. (1999). The asymptotic elasticity of utility functions and optimal investment in incomplete markets. *Annals of Applied Probability*, *9*(3), 904–950.

Kupper, M., & Schachermayer, W. (2009). Representation results for law invariant time consistent functions. *Mathematics and Financial Economics*, *2*(3), 189–210.

Leippold, M., Trojani, F., & Vanini, P. (2006). Equilibrium impact of value-at-risk regulation. *Journal of Economic Dynamics and Control*, *30*(8), 1277–1313.

Leitner, J. (2005). A short note on second-order stochastic dominance preserving coherent risk measures. *Mathematical Finance*, *15*(4), 649–651.

Leland, H. E. (1980). Who should buy portfolio insurance? *Journal of Finance*, *35*(2), 581–594.

Luskin, D. L. (1988). *Portfolio Insurance: A guide to dynamic hedging*. New York: Wiley.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance*, *7*(1), 77–91.

Milgrom, P., & Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, *70*(2), 583–601.

Pliska, S. R. (1986). A stochastic calculus model of continuous trading: Optimal portfolios. *Mathematics of Operations Research*, *11*(2), 371–382.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk*, *2*, 21–42.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking & Finance*, *26*(7), 1443–1471.

Rogers, L. (2009). Optimal and robust contracts for a risk-constrained principal. *Mathematics and Financial Economics*, *2*(3), 151–171.

Saunders, A. (2000). *Financial institutions management: A modern perspective*. Boston: McGraw-Hill College.

Schied, A. (2004). On the Neyman–Pearson problem for law-invariant risk measures and robust utility functionals. *Annals of Applied Probability*, *14*(3), 1398–1423.

SEC. (1997). Disclosure of accounting policies for derivative financial instruments and derivative commodity instruments and disclosure of quantitative and qualitative information about market risk inherent in derivative financial instruments, other financial instruments, and derivative commodity instruments. In *Other financial instruments, and derivative commodity instruments*. Washington, DC: Author.

Sereda, E. N., Bronshtein, E. M., Rachev, S. T., Fabozzi, F. J., Sun, W., & Stoyanov, S. V. (2010). Distortion risk measures in portfolio optimization. In J. Guerard (Ed.), *Handbook of portfolio construction* (pp. 649–673). New York: Springer.

Wang, S. (1995). Insurance pricing and increased limits ratemaking by proportional hazards transforms. *Insurance: Mathematics and Economics*, *17*(1), 43–54.

Wang, S. (1996). Premium calculation by transforming the layer premium density. *Astin Bulletin*, *26*(1), 71–92.

Wang, S. S. (2000). A class of distortion operators for pricing financial and insurance risks. *Journal of Risk and Insurance*, *67*(1), 15–36.

Wang, S. S., Young, V. R., & Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance: Mathematics and Economics*, *21*(2), 173–183.

Wei, P. (2017). *Equilibrium analysis of expected shortfall*. Working paper, University of Oxford.

Wirch, J. L., & Hardy, M. R. (1999). A synthesis of risk measures for capital adequacy. *Insurance: Mathematics and Economics*, *25*(3), 337–347.

Xia, J., & Zhou, X. Y. (2016). Arrow–Debreu equilibria for rank-dependent utilities. *Mathematical Finance*, *26*(3), 558–588.

Xu, Z. Q. (2014). A new characterization of comonotonicity and its application in behavioral finance. *Journal of Mathematical Analysis and Applications*, *418*(2), 612–625.

Xu, Z. Q. (2016). A note on the quantile formulation. *Mathematical Finance*, *26*(3), 589–601.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, *55*(1), 95–115.

Yiu, K.-F. C. (2004). Optimal portfolios under a value-at-risk constraint. *Journal of Economic Dynamics and Control*, *28*(7), 1317–1334.

How to cite this article: Wei P. Risk management with weighted VaR. *Mathematical Finance*. 2018;28:1020–1060. https://doi.org/10.1111/mafi.12160

APPENDIX A: QUANTILE FORMULATION

Quantile formulation, developed in a series of papers including Schied (2004), Carlier and Dana (2006), Jin and Zhou (2008), He and Zhou (2011), Carlier and Dana (2011), Xia and Zhou (2016), and Xu (2016), is a technique of solving optimal terminal payoff in portfolio choice problems. This technique can be applied once the objective function and constraints, except for the initial budget constraint, in a portfolio choice problem are law invariant and the objective function is improved with a higher level of the terminal wealth (i.e., the more the better). The basic idea of the quantile formulation is to choose quantile functions as the decision variables. The advantage of this formulation is that the quantile-based risk constraint can be directly embedded into the optimization, and hence traditional optimization techniques become applicable. Furthermore, there is a simple connection between the optimal solution to the portfolio choice problem and its quantile formulation.

We first show that in our problem, the objective function is improved with a higher level of the terminal wealth.

Lemma A.1. *If X^* is optimal to (3.4), then $\mathbb{E}[\xi X^*] = x$.*

Proof of Lemma A.1. If $E[\xi X^*] < x$, then let $X = X^* + \frac{x - E[\xi X^*]}{E[\xi]} > X^*$. We have $E[\xi X] = x$, $\rho_\Phi(X) \geq -\underline{x}$, and thus, X satisfies both constraints. Because $u(\cdot)$ is strictly increasing, $\mathbb{E}[u(X)] > \mathbb{E}[u(X^*)]$, which is a contradiction. $\square$

From Lemma A.1, we also see that the optimal value of (3.4), if it exists, is strictly increasing in x.

We also need the following lemma from Jin and Zhou (2008),

Lemma A.2 (Jin & Zhou, 2008). $\mathbb{E}[\xi G(1-F_\xi(\xi))] \le \mathbb{E}[\xi X]$ *for any lower bounded random variable* X *whose quantile function is* G. *Furthermore, if* $\mathbb{E}[\xi G(1-F_\xi(\xi))] < \infty$, *then the inequality becomes equality if and only if* $X = G(1-F_\xi(\xi))$, *a.s.*

In view of Lemmas A.2 and A.1, we can consider the quantile formulation of (3.4). Define

$$U(G(\cdot)) := \int_{[0,1)} u(G(z))dz,\ G(\cdot) \in \mathbb{G}.$$

We consider the following problem:

$$\begin{aligned} V(x,\underline{x}) := \max_{G(\cdot)\in\mathbb{G}} \ & U(G(\cdot)) \\ \text{subject to } & \int_{[0,1)} F_\xi^{-1}(1-z)G(z)dz = x, \\ & \int_{[0,1]} G(z)\Phi(dz) \ge \underline{x}. \end{aligned} \tag{A.1}$$

The following theorem verifies the equivalence of the portfolio selection problem (3.4) and the quantile formulation (A.1) in terms of the feasibility, well-posedness, existence, and uniqueness of the optimal solution. The proof is similar to that in He and Zhou (2011).

Theorem A.3. *We have the following assertions.*

1. *Problem (3.4) is feasible (well-posed) if and only if problem (A.1) is feasible (well-posed). Furthermore, they have the same optimal value.*
2. *The existence (uniqueness) of optimal solutions to (3.4) is equivalent to the existence (uniqueness) of optimal solutions to (A.1).*
3. *If* X^* *is optimal to (3.4), then* $G^*(\cdot)$ *is optimal to (A.1). If* $G^*(\cdot)$ *is optimal to (A.1), then* $X^* = G^*(1-F_\xi(\xi))$ *is optimal to (3.4).*

APPENDIX B: PROOFS

B.1 Proof of Proposition 4.4

Proof of Proposition 4.4. In view of the previous analysis, it is equivalent to consider the following problem:

$$\begin{aligned} C_\Phi(x) = \max_{G(\cdot)\in\mathbb{G}} & \int_{[0,1]} G(z)\Phi(dz) \\ \text{subject to} & \int_{[0,1]} F_\xi^{-1}(1-z)G(z)dz = x. \end{aligned}$$

1. If $\sup_{0<c<1}\kappa_\Phi((c,1]) > \frac{1}{\mathbb{E}\xi}$, then there exists $c \in (0,1)$ such that $\kappa_\Phi((c,1]) > \frac{1}{\mathbb{E}\xi} + \varepsilon$ for some $\varepsilon > 0$. Consider

$$G_n(z) = \frac{x - n\int_{[c,1]} F_\xi^{-1}(1-z)dz}{\mathbb{E}\xi} + n\mathbf{1}_{c\le z\le 1},$$

we have

$$\begin{aligned}\int_{[0,1]} G_n(z)\Phi(dz) &= \frac{x - n\int_{[c,1]} F_\xi^{-1}(1-z)dz}{\mathbb{E}\xi} + n\Phi([c,1]) \\ &> x + \varepsilon n\int_{[c,1]} F_\xi^{-1}(1-z)dz,\end{aligned}$$

and the claim follows easily.

2. If $\sup_{0<c<1}\kappa_\Phi((c,1]) \le \frac{1}{\mathbb{E}\xi}$, then for any $G(\cdot)\in\mathbb{G}$, we have

$$\begin{aligned}\int_{[0,1]} G(z)\Phi(dz) &= \int_{[0,1]} \Phi([z,1])dG(z) + G(0-) \\ &\le \frac{1}{\mathbb{E}\xi}\int_{[0,1]} F_\xi^{-1}(1-z)G(z)dz \\ &= \frac{x}{\mathbb{E}\xi},\end{aligned}$$

and the equality holds when $G(z) = \frac{x}{\mathbb{E}\xi}$. □

B.2 Proof of Theorem 4.7

We apply the Lagrange dual method to solve (A.1). The proof of Theorem 4.7 is decomposed into four steps: First, establish the relationship between the Lagrangian dual problem and the original problem; second, solve the Lagrangian dual problem; third, prove the existence of Lagrange multipliers; and fourth, solve (A.1) by recalling Theorem A.3.

B.2.1 Lagrange approach

Define the Lagrangian

$$\begin{aligned}U_{\lambda_1,\lambda_2}(G(\cdot)) := &\int_{[0,1)} u(G(z))dz - \lambda_1\int_{[0,1)} G(z)F_\xi^{-1}(1-z)dz \\ &+ \lambda_1\lambda_2\int_{[0,1)} G(z)\Phi(dz),\end{aligned}$$

and consider the problem

$$\overline{V}(\lambda_1,\lambda_2) = \max_{G(\cdot)\in\mathbb{G}} U_{\lambda_1,\lambda_2}(G(\cdot)). \tag{B.1}$$

Remark B.1. *In contrast to the convention, we define the second Lagrange multiplier in the form $\lambda_2' := \lambda_1\lambda_2$. This simplifies the notation in the analysis below.*

We now attempt to link (B.1) with (A.1).

Before we present the result, let us prove two useful lemmas.

Lemma B.2. *If* $(x, \underline{x}) \in \Delta_2$*, and* X^* *is optimal to (3.4), then*

$$\rho_\Phi(X^*) = -\underline{x}.$$

Proof of Lemma B.2. If $\rho_\Phi(X^*) < -\underline{x}$, then define

$$\varepsilon := \frac{-\underline{x} - \rho_\Phi(X^*)}{\rho_\Phi\left(X_0^*\right) - \rho_\Phi(X^*)} \in (0, 1).$$

Let $X = G_X(1 - F_\xi(\xi))$, where $G_X(z) = (\varepsilon G_{X_0^*} + (1 - \varepsilon)G_{X^*})(z) \in \mathbb{G}$. We have $\mathbb{E}[\xi X] = x$ and $\rho_\Phi(X) = -\underline{x}$, and thus, X satisfies both constraints. In view of the strict concavity of u and the fact that the optimal value of (3.1) is no less than that of (3.4), we have $\mathbb{E}[u(X)] > \mathbb{E}[u(X^*)]$, which contradicts the optimality of X^*. □

Lemma B.3. $V(\mathbf{x}) := V(x, \underline{x})$, $\mathbf{x} = (x, \underline{x}) \in \Delta_1$ *is a concave function, with a nonempty superdifferential* $\partial V(\mathbf{x})$ *at every* $\mathbf{x} \in \Delta_1$. $V(x, \underline{x})$ *is strictly increasing in* x*, decreasing in* $\underline{x}$*, and the monotonicity is strict whenever* $\mathbf{x} \in \Delta_2$.

Proof of Lemma B.3. In view of Lemmas A.1 and B.2, the monotonicity is obvious. Δ_1 is a convex, open set and $V(\mathbf{x})$, $\mathbf{x} \in \Delta_1$ is a concave function due to the strict concavity of $u(\cdot)$. Thus, the superdifferential of $V(\cdot)$ at any $\mathbf{x} \in \Delta_1$, $\partial V(\mathbf{x})$, is nonempty. □

The following proposition links (B.1) with (A.1).

Proposition B.4. *For* $(x, \underline{x}) \in \Delta_2$*,*

If (A.1) admits an optimal solution $\boldsymbol{G}^*(\cdot)$*, then there exists* $\lambda_1 > 0$ *and* $\lambda_2 > 0$ *such that* $\boldsymbol{G}^*(\cdot)$ *is also optimal for (B.1) and* $\boldsymbol{U}_{\lambda_1,\lambda_2}(\boldsymbol{G}^*(\cdot)) < \infty$*;*

Conversely, if $\boldsymbol{G}^*(\cdot)$ *solves (B.1) for some* λ_1 *and* λ_2 *satisfying*

$$\begin{cases} \displaystyle\int_{[0,1)} G^*(z) F_\xi^{-1}(1 - z)dz = x, \\ \displaystyle\int_{[0,1)} G^*(z)\Phi(dz) = \underline{x}, \end{cases} \tag{B.2}$$

then $\boldsymbol{G}^*(\cdot)$ *also solves (A.1) and* $U(\boldsymbol{G}^*(\cdot)) < \infty$.

Proof of Proposition B.4. Let $G^*(\cdot)$ solve (A.1) with $\mathbf{x} = (x, \underline{x}) \in \Delta_2$ and $U(G^*(\cdot)) < \infty$. For any $(\lambda_1, -\lambda_3) \in \partial V(\mathbf{x})$, i.e., a supergradient of V at $\mathbf{x}$, and any $\mathbf{y} \in \Delta_1$, we have $V(\mathbf{y}) \le V(\mathbf{x}) + (\lambda_1, -\lambda_3)^\top(\mathbf{y} - \mathbf{x})$, or equivalently $V(\mathbf{y}) - (\lambda_1, -\lambda_3)^\top \mathbf{y} \le V(\mathbf{x}) - (\lambda_1, -\lambda_3)^\top \mathbf{x}$. Because $V(\cdot)$ is strictly concave, strictly increasing in x, and strictly decreasing in $\underline{x}$ on Δ_2, $\lambda_1 > 0$ and (there is at least one) $\lambda_3 > 0$, and we can choose $\lambda_2 := \frac{\lambda_3}{\lambda_1}$. Next for any $G \in \mathbb{G}$, let $\mathbf{y} = (\int_{[0,1)} G(z)F_\xi^{-1}(1-z)dz, \int_{[0,1)} G(z)\Phi(dz))$. We have

$$\begin{aligned}
&U_{\lambda_1,\lambda_2}(G(\cdot)) \\
&= \int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)F_\xi^{-1}(1-z)dz + \lambda_1\lambda_2 \int_{[0,1)} G(z)\Phi(dz) \\
&\leq V(\mathbf{y}) - (\lambda_1, -\lambda_3)^\top \mathbf{y} \\
&\leq V(\mathbf{x}) - (\lambda_1, -\lambda_3)^\top \mathbf{x} \\
&= \int_{[0,1)} u(G^*(z))dz - \lambda_1 x + \lambda_1\lambda_2 \underline{x} \\
&= U_{\lambda_1,\lambda_2}(G^*(\cdot)) < \infty,
\end{aligned}$$

which implies that $G^*(\cdot)$ is also optimal for (B.1). Conversely, if $G^*(\cdot)$ solves (B.1) for some λ_1 and λ_2 satisfying (B.2), then for any $G(\cdot) \in \mathbb{G}$ that satisfies all the constrains in (A.1),

$$\begin{aligned}
&U(G(\cdot)) - \lambda_1 x + \lambda_1\lambda_2 \underline{x} \\
&\leq U_{\lambda_1,\lambda_2}(G(\cdot)) \\
&\leq U_{\lambda_1,\lambda_2}(G^*(\cdot)) \\
&= U(G^*(\cdot)) - \lambda_1 x + \lambda_1\lambda_2 \underline{x},
\end{aligned}$$

which thereby proves the desired result. □

Remark B.5. *If $\underline{x} = C_\Phi(x) = \frac{x}{\mathbb{E}\xi}$, then the superdifferential of $V(\cdot)$ at $(x, \frac{x}{\mathbb{E}\xi})$ could be empty and the Lagrange multipliers may not exist. This fact has been neglected by many authors, including Rogers (2009) and Cahuich and Hernández-Hernández (2013). In the continuous-time utility maximization literature, the Lagrange dual method is often employed to solve (3.1), and it is assumed a priori that the optimal solution exists, which can be found by solving the dual problem and finding a suitable Lagrange multiplier that corresponds to the budget constraint. For (3.4) or (A.1), if $\underline{x} = C_\Phi(x) = \frac{x}{\mathbb{E}\xi}$, then the Lagrange multiplier that corresponds to the risk constraint might not exist, but the optimal solution to the original problem might exist. In this case, the optimal solution to (3.4) or (A.1) is no longer given by the dual problem. We discuss this case in Appendix C.*

B.2.2 Lagrangian dual problem

We now solve (B.1). The technique used in this section is similar to Rogers (2009), who focuses on (B.1) when Φ admits a density.

Recall that

$$\varphi_{\lambda_2}(z) = -\int_{[z,1)} F_\xi^{-1}(1-s)ds + \lambda_2\Phi((z,1]),\ z \in [0,1),$$

with $\varphi_{\lambda_2}(1) = 0$, and

$$\varphi_{\lambda_2}(z-) = -\int_{[z,1)} F_\xi^{-1}(1-s)ds + \lambda_2\Phi([z,1])$$

$$= (\lambda_2 \kappa_\Phi([z,1)) - 1) \int_{[z,1)} F_\xi^{-1}(1-s)ds, z \in [0,1), \tag{B.3}$$

with $\varphi_{\lambda_2}(1-) = 0$.

The Lagrangian is now given by

$$U_{\lambda_1,\lambda_2}(G(\cdot)) = \int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)d\varphi_{\lambda_2}(z), \tag{B.4}$$

and we consider the problem

$$\overline{V}(\lambda_1,\lambda_2) = \max_{G(\cdot)\in\mathbb{G}} \int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)d\varphi_{\lambda_2}(z),\ \lambda_1 > 0,\ \lambda_2 > 0. \tag{B.5}$$

Note that $\varphi'_{\lambda_2}(z)$ may not exist, and it is not necessarily decreasing or nonnegative (provided it exists), and thus, point-wise optimization fails. Inspired by Rogers (2009), Xia and Zhou (2016), and Xu (2016), we replace $\varphi_{\lambda_2}(\cdot)$ with $\delta_{\lambda_2}(\cdot)$, the concave envelope of $\varphi_{\lambda_2}(\cdot-)$.

Recall that

$$\mathcal{A}_\Phi = \left\{ \lambda : \lambda > 0, \delta'_\lambda(z+) > 0,\ z \in [0,1) \right\}.$$

We now present the solution to (B.5).

Proposition B.6.

1. *If $\lambda_2 \in \mathcal{A}_\Phi$, the optimal solution to (B.5) is given by $G^*(\cdot) := (u')^{-1}(\lambda_1 \delta'_{\lambda_2}(\cdot+))$.*
2. *If $\lambda_2 > 0$ and $\lambda_2 \notin \mathcal{A}_\Phi$, $\overline{V}(\lambda_1,\lambda_2) = \infty,\ \forall \lambda_1 > 0$.*

Proof of Proposition B.6. $\delta_{\lambda_2}(\cdot)$ is concave and $\delta_{\lambda_2}(0) = \varphi_{\lambda_2}(0-)$, $\delta_{\lambda_2}(1) = \varphi_{\lambda_2}(1-) = 0$. For any $G(\cdot) \in \mathbb{G}$,

$$\int_{[0,1)} (\varphi_{\lambda_2}(z-) - \delta_{\lambda_2}(z))dG(z) \le 0,$$

and applying Fubini's theorem, we have

$$\begin{aligned} &\int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)d\varphi_{\lambda_2}(z) \\ &\le \int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)d\delta_{\lambda_2}(z) \\ &= \int_{[0,1)} u(G(z))dz - \lambda_1 \int_{[0,1)} G(z)\delta'_{\lambda_2}(z+)dz. \end{aligned}$$

For $\lambda_2 \in \mathcal{A}_\Phi$, we have

$$\int_{[0,1)} \left[u(G(z)) - \lambda_1 \delta'_{\lambda_2}(z+)G(z) \right] dz \le \int_{[0,1)} \left[u(\overline{G}(z)) - \lambda_1 \delta'_{\lambda_2}(z+)\overline{G}(z) \right] dz,$$

where $\overline{G}(z) := (u')^{-1}(\lambda_1 \delta'_{\lambda_2}(z+))$ is given by point-wise optimization.

It is now sufficient to show that

$$\int_{[0,1)} u(\overline{G}(z))dz - \lambda_1 \int_{[0,1)} \overline{G}(z)d\varphi_{\lambda_2}(z) = \int_{[0,1)} \left[u(\overline{G}(z)) - \lambda_1\delta'_{\lambda_2}(z+)\overline{G}(z)\right] dz,$$

or equivalently

$$\int_{[0,1)} (u')^{-1}\left(\lambda_1\delta'_{\lambda_2}(z+)\right) d\varphi_{\lambda_2}(z) - \int_{[0,1)} (u')^{-1}\left(\lambda_1\delta'_{\lambda_2}(z+)\right) d\delta_{\lambda_2}(z) = 0.$$

Applying Fubini's theorem again, the above identity is equivalent to

$$\int_{[0,1)} [\delta_{\lambda_2}(z) - \varphi_{\lambda_2}(z-)] \frac{1}{u''\left((u')^{-1}\left(\lambda_1\delta'_{\lambda_2}(z+)\right)\right)} d\delta'_{\lambda_2}(z+) = 0.$$

Because $\delta_{\lambda_2}(\cdot)$ dominates $\varphi_{\lambda_2}(\cdot-)$ on $[0, 1]$, $u''(\cdot) < 0$, and $\delta'_{\lambda_2}(\cdot+)$ is constant on any subinterval of $\{z \in (0,1) : \delta_{\lambda_2}(z) > \varphi_{\lambda_2}(z-)\}$, the above identity holds.

If $\lambda_2 \notin \mathcal{A}_\Phi$, then there exists $0 \le c < 1$ such that $\delta'_{\lambda_2}(z+) \le 0$, $z \in [c, 1)$. Let

$$\overline{G}_n(z) = (u')^{-1}\left(\lambda_1\left(\delta'_{\lambda_2}(z+) \vee \frac{1}{n}\right)\right),$$

and

$$\overline{V}_n = \int_{[0,1)} u(\overline{G}_n(z))dz - \lambda_1 \int_{[0,1)} \overline{G}_n(z)d\varphi_{\lambda_2}(z).$$

Similarly, we can show that

$$\begin{aligned}\overline{V}_n &= \int_{[0,1)} u(\overline{G}_n(z))dz - \lambda_1 \int_{[0,1)} \overline{G}_n(z)d\varphi_{\lambda_2}(z)\\ &= \int_{[0,1)} \left[u(\overline{G}_n(z)) - \lambda_1\delta'_{\lambda_2}(z+)\overline{G}_n(z)\right] dz.\end{aligned}$$

Because $u'(\cdot)$ is strictly decreasing, $\overline{V}_n \le \overline{V}_{n+1}$ and there exists $0 \le c < 1$ such that $\delta'_{\lambda_2}(z) \le 0$, $z \in [c, 1]$, we have

$$\lim_{n\to\infty} \overline{V}_n \ge \lim_{n\to\infty} \int_{[c,1)} \left[u(\overline{G}_n(z)) - \lambda_1\delta'_{\lambda_2}(z+)\overline{G}_n(z)\right] dz = \infty.$$

□

In view of Proposition (B.4), the optimal solution, if it exists, must be given by $(u')^{-1}(\lambda_1\delta'_{\lambda_2}(\cdot+))$, where $\lambda_2 \in \mathcal{A}_\Phi$. However, $\mathcal{A}_\Phi$ is not easy to obtain. Before we close this section, we provide an equivalent characterization.

Lemma B.7.

$$\mathcal{A}_\Phi = \{\lambda : \lambda > 0, \varphi_\lambda(z-) < 0,\ \forall z \in [0,1)\}.$$

Proof of Lemma B.7. $\forall \lambda \in \mathcal{A}_\Phi$, $\delta'_\lambda(z+) > 0$, $z \in [0, 1)$ implies $\varphi_\lambda(z-) \le \delta_\lambda(z) < \delta_\lambda(1) \le 0$, $z \in [0, 1)$.

$\forall \lambda \in \{\lambda : \lambda > 0, \varphi_\lambda(z-) < 0, \forall z \in [0,1)\}$, as $\varphi_\lambda(z-) \geq \varphi_\lambda(z+),\ z \in (0,1)$, $\varphi_\lambda(z-)$ is upper semi-continuous. We then have $\sup_{z\in[0,c)} \varphi_\lambda(z-) < 0,\ \forall c \in (0,1)$. From (4.5), we know $\delta_\lambda(z) < 0,\ z \in [0,1)$ and $\delta'_\lambda(1-) > 0$. Thus $\delta'_\lambda(z+) \geq \delta'_\lambda(1-) > 0,\ z \in [0,1)$. □

Proposition B.8.

1. *If* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$, *then* $\mathcal{A}_\Phi = \emptyset$.
2. *If* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) < \infty$, *then* $\mathcal{A}_\Phi \neq \emptyset$.

Proof of Proposition B.8.

1. $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = \infty$ implies that for any $\lambda_2 > 0$, there exists $z \in (0,1)$ such that $\kappa_\Phi([z,1)) > \frac{1}{\lambda_2}$ and $\varphi_{\lambda_2}(z-) \geq 0$. Thus, $\mathcal{A}_\Phi = \emptyset$.
2. $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) < \infty$ implies that there exists $a > 0$ such that $\kappa_\Phi([z,1)) < a,\ z \in [0,1)$. Then $\forall \lambda_2 \leq \frac{1}{a}$, $\varphi_{\lambda_2}(z-) < 0,\ z \in [0,1)$. Thus $\mathcal{A}_\Phi \neq \emptyset$. □

B.2.3 Existence of Lagrange multipliers

We now show the existence of Lagrange multipliers, which is the solution to the following system of equations

$$\begin{cases} \int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz = x, \\ \int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) \Phi(dz) = \underline{x}. \end{cases} \tag{B.6}$$

We impose the following integrability assumption.

Assumption B.9.

$$\mathbb{E}\left[(u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}((1-F_\xi(\xi))+)\right)\xi\right] = \int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz < \infty,$$

for all $\lambda_1 > 0$ *and* $\lambda_2 \in \mathcal{A}_\Phi$.

The following lemma provides a sufficient condition for Assumption B.9.

Lemma B.10. *If* $F_\xi^{-1}(\cdot)$ *is differentiable,* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$, *and there exists* $z_\Phi \in [0,1)$ *such that* $\Phi([z,1))$ *(as a function of z) is twice differentiable on* $(z_\Phi, 1)$, *then*

$$\int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz < \infty,$$

for all $\lambda_1 > 0$ *and* $\lambda_2 \in \mathcal{A}_\Phi$.

Proof of Lemma B.10. Fix $\lambda_1 > 0$ and $\lambda_2 \in \mathcal{A}_\Phi$. Because $\Phi([z,1))$ (as a function of z) is twice differentiable on $(z_\Phi, 1)$, we have $\Phi([z,1)) = \int_{[z,1)} \phi(s)ds$ for $z \in (z_\Phi, 1)$ and some $\phi(\cdot)$. $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$ implies $\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = 0$. According to L'Hopital's rule,

$$\lim_{z\uparrow 1} \kappa_\Phi([z,1)) = \lim_{z\uparrow 1} \frac{\phi'(z)}{\left(F_\xi^{-1}\right)'(1-z)} = 0.$$

There exists $b \in (z_\Phi, 1)$ such that $\lambda_2 \phi'(z) < (F_\xi^{-1})'(1-z),\ z \in (b,1)$. Consequently, $\varphi_{\lambda_2}(\cdot -)$ is strictly concave on $(b,1)$. We now show that there exists $c \in (b,1)$ such that $\delta_{\lambda_2}(z) = \varphi_{\lambda_2}(z),\ z \in [c,1]$. Otherwise, we can find $z_1 \in (b,1)$ such that $\delta_{\lambda_2}(z_1) > \varphi_{\lambda_2}(z_1)$. Let $z_2 := \inf\{z > z_1 | \delta_{\lambda_2}(z) = \varphi_{\lambda_2}(z)\}$. Because

$$
\begin{aligned}
0 \le \delta'_{\lambda_2}(1-) &\le \lim_{z\uparrow 1} \frac{\varphi_{\lambda_2}(1-) - \varphi_{\lambda_2}(z-)}{1-z} \\
&= \lim_{z\uparrow 1} \frac{(1-\lambda_2 \kappa_\Phi([z,1))) \int_{[z,1)} F_\xi^{-1}(1-s)ds}{1-z} \\
&= \lim_{z\uparrow 1} F_\xi^{-1}(1-z) \\
&= 0,
\end{aligned}
$$

we must have $z_2 < 1$. We can then find $z_3 \in (z_2, 1)$ such that $\delta_{\lambda_2}(z_3) > \varphi_{\lambda_2}(z_3)$. We have $z_4 := \sup\{z < z_3 | \delta_{\lambda_2}(z) = \varphi_{\lambda_2}(z)\} \in (z_2, z_3)$ and $z_5 := \inf\{z > z_3 | \delta_{\lambda_2}(z) = \varphi_{\lambda_2}(z)\} \in (z_3, 1)$. Then, $\varphi'_{\lambda_2}(z_4) = \delta'_{\lambda_2}(z_4+) = \delta'_{\lambda_2}(z_5-) = \varphi'_{\lambda_2}(z_5)$, which is a contradiction.

Next, we can find $d \in (c,1)$ such that $\delta'_{\lambda_2}(z+) = F_\xi^{-1}(1-z) - \lambda_2\phi(z) > \frac{1}{2}F_\xi^{-1}(1-z),\ z \in [d,1)$. We then have

$$
\begin{aligned}
&\int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz \\
&= \int_{[0,d)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz + \int_{[d,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz \\
&< (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(d+)\right) \int_{[0,d)} F_\xi^{-1}(1-z)dz + \int_{[d,1)} (u')^{-1}\left(\frac{\lambda_1}{2} F_\xi^{-1}(1-z)\right) F_\xi^{-1}(1-z)dz \\
&\le (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(d+)\right) \mathbb{E}\xi + \mathbb{E}\left[(u')^{-1}\left(\frac{\lambda_1}{2}\xi\right)\xi\right] < \infty,
\end{aligned}
$$

and the claim follows easily. □

Under Assumption B.9, we can show the integrability of the second equation of (B.6).

Lemma B.11. *If* $\limsup_{z\uparrow 1} \kappa_\Phi([z,1)) < \infty$ *and Assumption B.9 holds, then*

$$
\int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) \Phi(dz) < \infty,
$$

for all $\lambda_1 > 0$ *and* $\lambda_2 \in \mathcal{A}_\Phi$.

Proof of Lemma B.11. For any $K > \limsup_{z\uparrow 1} \kappa_\Phi([z,1))$, there exists $0 < b < 1$ such that $\kappa_\Phi((z,1)) < K,\ z \in [b,1)$. We then have

$$
\begin{aligned}
&\int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right)\Phi(dz) \\
&\le (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right) + \int_{(b,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right)\Phi(dz) \\
&= (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right) + (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right)\Phi((b,1)) \\
&\quad + \int_{(b,1)} \Phi((z,1)) d\left((u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right)\right) \\
&< (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right) + K(u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right)\int_{(b,1)} F_\xi^{-1}(1-z)dz \\
&\quad + K\int_{(b,1)}\left(\int_{(z,1)} F_\xi^{-1}(1-s)ds\right) d\left((u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right)\right) \\
&= (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(b+)\right) + K\int_{(b,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz < \infty,
\end{aligned}
$$

for all $\lambda_1 > 0$ and $\lambda_2 \in \mathcal{A}_\Phi$, and the claim follows easily. □

Before we proceed, let us show a very useful lemma.

Lemma B.12.

1. *Denote the convex conjugate of $-\varphi_\lambda(z-)$ by*

$$\hat{\varphi}_\lambda(x) := \sup_{z\in[0,1]} \{xz + \varphi_\lambda(z-)\},\ x \in R,$$

and let $A(x,\lambda) := \{z : \hat{\varphi}_\lambda(x) = xz + \varphi_\lambda(z-)\}$. The right (left) derivatives of $\hat{\varphi}_\lambda(x)$ (with respect to x) are given by

$$\hat{\varphi}'_\lambda(x+) = \max_{z\in A(x,\lambda)} z,$$

$$\hat{\varphi}'_\lambda(x-) = \min_{z\in A(x,\lambda)} z,$$

respectively.

2. *For $\lambda_1 < \lambda_2$,*

$$\hat{\varphi}'_{\lambda_2}(x+) \le \hat{\varphi}'_{\lambda_1}(x+),$$

$$\hat{\varphi}'_{\lambda_2}(x-) \le \hat{\varphi}'_{\lambda_1}(x-),\ \forall x \in R.$$

Moreover, $\forall \varepsilon > 0$, $\exists \delta > 0$, such that whenever $|\lambda_1 - \lambda_2| < \delta$, we have

$$\left|\hat{\varphi}'_{\lambda_2}(x+) - \hat{\varphi}'_{\lambda_1}(x+)\right| < \varepsilon,$$

$$\left|\hat{\varphi}'_{\lambda_2}(x-) - \hat{\varphi}'_{\lambda_1}(x-)\right| < \varepsilon,\ \forall x \in R.$$

3.

$$\delta_\lambda(z) = -\sup_{x\in R}\{xz - \hat{\varphi}_\lambda(x)\},\ z \in [0,1]$$

and the set $B(z,\lambda) := \{x : \delta_\lambda(z) = \hat{\varphi}_\lambda(x) - xz\}$ *is nonempty for* $z \in (0,1)$. *The right (left) derivatives of* $\delta_\lambda(z)$ *(with respect to* $z \in (0,1)$*) are given by*

$$\delta_\lambda'(z+) = -\max_{x\in B(z,\lambda)} x,$$

$$\delta_\lambda'(z-) = -\min_{x\in B(z,\lambda)} x,$$

respectively.

4. *For* $z \in [0,1)$, $\delta_\lambda'(z+)$ *is continuous and nonincreasing in* λ.

Proof of Lemma B.12.

1. Note that $\varphi_\lambda(\cdot-)$ is upper semicontinuous, the proof is similar to that of Corollary 4 in Milgrom and Segal (2002).

2. Now for $\lambda_1 < \lambda_2$, let

$$z_1 = \max_{z\in A(x,\lambda_1)} z = \hat{\varphi}_{\lambda_1}'(x+),$$

$$z_2 = \max_{z\in A(x,\lambda_2)} z = \hat{\varphi}_{\lambda_2}'(x+).$$

$\forall z > z_1$,

$$\begin{aligned}
&xz + \varphi_{\lambda_2}(z-)\\
&\quad= xz + \varphi_{\lambda_1}(z-) + (\lambda_2 - \lambda_1)\Phi([z,1))\\
&\quad< xz_1 + \varphi_{\lambda_1}(z_1-) + (\lambda_2 - \lambda_1)\Phi([z_1,1))\\
&\quad= xz_1 + \varphi_{\lambda_2}(z_1-).
\end{aligned}$$

Thus, $z_2 \le z_1$ and $\hat{\varphi}_{\lambda_2}'(x+) \le \hat{\varphi}_{\lambda_1}'(x+)$. Similarly, $\hat{\varphi}_{\lambda_2}'(x-) \le \hat{\varphi}_{\lambda_1}'(x-)$.
Without loss of generality, let us assume $\lambda_1 < \lambda_2$. $\forall z > z_2$, $xz + \varphi_{\lambda_2}(z) < xz_2 + \varphi_{\lambda_2}(z_2)$, and $\forall \varepsilon > 0$, $\exists \delta > 0$ such as whenever $z - z_2 > \varepsilon$, we have

$$xz + \varphi_{\lambda_2}(z-) < xz_2 + \varphi_{\lambda_2}(z_2-) - \delta.$$

Because $\Phi([z_2,z)) \le 1$, we have for $z - z_2 > \varepsilon$

$$\begin{aligned}
&xz + \varphi_{\lambda_1}(z-) - xz_2 - \varphi_{\lambda_1}(z_2-)\\
&\quad= xz + \varphi_{\lambda_2}(z-) - xz_2 - \varphi_{\lambda_2}(z_2-) + (\lambda_1 - \lambda_2)\Phi([z_2,z))\\
&\quad< -\delta + (\lambda_2 - \lambda_1)\Phi([z_2,z))\\
&\quad< -\frac{1}{2}\delta,
\end{aligned}$$

whenever $|\lambda_1 - \lambda_2| < \frac{1}{2}\delta$. Thus, $0 \le z_1 - z_2 < \varepsilon$, i.e., $0 \le \hat{\varphi}'_{\lambda_1}(x+) - \hat{\varphi}'_{\lambda_2}(x+) < \varepsilon$. Similarly, we can show that for the left derivatives.

3. First,

$$\begin{aligned}\underline{\delta}_\lambda(z) := & -\sup_{x\in R}\{xz - \hat{\varphi}_\lambda(x)\} \\ = & -\sup_{x\in R}\left\{xz - \sup_{y\in[0,1]}\{xy + \varphi_\lambda(y-)\}\right\} \\ \ge & -\sup_{x\in R}\{xz - xz - \varphi_\lambda(z-)\} \\ = & \ \varphi_\lambda(z-),\ z \in [0,1].\end{aligned}$$

Because $\hat{\varphi}_\lambda(x) \ge x$, and $\hat{\varphi}_\lambda(x) = x$ for $x > 0$, $\underline{\delta}_\lambda(1) = -\sup_{x\in R}\{x - \hat{\varphi}_\lambda(x)\} = 0 = \varphi_\lambda(1-)$.

$$\begin{aligned}\underline{\delta}_\lambda(0) = & \inf_{x\in R}\left\{\sup_{z\in[0,1]}\{xz + \varphi_\lambda(z-)\}\right\} \\ \le & \inf_n\left\{\sup_{z\in[0,1]}\{-nz + \varphi_\lambda(z-)\}\right\} \\ = & \inf_n\{-nz_n + \varphi_\lambda(z_n-)\},\end{aligned}$$

where z_n can be any z that achieves the supremum above. Now $\{z_n\}$ has a subsequence $z_{n_k} \to z_\infty$. If $z_\infty > 0$, then $\delta_\lambda(0) \le \liminf_{n\to\infty}\{-nz_n + \varphi_\lambda(z_n)\} = -\infty$, which is a contradiction. Thus $z_\infty = 0$ and $\underline{\delta}_\lambda(0) \le \lim_{k\to\infty}\{-n_k z_{n_k} + \varphi_\lambda(z_{n_k}-)\} \le \varphi_\lambda(0-)$, and we have $\underline{\delta}_\lambda(0) = \varphi_\lambda(0-)$. It is then a simple exercise to show that $\underline{\delta}_\lambda$ defined in this way is indeed the concave envelope of $\varphi_\lambda(\cdot-)$.
We show that $x = +\infty$ or $-\infty$ is never optimal for $xz - \hat{\varphi}_\lambda(x)$, when $z \in (0,1)$. First, as $\hat{\varphi}_\lambda(x) = x$ for $x > 0$, $x = +\infty$ is not optimal when $z < 1$. Because $\hat{\varphi}_\lambda(x) \ge \varphi_\lambda(0-)$ for all x, $x = -\infty$ is not optimal for $z > 0$.
The expression for the right (left) derivatives follows from a similar argument as in 2.

4. Fix $z \in [0,1)$, let $x_\lambda := \max_{x\in B(z,\lambda)} x$, and we have

$$\hat{\varphi}'_\lambda(x_\lambda-) \le z \le \hat{\varphi}'_\lambda(x_\lambda+),$$

and $\forall\varepsilon > 0$,

$$z < \hat{\varphi}'_\lambda((x_\lambda + \varepsilon)-).$$

Now for $\lambda_1 < \lambda_2$, if $x_{\lambda_2} < x_{\lambda_1}$,

$$z < \hat{\varphi}'_{\lambda_2}(x_{\lambda_1}-) \le \hat{\varphi}'_{\lambda_1}(x_{\lambda_1}-),$$

which is a contradiction. Thus, $x_{\lambda_2} \ge x_{\lambda_1}$.
Without loss of generality, let us assume $\lambda_1 < \lambda_2$. Suppose $\exists\varepsilon_0 > 0$ such that for all $\delta > 0$ there exists λ_1, λ_2 satisfying $0 < \lambda_2 - \lambda_1 < \delta$ and $x_{\lambda_2} > x_{\lambda_1} + \varepsilon_0$. We have $z < \hat{\varphi}'_{\lambda_1}(x_{\lambda_2}-)$. From 2 we know, for $\varepsilon_1 = \frac{1}{2}(\hat{\varphi}'_{\lambda_1}(x_{\lambda_2}-) - z)$, $\exists\delta_1 > 0$ such that whenever $0 < \lambda_2 - \lambda_1 < \delta_1$,

$$z < \hat{\varphi}'_{\lambda_1}(x_{\lambda_2}-) - \varepsilon_1 < \hat{\varphi}'_{\lambda_1}(x_{\lambda_1}-),$$

which is a contradiction, thereby proving the lemma. □

Recall that

$$f_{\lambda_2}(\lambda_1) = \int_{[0,1)} (u')^{-1}\left(\lambda_1 \delta'_{\lambda_2}(z+)\right) F_\xi^{-1}(1-z)dz,\ \lambda_1 > 0.$$

For any given $\lambda_2 \in \{0\} \cup \mathcal{A}_\Phi$, $f_{\lambda_2}(\cdot)$ is continuous, strictly increasing on $(0, \infty)$, and

$$\lim_{\lambda_1 \to 0+} f_{\lambda_2}(\lambda_1) = \infty,\ \lim_{\lambda_1 \to +\infty} f_{\lambda_2}(\lambda_1) = 0.$$

For any $x > 0$, there exists a unique $\lambda_1^* > 0$ such that $f_{\lambda_2}(\lambda_1^*) = x$. Therefore, $g(\lambda_2, x) = f_{\lambda_2}^{-1}(x)$ is well defined, and we know that $g(\cdot, x)$ is continuous, nondecreasing for all $x > 0$.

Recall that

$$h(\lambda_2, x) = \int_{[0,1)} (u')^{-1}\left(g(\lambda_2, x)\delta'_{\lambda_2}(z+)\right) \Phi(dz).$$

We have

Lemma B.13. *For all $x > 0$, $h(\cdot, x)$ is continuous on* $\{0\} \cup \mathcal{A}_\Phi$.

Proof of Lemma B.13. We want to show that $\forall \lambda \in \{0\} \cup \mathcal{A}_\Phi$ and $\lambda_n \to \lambda$, $\lim_{n\to\infty} h(\lambda_n, x) = h(\lambda, x)$. By Fatou's lemma, we have $h(\lambda, x) \le \liminf_{n\to\infty} h(\lambda_n, x)$. Next, let

$$h_0(\lambda_1, \lambda_2, x) := \int_{[0,1)} (u')^{-1}\left(g(\lambda_1, x)\delta'_{\lambda_2}(z+)\right) \Phi(dz).$$

We have that $h(\lambda, x) = h_0(\lambda, \lambda, x)$ and $h_0(\lambda_1, \lambda_2, x)$ is decreasing in λ_1 and nondecreasing in λ_2. For $\lambda_n \uparrow \lambda$ (except for $\lambda = 0$),

$$\begin{aligned} h_0(\lambda_n, \lambda_n, x) &= h_0(\lambda_n, \lambda, x) + h_0(\lambda_n, \lambda_n, x) - h_0(\lambda_n, \lambda, x) \\ &\le h_0(\lambda_n, \lambda, x) \to h_0(\lambda, \lambda, x),\ n \to \infty, \end{aligned}$$

where the convergence follows from the dominated convergence theorem. Thus, $h(\lambda, x) \ge \limsup_{n\to\infty} h(\lambda_n, x)$ and $h(\lambda, x) = \lim_{n\to\infty} h(\lambda_n, x)$.

For $\lambda_n \downarrow \lambda$ (except for $\lambda = \sup \mathcal{A}_\Phi$),

$$\begin{aligned} h_0(\lambda_n, \lambda_n, x) &= h_0(\lambda, \lambda_n, x) + h_0(\lambda_n, \lambda_n, x) - h_0(\lambda, \lambda_n, x) \\ &\le h_0(\lambda, \lambda_n, x) \to h_0(\lambda, \lambda, x),\ n \to \infty. \end{aligned}$$

Thus, $h(\lambda, x) \ge \limsup_{n\to\infty} h(\lambda_n, x)$ and $h(\lambda, x) = \lim_{n\to\infty} h(\lambda_n, x)$, which proves the lemma. □

Note that for all $x > 0$, $h(\cdot, x)$ is continuous and $h(0, x) = R_\Phi(x)$. Because

$$S_\Phi(x) = \sup_{\lambda_2 \in \mathcal{A}_\Phi} h(\lambda_2, x),$$

we conclude that for any $\underline{x} \in (R_\Phi(x), S_\Phi(x))$ there exists at least one $\lambda_2^* > 0$ such that $h(\lambda_2^*, x) = \underline{x}$. Moreover, there is no solution if $\underline{x} > S_\Phi(x)$.

Define

$$\Delta_3 := \{(x, \underline{x}) : \underline{x} < S_\Phi(x))\} \cap \Delta_2,$$

$$\Delta_4 := \{(x, \underline{x}) : \underline{x} > S_\Phi(x))\} \cap \Delta_1.$$

We now have

Proposition B.14. $\forall (x, \underline{x}) \in \Delta_3$, *there exists at least one pair of* $\lambda_1^* > 0, \lambda_2^* > 0$ *that solves (B.6).* $\forall (x, \underline{x}) \in \Delta_4$, *(B.6) admits no solution.*

Proof of Proposition B.14. In view of the above analysis, there exists at least one $\lambda_2^* > 0$ such that $h(\lambda_2^*, x) = \underline{x}$. Now $(g(\lambda_2^*, x), \lambda_2^*)$ solves (B.6). ☐

Remark B.15. *For* $\underline{x} = S_\Phi(x)$, *there exists a solution if and only if there exists* $\lambda_2^* \in \mathcal{A}_\Phi$ *such that* $h(\lambda_2^*, x) = \sup_{\lambda_2 \in \mathcal{A}_\Phi} h(\lambda_2, x)$.

B.2.4 Optimal solution

In view of previous analysis, we can now characterize the optimal solution to (3.4).

Proof of Theorem 4.7. The claim follows from Proposition B.8, Proposition B.6, Proposition B.14 and Theorem A.3. ☐

B.3 Proof of Proposition 5.1

Proof of Proposition 5.1. 1–5 is trivial. Now assume that $F_\xi(x) = \Phi_N(\frac{\ln x - \mu_\xi}{\sigma_\xi})$. It is a simple exercise to show that $\lim_{z \downarrow 0} \frac{\Phi_N^{-1}(z)}{\ln z} = 0$. We then have

6. For the power spectral risk measures with $\gamma > 1$,

$$\lim_{z \uparrow 1} \kappa_\Phi([z, 1)) = \lim_{z \downarrow 0} \gamma e^{\left(\gamma - 1 - \sigma_\xi \frac{\Phi_N^{-1}(z)}{\ln z}\right) \ln z - \mu_\xi} = 0.$$

7. For the Wang (2000) risk measure,

$$\begin{aligned}
\lim_{z \uparrow 1} \kappa_\Phi([z, 1)) &= \lim_{z \uparrow 1} \frac{e^{\Phi_N^{-1}(q)\Phi_N^{-1}(z) - \frac{1}{2}\left(\Phi_N^{-1}(q)\right)^2}}{e^{-\sigma_\xi \Phi_N^{-1}(z) + \mu_\xi}} \\
&= \lim_{z \uparrow 1} e^{\left(\sigma_\xi + \Phi_N^{-1}(q)\right)\Phi_N^{-1}(z) - \frac{1}{2}\left(\Phi_N^{-1}(q)\right)^2 - \mu_\xi} \\
&= \begin{cases} 0, & q < \Phi_N(-\sigma_\xi); \\ e^{-\frac{1}{2}\left(\Phi_N^{-1}(q)\right)^2 - \mu_\xi}, & q = \Phi_N(-\sigma_\xi); \\ \infty, & q > \Phi_N(-\sigma_\xi). \end{cases}
\end{aligned}$$

8. For the beta family of distortion risk measures,

$$\begin{aligned}\lim_{z\uparrow 1}\kappa_\Phi([z,1)) &= \lim_{z\uparrow 1}\frac{\frac{1}{\beta(a,b)}z^{a-1}(1-z)^{b-1}}{e^{\sigma_\xi\Phi_N^{-1}(1-z)+\mu_\xi}}\\ &= \lim_{z\downarrow 0}\frac{1}{\beta(a,b)}(1-z)^{a-1}e^{(b-1)\ln z-\sigma_\xi\Phi_N^{-1}(z)-\mu_\xi}\\ &= \begin{cases}0, & b>1;\\ +\infty, & 0<b\le 1.\end{cases}\end{aligned}$$

□

B.4 Proof of Proposition 5.2

Proof of Proposition 5.2. From the proof of Lemma B.10, we have $\delta'_{\lambda_2}(1-)=0$. Thus

$$\begin{aligned}\lim_{\xi\downarrow 0}X^* &= \lim_{\xi\downarrow 0}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}((1-F_\xi(\xi))+)\right)\\ &= \lim_{z\uparrow 1}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}(z+)\right)\\ &\ge \lim_{z\uparrow 1}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}(z-)\right)\\ &= +\infty.\end{aligned}$$

□

B.5 Proof of Proposition 5.3

We first show a lemma.

Lemma B.15.

1. *If* $\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}<+\infty$, *then* $\delta'_{\lambda_2^*}(0+)<+\infty$ *and* $\operatorname{ess\,inf} X^*>0$.
2. *If* $\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}=+\infty$, *then* $\delta'_{\lambda_2^*}(0+)=+\infty$ *and* $\operatorname{ess\,inf} X^*=0$.

Proof of Lemma B.15.

1. If $\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}<+\infty$, we claim $\delta'_{\lambda_2^*}(0+)<+\infty$. Otherwise, $\delta'_{\lambda_2^*}(0+)=+\infty$. For any $a>\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}$, there exists $b\in(0,1]$ such that $\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}<a<\frac{\delta_{\lambda_2^*}(z)-\delta_{\lambda_2^*}(0)}{z}$, $z\in(0,b]$. Thus, on $(0,b)$, $\delta_{\lambda_2^*}(z)>\varphi_{\lambda_2^*}(z-)$, $\delta_{\lambda_2^*}(z)$ is affine, but $\delta'_{\lambda_2^*}(z+)=+\infty$, which is a contradiction. Thus, we have

$$\begin{aligned}\lim_{\xi\uparrow\infty}X^* &= \lim_{\xi\uparrow\infty}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}((1-F_\xi(\xi))+)\right)\\ &= \lim_{z\downarrow 0}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}(z+)\right)\\ &> 0.\end{aligned}$$

2. If $\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}=+\infty$,

$$\begin{aligned}\delta'_{\lambda_2^*}(0+) &= \limsup_{z\downarrow 0}\frac{\delta_{\lambda_2^*}(z)-\delta_{\lambda_2^*}(0)}{z}\\ &\geq \limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}\\ &= \infty,\end{aligned}$$

and

$$\begin{aligned}\lim_{\xi\uparrow+\infty}X^* &= \lim_{\xi\uparrow\infty}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}((1-F_\xi(\xi))+)\right)\\ &= \lim_{z\downarrow 0}(u')^{-1}\left(\lambda_1^*\delta'_{\lambda_2^*}(z+)\right)\\ &= 0.\end{aligned}$$

$\square$

Proof of Proposition 5.3. We have

$$\begin{aligned}&\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}\\ &= \limsup_{z\downarrow 0}\frac{\left(1-\lambda_2^*\kappa_\Phi([0,z))\right)\int_{[0,z)}F_\xi^{-1}(1-s)ds}{z}\\ &\leq \limsup_{z\downarrow 0}F_\xi^{-1}(1-z)\cdot\left(1-\liminf_{z\downarrow 0}\lambda_2^*\kappa_\Phi([0,z))\right)\\ &< 0.\end{aligned}$$

The claim then follows from Lemma B.15. $\square$

B.6 Proof of Proposition 5.4

Proof of Proposition 5.4.

1. Because $\liminf_{z\downarrow 0}\kappa_\Phi([0,z))=0$, we can find a sequence $z_n\downarrow 0$ such that $\lim_{n\to\infty}\kappa_\Phi([0,z_n))=0$. We then have

$$\begin{aligned}&\limsup_{z\downarrow 0}\frac{\varphi_{\lambda_2^*}(z-)-\varphi_{\lambda_2^*}(0-)}{z}\\ &= \limsup_{z\downarrow 0}\frac{\left(1-\lambda_2^*\kappa_\Phi([0,z))\right)\int_{[0,z)}F_\xi^{-1}(1-s)ds}{z}\\ &\geq \lim_{n\to\infty}\frac{\left(1-\lambda_2^*\kappa_\Phi([0,z_n))\right)\int_{[0,z_n)}F_\xi^{-1}(1-s)ds}{z_n}\\ &> \frac{1}{2}\lim_{z\downarrow 0}F_\xi^{-1}(1-z)\\ &= +\infty,\end{aligned}$$

and the first claim follows from Lemma B.15.

2. From Lemma B.12, we have $\lambda_1^* \geq \lambda_0^*$. If $\lambda_1^* = \lambda_0^*$, then the budget/risk constraint cannot be satisfied simultaneously. Thus, $\lambda_1^* > \lambda_0^*$.

From Lemma B.15, $\delta'_{\lambda_2^*}(0+) = +\infty$. Because $\Phi([0,z))$ (as a function of z) is twice differentiable on $(0, \underline{z}_\Phi)$, we have $\Phi([0,z)) = \int_{[0,z)} \phi(s)ds$ for $z \in (0, \underline{z}_\Phi)$ and some $\phi(\cdot)$. Similar to the proof of Lemma B.10, we can show that there exists $b \in (0,1]$ such that $\delta_{\lambda_2^*}(z) = \varphi_{\lambda_2^*}(z-),\ z \in (0,b)$. Next, we can find $c \in (0,b)$ such that $\frac{\phi(z)}{F_\xi^{-1}(1-z)} < (1-\frac{\lambda_0^*}{\lambda_1^*})\frac{1}{\lambda_2^*},\ z \in (0,c)$. For $z \in (0,c)$, $\delta'_{\lambda_2^*}(z+) = \varphi'_{\lambda_2^*}(z) = F_\xi^{-1}(1-z) - \lambda_2^*\phi(z) > \frac{\lambda_0^*}{\lambda_1^*}F_\xi^{-1}(1-z)$. Thus, for $\xi > \bar{\xi} := F_\xi^{-1}(1-c)$, $X^* < X_0^*$. □

Remark B.16. *If* $\lim_{z\downarrow 0} \kappa_\Phi([0,z)) = a \in (0,+\infty)$*, whether there is portfolio insurance depends on* λ_2^**. Note that*

$$\lim_{z\downarrow 0} \frac{\varphi_{\lambda_2^*}(z-) - \varphi_{\lambda_2^*}(0-)}{z} = \lim_{z\downarrow 0} \frac{\left(1 - \lambda_2^*\kappa_\Phi([0,z))\right)\int_{[0,z)} F_\xi^{-1}(1-s)ds}{z}$$

$$= \begin{cases} \infty, & \lambda_2^* < \frac{1}{a}; \\ -\infty, & \lambda_2^* > \frac{1}{a}. \end{cases}$$

B.7 Proof of Proposition 5.5

Proof of Proposition 5.5. 1–3 is trivial. Now assume that $F_\xi(x) = \Phi_N(\frac{\ln x - \mu_\xi}{\sigma_\xi})$. Recall that $\lim_{z\downarrow 0} \frac{\Phi_N^{-1}(z)}{\ln z} = 0$. We have the following:

4. For the Wang (2000) risk measure with $q < \Phi_N(-\sigma_\xi)$,

$$\liminf_{z\downarrow 0} \kappa_\Phi([0,z)) = \lim_{z\downarrow 0} \frac{e^{\Phi_N^{-1}(q)\Phi_N^{-1}(z) - \frac{1}{2}(\Phi_N^{-1}(q))^2}}{e^{-\sigma_\xi \Phi_N^{-1}(z) + \mu_\xi}} = +\infty.$$

5. For the beta family of distortion risk measures with $b > 1$,

$$\liminf_{z\downarrow 0} \kappa_\Phi([0,z)) = \lim_{z\downarrow 0} \frac{\frac{1}{\beta(a,b)} z^{a-1}(1-z)^{b-1}}{e^{-\sigma_\xi \Phi_N^{-1}(z) + \mu_\xi}}$$

$$= \lim_{z\downarrow 0} \frac{1}{\beta(a,b)}(1-z)^{b-1} e^{(a-1)\ln z + \sigma_\xi \Phi_N^{-1}(z) - \mu_\xi}$$

$$= \begin{cases} 0, & a \geq 1; \\ +\infty, & 0 < a < 1. \end{cases}$$

□

APPENDIX C: BOUNDARY SOLUTION

We now study the optimal solution to (3.4) when $(x, \underline{x})$ is on the boundary of the feasible set. Suppose that $\sup_{0<c<1} \kappa_\Phi((c,1]) = \frac{1}{\mathbb{E}\xi}$ and let $C := \{c \in (0,1) | \kappa_\Phi((c,1]) = \frac{1}{\mathbb{E}\xi}\}$. There are two cases:

1. C is empty, then $X = \frac{x}{\mathbb{E}\xi}$ is the unique terminal wealth that satisfies both $\mathbb{E}[\xi X] = x$ and $\rho_\Phi(X) = -\frac{x}{\mathbb{E}\xi}$, which is also optimal to (3.4). Because of Proposition 5.2, we cannot find this optimal solution by solving the Lagrange dual problem.
2. C is nonempty. From the proof of Proposition 4.1, the set of all feasible quantile functions is given by

$$\mathbb{S} := \left\{ G(\cdot) \in \mathbb{G} \middle| \int_{[0,1)} \mathbf{1}_{z \notin C} dG(z) = 0 \quad \text{and} \quad \int_{[0,1)} F_\xi^{-1}(1-z)G(z)dz = x \right\}.$$

If C is finite, then it is easy to see that

$$\mathbb{S} = \left\{ G(\cdot) \in \mathbb{G} \middle| G(z) = a + \sum_{c_i \in C} b_i \mathbf{1}_{c_i \le z \le 1},\ a \in \mathbb{R}, b_i \ge 0, \quad \text{and} \quad \int_{[0,1)} F_\xi^{-1}(1-z)G(z)dz = x \right\}.$$

It is then a finite-dimensional optimization problem to find the optimal solution to (3.4). If the optimal solution exists, then it cannot be given by solving the Lagrange dual problem due to Proposition 5.2.

If C is infinite, then the problem is subtler. Let $\overline{C}$ be the closure of C. Define

$$\mathbb{S}(C) := \left\{ G(\cdot) \in \mathbb{G} \middle| G(z) = \frac{x - b\int_{(c,1)} F_\xi^{-1}(1-z)dz}{\mathbb{E}\xi} + b\mathbf{1}_{c \le z \le 1},\ b \ge 0, c \in C \right\},$$

and let $\overline{\text{conv}}(\mathbb{S}(\overline{C}))$ be the closed convex hull of $\mathbb{S}(\overline{C})$. It is not difficult to show that

$$\mathbb{S}(C) \subset \mathbb{S} \subset \overline{\text{conv}}(\mathbb{S}(\overline{C})).$$

We can search over $\overline{\text{conv}}(\mathbb{S}(\overline{C}))$ to find the optimal solution to (3.4), which remains an interesting topic for future research.