

Multi-market Portfolio Optimization with Conditional Value at Risk

Stefano Nasini^{a,b}, Martine Labbé^{c,*}, Luce Brotcorne^{d,*}

^aIESEG School of Management, 3 rue de la Digue, 59000 Lille, France.

^bLEM-CNRS 9221, 3 rue de la Digue, 59000 Lille, France.

^cUniversité Libre de Bruxelles, CP212, boulevard du Triomphe, 1050 Bruxelles, Belgium.

^dINRIA Lille Nord-Europe, 40 Avenue Halley, 59000 Lille, France.

Abstract

In this paper we propose an optimization framework for multi-markets portfolio management, where a central headquarter relies upon local affiliates for the market-wise selection of investment options. Being averse to risk, the headquarter endogenously selects the maximum expected loss (conditional value at risk) for the affiliates, who respond designing portfolios and selecting management fees. In its essence, this problem constitutes a single-leader-multi-follower game resulting from the decentralized design of the investment selection. Starting from a bilevel formulation, our results build on the equivalence with the high point relaxation and provide theoretical insights about the decentralized portfolio properties and numerical solution approaches. We show that the problem is NP-Hard and propose a decomposition procedure and strong valid inequalities, capable of substantially boosting the efficiency of the computational solution, when instances become large. In the same line, optimality bounds exploiting overlooked properties of the conditional value at risk are deduced, to provide *almost exact* solutions with few seconds of computation. Building on this theoretical development, we conduct computational tests to validate and compare the proposed investment framework, using comprehensive firm-level data from 1999 to 2014 on 7256 U.S. listed enterprises. The numerical tests supports the effectiveness of the decomposition procedure, as well as the one resulting from the inclusion of strong valid inequalities, improving the LP relaxation by up to 98%.

Keywords: Portfolio optimization; Multi-market investment selection; Conditional value at risk; Polyhedral representation; Valid inequalities

1. Introduction

As noted by Hausler (2002), *financial globalization has brought considerable benefits to national economies and to investors and savers, but it has also changed the structure of markets, creating new risks and challenges for market participants and policymakers.*

Multinational investment companies (hereafter referred to as MIC) are the main players in these cross-border capital movements (Doukas & Lang 2003). In the USA, they are regulated under the *Securities Act of 1933* and the *Investment Company Act of 1940*, which allow them to own other companies' outstanding stocks and manage securities for investment purposes. Over the last decades, they have relied upon specialized agents (hereafter referred to as SA) offering intermediation services. Reasons in favour to a decentralization of financial activities encompass not only the overload of information that is more easily accessible through the overseas business network in which SAs are embedded, but also the division of labour

*Corresponding author

Email addresses: s.nasini@ieseg.fr (Stefano Nasini), mlabbe@ulb.ac.be (Martine Labbé), luce.brotcorne@inria.fr (Luce Brotcorne)

and specialization within a large multinational corporation. SAs are then in charge for the specific investment selection at each market, conditional upon specific guidelines and requirements by the MIC (Doukas & Lang 2003, Barry & Kearney 2006, Ambos et al. 2010, Kawai & Strange 2014).

In this work we present an optimization framework (i.e., modelling design, optimality properties and solution approach) for a multi-market portfolio optimization, which allows endogenizing the management costs that the financial intermediation entails, as well as the centralized directives of the MIC about budget allocation and risk regulation.

Formally, building on a leader-follower game, we consider a collection of m local markets. The MIC has incomplete information about them, so that it assigns probabilities to a collection of possible market types. Next, we consider the proportion x_{ikt} of a fixed budget invested in the i^{th} asset belonging to the k^{th} local market, when the latter is of type t , and the corresponding rate of return R_{ikt} . The MIC (acting as the leader) selects the desired proportion of budget to invest in each market $z_1 \dots z_m$, as well as the maximum level of tolerable expected loss Θ from the markets. As a response to this centralized decision, SAs (acting as risk neutral followers) select individual assets by maximizing the expected local portfolio return, subject to the risk regulation imposed by the MIC, and charging a percentage of the expected investment return as a fee.¹ We assume that the risk regulation is imposed by the MIC in the form of the *Conditional Value-at-Risk* (hereafter referred to as CVaR) of potential losses.²

While extending the traditional literature of mean-CVaR portfolio optimization towards a broader class of multi-market problems, the proposed modelling framework opens a collection a new methodological challenges concerning its actual algorithmic solution. Contextually, we offer insights on three aspects:

- Starting from a mathematical programming formulation of a single-leader-multiple-follower game (bilevel program), we prove the existence of an optimal solution and its equivalence with its high point relaxation.³ We show that the computation of the latter is NP-Hard.
- On the algorithmic side, we tailor a solution procedure for the overall problem that combines a decomposition strategy (for the CVaR constraints) and the generation of strong valid inequalities (for the management fees). The decomposition strategy reduces the number of decision variables and constraints by one order of magnitude, by sequentially appending feasibility cuts on the branch-and-bound tree.⁴ This provides exact solutions with a substantially lower computational effort (both in terms of memory consumption and CPU-time).
- Building on newly uncovered features of the expected loss, tight bounds on the optimal MIC payoff are constructed to provide *almost exact* solutions with few seconds of computation.

¹The proposed model focuses on a linear sharing rule on the portfolio returns. As noted by Stoughton (1993), this rule implies that the SA's private information is indirectly revealed through its portfolio choice.

²Key sources of uncertainty (Bacells & Borgonovo 2013) and risk measures (Fábián 2008, Künzi-Bay & Mayer 2006) have been studied in the portfolio optimization literature, ranging from *value-at-risk*, *conditional value-at-risk* and *maximum drawdown* to *variance*. While the variance is a quadratic risk measures, linear programming computable measures have been studied by Mansini et al. (2007). The value-at-risk is one of the most accepted among them, consisting in quantile of the loss distribution. As shown by (Prékopa 1974), this can be formulated as probabilistic constraints. Current regulations in several western countries establish a maximum value-at-risk, as a proportion to the investment (Hull 2012). This can be enforced using the conditional-value-at-risk on the random losses, whose value represents an upper bound to the value-at-risk.

³The high point relaxation problem of a bilevel problem provides a best first level solution under the assumption that the second level reacts in the most favorable way for the leader (Dempe 2002).

⁴In practice, the implementation of this procedure is made possible by the availability of callback functions in recent versions of ILOG CPLEX, enabling the user to alter the behavior of the branch-and-bound tree during the solve.

On the empirical side, the proposed methodology is validated using a large data set comprising the entire population of 7256 U.S. listed enterprises within the period 1999–2014 (available from the *Center for Research in Security Prices*). We consider a collection of 73 industries classified by the grouped-industry code from the North American Industry Classification System. Building on this data set, a computational test is conducted, to explore the main figures of the investment solution and the computational performance over more than one thousand instances.

The results show the remarkable impact of the combined decomposition procedure and the strong valid inequalities. The latter improve the GAP between the LP relaxation and the optimal value by up to 98%. Next, the tightness of the proposed optimality bounds (with a GAP smaller than $1.0e - 2$) is observed in all the instances.

The remainder of the paper is organized as follows: Section 2 surveys the relevant studies on portfolio optimization, focusing on decentralized approaches and conditional value at risk. Section 3 introduces our baseline model. Section 4 establishes a collection of fundamental properties that support the computational solvability of the problem. Section 5 presents reformulation approaches that allow building our main solution strategy, based on the problem decomposition and the generation of strong valid inequalities. Section 6 presents tight lower and upper bounds on the optimal MIC payoff. Section 7 provides computational tests analyzing the solution and the performance of the proposed approaches under the different risk specifications. It also brings into the limelight the impact of the decomposition and the inclusion of valid inequalities, within the branch-and-cut method. Section 8 concludes the paper with some directions for further research. All the mathematical proofs of propositions are reported in Appendix A.

2. Literature review

This work is connected to three main streams of literature that we summarize in this section, highlighting the relationship with our contribution.

Decentralized portfolio optimization. The initial interest towards decentralized investment focused on the principal-agent relationship that results from the delegated decisions on portfolio composition (Ou-Yang 2003, Stracca 2006, Maug & Naik 2011). This stream of literature pivots on the study of compensation contracts and their implications for the asset pricing. From a different angle, this problem has been rephrased in the operations research community by Liou & Yao (2005), Thi et al. (2012), Li et al. (2014), who proposed different bilevel optimization models, extending the traditional mean-variance portfolio approach. Their modeling framework considers an investor (leader) who determines how to distribute a fixed budget between a unique risk-free asset and a set of financial securities, whereas a subsidiary (follower) builds the portfolio of securities by maximizing its expected return, conditional upon some measures of risk. Recently, Benita et al. (2018) explored a number of theoretical properties of this bilevel portfolio formulation, boosting its algorithmic solution. Two years later this work has been followed by the work of Leal et al. (2020), who endogenized transaction costs and integrated the CVaR risk measure into the aforementioned bilevel decision-making. Hence, while the idea of decentralized portfolio optimization has been introduced in the operations research literature by the last decade, our approach allows departing from the original bilevel framework (whose globally optimal solutions are difficult to characterize), without losing the decentralization features (such as the determination of the compensation contract, as a management fee).

Risk regulation and aggregation. The use of CVaR in portfolio optimization arose as a risk coherent alternative (satisfying sub-additivity) to the Value at Risk, as first studied by Artzner et al. (1999) and by Rockafellar et al. (2000). Its fundamental role in portfolio optimization is drawn from the results of Prekopa

(1973), who developed an exact mathematical programming framework for constraints involving conditional expectations, as an extension to the chance-constrained model introduced by Charnes et al. (1958). Contextually, Rockafellar et al. (2000) derived a representation of CVaR as a minimization problem, which can be approached by linear programming methods, under a multi-scenario formulation of the stochastic terms. It has been also shown that this risk measure facilitates the construction of approximations for general chance-constrained problems (Chen et al. 2010). For its theoretical properties and for its interpretation in terms of expected loss, a long stream of financial applications and computational studies have appeared in the last two decades, establishing the CVaR measure as a standard benchmark in portfolio management (Andersson et al. 2001, Rockafellar & Uryasev 2002, Fábíán 2008, Zymler et al. 2013, Du & Escanciano 2016, Ban et al. 2016, Lux & Rüschendorf 2019). Contextually, beside the selection of a proper risk measure, MICs have to decide about alternative risk aggregation criteria, when potential losses come from multiple autonomous investments. As noted by Uryasev et al. (2010) and by Hull (2012), financial institutions typically calculate the loss distributions for different business units on a standalone basis and it combines the loss distributions, based on a simple aggregation form. Uryasev et al. (2010) studied and proposed a number of aggregation alternatives. In this vein, we assume that the MIC targets a minimization of the worst-case expected loss among local markets, by setting a threshold level about the maximum expected loss for each SA. Then, from an optimization perspective, the risk neutrality of the SAs (reflecting a morally hazardous behavior) is counteracted by a CVaR regulation which translates an optimality criterion of the MIC (leader) into a feasibility criterion for the SAs (followers). Along the same lines, we exploit overlooked properties of the CVaR constraints to define exact optimality bounds for our problem, as well as to derive a specialized decomposition method.

Linearization and reformulation. While mixed-integer linear programs can be solved with the computational power of existing machines, nonlinear and bi-linear terms commonly appear in mathematical programming formulations of games. The analysis of linearization and reformulation strategies for these terms is an active field of research (Pshenichnyj 2012), having direct bearing on the way in which the endogenous management fee is treated in our problem. Our contribution to this aspect turns on the work of Croxton et al. (2003) (later extended by Vielma et al. (2010)), which allows approximating with arbitrary precision the endogenous management fee by piecewise constant functions. Specifically, Croxton et al. (2003) provided a theoretical comparison between three linearization strategies for piecewise constant functions: the *incremental model* (originally introduced by (Dantzig 1960, Vajda 1964)), the *multiple choice model* (originally introduced by (Balakrishnan & Graves 1989)) and the *convex combination model* (originally introduced by (Dantzig 1998)). Exploiting the combinatorial properties of the piecewise constant management fees, we refine these linearization approaches to derive strong valid inequalities for our problem.

3. Baseline model

Consider an investment context where a single MIC (leader) aims to optimize a combination of expected return and loss (portfolio shortfall) by determining the budget proportion and the level of maximum expected loss in each market. In response, SAs (followers) assemble local portfolios of investment options using the assigned budget. We assume that the MIC ignores the actual probability distribution of the individual rate of returns, so that he/she is unable to deterministically characterize the SAs best responses. Therefore, building on an incomplete information setting, the MIC might consider a collection of possible types of SAs, so that a given SA can be of a certain type or another with a given probability.

Before presenting our model, we introduce the symbols that will be used throughout the paper.

Sets:

- \mathcal{M} the set of local markets (with $|\mathcal{M}| = m$);
- \mathcal{N}_k the set of investment options in the k^{th} local market (with $|\mathcal{N}_k| = n_k$);
- \mathcal{T}_k the set of possible types in the k^{th} local market (with $|\mathcal{T}_k| = \tau_k$).

Parameters:

- β the risk aversion parameter, quantifying the monetary cost of expected losses;
- R_{ikt} a random rate of return of asset i in market k , when the market is of type t ;
- p_{kt} the probability of type t in market k .

Decision variables:

- Θ the tolerable worst-case expected loss from local markets;
- z_k the budget proportion invested in the k^{th} local market;
- x_{ikt} the budget proportion invested in asset i in the k^{th} local market by a SA of type t .

Functions:

- δ_k the proportion of the investment return retained by the k^{th} SA;
- U_{kt} the expected portfolio return in market k when the SA is of type t .

The rates of return are defined on a probability space $(\Xi, \mathcal{F}, \mathbb{P})$, whose specification is not critical for our results to be valid. Their marginal distributions are denoted as $G_{R_{ikt}}$.⁵ Given a random quantity R , we define the expectation $\hat{r} = \mathbb{E}[R]$ and the conditional expectation $\tilde{r}^{(\alpha)} = \mathbb{E}[R \mid R \leq G_R^{-1}(1 - \alpha)]$, where $G_R^{-1}(1 - \alpha)$ is the image of $1 - \alpha$ by the inverse cumulative distribution function of R . In the rest of this paper, $\tilde{r}^{(\alpha)}$ is assumed to be strictly negative for a sufficiently large value of α , as it is the case for the vast majority of applications. When clear from the context, the notation \tilde{r} is used instead of $\tilde{r}^{(\alpha)}$. In vector form, we use the following boldface characters: $\mathbf{z} = [z_1 \dots z_m]^\top$, $\mathbf{x}_{kt} = [x_{1,kt} \dots x_{n_k,kt}]^\top$, $\mathbf{R}_{kt} = [R_{1,kt} \dots R_{n_k,kt}]^\top$, $\hat{\mathbf{r}}_{kt} = [\hat{r}_{1,kt} \dots \hat{r}_{n_k,kt}]^\top$ and $\tilde{\mathbf{r}}_{kt}^{(\alpha)} = [\tilde{r}_{1,kt}^{(\alpha)} \dots \tilde{r}_{n_k,kt}^{(\alpha)}]^\top$, for each $k \in \mathcal{M}$ and $t \in \mathcal{T}_k$.⁶

To define the leader-follower game, the payoff for both types of players is expressed as:

$$\begin{aligned} \text{MIC payoff} &\equiv \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) U_{kt}(\mathbf{x}_{kt}) - (1 - \beta)\Theta, \\ k^{th} \text{ SA payoff of type } t &\equiv \delta_k(z_k) U_{kt}(\mathbf{x}_{kt}), \end{aligned}$$

where $U_{kt}(\mathbf{x}_{kt}) = \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt}$. Thus, the MIC payoff is a convex combination between the expected portfolio return (averaged over the SA's types) and the worst-case expected loss. The function δ_k captures a form of transaction cost, modeled as a proportion of the investment return retained by each SA, as a response to the budget allocation z_k by the MIC in the k^{th} local market.⁷

Different ways of accounting for the transaction costs into portfolio models have been studied by the last two decades (Best & Hlouskova 2003, Lobo et al. 2007, Potapchik et al. 2008, Leal et al. 2020). Most of them assume that transaction costs are subtracted from the expected return. Our model departs from this logic to capture a different form of partnership between a central headquarter and local affiliates, who

⁵Note that in the incomplete information the MIC knowledge of the rate of return of asset i in market k is a mixture of distributions: $G_{R_{ik}} = \sum_t p_{kt} G_{R_{ikt}}$.

⁶To facilitate the readability, the notation adopted for subscript indexes is that letters are written consecutively, whereas numbers are separated by a comma.

⁷This function plays a fundamental role in the complexity of the problem and its solution strategy, as discussed next in this section and in Section 5.2.

agreed upon sharing rules on the portfolio return, reflecting management fees (Doukas & Lang 2003, Barry & Kearney 2006, Ambos et al. 2010, Kawai & Strange 2014).

In a similar vein, the MIC decisions on z_k and Θ can be seen as a regulatory system over decentralized affiliates, forcing the expected losses of the investment composition $x_{1,kt} \dots x_{n_k,kt}$ (selected by the k^{th} SA of type t) below a tolerable level (which reflects the MIC sensitivity β). Contextually, for a given local portfolio \mathbf{x} associated to returns \mathbf{R} (we can drop the sub-indexes k and t as the argument is valid for all markets and all SA types), we define a random variable capturing each market loss as:

$$L(z, \mathbf{x}, \mathbf{R}) = z - \mathbf{R}^\top \mathbf{x}. \quad (1)$$

The CVaR of the local portfolio loss with shortfall probability (confidence level) α is derived from its left-continuous quantile function: $F^{-1}(z, \mathbf{x}, \alpha) = \inf \{q : F(z, \mathbf{x}, q) \geq \alpha\}$, where $F(z, \mathbf{x}, \eta) = \mathbb{P} \{z - \mathbf{R}^\top \mathbf{x} \leq \eta\}$. At the k^{th} local market, the probability distribution of portfolio loss when the SA is of type t is denoted as $F_{kt}(z, \mathbf{x}_{kt}, \eta)$. Then, the CVaR with shortfall probability α (frequently denoted as α -CVaR) comes out as the conditional expectation of the local loss to that loss being non-smaller than its α -quantile:

$$\langle\langle L(z, \mathbf{x}, \mathbf{R}) \rangle\rangle_\alpha = \mathbb{E} [L(z, \mathbf{x}, \mathbf{R}) \mid L(z, \mathbf{x}, \mathbf{R}) \geq F^{-1}(z, \mathbf{x}, \alpha)]. \quad (2)$$

Rockafellar et al. (2000) showed that for smooth F the following CVaR representation is valid:

$$\langle\langle L(\mathbf{x}, \mathbf{R}) \rangle\rangle_\alpha = \min_{\eta} \left\{ \eta + \frac{1}{(1-\alpha)} \mathbb{E}_F [(L(z, \mathbf{x}, \mathbf{R}) - \eta)^+] \right\}, \quad (3)$$

where $(s)^+ = \max\{s, 0\}$ and $\mathbb{E}_F[W]$ is the expectation of W based on the distribution F . This representation plays an important role in portfolio optimization, as it entails that $\langle\langle L(\mathbf{x}, \mathbf{R}) \rangle\rangle_\alpha$ is a convex function of the portfolio weights \mathbf{x} .⁸ Hence, the constraint $\langle\langle L(z, \mathbf{x}, \mathbf{R}) \rangle\rangle_\alpha \leq \Theta$ imposes a maximum level of tolerable risk, common to all local market and agent types.

The SA best response $\Psi_{kt}(z, \Theta) : [0, 1] \times \mathbb{R} \rightarrow 2^{[0,1]^{n_k}}$ is the solution set mapping of the k^{th} SA problem of type t for fixed decisions (z, Θ) by the investor.⁹ This is expressed as

$$\Psi_{kt}(z_k, \Theta) = \underset{\mathbf{x}_k}{\operatorname{argmax}} \delta_k(z_k) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt}, \quad \text{subj. to } \mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta),$$

where

$$\Lambda_{kt}(z, \Theta) = \{\mathbf{x} \geq 0 \mid \mathbf{1}^\top \mathbf{x} \leq z, \langle\langle L(z, \mathbf{x}, \mathbf{R}_{kt}) \rangle\rangle_\alpha \leq \Theta\},$$

The Perfect Bayesian Equilibrium of this sequential game is obtained by solving a bilevel nonlinear problem with $m\tau$ convex follower problems (where τ is the average number of types per market) and a nonlinear leader problem:¹⁰

$$\phi = \begin{cases} \max_{\mathbf{z}, \mathbf{x}, \Theta} \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) U_{kt}(\mathbf{x}_{kt}) - (1 - \beta) \Theta & (4a) \\ \text{subj. to } \mathbf{x}_{kt} \in \Psi_{kt}(z_k, \Theta) & k = 1 \dots m, t = 1 \dots \tau_k & (4b) \\ \mathbf{z} \in \Xi, \Theta \geq 0 & (4c) \end{cases}$$

⁸For detailed discussions about the properties of CVaR, see the works of Rockafellar et al. (2000), Rockafellar & Uryasev (2002), Pflug (2000) and Acerbi & Tasche (2002).

⁹Here $2^{[0,1]^{n_k}}$ is the family of all non-empty subsets of $[0, 1]^{n_k}$.

¹⁰The convexity of the follower problems arises from the fact that $\langle\langle L(z, \mathbf{x}, \mathbf{R}) \rangle\rangle_\alpha$ is convex (and continuously differentiable when the distribution of \mathbf{R} is continuous), as shown by Andersson et al. (2001) and Rockafellar et al. (2000).

where $\Xi = \{\mathbf{z} \in [0, 1]^m \mid \mathbf{1}^\top \mathbf{z} = 1\}$. The formulation (4a)-(4c) adopts the *optimistic approach*, in which the MIC is able to influence the SAs, when multiple best responses are available.¹¹

4. Dominance properties and solvability

This section focuses on determining fundamental properties of the introduced leader-follower game for multi-market portfolio optimization with conditional value at risk.

Proposition 1 (Existence). *Consider the bilevel problem (4a)-(4c). The investor can always find at least one optimal budget allocation \mathbf{z} and shortfall Θ .*

The proof of Proposition 1 is in Appendix A.

Proposition 2 (Loss bounds). *For any random vector \mathbf{R} , the following bounds hold:*

$$z - \hat{\mathbf{r}}^\top \mathbf{x} \leq \langle \langle z - \mathbf{R}^\top \mathbf{x}, F \rangle \rangle_\alpha \leq z - \tilde{\mathbf{r}}^\top \mathbf{x} \quad (5)$$

The proof of Proposition 2 is in Appendix A.¹²

A direct consequence of Proposition 2 is the following bounds on the tolerable expected loss, that must be satisfied by all solutions of (4a)-(4c):

$$\max\{z_k - \hat{\mathbf{r}}_{kt}^\top \mathbf{x}_{kt} : k \in \mathcal{M}, t \in \mathcal{T}_k\} \leq \Theta \leq \min\{z_k - \tilde{\mathbf{r}}_{kt}^\top \mathbf{x}_{kt} : k \in \mathcal{M}, t \in \mathcal{T}_k\}. \quad (6)$$

For a complete characterization of the solutions of (4a)-(4c), we first consider its high point relaxation:

$$\left\{ \begin{array}{l} \max_{\mathbf{z}, \mathbf{x}_k, \Theta} \quad \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) U_{kt}(\mathbf{x}_{kt}) - (1 - \beta)\Theta \quad (7a) \\ \text{subj. to} \quad \mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta) \quad k = 1 \dots m, t = 1 \dots \tau_k \quad (7b) \\ \mathbf{z} \in \Xi, \Theta \geq 0 \quad (7c) \end{array} \right.$$

This problem provides a best first level solution, under the assumption that the second level reacts in the most favorable way for the leader (Dempe 2002).

Proposition 3 (Equivalence). *The solutions of (4a)-(4c) coincide to those of (7a)-(7c).*

The proof of Proposition 3 is in Appendix A.

From the economic viewpoint, Proposition 3 is in consonance with certain aligned incentives between the MIC and the SAs (captured by the expected portfolios' returns U_{kt}). However, the fundamental divergence between the MIC and SAs' strategic decisions is driven by the role that the CVaR risk plays in their respective problems. While the MIC targets a minimization of the worst-case expected loss among local markets, by setting a threshold level Θ , the SAs receive this expected loss threshold as a regulation (a constraint on their range of decisions). Therefore, the risk neutrality of the SAs (reflecting a morally hazardous behavior) is counteracted by a CVaR regulation which translates an optimality criterion of the leader into a feasibility criterion for the followers.

¹¹Behind the modeling advantages of optimistic approaches, it must be stressed that this assumption is of little practical importance, due to the random nature of the objective function and constraints, entailing that alternative optimal solutions are unlikely to happen at the follower level.

¹²Note that the right-hand side of (5) follows from the second principles of coherent risk measures (sub-additivity), as first introduced by Artzner et al. (1999).

On the computational side, the main consequence of Proposition 3 is that problem (7a)-(7c) can be equivalently solved, instead of problem (4a)-(4c), paving the way for a fundamental departure between the solution approaches applied by Benita et al. (2018) and by Leal et al. (2020), for related classes of decentralized investment problems. That having been said, the high point relaxation problem (7a)-(7c) is still a difficult problem to solve. In fact, the following proposition implies its NP-Hardness.¹³

Proposition 4 (NP completeness). *Let $\Theta > 1$ and consider the case when $n_k = \tau_k = 1$ for all $k \in \mathcal{M}$ (one asset and one type per market). Let $MICP(\Theta)$ be the recognition version of the MIC problem which asks whether there exists a solution to (7b)-(7c) for which the objective function (7a) is larger than or equal to a given threshold, for fixed Θ . We claim that*

$$\text{if } \delta_k(z) = 1 - \frac{\min\{z(1 - \tilde{r}_{k,1})/\Theta, 1\}}{\hat{r}_{k,1}}, \text{ then } MICP(\Theta) \text{ is NP complete.}$$

The proof of Proposition 4 is in Appendix A.

The overall picture of the introduced leader-follower game resulting from the above propositions gives a glimpse of exact solution strategies, optimality bounds and approximations, as extensively explored in the rest of this paper.

5. Numerical solution strategies

Focusing on the equivalent problem (7a)-(7c), this section provides a mixed integer programming reformulation, that can be efficiently tackled by a decomposition procedure and the generation of strong valid inequalities.

5.1. Multi-scenarios reformulation and decomposition

Using the CVaR representation (3), a multi-scenarios reformulation is presented hereafter. Firstly, the problem of the k^{th} SA of type t can be expressed as

$$\left\{ \begin{array}{l} \max_{\mathbf{x}_{kt}} \delta_k(z_k) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt} \\ \text{subj. to } \mathbf{1}^\top \mathbf{x}_{kt} \leq z_k \\ \min_{\eta_{kt}} \{ \eta_{kt} + (1 - \alpha)^{-1} \mathbb{E}_{F_{kt}} [(L(z_k, \mathbf{x}_{kt}, \mathbf{R}_{kt}) - \eta_{kt})^+] \} \leq \Theta \\ \mathbf{x}_{kt} \geq 0. \end{array} \right. \quad \begin{array}{l} (8a) \\ (8b) \\ (8c) \\ (8d) \end{array}$$

Note that constraint (8c) is satisfied iff there exist η_{kt} satisfying

$$\eta_{kt} + (1 - \alpha)^{-1} \mathbb{E}_{F_{kt}} [(L(z_k, \mathbf{x}_{kt}, \mathbf{R}_{kt}) - \eta_{kt})^+] \leq \Theta.$$

¹³Next to this result, it must be noticed that when $\delta_k(z)$ is twice continuously differentiable, the Hessian matrix of the objective function of problem (4a)-(4c) becomes indefinite. Sahni (1974) showed that the maximization of quadratic functions with positive definite Hessian matrices over a compact polyhedron is NP-hard, while Pardalos & Vavasis (1991) proved a stronger condition by showing that even one negative eigenvalue makes the problem NP-hard.

An exhaustive proof of this equivalence is provided by [Rockafellar & Uryasev \(2002\)](#). Therefore, the solutions of the following problem coincide with the ones of problem (8a)-(8d):

$$\left\{ \begin{array}{l} \max_{\mathbf{x}_{kt}, \eta_{kt}} \delta_k(z_k) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt} \quad (9a) \\ \text{subj. to } \mathbf{1}^\top \mathbf{x}_{kt} \leq z_k \quad (9b) \\ \eta_{kt} + (1 - \alpha)^{-1} \mathbb{E}_{F_{kt}} [(L(z_k, \mathbf{x}_{kt}, \mathbf{R}_{kt}) - \eta_{kt})^+] \leq \Theta \quad (9c) \\ \mathbf{x}_k \geq 0. \quad (9d) \end{array} \right.$$

Secondly, problem (9a)-(9d) can be reformulated by approaching the expectation in (9c), by sampling a collection of realizations of \mathbf{R}_{kt} , as suggested by [Fábíán \(2008\)](#). Let \mathcal{S} be a set of indexes of s realizations of \mathbf{R}_{kt} (with $|\mathcal{S}| = s$). The multi-scenario reformulation of (7a)-(7c) is then expressed as follows

$$\phi'' = \left\{ \begin{array}{l} \max_{\mathbf{z}, \mathbf{x}, \mathbf{w}, \Theta, \eta} \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt} - (1 - \beta)\Theta, \quad (10a) \\ \text{subj. to } \mathbf{1}^\top \mathbf{x}_{kt} \leq z_k \quad k \in \mathcal{M}, t \in \mathcal{T}_k \quad (10b) \\ \mathbf{1}^\top \mathbf{z} \leq 1, \quad (10c) \\ \eta_{kt} + (1 - \alpha)^{-1} \sum_{j \in \mathcal{S}} \pi_j w_{jkt} \leq \Theta \quad k \in \mathcal{M}, t \in \mathcal{T}_k \quad (10d) \\ w_{jkt} \geq z_k - (\mathbf{r}_{kt}^j)^\top \mathbf{x}_{kt} - \eta_{kt} \quad j \in \mathcal{S}, k \in \mathcal{M}, t \in \mathcal{T}_k \quad (10e) \\ \mathbf{x}_{kt} \geq 0 \quad k \in \mathcal{M}, t \in \mathcal{T}_k \quad (10f) \\ w_{jkt} \geq 0 \quad j \in \mathcal{S}, k \in \mathcal{M}, t \in \mathcal{T}_k \quad (10g) \\ \mathbf{z} \geq 0, \Theta \geq 0 \quad (10h) \end{array} \right.$$

where $(\mathbf{r}_{1,1}^j \dots \mathbf{r}_{m,\tau_m}^j)$ is a realization of the $(m\tau)$ -dimensional vector \mathbf{R} , whose probability is π_j . From the Glivenko–Cantelli theorem, the CVar constraints in (10a)-(10h) becomes asymptotically correct when s grows large. However, as the multi-scenario reformulation requires $m + 1 + \sum_{k=1}^m ((s + 1)\tau_k + \sum_{t=1}^{\tau_k} n_k)$ continuous variables, this correctness comes at the expense of the computational effort to solve (10a)-(10h) to approach (7a)-(7c).

This large number of variables is mainly driven by the presence of the newly introduced variables w_{jtk} . However, these variables can be eliminated using the Fourier-Motzkin procedure ([Dal Sasso et al. 2019](#)). By doing so, inequalities (10d)-(10e) can be replaced by the following (exponential) family of constraints:

$$\eta_{kt} + (1 - \alpha)^{-1} \sum_{j \in \mathcal{S}'_{kt}} \pi_j (z_k - (\mathbf{r}_{kt}^j)^\top \mathbf{x}_{kt} - \eta_{kt}) \leq \Theta \quad \text{for } k \in \mathcal{M}, t \in \mathcal{T}_k, \quad (11)$$

where $\mathcal{S}'_{kt} \subseteq \mathcal{S}$. Given a subset \mathcal{H}_{kt} of these 2^s candidate constraints, one can generate an upper bound to

ϕ'' as follows

$$\phi''(\tilde{\mathcal{H}}) = \begin{cases} \max_{\mathbf{z}, \mathbf{x}, \Theta} & \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt} - (1 - \beta)\Theta, & (12a) \\ \text{s. to} & \mathbf{1}^\top \mathbf{x}_{kt} \leq z_k & k \in \mathcal{M}, t \in \mathcal{T}_k & (12b) \\ & \mathbf{1}^\top \mathbf{z} \leq 1, & (12c) \\ & A_{kth}^\eta \eta_{kt} + \sum_{i \in \mathcal{N}_k} A_{ikth}^x x_{ikt} \leq \Theta & k \in \mathcal{M}, t \in \mathcal{T}_k, h \in \tilde{\mathcal{H}} & (12d) \\ & \mathbf{x}_{kt} \geq 0 & k \in \mathcal{M}, t \in \mathcal{T}_k & (12e) \\ & \mathbf{z} \geq 0, \Theta \geq 0 & (12f) \end{cases}$$

where $\tilde{\mathcal{H}} = (\mathcal{H}_{11} \dots \mathcal{H}_{m\tau_m})$ and

$$A_{kth}^\eta = 1 - (1 - \alpha)^{-1} \sum_{j \in \mathcal{S}_{kt}^h} \pi_j \quad \text{and} \quad A_{ikth}^x = 1 - (1 - \alpha)^{-1} \sum_{j \in \mathcal{S}_{kt}^h} \pi_j r_{ikt}^j. \quad (13)$$

Therefore, $\phi''(\tilde{\mathcal{H}}) \geq \phi''(\tilde{\mathcal{H}}')$ for all $\tilde{\mathcal{H}} \subseteq \tilde{\mathcal{H}}'$. The computational challenge is then the iterative selection of constraints to be included. We establish the following generation procedure:

0. Initialization: $\mathcal{H}_{kt} = \emptyset$.
1. Solve (12a)-(12f) and let $(z_k, \mathbf{x}_{kt}, \eta_{kt})$ be the corresponding solution.
2. If there exists a $k \in \mathcal{M}$ and $t \in \mathcal{T}_k$ for which (11) is not verified, go to 3. Else, stop.
 - 3.a. Define $\mathcal{S}_{kt} \subseteq \{j \in \mathcal{S} \mid z_k > (\mathbf{r}_{kt}^j)^\top \mathbf{x}_{kt} + \eta_{kt}\}$ and the corresponding coefficients in (13).
 - 3.b. Set $\mathcal{H}_{kt} = \mathcal{H}_{kt} \cup \mathcal{S}_{kt}$.
 - 3.c. Go to 1.

The efficiency of this procedure is numerically studied in Section 7.

5.2. Discretisation and linearisation of the intermediation fees

The exact solution of problem (12a)-(12f) at each iteration of the above decomposition procedure is still a challenging computational task, due to the presence of δ_k in the objective function. Taking advantage of applications where management fees are set at a finite collection of intervals, this subsection casts a closer look at piecewise constant intermediation fees δ_k , with a view to deducing strong valid inequalities (hereafter referred to as VI), for the algorithmic solvability of large instances of (12a)-(12f) and (10d)-(10e).

When δ_k is constant (exogenous intermediation fees), problems (12a)-(12f) and (10d)-(10e) are solvable using standard linear programming approaches (although the presence of the CVaR constraints at multiple markets and multiple type makes this linear program a large-scale and potentially challenging one). However, intermediation fee policies are in general monotonically decreasing with respect to the budget allocation.

On the one hand, any specification of δ_k can be approximated by a piecewise constant function. On the other hand, the opposite might also be the case, as δ_k can be originally defined as a piecewise constant function, where SAs establish a collection $\mathcal{L} = \{1 \dots L\}$ of budget allocation thresholds $b_1 \dots b_L$ and corresponding intermediation fees $d_1 > d_2 > \dots > d_L$. In other words, $\delta_k(z_k) = d_\ell$ when z_k is in the interval $[b_\ell, b_{\ell+1}]$. In both cases, the MIC objective function involves nonlinear terms $(1 - \delta(z_k))x_{i,k}$ for which we present two different MIP representations, related to the linearization strategies studied by Croxton et al. (2003) and Vielma et al. (2010): the *incremental model* and the *multiple choice model*. In the rest of this subsection, we drop the subindexes k and t , as the presented results are valid for all SAs.

Linearization based on the incremental model. Adapting this linearization strategy to $\delta(z)$ requires the inclusion of continuous variables $\tilde{z}_1 \dots \tilde{z}_L$ representing the segment loads and binary variables $s_1 \dots s_L$ indicating whether $z \leq b_l$. The following model allows representing the different values that $\delta(z)$ can take:

$$\left\{ \begin{array}{ll} \sum_{l \in \mathcal{L}} \tilde{z}_l = z & \\ (b_\ell - b_{\ell-1})s_{\ell+1} \leq \tilde{z}_\ell & \ell = 2 \dots L-1 \\ (b_\ell - b_{\ell-1})s_\ell \geq \tilde{z}_\ell & \ell = 2 \dots L \\ \tilde{z}_\ell \geq 0 & \ell = 1 \dots L \\ s_\ell \text{ binary} & \ell = 1 \dots L \end{array} \right. \quad (\text{lin-IM})$$

Thus, s_ℓ is equal to zero for the rightmost segments. This formulation requires the inclusion of L binary variables, $L + 1$ continuous variables and $3L + 1$ constraints. Next, let $\varphi = (1 - \delta(z))x$. Together with (lin-IM), the following set of constraints provides a valid representation of this function. Given that in the MIC objective function the coefficient of φ is nonnegative, a valid model only needs to bound φ from above:

$$\left\{ \begin{array}{ll} [(1 - d_\ell)x + s_{\ell+1}] \geq \varphi & \ell = 1 \dots L-1 \\ (1 - d_L)x \geq \varphi & \end{array} \right. \quad (\text{lin-of-IM})$$

Linearization based on the multiple choice model. As in the incremental model, we include the continuous variables $\tilde{z}_1 \dots \tilde{z}_L$ (representing the segment loads) and binary variables $s'_1 \dots s'_L$ indicating whether $b_{\ell-1} \leq z \leq b_\ell$.

$$\left\{ \begin{array}{ll} \sum_{\ell \in \mathcal{L}} \tilde{z}_\ell = z & \\ \sum_{\ell \in \mathcal{L}} s'_\ell = 1 & \\ b_{\ell-1}s'_\ell \leq \tilde{z}_\ell & \ell = 2 \dots L \\ b_\ell s'_\ell \geq \tilde{z}_\ell & \ell = 1 \dots L \\ \tilde{z}_\ell \geq 0 & \ell = 1 \dots L \\ s'_\ell \text{ binary} & \ell = 1 \dots L \end{array} \right. \quad (\text{lin-MCM})$$

In this case s'_ℓ is equal to zero for all the segments except the selected one. As for the incremental model, this formulation requires the same collection of additional decision variables, but a larger set of constraints (i.e., $4L + 1$ constraints). Further, function $\varphi = (1 - \delta(z))x$ can be represented by adding the following set of constraints to (lin-MCM):

$$\left\{ \begin{array}{ll} [(1 - d_\ell)x + (1 - s'_\ell)] \geq \varphi & \ell = 1 \dots L \\ (1 - d_\ell)x - (1 - s'_\ell) \leq \varphi & \ell = 1 \dots L \end{array} \right. \quad (\text{lin-of-MCM})$$

Comparing the two models. Since (lin-IM) and (lin-MCM) are valid, Croxton et al. (2003) answered the question about whether one is better than the other. They found that *the LP relaxations of the incremental and multiple choice formulations are equivalent, in the sense that any feasible solution of one LP relaxation corresponds to a feasible solution of the other with the same cost*. However in our problem, the nonlinear function φ has a particular structure : it is the product of a piecewise constant function depending on z and another variable x . As a consequence, one representation is stronger than the other.

Proposition 5 (Bound on the LP relaxation). *The upper bound on φ from the LP relaxation of (lin-of-IM) is smaller than or equal to the one from the LP relaxation of (lin-of-MCM).*

The proof of Proposition 5 is in Appendix A.

Both formulations, (lin-of-IM) and (lin-of-MCM), can be strengthened by some families of VIs. We present them using the variables of (lin-MCM).

Proposition 6 (Strong VIs for (lin-of-MCM)). *Let φ be the MIC profit from a given stock in a given local market. The following is a collection of strong VIs for the linearized problem:*

$$\varphi \leq \hat{r} \left[(1 - d_\ell)x + \sum_{\ell'=\ell+1}^L (d_\ell - d_{\ell'})b_{\ell'}s'_{\ell'} \right], \quad \text{for } \ell = 1 \dots L. \quad (14)$$

The proof of Proposition 6 is in Appendix A.

The impact of Proposition 6 for the solution of large-scale instances of (12a)-(12f) and (10d)-(10e) is numerically studied in Section 7.

6. Optimality bounds

As an alternative to the exact (but computationally challenging) solution of the equivalent problem (7a)-(7c), optimality bounds can be computed, exploiting some of the previously studied dominance properties. In fact, building on Proposition 2, we define $g_{ikt}(\gamma) = -(\gamma\tilde{r}_{ikt} + (1-\gamma)\hat{r}_{ikt})$ and consider a γ -bound version of the MIC problem:

$$\phi'''(\gamma) = \begin{cases} \max_{\mathbf{z}, \mathbf{x}_k, \Theta} & \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) \sum_{i=1}^{n_k} \hat{r}_{ikt} x_{ikt} - (1 - \beta)\Theta, & (15a) \\ \text{subj. to} & \mathbf{1}^\top \mathbf{x}_{kt} \leq z_k & k \in \mathcal{M}, t \in \mathcal{T}_k & (15b) \\ & \mathbf{1}^\top \mathbf{z} \leq 1, & & (15c) \\ & \mathbf{g}_{kt}(\gamma)^\top \mathbf{x}_{kt} \leq \Theta - z_k & k \in \mathcal{M}, t \in \mathcal{T}_k & (15d) \\ & \mathbf{x}_{kt} \geq 0 & k \in \mathcal{M}, t \in \mathcal{T}_k & (15e) \\ & \mathbf{z} \geq 0, \Theta \geq 0 & & (15f) \end{cases}$$

Problem (15a)-(15f) has $m + \sum_{k=1}^m n_k \tau_k + 1$ continuous variables and verifies properties which are important to uncover its relationship with the original problem (4a)-(4c).

Proposition 7. *We have the following cases:*

- (i) $\phi'''(1) \leq \phi$;
- (ii) $\phi'''(0) \geq \phi$;
- (iii) $\phi'''(\gamma) \geq \phi'''(\gamma')$, for any $\gamma \in [0, 1]$, such that $\gamma \leq \gamma'$.

The proof of Proposition 7 is in Appendix A.

For this relationship with the original problem and the substantial simplification that it entails on a computational level, problem (15a)-(15f) represents a useful alternative to (4a)-(4c), when large-scale instances are taken into account. We describe hereafter closed-form properties of (15a)-(15f), which facilitate the interpretability of its solution.

Proposition 8. Let $i(k, t)$ be the index $i \in \mathcal{N}_k$ corresponding to the maximum value of \hat{r}_{ikt} and assume that $\hat{r}_{i(k,t)kt} \geq 0$. If for all $i \in \mathcal{N}_k$

$$\hat{r}_{ikt} \leq \frac{\hat{r}_{i(k,t)kt}}{g_{i(k,t)kt}(\gamma)} g_{ikt}(\gamma)$$

then the γ -bound version of the SA problem admits of the following solution for any budget proportion z_k and risk level Θ :

$$x_{ikt} = \begin{cases} \min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} & \text{if } i = i(k, t) \\ 0 & \text{otherwise} \end{cases} \quad (16)$$

The intuition behind Proposition 8 is the fact that the selection of a unique investment option per market only depends on the structure of the data set (i.e., the expected returns \hat{r}_{ik} and expected shortfalls \tilde{r}_{ik}), not on the leader decision on z_k and Θ . Therefore, when the data set verifies the condition in Proposition 8 and the γ -bound is used to approximate the CVaR measure, the MIC problem can be written as

$$\phi'(\gamma) = \begin{cases} \max_{\mathbf{z}, \Theta} \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) \hat{r}_{i(k,t)kt} \min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} - (1 - \beta)\Theta, & (17a) \\ \text{subj. to } \mathbf{1}^\top \mathbf{z} = 1, & (17b) \\ \mathbf{z} \geq 0, \Theta \geq 0 & (17c) \end{cases}$$

This problem has $m + 1$ continuous variables and one constraint. The following result characterizes the optimal Θ solution of (17a)-(17c). We denote the objective function (17a) by $\xi(\Theta, z_1, \dots, z_m)$.

Corollary 1. Let $\tau_k = 1$ and $i(k, 1)$ be the index $i \in \mathcal{N}_k$ corresponding to the maximum value of $\hat{r}_{ik,1}$, for all $k \in \mathcal{M}$. For any fixed $z_1 \dots z_m$, let $j_1 \dots j_m$ be a permutation of the indices in such a way that $z_{j_1}(g_{i(j_1,1),1} + 1) < z_{j_2}(g_{i(j_2,1),1} + 1) < \dots < z_{j_m}(g_{i(j_m,1),1} + 1)$. Under the assumptions of Proposition 8, we claim that $\xi(\Theta, z_1, \dots, z_m)$ is concave and piecewise linear in Θ . Further, the unique optimal solution Θ^* of $\xi(\Theta, z_1, \dots, z_m)$ is

$$\Theta^* = \begin{cases} 1, & \text{if } \beta \sum_{k=1}^m \hat{r}_{i_k, k, 1} (1 - \delta_k(z_k)) / g_{i(k,1),1}(\gamma) \leq (1 - \beta) \\ z_h (g_{i(h,1),1}(\gamma) + 1) & \text{otherwise} \end{cases}$$

where the sub-index h is the largest value $\ell \in \{1, \dots, m\}$ satisfying the inequality $\beta \sum_{k=\ell+1}^m \hat{r}_{i_{j_k}, j_k} (1 - \delta_k(z_{j_k})) / g_{i(j_k), j_k}(\gamma) > (1 - \beta)$.

The proof of Corollary 1 is in Appendix A.

Consistently with (6), the intuition behind Corollary 1 defines the leading factor of the MIC decision on tolerable expected loss, which is the tail behavior of returns in the main local market (i.e. the local market in which the MIC invests the most).

7. Empirical analysis and computational experiments

In this section, large-scale instances of problem (12a)-(12f) are constructed and solved based on the decomposition procedure presented in Subection 5.1, the linearization and strong valid inequality presented in Subsection 5.2 and the γ -bounds presented in Section 6.

Instances reflects an investment setting in which a MIC attempts to invest in the U.S. stock market relying on local brokers who are tasked with selecting stocks from listed enterprises belonging to different industries (i.e., our set \mathcal{M}). We use financial data from US listed enterprises in the time horizon 1999–2014

(as described in Subsection 7.1). With a view to assessing the impact of the incomplete information, the different follower types are built by assuming increasing order of uncertainties. Let r_{ikt}^j be an observed return for a given company at a given market under a given scenario. The return at type t is defined by keeping the expectation invariant with respect to the one observed in the data set: $r_{ikt}^j = r_{ik}^j/t - (1/t - 1)\hat{r}_{ik}$. Therefore, the first follower type has an uncertainty analogous to the one observed in the data set and the subsequent types (ordered from 1 to τ_k) have smaller and smaller level of uncertainty. The MIC has a probability distribution over the different follower types, that we assume to be uniform: $p_{tk} = 1/\tau_k$.

Three types of computational tests are provided in this section.

- In Subsection 7.2 the impact of solving problem (7a)-(7c) is studied, by comparing its solution to a market portfolio (obtained by setting $z_k = 1/m$, for all $k \in \mathcal{M}$, and solving (7a)-(7c) with respect to the remaining variables).
- In Subsection 7.3 the relationship between the γ -bound transformation (15a)-(15f) and the exact problem (7a)-(7c) is examined.
- In Subsection 7.4, the efficiency of different solution methods is analyzed, using the decomposition procedure presented in Subection 5.1, the linearization and strong valid inequality presented in Subsection 5.2.

All optimization procedures are solved using cutting edge solvers for mixed-integer optimization (i.e. the IBM ILOG CPLEX 12.9 implementation of the Branch-and-cut algorithm and the Benders decomposition) on an R5500 work-station with processor Intel(R) Xeon(R) CPU E5645 2.40 GHz, and 48 Gbytes of RAM, under a Windows Server 2012 operating system.

7.1. Data selection and processing

Using data from the *Center for Research in Security Prices*, stock prices of 7256 U.S. listed enterprises within the period 1999–2014 are considered. Our collection of local markets \mathcal{M} is composed of $m = 74$ industries, from the Global Industry Classification Standard.

Table 1 reports relevant summary information of the data set. From top to bottom the rows contain information about the mean, the standard deviation, the minimum and the maximum number of listed enterprises and stock return per industry. These 16 year information are grouped in two period (corresponding to different market regimes): before and after the 2008 crisis.

	n_k		R	
	Before crisis	After crisis	Before crisis	After crisis
mean	127.0548	94.2714	0.014	0.009
sd	132.1056	97.0266	0.033	0.029
min	2.0000	7.000	-0.141	-0.099
max	781.0000	514.00	0.277	0.134

Table 1: Historical information concerning the number of listed enterprises per industry and stock returns. The summary statistics are taken over monthly averages per industry.

Aggregate information of four leading industries are reported in Table 2, including the number of stocks and the α -CVaR.

	# stocks	0.5-CVaR	0.5-CVaR
Banks	869	-0.10	-0.03
Software	545	-0.21	-0.06
Consumable Fuels	512	-0.15	-0.05
Biotechnology	471	-0.25	-0.08

Table 2: Aggregate information of four among the 74 industries.

Further descriptions of this stock return data set are provided by [Nasini & Erdemlioglu \(2019\)](#).

7.2. The impact of the endogenous multi-market partitioning

In this subsection the impact of solving model (7a)-(7c) is studied, by comparing the expected return of the global portfolio, the number of stocks and the value of the tolerable worst-case expected loss from local markets. As a benchmark for the comparison, we consider the market portfolio (obtained by setting $z_k = 1/m$, for all $k \in \mathcal{M}$, and solving with respect to the remaining variables).

These computational test involves $16 \times 2^4 = 256$ instances for the multi-scenario reformulation (12a)-(12f), obtained by replicating the 16 years of observations presented in Subsection 7.1, based on combinations of four problem parameters. The four parameters of the experiment are the risk aversion parameter β (fixed at two levels, $\beta \in \{0.1, 0.9\}$), the quantile parameter α (fixed at two levels, $\alpha \in \{0.5, 0.9\}$), the number of piecewise constant steps L for the linearization of δ (fixed at two levels, $L \in \{2, 10\}$), the number of types τ_k (fixed at two levels, $\tau_k \in \{2, 10\}$). Therefore, the combinations of β , α , L and T give rise to 16 instances for each of the 16 years of historical data. For the decomposition procedure, each instance is replicated for $\kappa \in \{1, 5\}$ and $\varepsilon \in \{0, 0.01\}$. Overall, a collection of $256 \times 2 + 256 \times 2 \times 4 = 2560$ instances are solved.

The summary reports in Table 3 (for $\beta = 0.1$ and $\alpha = 0.5$), Table 4 (for $\beta = 0.1$ and $\alpha = 0.9$), Table 5 (for $\beta = 0.9$ and $\alpha = 0.5$) and 6 (for $\beta = 0.9$ and $\alpha = 0.9$) allows comparing the solution of problem (7a)-(7c) with the market portfolio. These are calculated over the pointwise differences between both solutions, which are taken with respect to 16×4 points (corresponding to 16 years observations at 2 levels of L and T).

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
$\sum_{k,t} p_t U_{kt}$	0.008	0.021	0.551	0.956	1.586	3.872
# Stocks	-142.513	-108.782	-93.323	-75.538	-33.642	-8.067
Θ	-0.002	0.001	0.056	0.108	0.171	0.444

Table 3: Summary statistics for the case $\beta = 0.1$ and $\alpha = 0.5$.

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
$\sum_{k,t} p_t U_{kt}$	0.011	0.047	0.765	1.197	2.172	4.332
# Stocks	-286.567	-168.492	-147.767	-153.873	-118.57	-39.367
Θ	-0.003	0	0.08	0.131	0.24	0.478

Table 4: Summary statistics for the case $\beta = 0.1$ and $\alpha = 0.9$.

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
$\sum_{k,t} p_t U_{kt}$	0.023	0.051	0.426	0.374	0.615	0.929
# Stocks	-72.64	-72.258	-71.517	-68.349	-65.075	-58.4
Θ	0.07	0.17	0.524	0.71	1.165	1.991

Table 5: Summary statistics for the case $\beta = 0.9$ and $\alpha = 0.5$.

	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum
$\sum_{k,t} p_t U_{kt}$	0.023	0.073	0.384	0.341	0.538	0.847
# Stocks	-73.18	-72.15	-71.097	-69.079	-66.808	-58.473
Θ	0.07	0.198	0.898	0.828	1.108	2.449

Table 6: Summary statistics for the case $\beta = 0.9$ and $\alpha = 0.9$.

Tables 3-6 suggest that under all configuration of β and α , the endogenous multi-market partitioning (obtained by solving problem (7a)-(7c)) gives rise to a higher expected portfolio return and a higher tolerance to worst-case expected loss from local markets. The latter is a direct consequence of the compensation that a higher $\sum_{k,t} p_t U_{kt}$ might have on the willingness of the MIC to tolerate higher values of Θ . Besides, this figures are obtained with a smaller diversification, i.e., by selecting a smaller number of stocks.

7.3. On the empirical assessment of the γ -bound transformation

In this subsection we cast a closer look into the usage of the γ -bound transformation to approximate the MIC payoff. This computational tests rely upon a collection of $16 \times 3^3 = 432$ instances for the multi-scenario reformulation (12a)-(12f), which are replicated over three levels of γ (thus, becoming 1296 instances) for the γ -bound transformation (15a)-(15f). The three factors of the computational experiment are the risk aversion parameter β (fixed at three levels, $\beta \in \{0.01, 0.5, 0.99\}$), the quantile parameter α (fixed at three levels, $\alpha \in \{0.85, 0.90, 0.95\}$) and the number of piecewise constant steps L for the linearization of δ (fixed at three levels, $L \in \{1, 2, 10\}$), $\tau_k = 1$, for all $k \in \mathcal{M}$. Therefore, the combination of β , α and L gives rise to 27 instances for each of the 16 years of historical data.

The plots in Figure 1 illustrate the values of the MIC payoff for formulation (7a)-(7c) and the γ -bound transformation (15a)-(15f), based on 27 instances generated by the cross-combinations of β , α and L .

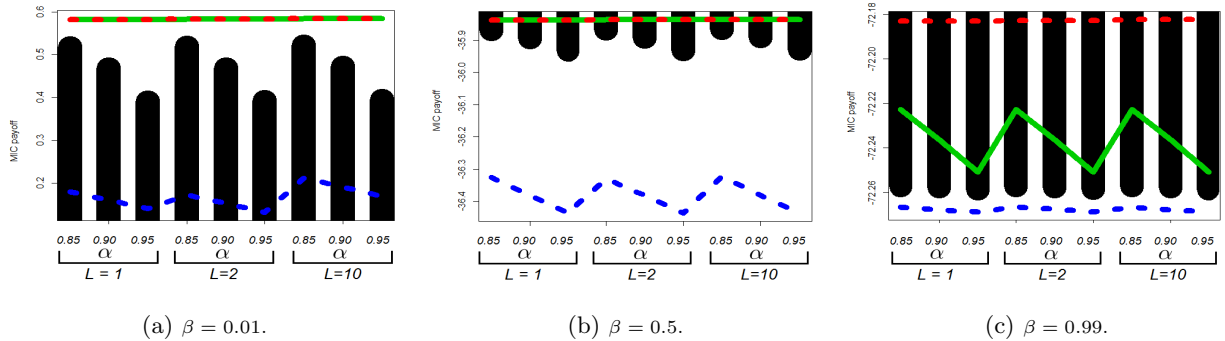


Figure 1: Comparison between the multi-scenario formulation and the γ -bound transformation. The black bars depict the MIC payoff under the optimal solution of problem (7a)-(7c), whereas the dotted red, solid green and dotted blue lines depict the MIC payoff under the optimal solution of (15a)-(15f), for $\gamma = 0.01$, $\gamma = 0.5$ and $\gamma = 0.99$ respectively. From left to right, the three plots correspond to $\beta = 0.01$, $\beta = 0.5$ and $\beta = 0.99$. Each of these three plots contains nine instances built out of the cross combinations of α and L and averaged over 16 year observations, as summarized in Table 1.

When $\beta \geq 0.5$ (reflecting MICs with a minimum level of tolerance to worst-case expected losses) the computed γ -bounds have a gap smaller than 0.01 from the optimal MIC payoff, regardless of the values of α and L . An important insight from Figure 1 is that the correct value of γ (the one that boost the goodness of the approximation) is positively related to β . As a rule of thumb, for $\beta \in (0, 0.5]$, one should set $\gamma \in (0, 0.5]$. From constraint (15d), this mirrors the need for a minimum expected portfolio return from each local market $k \in \mathcal{M}$ (that depends linearly on z_k). By contrast, for $\beta \in (0.5, 1]$, one should set $\gamma \in (0.5, 1]$.

Table 7 report the total number of selected markets based on the multi-scenario reformulation and the γ -bound transformation. These quantities are averaged over the 16 year horizon of the analyzed data set.

L	α	$\beta = 0.01$			Exact	$\beta = 0.5$			Exact	$\beta = 0.99$					
		Exact	γ			Exact	γ			Exact	γ				
			.01	.5			.99	.01			.5	.99	.01	.5	.99
1	.85	1.00	1.00	1.00	3.00	2.00	2.00	2.00	4.00	4.50	73.00	73.00	5.00		
	.90	1.00	1.00	1.00	3.00	2.00	2.00	2.00	4.00	3.06	73.00	73.00	5.00		
	.95	1.00	1.00	1.00	3.25	2.00	2.00	2.00	4.25	3.00	73.00	73.00	5.25		
2	.85	1.00	1.00	1.00	3.00	2.12	2.00	2.00	4.00	3.37	73.00	73.00	5.00		
	.90	1.00	1.00	1.00	3.25	2.00	2.00	2.00	4.25	3.25	73.00	73.00	5.25		
	.95	1.00	1.00	1.00	3.43	2.06	2.00	2.00	4.44	3.37	73.00	73.00	5.44		
10	.85	1.00	1.00	1.00	3.00	2.44	2.00	2.00	4.00	3.81	73.00	73.00	5.00		
	.90	1.00	1.00	1.00	3.00	2.06	2.00	2.00	4.00	3.12	73.00	73.00	5.00		
	.95	1.00	1.00	1.00	3.19	2.00	2.00	2.00	4.18	3.12	73.00	73.00	5.19		

Table 7: Difference with respect to the total number of selected markets.

Figures suggest that setting $\gamma < 0.5$ is a good strategy when using the γ -bound approximation for highly risk averse MICs (i.e. when $\beta < 0.5$). By contrast, when the risk aversion is low (i.e. when $\beta = 0.99$), the best γ -bound approximation is obtained with $\gamma = 0.99$. Therefore, to maximize the goodness of the approximation, Table 7 suggest that γ should be increased with β .

7.4. Computational performance of solution methods

This subsection digs into the efficiency of solving large-scale instances of problem (12a)-(12f), using the decomposition procedure presented in Subection 5.1 and the linearization and strong valid inequality presented in Subsection 5.2.

The implementation of the decomposition strategy consists in appending violated CVaR constraints at the branch-and-bound nodes, using the ILOG CPLEX callback functions. Two rules for constraint generation strategies are considered:

- only the $\kappa \leq m\tau$ most violated constraints are appended at each branch-and-bound node;
- an iteration dependent tolerance ε is set to assess the constraint violation.

As for the analysis in Subsection 7.2, these computational tests rely upon a factorial experiment with four factors involving a collection of $16 \times 2^4 = 256$ instances for the multi-scenario reformulation (12a)-(12f), that are solved with and without the strong valid inequalities (14). The four factors of the experiment coincides with the ones discussed in Subsection 7.2: the risk aversion parameter β (fixed at two levels, $\beta \in \{0.1, 0.9\}$), the quantile parameter α (fixed at two levels, $\alpha \in \{0.5, 0.9\}$), the number of piecewise constant steps L for the linearization of δ (fixed at two levels, $L \in \{2, 10\}$), the number of types τ_k (fixed at two levels, $\tau_k \in \{2, 10\}$).

Table 8 reports the performance and objective value of (7a)-(7c) (as well as its LP relaxation) to solve problem (12a)-(12f) with the CPLEX Branch-and-cut method and with the proposed decomposition method. The same collection of instances are considered for the results in Table 9, with the difference that the strong valid inequalities (14) are included.

T	L	β	α	Branch-and-cut		Decomposition through CVaR constraints separation				LP relax	
				CPU	OF	CPU (best)	CPU (average)	CPU (worst)	OF	OF	
2	2	0.1	0.5	4780.1 (116.8)	-0.007	715.7 (1.8, 502.0)	1121.6 (5.8, 867.2)	1901.305 (16.2, 1403.6)	-0.006	1.651	
10	2	0.1	0.5	7200.0 (1.0)	-0.025	3362.2 (19.0, 2007.4)	4877.7 (10.2, 2430.3)	6945.017 (1.8, 1951.8)	-0.303	1.654	
2	10	0.1	0.5	7200.0 (5949.6)	-0.030	7200.0 (121885.4, 3443.6)	7200.0 (188438.5, 2692.8)	7200.000 (191027.2, 3362.0)	-0.010	4.396	
10	10	0.1	0.5	7146.8 (1.0)	-0.209	6912.2 (3314.2, 4519.2)	7166.0 (3644.6, 3427.8)	7200.000 (2024.6, 2058.4)	-0.077	4.398	
2	2	0.9	0.5	762.1 (22.2)	0.270	456.8 (12.4, 58.2)	514.7 (10.4, 76.8)	577.759 (7.2, 100.8)	0.271	16.815	
10	2	0.9	0.5	6870.0 (5.8)	0.278	795.5 (13.8, 284.6)	1178.9 (16.4, 455.8)	1641.297 (17.2, 695.8)	0.279	16.817	
2	10	0.9	0.5	7200.0 (24348.4)	0.338	4651.8 (261598.2, 266.2)	6570.4 (368482.4, 359.0)	7200.000 (404583.6, 79.4)	0.342	40.497	
10	10	0.9	0.5	7200.0 (1.0)	0.206	7200.0 (8902.0, 1282.0)	7200.0 (10897.4, 969.8)	7200.000 (11664.4, 1294.0)	0.312	40.498	
2	2	0.1	0.9	4268.1 (86.4)	-0.009	722.6 (3.8, 414.6)	1277.6 (49.0, 1136.0)	1945.889 (134.4, 1424.2)	-0.006	1.640	
10	2	0.1	0.9	7200.0 (1.0)	-0.356	3371.9 (5.6, 1942.0)	5110.8 (3.2, 2284.8)	7200.000 (1.0, 2130.6)	-0.354	1.646	
2	10	0.1	0.9	7200.0 (5730.0)	-0.040	5178.4 (50514.6, 2745.2)	6704.1 (111222.5, 5532.2)	7200.000 (100761.0, 6106.4)	-0.017	4.392	
10	10	0.1	0.9	7200.0 (1.0)	-0.048	7200.0 (3972.0, 7886.2)	7200.0 (1811.6, 5084.6)	7200.000 (285.8, 4811.8)	-0.102	4.395	
2	2	0.9	0.9	803.3 (25.0)	0.208	622.5 (12.2, 97.6)	702.2 (10.9, 107.0)	794.514 (8.4, 124.0)	0.210	16.804	
10	2	0.9	0.9	6614.5 (7.0)	0.249	1189.4 (14.0, 359.6)	1596.9 (16.4, 591.2)	1960.110 (13.2, 569.2)	0.266	16.811	
2	10	0.9	0.9	7200.0 (22403.8)	0.242	7200.0 (350903.0, 2036.6)	7200.0 (382861.8, 943.4)	7200.000 (408031.8, 192.6)	0.271	40.492	
10	10	0.9	0.9	7200.0 (2.2)	0.178	7200.0 (9434.6, 1856.8)	7200.0 (7708.6, 1821.6)	7200.000 (4117.6, 2734.8)	0.280	40.496	

Table 8: Solution without the inclusion of the strong valid inequalities (14).

T	L	β	α	Branch-and-cut		Decomposition through CVaR constraints separation				LP relax	
				CPU	OF	CPU (best)	CPU (average)	CPU (worst)	OF	OF	
2	2	0.1	0.5	4188.363 (144.8)	-0.007	662.6 (1.8, 409.0)	1061.7 (48.1, 785.6)	1683.266 (138.4, 1107.4)	-0.006	-0.004	
10	2	0.1	0.5	7200.000 (1.0)	-0.750	4503.9 (172.0, 3105.2)	5684.5 (153.8, 2662.7)	7095.087 (143.8, 2147.0)	-0.750	-0.750	
2	10	0.1	0.5	7200.000 (1.4)	-0.087	988.0 (43.4, 586.8)	1637.8 (140.6, 863.5)	2569.464 (230.6, 1130.0)	-0.006	-0.004	
10	10	0.1	0.5	7200.000 (1.0)	-0.574	6645.5 (52.4, 1909.0)	6916.4 (52.9, 1901.0)	7200.000 (43.4, 1778.0)	-0.579	-0.574	
2	2	0.9	0.5	267.827 (16.8)	0.270	108.6 (9.6, 67.8)	122.5 (9.7, 66.0)	139.007 (10.4, 82.8)	0.271	0.281	
10	2	0.9	0.5	2990.373 (19.0)	0.278	532.7 (9.2, 270.8)	741.7 (9.4, 375.5)	987.011 (9.0, 543.8)	0.279	0.289	
2	10	0.9	0.5	804.953 (30.8)	0.356	215.6 (8.2, 72.0)	238.1 (9.8, 73.3)	263.745 (11.2, 100.4)	0.356	0.377	
10	10	0.9	0.5	5738.885 (1.0)	0.311	2566.2 (8.8, 342.0)	2952.6 (25.0, 416.4)	3345.026 (34.4, 300.2)	0.332	0.332	
2	2	0.1	0.9	3287.817 (111.0)	-0.009	658.3 (2.2, 416.0)	1445.2 (117.0, 1147.6)	2427.835 (228.6, 1755.4)	-0.006	-0.006	
10	2	0.1	0.9	7200.000 (1.0)	-0.394	3692.2 (100.2, 2313.6)	5766.8 (131.1, 2546.5)	7200.000 (163.4, 1752.8)	-0.366	-0.366	
2	10	0.1	0.9	6467.369 (171.8)	-0.009	861.3 (26.2, 462.2)	2004.9 (191.8, 1269.3)	3398.286 (319.8, 1873.8)	-0.006	-0.004	
10	10	0.1	0.9	7200.000 (1.0)	-0.468	6875.2 (100.4, 3797.0)	7142.9 (69.3, 2706.4)	7200.000 (64.4, 2887.6)	-0.653	-0.468	
2	2	0.9	0.9	196.380 (20.4)	0.208	142.3 (7.4, 52.6)	190.4 (8.0, 87.2)	242.101 (10.2, 108.0)	0.210	0.221	
10	2	0.9	0.9	1655.705 (12.8)	0.249	994.0 (10.8, 261.0)	1363.8 (10.6, 621.0)	2077.676 (19.6, 719.8)	0.259	0.259	
2	10	0.9	0.9	493.106 (48.4)	0.269	206.1 (3.8, 57.0)	251.4 (8.4, 85.2)	296.493 (10.2, 100.8)	0.271	0.300	
10	10	0.9	0.9	4997.600 (35.0)	0.319	2937.6 (3.6, 295.4)	4279.4 (97.2, 1099.1)	5245.974 (80.4, 694.6)	0.333	0.348	

Table 9: Solution with the inclusion of the strong valid inequalities (14).

In both tables each row is averaged over 16 instances, corresponding to the period 1999–2014. The first four columns reports the information about the specification of T , L , β , and α . The subsequent columns are partitions in three groups: the ones associated to the the branch-and-cut algorithm, the ones associated to the decomposition procedure, the ones associated to the LP relaxation. For all of them the CPU time (in sec) and the value of the objective function are reported. Into parenthesis the average number of branch-and-bound nodes and the appended constraints are noted. For the decomposition procedure, we report the best, the average and the worst performance case among the four combinations of κ and ϵ . A time limit of

7200 seconds (namely, two hours) is set.

The main figure in tables 8 and 9 can be summarized by noticing that both the use of the decomposition procedure and the inclusion of the strong valid inequalities (14) have a strong impact in all the examined instances. Specifically, the decomposition approach reduces the CPU time by approximately a factor of four (in the best case) and by a half in the average case. Note that even in the worse case the decomposition procedure still outperforms the global resolution by the ILOG CPLEX branch-and-cut. These figures are strengthened by the inclusion of the strong valid inequalities (14), through the improvement of the LP relaxation at the root node. This constitutes a numerical assessment of Proposition 5.

To cast a closer look into the leading effects behind these discrepancies, Table 10 reports elasticities of the tested parameters (assuming linear relationships with order-two interaction factors) with respect to the CPU times and the GAPs between the optimal solution and the LP relaxation.

Factor	Regressions					
	with valid inequalities			without valid inequalities		
	branch-and-cut	decomposition	LP gap	branch-and-cut	decomposition	LP gap
T	215.91	448.89	0.36	530.00	448.89	-0.01
L	222.47	695.42	-0.67	495.21	695.42	-0.02
β	-5286.60	-2595.45	4.57	-4982.40	-2595.45	-0.07
α	-1587.75	118.00	5.08	-809.57	118.00	-0.24
$T : L$	5.66	-31.23	0.01	-67.66	-31.23	0.00
$T : \beta$	137.47	-229.63	-0.26	259.24	-229.63	0.02
$T : \alpha$	74.38	20.40	-0.28	22.33	20.40	-0.02
$L : \beta$	-93.30	323.45	0.33	331.16	323.45	-0.01
$L : \alpha$	62.79	-60.39	0.42	62.03	-60.39	0.03
$\beta : \alpha$	-858.11	1004.09	-6.49	196.94	1004.09	0.21

Table 10: Elasticities of the tested parameters with respect to the CPU times and the GAPs between the optimal solution and the LP relaxation. The coefficients are estimated using ordinary least square on standardized data.

The risk aversion parameter β and the CVaR confidence level α seem to be the leading factors in determining the CPU time and the LP gap, so that the difficulty in solving problem (12a)-(12f) can vary substantially depending on the specific risk configuration. However, this dependency drop substantially when the decomposition method is used.

Finally, Table 11 report the percentage of solved instances using the CPLEX branch-and-bound algorithm directly to problem (12a)-(12f) (first line) or the constraint separation approach appending on constraint per node (second line) or five constraints per node (third line).

	without valid inequalities	with valid inequalities
Solve (12a)-(12f)	28.75	61.25
Decomposition with 1 cut per node	48.12	83.75
Decomposition with 5 cuts per node	46.88	81.88

Table 11: Percentage of solved instances.

The overall picture reveal that the joint combination of the strong valid inequalities and the constraint separation can almost triple the percentage of solved instances, passing from 28.75% to 81.88%.

8. Conclusions

In this work we presented an optimization framework (i.e., modelling design, optimality properties and resolution approach) for multi-market investment selection. The model endogenizes the transaction costs that the financial intermediation entails, as well as the centralized directives of the headquarter about budget allocation and risk regulation.

In its essence, this problem constitutes a single-leader-multi-follower game. Starting from a bilevel formulation, our results build on the equivalence with the high point relaxation and allow deriving a number of theoretical properties having important consequence on the computational resolution of the problem. The main theoretical derivations of this paper can be summarized as follows.

- We show that the multi-market portfolio optimization problem with CVaR constraints on local markets is NP-Hard and propose a solution strategy that combines a decomposition strategy (for the CVaR constraints) and the generation of strong valid inequalities (for the management fees), capable of reducing the number of decision variables and constraints by one order of magnitude, by sequentially appending feasibility cuts on the branch-and-bound tree. Next, Section 5.2 provide a polyhedral representation to approach the resulting non-linearity by a set of inequalities. On the computational side, we showed in Subsection 7.4 that these inequalities are capable of substantially improving the bound of the LP relaxation and boosting the efficiency of the computational resolution.
- Based on the high point relaxation, a closed-form characterization of the lower and upper bounds on the optimal MIC payoff is analytically derived in Section 6, using newly uncovered properties of the CVaR constraints. Subsection 7.3 provides an empirical support for the goodness of the γ -bound approximation. Due to the efficiency of its computation, the use of these bounds might represent a tool to circumvent the solution of problem (12a)-(12f).

As a way to further contribute to the work presented in this paper, the following lines of research can be explored:

- The extension of the proposed modeling approach to multi-period settings (Gülpinar & Rustem 2007), to other risk constraints (beyond the CVaR (Fábián 2008)) and to other forms of intermediation fees and transaction costs.
- The algorithmic improvements to target large-scale portfolio instances (Perold 1984), based on specialized decomposition approaches (Rahmaniani et al. 2017, Castro et al. 2017), for the case of piecewise constant management fees formulation, as well as specialized interior-point methods (Castro & Nasini 2017, 2021), for the case of continuous management fees.

Lastly, the research presented in this paper contributes to the enlargement of the numerical optimization horizon to the topics of multi-market portfolio management. It offers a well-functioning approach to address the solution of a class of complex NP-Hard problems, opening new possibilities for solving realistic and large-scale investment problems.

References

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

- Ambos, T. C., Andersson, U., & Birkinshaw, J. (2010). What are the consequences of initiative-taking in multinational subsidiaries? *Journal of International Business Studies*, *41*, 1099–1118.
- Andersson, F., Mausser, H., Rosen, D., & Uryasev, S. (2001). Credit risk optimization with conditional value-at-risk criterion. *Mathematical Programming*, *89*, 273–291.
- Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical finance*, *9*, 203–228.
- Balakrishnan, A., & Graves, S. C. (1989). A composite algorithm for a concave-cost network flow problem. *Networks*, *19*, 175–202.
- Ban, G.-Y., El Karoui, N., & Lim, A. E. (2016). Machine learning and portfolio optimization. *Management Science*, *64*, 1136–1154.
- Barry, F., & Kearney, C. (2006). MNEs and industrial structure in host countries: a portfolio analysis of Irish manufacturing. *Journal of International Business Studies*, *37*, 392–406.
- Baucells, M., & Borgonovo, E. (2013). Invariant probabilistic sensitivity analysis. *Management Science*, *59*, 2536–2549.
- Benita, F., Lopez-Ramos, F., & Nasini, S. (2018). A bi-level programming approach for global investment strategies with financial intermediation. *European Journal of Operational Research*, *274*, 375 – 390.
- Berge, C. (1997). *Topological Spaces: including a treatment of multi-valued functions, vector spaces, and convexity*. Courier Corporation.
- Best, M. J., & Hlouskova, J. (2003). Portfolio selection and transactions costs. *Computational Optimization and Applications*, *24*, 95–116.
- Castro, J., & Nasini, S. (2017). On geometrical properties of preconditioners in ipms for classes of block-angular problems. *SIAM Journal on Optimization*, *27*, 1666–1693.
- Castro, J., & Nasini, S. (2021). A specialized interior-point algorithm for huge minimum convex cost flows in bipartite networks. *European Journal of Operational Research*, *290*, 857 – 869.
- Castro, J., Nasini, S., & Saldanha-da Gama, F. (2017). A cutting-plane approach for large-scale capacitated multi-period facility location using a specialized interior-point method. *Mathematical Programming*, *163*, 411–444.
- Charnes, A., Cooper, W. W., & Symonds, G. H. (1958). Cost horizons and certainty equivalents: an approach to stochastic programming of heating oil. *Management science*, *4*, 235–263.
- Chen, W., Sim, M., Sun, J., & Teo, C.-P. (2010). From cvar to uncertainty set: Implications in joint chance-constrained optimization. *Operations research*, *58*, 470–485.
- Croxtan, K. L., Gendron, B., & Magnanti, T. L. (2003). A comparison of mixed-integer programming models for nonconvex piecewise linear cost minimization problems. *Management Science*, *49*, 1268–1273.
- Dal Sasso, V., De Giovanni, L., & Labbé, M. (2019). Strengthened formulations and valid inequalities for single delay management in public transportation. *Transportation Science*, *53*, 1271–1286.
- Dantzig, G. B. (1960). On the significance of solving linear programming problems with some integer variables. *Econometrica, Journal of the Econometric Society*, (pp. 30–44).
- Dantzig, G. B. (1998). *Linear programming and extensions*. Princeton university press.
- Dempe, S. (2002). *Foundations of bilevel programming*. Springer Science & Business Media.
- Doukas, J. A., & Lang, L. H. (2003). Foreign direct investment, diversification and firm performance. *Journal of International Business Studies*, *34*, 153–172.
- Du, Z., & Escanciano, J. C. (2016). Backtesting expected shortfall: accounting for tail risk. *Management Science*, *63*, 940–958.
- Fabián, C. I. (2008). Handling cvar objectives and constraints in two-stage stochastic models. *European Journal of Operational Research*, *191*, 888–911.

- Gülpınar, N., & Rustem, B. (2007). Worst-case robust decisions for multi-period mean–variance portfolio optimization. *European Journal of Operational Research*, *183*, 981–1000.
- Hausler, G. (2002). The globalization of finance. *Finance and Development*, *39*, 10–12.
- Hull, J. (2012). *Risk management and financial institutions* volume 733. John Wiley & Sons.
- Kawai, N., & Strange, R. (2014). Subsidiary autonomy and performance in Japanese multinationals in Europe. *International Business Review*, *23*, 504–515.
- Künzi-Bay, A., & Mayer, J. (2006). Computational aspects of minimizing conditional value-at-risk. *Computational Management Science*, *3*, 3–27.
- Leal, M., Ponce, D., & Puerto, J. (2020). Portfolio problems with two levels decision-makers: Optimal portfolio selection with pricing decisions on transaction costs. *European Journal of Operational Research*, *284*, 712–727.
- Li, J., Li, C., Wu, Z., & Huang, J. (2014). A feedback neural network for solving convex quadratic bi-level programming problems. *Neural Computing and Applications*, *25*, 603–611.
- Liou, Y.-C., & Yao, J.-C. (2005). Bilevel decision via variational inequalities. *Computers and Mathematics with Applications*, *49*, 1243–1253.
- Lobo, M. S., Fazel, M., & Boyd, S. (2007). Portfolio optimization with linear and fixed transaction costs. *Annals of Operations Research*, *152*, 341–365.
- Lux, T., & Rüschendorf, L. (2019). Value-at-risk bounds with two-sided dependence information. *Mathematical Finance*, *29*, 967–1000.
- Mansini, R., Ogryczak, W., & Speranza, M. G. (2007). Conditional value at risk and related linear programming models for portfolio optimization. *Annals of Operations Research*, *152*, 227–256.
- Maug, E., & Naik, N. (2011). Herding and delegated portfolio management: The impact of relative performance evaluation on asset allocation. *The Quarterly Journal of Finance*, *1*, 265–292.
- Nasini, S., & Erdemlioglu, D. (2019). Multiple channels of financial contagion: an empirical analysis of stock price dynamics. *Finance*, *40*, 87–133.
- Ou-Yang, H. (2003). Optimal contracts in a continuous-time delegated portfolio management problem. *The Review of Financial Studies*, *16*, 173–208.
- Pardalos, P. M., & Vavasis, S. A. (1991). Quadratic programming with one negative eigenvalue is np-hard. *Journal of Global Optimization*, *1*, 15–22.
- Perold, A. F. (1984). Large-scale portfolio optimization. *Management science*, *30*, 1143–1160.
- Pflug, G. C. (2000). Some remarks on the value-at-risk and the conditional value-at-risk. In *Probabilistic constrained optimization* (pp. 272–281). Springer.
- Potapchik, M., Tunçel, L., & Wolkowicz, H. (2008). Large scale portfolio optimization with piecewise linear transaction costs. *Optimization Methods & Software*, *23*, 929–952.
- Prekopa, A. (1973). Contributions to the theory of stochastic programming. *Mathematical Programming*, *4*, 202–221.
- Prékopa, A. (1974). Programming under probabilistic constraints with a random technology matrix. *Statistics: A Journal of Theoretical and Applied Statistics*, *5*, 109–116.
- Pshenichnyj, B. N. (2012). *The linearization method for constrained optimization* volume 22. Springer Science & Business Media.
- Rahmaniani, R., Crainic, T. G., Gendreau, M., & Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, *259*, 801–817.
- Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of banking & finance*, *26*, 1443–1471.
- Rockafellar, R. T., Uryasev, S. et al. (2000). Optimization of conditional value-at-risk. *Journal of risk*, *2*, 21–42.
- Sahni, S. (1974). Computationally related problems. *SIAM Journal on Computing*, *3*, 262–279.

- Stoughton, N. M. (1993). Moral hazard and the portfolio management problem. *The journal of finance*, *48*, 2009–2028.
- Stracca, L. (2006). Delegated portfolio management: A survey of the theoretical literature. *Journal of Economic surveys*, *20*, 823–848.
- Thi, L., Duc, T., & Dinh, P. (2012). A dc programming approach for a class of bilevel programming problems and its application in portfolio selection. *Numerical Algebra, Control and Optimization*, *2*, 167–185.
- Uryasev, S., Theiler, U. A., & Serraino, G. (2010). Risk-return optimization with different risk-aggregation strategies. *The journal of risk finance*, .
- Vajda, S. (1964). *Mathematical programming*. Addison-Welsey.
- Vielma, J. P., Ahmed, S., & Nemhauser, G. (2010). Mixed-integer models for nonseparable piecewise-linear optimization: Unifying framework and extensions. *Operations research*, *58*, 303–315.
- Zymler, S., Kuhn, D., & Rustem, B. (2013). Worst-case value at risk of nonlinear portfolios. *Management Science*, *59*, 172–188.

Appendix A: Proofs

Proof of Proposition 1.

Firstly, we note that for any $k \in \mathcal{M}$, $t \in \mathcal{T}_k$ and $\varepsilon > 0$ we have

$$\Psi_{kt}(z_k, 1 + n_k) = \Psi_{kt}(z_k, 1 + n_k + \varepsilon)$$

In fact, as returns cannot be smaller than -1 , the expected shortfall at the k^{th} market cannot be greater than $1 + n_k$, for any type $t \in \mathcal{T}_k$. In fact, when $\Theta > 1 + n_k$ the CVaR constraints are never active. Hence, there exists an implicit upper bound on Θ in the bilevel problem, i.e., $\Theta \leq 1 + \max\{n_k : k \in \mathcal{M}\}$.

Then, note that $|\Lambda_{kt}(\mathbf{z}, \Theta)| > 0$, for any $(\mathbf{z}, \Theta) \in \Xi \times [0, 1 + \max\{n_k : k \in \mathcal{M}\}]$, as $\mathbf{0} \in \Lambda_{kt}(0, 0)$. Therefore, for any $(\mathbf{z}, \Theta) \in \Xi \times [0, 1 + \max\{n_k : k \in \mathcal{M}\}]$, the feasible set $\Lambda_{kt}(z_k, \Theta) \subseteq [0, 1]^{n_k+1}$ is non-empty and compact. Consequently, the bilevel problem (4a)-(4c) possesses at least one feasible point, i.e. the inducible region is nonempty.

Next, we define the joint best responses of the SAs as the following Cartesian product

$$\tilde{\Psi}(\mathbf{z}, \Theta) = \prod_{k \in \mathcal{M}} \prod_{t \in \mathcal{T}_k} \Psi_{kt}(z_k, \Theta).$$

It is well-known from the the maximum theorem of Berge (1997) that $\tilde{\Psi}(\mathbf{z}, \Theta)$ is closed, compact valued and upper hemicontinuous. Consequently, the inducible region is closed and bounded. In conclusion, the inducible region is non-empty and the objective function of (4a)-(4c) is continuous, thus following Weierstrass theorem, (4a)-(4c) admits a global optimal solution. □

Proof of Proposition 2.

First, define the following quantities

$$c_I = \underset{\eta}{\operatorname{argmin}} \left\{ \eta + \frac{1}{(1-\alpha)} \mathbb{E}[(z - \eta)^+] \right\} \quad \text{and} \quad c_j = \underset{\eta}{\operatorname{argmin}} \left\{ \eta + \frac{1}{(1-\alpha)} \mathbb{E}[(R_j x_j - \eta)^+] \right\}$$

Based on (3), we have

$$\begin{aligned}
\langle\langle z - \mathbf{R}^\top \mathbf{x} \rangle\rangle_\alpha &\leq \left(c_I + \sum_j c_j \right) + \frac{1}{(1-\alpha)} \mathbb{E} \left[\left(z - \mathbf{R}^\top \mathbf{x} - \left(c_I + \sum_j c_j \right) \right)^+ \right] \\
&\leq c_I + \frac{1}{(1-\alpha)} \mathbb{E} \left[(z - c_I)^+ \right] + \sum_j \left(c_j + \frac{1}{(1-\alpha)} \mathbb{E} \left[(-R_j x_j - c_j)^+ \right] \right) \\
&\leq c_I + \frac{1}{(1-\alpha)} \mathbb{E} \left[(z - c_I)^+ \right] + \sum_j x_j \min_\eta \left\{ \eta/x_j + \frac{1}{(1-\alpha)} \mathbb{E} \left[(-R_j - \eta/x_j)^+ \right] \right\} \\
&\leq \langle\langle z \rangle\rangle_\alpha + \sum_j x_j \langle\langle -R_j \rangle\rangle_\alpha
\end{aligned}$$

where the second and third inequalities are based on the fact that $(a+b)^+ \leq (a)^+ + (b)^+$. As a result, by letting $\tilde{r}_j = \mathbb{E}[R_j | R_j \leq G_{R_j}^{-1}(1-\alpha)]$, we have

$$\langle\langle z - \mathbf{R}^\top \mathbf{x} \rangle\rangle_\alpha \leq \langle\langle z \rangle\rangle_\alpha - \sum_j x_j \langle\langle R_j \rangle\rangle_\alpha = z - \hat{\mathbf{r}}^\top \mathbf{x}$$

Secondly, we prove the lower bound. Building on the Jensen's inequality we have

$$\min_\eta \left\{ \eta + \frac{1}{(1-\alpha)} (z - \hat{\mathbf{r}}^\top \mathbf{x} - \eta)^+ \right\} \leq \min_\eta \left\{ \eta + \frac{1}{(1-\alpha)} \mathbb{E} \left[(z - \mathbf{R}^\top \mathbf{x} - \eta)^+ \right] \right\},$$

where $\hat{r}_j = \mathbb{E}[r_j]$. Note that

$$\eta + \frac{1}{(1-\alpha)} (z - \hat{\mathbf{r}}^\top \mathbf{x} - \eta)^+ = \begin{cases} -\eta \frac{\alpha}{(1-\alpha)} + \frac{1}{(1-\alpha)} (z - \hat{\mathbf{r}}^\top \mathbf{x}) & \text{if } \eta \leq z - \hat{\mathbf{r}}^\top \mathbf{x} \\ \eta & \text{otherwise} \end{cases}$$

which implies that the minimizer is $\eta = 1 - \hat{\mathbf{r}}^\top \mathbf{x}$, so that $\langle\langle z - \mathbf{r}^\top \mathbf{x} \rangle\rangle_\alpha \geq z - \hat{\mathbf{r}}^\top \mathbf{x}$

□

Proof of Proposition 3.

Problem (4a)-(4c) can be written as

$$\max \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} \Gamma_{kt}(z_k, \Theta) - (1-\beta)\Theta, \quad \text{subj. to } \mathbf{z} \in \Xi, \Theta \geq 0, \quad (18)$$

where for any $k = 1 \dots m$ the inner problem is defined as

$$\Gamma_{kt}(z_k, \Theta) = \begin{cases} \sup_{\mathbf{x}_{kt}} \mathbb{E} \left[(1 - \delta_k(z_k)) \hat{\mathbf{R}}_{kt}^\top \mathbf{x}_{kt} \right] \\ \text{s. to } \mathbf{x}_{kt} \in \begin{cases} \operatorname{argmax}_{\mathbf{x}_{kt}} \mathbb{E} \left[\delta_k(z_k) \mathbf{R}_{kt}^\top \mathbf{x}_{kt} \right] \\ \text{subj. to } \mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta) \end{cases} \end{cases} = \begin{cases} \sup_{\mathbf{x}_{kt}} \mathbb{E} \left[(1 - \delta_k(z_k)) \mathbf{R}_{kt}^\top \mathbf{x}_{kt} \right] \\ \text{subj. to } \mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta) \end{cases}$$

The equality comes from the fact that $(1 - \delta_k(z_k))$ and $\delta_k(z_k)$ are multiplicative constants for the leader and follower objectives respectively. This implies that problem (18) is equivalent to

$$\max_{\mathbf{z}, \Theta} \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} \left(\sup_{\mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta)} \mathbb{E} \left[(1 - \delta_k(z_k)) \mathbf{R}_{kt}^\top \mathbf{x}_{kt} \right] \right) - (1-\beta)\Theta, \quad \text{subj. to } \mathbf{z} \in \Xi, \Theta \geq 0,$$

which is equivalent to

$$\begin{cases} \max_{\mathbf{z}, \mathbf{x}, \Theta} & \beta \sum_{k=1}^m \sum_{t=1}^{\tau_k} p_{kt} (1 - \delta_k(z_k)) U_{kt}(\mathbf{x}_{kt}) - (1-\beta)\Theta \\ \text{subj. to} & \mathbf{x}_{kt} \in \Lambda_{kt}(z_k, \Theta) \\ & \mathbf{z} \in \Xi \\ & \Theta \geq 0 \end{cases} \quad k = 1 \dots m, t = 1 \dots \tau_k$$

□

Proof of Proposition 4.

We first note that when $n_k = \tau_k = 1$, we have

$$\langle \langle L(z, x, R) \rangle \rangle_\alpha = z - \mathbb{E}[xR \mid xR \leq G_{xR}^{-1}(1 - \alpha)] = z - x\tilde{r} \leq \Theta$$

We will show that for some parameter values, MICP(Θ) is equivalent to problem SUBSET SUM, say SSP. Given $a_1 \dots a_m > 0$ and $b > 0$, SSP consists in determining if there exists a subset $S \subset \{1 \dots m\}$ such that

$$\sum_{k \in S} a_k = b.$$

The recognition version of problem (7a)-(7c) (for the case $n_k = \tau_k = 1$) that can be written as

$$\left\{ \begin{array}{l} \sum_{k=1}^m \hat{r}_{1,k} x_k (1 - \delta_k(z_k)) \geq 1 \\ \sum_{k=1}^m z_k = 1 \\ x_k \leq \frac{z_k - \Theta}{\tilde{r}_k} \quad \text{for } k \in \mathcal{M} \\ x_{1,k} \leq z_k \quad \text{for } k \in \mathcal{M} \\ x_k, z_k \geq 0 \quad \text{for } k \in \mathcal{M}. \end{array} \right. \quad (19)$$

Let us consider the following case:

$$\delta_k(z) = 1 - \frac{1}{\hat{r}_k} \min \left\{ z \frac{b}{a_k}, 1 \right\}, \quad \text{where} \quad \frac{a_k}{b} = \frac{\Theta}{1 - \tilde{r}_{k,1}}.$$

On the one hand, if the answer of the SPP is yes, then (19) admits a solution. In fact, it is sufficient to note that the following is a feasible solution of (19):

$$x_k = z_k = \begin{cases} a_k/b & \text{for all } k \in S, \\ 0 & \text{for all } k \notin S. \end{cases}$$

In fact, as long as \tilde{r}_k is negative, we have

$$\sum_{k=1}^m \hat{r}_{1,k} x_k (1 - \delta_k(z_k)) = \sum_{k \in S} \min \left\{ z_k, \frac{a_k}{b} \right\} \min \left\{ z_k \frac{b}{a_k}, 1 \right\} = 1.$$

On the other hand, for each $j = 1, \dots, m$, let x'_j and z'_j be a solution to (19). Given that $1 - \delta(z'_j) \geq 0$, we may assume that x'_j is as large as possible, i.e.

$$x'_k = g_k(z'_k) = \min \left\{ \frac{z'_k - \Theta}{\tilde{r}_k}, z'_k \right\}.$$

Then the first constraint of (19) becomes:

$$\sum_{k=1}^m \min \left\{ \frac{z'_k - \Theta}{\tilde{r}_k}, z'_k \right\} \min \left\{ z'_k \frac{b}{a_k}, 1 \right\} \geq 1,$$

Note that $g_k(z)$ is piecewise linear with slop 1 for $z \in [0, a_k/b)$ and slop $1/\tilde{r}_k$ for $z \in [a_k/b, 1)$. Furthermore,

$$\max g_k(z) = \frac{a_k}{b} \quad \text{and} \quad \operatorname{argmax} g_k(z) = \frac{a_k}{b}$$

Therefore, we can consider two cases. If $z'_k \in [0, a_k/b)$, we have

$$\min \left\{ \frac{z'_k - \Theta}{\tilde{r}_k}, z'_k \right\} \min \left\{ z'_k \frac{b}{a_k}, 1 \right\} = \frac{b}{a_k} (z'_k)^2 \leq z'_k \leq \frac{a_k}{b}$$

If $z'_k \in [a_k/b, 1)$, we have

$$\min \left\{ \frac{z'_k - \Theta}{\tilde{r}_k}, z'_k \right\} \min \left\{ z'_k \frac{b}{a_k}, 1 \right\} = \frac{z'_k - \Theta}{\tilde{r}_k} \leq z'_k \leq \frac{a_k}{b}$$

Thus, the left-hand-side of the first constraint of (19), can be rewritten as

$$G(z'_1, \dots, z'_m) = \sum_{k: z'_k < a_k/b} \frac{b}{a_k} (z'_k)^2 + \sum_{k: z'_k \geq a_k/b} \frac{z'_k - \Theta}{\tilde{r}_k} \leq \sum_{k: z'_k < a_k/b} z'_k + \sum_{k: z'_k \geq a_k/b} \leq z'_k$$

Note that $\max G(z'_1, \dots, z'_m) = 1$ and the maximum can only be attained when $z'_j \in \{0, \frac{a_j}{b}\}$, for all $j = 1, \dots, m$. Then, the first constraint of (19) can only be satisfied when $z'_j \in \{0, \frac{a_j}{b}\}$, for all $j = 1, \dots, m$. Hence, z' provides a positive answer to MIC (Θ) only if $z'_j \in \{0, \frac{a_j}{b}\}$ for all $j = 1, \dots, m$ and $\sum_j z'_j = 1$. Consequently, $S = \{j : z'_j = \frac{a_j}{b}\}$ provides a certificate of positive answer to SSP.

In conclusion, if $\delta_k(z) = 1 - \min\{z\tilde{r}_{k,1}/(1 - \Theta), 1\}/\hat{r}_{1,k}$, then MICP(Θ) is NP complete. \square

Proof of Proposition 5.

Note that s_ℓ can be expressed in terms of s'_ℓ as:

$$s_\ell = \sum_{\ell'=\ell}^L s'_{\ell'}, \quad \text{so that} \quad s'_\ell = s_\ell - s_{\ell+1}, \quad \text{with } s_1 = 1.$$

Thus, using (lin-of-IM) we can write

$$\varphi \leq \hat{r}[(1 - d_\ell)x + s_{\ell+1}] = \hat{r} \left[(1 - d_\ell)x + \sum_{\ell'=\ell+1}^L s'_{\ell'} \right]$$

and using (lin-of-MCM) we can write

$$\varphi \leq \hat{r}[(1 - d_\ell)x + (1 - s'_\ell)] = \hat{r} \left[(1 - d_\ell)x + \sum_{\ell' \neq \ell} s'_{\ell'} \right]$$

Thus, (lin-of-IM) is stronger than (lin-of-MCM). \square

Proof of Proposition 6.

We show that (14) is valid by noticing that if $s'_{\ell'} = 0$ for all $\ell' \geq \ell + 1$ then,

$$\varphi \leq \hat{r}(1 - d_\ell)x = \hat{r} \left[(1 - d_\ell)x + \sum_{\ell'=\ell+1}^L (d_\ell - d_{\ell'})b_{\ell'}s'_{\ell'} \right]$$

Besides, if there exist $\ell' \geq \ell + 1$ such that $s'_{\ell'} = 1$, then

$$\varphi \leq \hat{r} \left[(1 - d_\ell)x + \sum_{\ell'=\ell+1}^L (d_\ell - d_{\ell'})b_{\ell'}s'_{\ell'} \right] \tag{20}$$

\square

Proof of Proposition 7.

Let Ξ be the feasible region of problem (10a)-(10h) and Ξ_γ the feasible region of problem (15a)-(15f) for a specified value of γ .

To prove (i) we first note that when $\gamma = 1$ the CVaR constraint is replaced with a stronger one. In fact, from Proposition 2 we have $\langle\langle z - \mathbf{R}^\top \mathbf{x}, F \rangle\rangle_\alpha \leq z - \hat{\mathbf{r}}^\top \mathbf{x}$, for any $\mathbf{x} \geq 1$. Thus $\Xi_1 \subseteq \Xi$, so that $\phi'''(1) \leq \phi$.

To prove (ii) we note that when $\gamma = 0$ the CVaR constraint is replaced with a weaker one. In fact, from Proposition 2 we have that $\langle\langle z - \mathbf{R}^\top \mathbf{x}, F \rangle\rangle_\alpha \geq z - \hat{\mathbf{r}}^\top \mathbf{x}$, for any $\mathbf{x} \geq 1$. Thus $\Xi \subseteq \Xi_0$, so that $\phi'''(0) \geq \phi$.

To prove (iii), we need to note that $\tilde{\mathbf{r}} \leq \hat{\mathbf{r}}$, for every $\tau \in [0, 1]$. This implies that for any non negative \mathbf{x} we have

$$[-\gamma \tilde{\mathbf{r}}(\tau)^\top - (1 - \gamma) \hat{\mathbf{r}}^\top] \mathbf{x} \leq [-\gamma' \tilde{\mathbf{r}}(\tau)^\top - (1 - \gamma') \hat{\mathbf{r}}^\top] \mathbf{x}$$

as long as $\gamma \leq \gamma'$. This implies that $\Xi_{\gamma'} \subseteq \Xi_\gamma$, so that $\phi'''(\gamma) \geq \phi'''(\gamma')$. □

Proof of Proposition 8.

Consider the follower problem associated to market k of type t in the γ -bound problem (15a)-(15f), which is obtained by fixing z_k and Θ .

For a given follower problem, the stationarity and complementary slackness from the first order Karush–Kuhn–Tucker conditions requires that

$$\left\{ \begin{array}{ll} \lambda_1 + g_{ikt}(\gamma)\lambda_2 - \mu_i = \hat{r}_{ikt} & i \in \mathcal{N}_k \\ \mu_i x_{ikt} = 0 & i \in \mathcal{N}_k \\ \lambda_1 \left(\sum_{i=1}^{n_k} x_{ikt} - z_k \right) = 0 & \\ \lambda_2 \left(\sum_{i=1}^{n_k} g_{ikt}(\gamma)x_{ikt} - \Theta + z_k \right) = 0 & \end{array} \right. \quad (21)$$

Since (16) is a feasible solution for (15a)-(15f), its optimality only depends on whether it verifies (21), for some $\lambda_1, \lambda_2, \mu_i \geq 0$. Since (16) contains a unique positive variable, from the complementary slackness we know that $\mu_{i(k,t)} = 0$ and $\mu_i \geq 0$, for all $i \in \mathcal{N}_k / \{i(k,t)\}$. Therefore, we have $\lambda_1 = \hat{r}_{i(k,t)kt} - g_{i(k,t)kt}(\gamma)\lambda_2$ and

$$\left\{ \begin{array}{ll} \hat{r}_{i(k,t)kt} - g_{i(k,t)kt}(\gamma)\lambda_2 + g_{i(k,t)kt}(\gamma)\lambda_2 \geq \hat{r}_{ikt} & i \in \mathcal{N}_k / \{i(k,t)\} \\ (\hat{r}_{i(k,t)kt} - g_{i(k,t)kt}(\gamma)\lambda_2) \left(\min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} - z_k \right) = 0 & \\ \lambda_2 \left(g_{ikt}(\gamma) \min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} - \Theta + 1 \right) = 0 & \end{array} \right.$$

We consider the following three cases to determine the value of λ_2 :

- i if $z_k < \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, then $\lambda_2 = 0$;
- ii if $z_k > \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, then $\lambda_2 = \frac{\hat{r}_{i(k,t)kt}}{g_{i(k,t)kt}(\gamma)}$;
- i if $z_k = \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, then λ_2 is free.

Therefore, substituting back, we obtain the following cases:

i when $z_k < \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, the feasible solution (16) is optimal iff

$$\hat{r}_{i(k,t)kt} \geq \hat{r}_{ikt} \quad \text{for each } i \in \mathcal{N}_k / \{i(k,t)\};$$

ii when $z_k > \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, the feasible solution (16) is optimal iff

$$\hat{r}_{ikt} \leq \frac{\hat{r}_{i(k,t)kt}}{g_{i(k,t)kt}(\gamma)} g_{ikt}(\gamma)$$

i when $z_k = \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)}$, the feasible solution (16) is optimal if (sufficient condition) either of the two previous cases are verified.

□

Proof of Corollary 1.

Once $z_1 \dots z_m$ are fixed, the first term of function

$$\begin{aligned} \xi(\Theta) &= \beta \sum_{k=1}^m (1 - \delta_k(z_k)) \hat{r}_{i(k,t)kt} \min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} - (1 - \beta)\Theta \\ &= \beta \sum_{k=1}^m w_{kt} \min \left\{ z_k, \frac{\Theta - z_k}{g_{i(k,t)kt}(\gamma)} \right\} - (1 - \beta)\Theta \end{aligned}$$

is a weighted sum of minimum of linear functions and the second term of $\xi(\Theta)$ is linear. Hence $\xi(\Theta)$ is concave and piecewise linear. As a consequence, its maximum is achieved at 0, 1 or at one of its breakpoints $z_k [g_{i(k,t)kt}(\gamma) + 1]$, for $k \in \{1, \dots, m\}$.

The indices of local markets are sorted in accordance to the increasing order of $z_k g_{i(k),k}(\gamma)$. Therefore, within the interval $\Theta \in [0, z_{j_1} (g_{i(j_1),j_1}(\gamma) + 1))$, function $\xi(\Theta)$ is linear with slope $\beta \sum_{k=1}^m \hat{r}_{i(j_k),j_k} (1 - \delta_{j_k}(z_{j_k})) / g_{i(j_k),j_k}(\gamma) - (1 - \beta)$. Within the interval $\Theta \in [z_{j_\ell} (g_{i(j_\ell),j_\ell}(\gamma) + 1), z_{j_{\ell+1}} (g_{i(j_{\ell+1}),j_{\ell+1}}(\gamma) + 1))$, the function $\xi(\Theta)$ is linear with slope $\beta \sum_{k=\ell+1}^m \hat{r}_{i(j_k),j_k} (1 - \delta_{j_k}(z_{j_k})) / g_{i(j_k),j_k}(\gamma) - (1 - \beta)$. Therefore, the unique optimal solution Θ^* of $\xi(\Theta, z_1, \dots, z_m)$ is equal to 1, if $\beta \sum_{k=1}^m \hat{r}_{i_k,k} (1 - \delta_k(z_k)) / g_{i(k),k}(\gamma) \leq (1 - \beta)$. Otherwise,

$$\Theta^* = z_t [g_{i(t),t}(\gamma) + 1],$$

where the sub-index t is the largest value $\ell \in \{1, \dots, m\}$ satisfying the inequality $\beta \sum_{k=\ell+1}^m \hat{r}_{i_{j_k},j_k} (1 - \delta_k(z_{j_k})) / g_{i(j_k),j_k}(\gamma) > (1 - \beta)$.

□