Robust Replicating Portfolio via Risk-Averse Stochastic Optimization

Fernanda Maria Müller[a*] Marcelo Brutti Righi[a]

fernanda.muller@ufrgs.br marcelo.righi@ufrgs.br

João Ramos Jungblut[a]

joaoramosjungblut@hotmail.com

[a]Federal University of Rio Grande do Sul, Washington Luiz, 855, Porto Alegre, Brazil, zip 90010-460

Abstract

We propose a conditional framework for futures-based index replication that links a risk-averse stochastic program (RASP) with linear regression models. The approach nests a rich family of risk–score pairs—Expected Loss (EL), Mean plus Semi-Deviation (MSD), Expected Shortfall (ES), Expectile, and Maximum Loss (ML)—with Squared Error (SE), Absolute Error (AE), Quantile, Expectile, and Linear-Exponential Loss (LINEX). Using S&P 100 constituents to replicate S&P 500 futures over 2010–2022, with rolling windows and semiannual rebalancing, we compare 25 specifications under Monte Carlo resampling. Parsimonious, symmetric combinations (EL/MSD with AE/SE) yield the lowest absolute tracking error and risk. In contrast, aggressive RASP (ML with LINEX/Quantile) results in higher returns but higher absolute tracking error and risk. Results are robust across market regimes.

[*]Corresponding author. The authors are grateful for the financial support of the Brazilian Research Council (CNPq) under grant number 307779/2022-0 and 302869/2024-7 and the Coordination for the Improvement of Higher Education Personnel under process number 88887.959969/2024-00.

Keywords: Robust measures; Replicating portfolio; Tracking error; Risk-averse.

1 Introduction

Replicating contingent claims in financial markets is one of the cornerstones of mathematical finance. In the canonical framework of complete markets, any contingent claim can be perfectly reproduced through an appropriate dynamic trading strategy, and the associated replication cost equals the expected discounted payoff under the risk-neutral measure (Huang & Guo 2013). However, real markets diverge from this idealized scenario. These markets are typically incomplete due to informational asymmetries, transaction costs, and commercial restrictions that hinder the exact replication of the asset (Balter & Pelsser 2020). Replication in incomplete markets is therefore often cast as an optimization problem that aims to approximate a target payoff or benchmark as closely as possible subject to realistic constraints—an approach consistent with the literature on index tracking[1] and tracking-error minimization (Beasley et al. 2003, Meihua et al. 2011, Sant'Anna et al. 2017).

More recent developments in the index-tracking literature have placed substantially greater emphasis on risk control. Sant'Anna et al. (2022) highlights that an 'ideal' tracking portfolio or replicating portfolio, as we will use in this study, should follow the benchmark, in our study, a derivative, in terms of returns, while also keeping risk close to or below that of the benchmark. Meihua et al. (2011) introduce into a general index-tracking model a risk constraint based on the Expected Shortfall (ES) to control the downside risk of tracking portfolios consisting of a subset of component stocks in the given index. Sant'Anna et al. (2022) expand this proposal and evaluate seven different risk measures as constraints for estimating portfolios that follow the market benchmark. Other proposals in this regard can be found in Gaivoronski et al. (2005) and Guastaroba et al. (2020).

In a novel method, Anis et al. (2023) minimize the squared difference between the percentage risk contributions of the reference sector and the candidate portfolio. Based on this proposal, the authors control the individual risk contribution of each constituent in

[1]Index tracking (IT) is a passive strategy in which portfolios are constructed to replicate a market index over time (Sant'Anna et al. 2022).

tracking portfolio. This method produces a closer approximation to the benchmark asset by reducing deviations in both volatility and returns, thereby yielding risk-adjusted returns that are closer to the benchmark. Despite advances, the lack of control over the tracking error's tail risk and the lack of evaluations using error metrics that highlight tail behavior and asymmetry remain open issues in the literature.

We address this gap by connecting the Robust Risk-Averse Stochastic Problem (RASP) (Righi et al. 2025) – which links coherent risk measures ρ with strictly consistent scoring functions S via the problem $\inf_{y\in\mathbb{R}} \rho\left(-S(X, y)\right)$ - to linear regression models. By this connection, we present a conditional framework for the optimal replicating portfolio problem. We provide results that guarantee the existence of a solution to the optimal replicating portfolio problem. The key distinction between our approach and prior literature is its ability to accommodate a variety of scoring functions and risk measures, thereby capturing different risk perceptions in the solution of the problem.

For empirical assessment, we conduct a simulation study comparing multiple RASP specifications to determine replicating portfolios for S&P 500 index futures. In the estimation, we use the most common coherent risk measures (Expected Loss (EL), Mean plus Semi-Deviation (MSD), Expected Shortfall (ES), Expectile, and Maximum Loss (ML)) and score functions (Squared Error (SE), Absolute Error (AE), Quantile, Expectile, and linear-exponential loss function (LINEX)) from the finance literature. For the estimation, we consider daily returns from stocks in the S&P 100 index and construct portfolios with different numbers of stocks, following a Monte Carlo simulation procedure inspired by Righi & Borenstein (2018). We evaluate different RASP's across two rolling window estimations (250 and 1000), three significance levels (1%, 2.5%, and 5%), portfolios with and without rebalancing, and transaction costs. We analyze the behavior of the optimal replicating portfolio with descriptive statistics, absolute error portfolio, and standard performance metrics.

To the best of our knowledge, this is the first study to present a solution to the portfolio replication problem that integrates robust measures with linear regression models. Relative to prior studies that use regression to estimate portfolio weights (Britten-Jones 1999, Li 2015, Shalit 2020), our contribution is a solution method that admits a unique solution while accommodating different scoring functions and coherent risk measures. The advantage of using a robust framework (RASP) to solve the replicating portfolio problem is that the risk measures are not sensitive to the choice of a single probability measure representing a particular belief about the world. This class of risk measures is aligned with the proposals of Righi et al. (2020) and Bellini et al. (2021), for instance.

Our second contribution is a comprehensive comparison of different combinations of ρ and S for obtaining the weights of the optimal replicating portfolio. We also provide a full characterization of optimal replication, which allows us to identify the most appropriate RASP specification for our problem. Although the literature compares various risk measures for portfolio optimization (Pflug et al. 2012, Righi & Borenstein 2018), we find no evidence of a broad comparison of conditionally robust risk measures for constructing replicating portfolios. Sant'Anna et al. (2022), for example, compare different risk measures within a standard index-tracking optimization model. Unlike our approach, they use a quadratic objective that penalizes the squared distance between the portfolio and the underlying asset, with risk measures imposed as constraints.

The remainder of this article is organized as follows. Section 2 provides a theoretical review of the core concepts and related works and the formalization regarding the optimal replicating portfolio problem. Section 3 presents the empirical methodology, including the dataset, the estimation process, and the performance evaluation metrics. Section 4 presents and discusses the empirical findings. Finally, Section 5 concludes the study with final remarks and suggestions for future research.

2 Proposed approach

Consider X a financial random variable, where $X \geq 0$ denotes a gain and $X < 0$ represents a loss. The variable X is defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and all equalities and inequalities should be interpreted as holding $\mathbb{P}$-almost surely. We define $X^- := \max(-X, 0)$. For $p \in [1, \infty)$, let $L^p := L^p(\Omega, \mathcal{F}, \mathbb{Q})$ denote the space of (equivalence classes of) random variables such that $\|X\|_p^p = E[|X|^p] < \infty$, and for $p = \infty$, $\|X\|_\infty = \operatorname{ess\,sup} |X| < \infty$, where E denotes the expectation. Unless explicitly stated otherwise, all definitions and results are valid for any fixed L^p with $p \in [1, \infty]$. Let $\mathcal{Q}$ denote the set of probability measures $\mathbb{Q}$ in

$(\Omega, \mathcal{F})$ that are absolutely continuous with respect to $\mathbb{P}$, with Radon–Nikodym derivative $\frac{d\mathbb{Q}}{d\mathbb{P}} \in L^q$, where $\frac{1}{p} + \frac{1}{q} = 1$, adopting the convention $q = \infty$ when $p = 1$ and $q = 1$ when $p = \infty$. The expected value, cumulative distribution function, and quantile function (left) of X under $\mathbb{Q}$ are given, respectively, by $E_{\mathbb{Q}}[X] = \int_{\Omega} X \, d\mathbb{Q}$, $F_{X,\mathbb{Q}}(x) = \mathbb{Q}(X \leq x)$, and $F^{-1}_{X,\mathbb{Q}}(\alpha) = \inf\{x \in \mathbb{R} : F_{X,\mathbb{Q}}(x) \geq \alpha\}$. When these quantities are calculated concerning $\mathbb{P}$, the subscript is omitted.

A risk measure is a functional $\rho : L^p \to \mathbb{R}$ that assigns to each financial position $X \in L^p$ a real value $\rho(X)$, representing the level of risk associated with X. A risk measure ρ is called *coherent* if it satisfies the following properties (Artzner et al. 1999):

(i) Monotonicity: If $X \leq Y$ almost surely, then $\rho(X) \geq \rho(Y), X, Y \in L^p$;

(ii) Translation invariance: For any constant $m \in \mathbb{R}$, $\rho(X + m) = \rho(X) - m, X \in L^p$;

(iii) Subadditivity $\rho(X + Y) \leq \rho(X) + \rho(Y), X, Y \in L^p$;

(iv) Positive homogeneity: $\rho(\lambda X) = \lambda \rho(X), X \in L^p$, $\forall \, \lambda \geq 0$.

According to Theorems 2.11 and 3.1 in Kaina & Rüschendorf (2009), a mapping $\rho : L^p \to \mathbb{R}$, with $p \in [1, \infty)$, constitutes a coherent risk measure if and only if it admits the dual representation

$$\rho(X) = \max_{\mathbb{Q} \in \mathcal{Q}_\rho} E_{\mathbb{Q}}[-X], \quad \forall \, X \in L^p,$$

where $\mathcal{Q}_\rho \subseteq \mathcal{Q}$ is a set that is not empty, closed, and convex.

To illustrate examples of coherent risk measures, in the sense of Artzner et al. (1999), we present Example 1.

Example 1. Consider that $X \in L^p$. Examples of coherent risk measures, in the sense of Artzner et al. (1999), include the following functionals:

- **Expected Loss (EL):** This is the simplest coherent risk measure, corresponding to the negative of the expectation (mean) of returns. We define $EL(X) = -E[X]$. For EL, the dual set is a singleton $\mathcal{Q}_{EL} = \{\mathbb{P}\}$.
- **Mean plus Semi-Deviation (MSD):** The Mean plus Semi-Deviation (MSD) is a coherent risk measure defined by $MSD^\beta(X) = -E[X] + \beta\sqrt{E[((X - E[X])^-)^2]}$, where $\beta \in [0, 1]$ represents the proportion of semi-deviation considered to penalize the EL. This measure belongs to the class of loss-deviation measures discussed by Righi (2019). Its dual

set is given by $\mathcal{Q}_{MSD} = \left\{ \mathbb{Q} \in \mathcal{Q} \colon \frac{d\mathbb{Q}}{d\mathbb{P}} = 1 + \beta(W + E[W]), W \leq 0, ||W||_q \leq 1 \right\}$.

- **Expectile:** The Expectile is a generalization of the quantile function used for VaR estimation. Expectile is defined as $Expectile^{\alpha}(X) = -\arg\min_{x \in \mathbb{R}} E[\alpha[(X - x)^+]^2 + (1 - \alpha)[(X - x)^-]^2]$, where $\alpha \in (0, 1)$. For $\alpha \leq 0.5$, the Expectile is a law-invariant coherent risk measure. Taking $\alpha = 0.5$ results in the recovery of traditional least squares, and the Expectile coincides with the EL. The dual set of the Expectile is given by $\mathcal{Q}_{Expectile^{\alpha}} = \left\{ \mathbb{Q} \in \mathcal{Q} \colon \exists\, a > 0,\ a \leq \frac{d\mathbb{Q}}{d\mathbb{P}} \leq a \frac{1-\alpha}{\alpha} \right\}$.
- **Expected Shortfall (ES):** Expected Shortfall (ES) stands as the canonical example of a coherent risk measure, forming the basis for many representation theorems in this field. It is defined as $ES^{\alpha}(X) = -\frac{1}{\alpha} \int_0^{\alpha} F_X^{-1}(s)\, ds$, where $\alpha \in (0, 1)$, and its dual set is given by $\mathcal{Q}_{ES^{\alpha}} = \left\{ \mathbb{Q} \in \mathcal{Q} : \frac{d\mathbb{Q}}{d\mathbb{P}} \leq \frac{1}{\alpha} \right\}$..
- **Maximum Loss (ML):** Maximum Loss (ML) is defined as $ML(X) = -\operatorname{ess\,inf} X$. ML stands out as the most extreme coherent risk measure, leading to more protective situations, such as $ML(Y) \geq \rho(Y)$ for any coherent risk measure ρ. The dual set of ML is given by $\mathcal{Q}_{ML} = \mathcal{Q}$.

A mapping $S : \mathbb{R}^2 \to \mathbb{R}_+$ is called a *scoring function* if, for any $X, Y \in L^p$, the mapping $\omega \mapsto S(X(\omega), Y(\omega))$ belongs to L^1, and if it satisfies the following properties for all $x, y \in \mathbb{R}$:

(i) $S(x, y) \geq 0$ and $S(x, y) = 0$ if and only if $x = y$.

(ii) There exists a function $f_S : \mathbb{R} \to \mathbb{R}$ such that $S(x, y) = f_S(x - y)$.

(iii) The function $y \mapsto S(x, y)$ is convex and continuous, non-decreasing for $y > x$, and non-increasing for $y < x$.

Example 2. Given a mapping $S : \mathbb{R}^2 \to \mathbb{R}_+$, are examples of the scoring functions, the following functionals:

- **Squared Error (SE)**: SE is expressed as $S_{SE}(x, y) = (x - y)^2$. SE is one of the most commonly used scoring functions and is directly associated with standard least squares regression tools. The EL is elicitable under this loss function.
- **Quantile**: This score is expressed as $S_{Quantile^{\alpha}}(x, y) = \alpha(x - y)^+ + (1 - \alpha)(x - y)^-$. The argument that minimizes $S_{Quantile^{\alpha}}$ is VaR. This score is directly associated with quantile regression.

- **Absolute Error (AE)**: This score is given by $S_{AE}(x, y) = |x - y|$. The Median minimizes this scoring function and represents a special case of Value-at-Risk with $\alpha = 0.5$, i.e., $VaR^{50\%}$.
- **Expectile**: Expectile scoring function is given by $S_{Expectile^\alpha}(x, y) = \alpha[(x - y)^+]^2 + (1 - \alpha)[(x - y)^-]^2$. The $Expectile^\alpha$ is the argument that minimizes $S_{Expectile^\alpha}$.
- **Linear exponential (LINEX)**: This score function can be written as $S_\gamma(x, y) = e^{\gamma(x-y)} - \gamma(x - y) - 1$. $\gamma > 0$ representing a risk aversion parameter. This function is the scoring function associated with the entropic risk measure.

The mapping $\mathcal{D} : L^p \to \mathbb{R}_+$ represents a generalized deviation measure in the sense of Rockafellar et al. (2006). Let $S : \mathbb{R}^2 \to \mathbb{R}$ be a scoring function, and let $\mathcal{Q}_\rho$ denote the dual set of a coherent risk measure ρ. Following Righi et al. (2025), the risk and deviation functionals $R_{\rho,S}, D_{\rho,S} : L^p \to \mathbb{R}$ are defined as follows:

$$R_{\rho,S}(X) := -\min\left\{\arg\min_{y\in\mathbb{R}} \rho(-S(X, y))\right\}, \quad D_{\rho,S}(X) := \min_{y\in\mathbb{R}} \rho(-S(X, y)),$$

where $\rho(X) := \sup_{\mathbb{Q}\in\mathcal{Q}_\rho} \mathbb{E}_\mathbb{Q}[-X]$. We denote the set of minimizers by $B_X := \arg\min_{y\in\mathbb{R}} \rho(-S(X, y))$. $R_{\rho,S}$ is interpreted as the monetary cost derived from a scoring function. The functional $D_{\rho,S}$ represents the minimal generalized shortfall deviation measure associated with a given scoring function and risk measure.

We now formalize the optimal replicating portfolio problem based on conditional risk measures, obtained by combining the proposal of Righi et al. (2025) with linear regression models. We also provide a simple result to guarantee that the problem is well-posed and has a closed-form for the weights of the optimal replicating portfolio.

Definition 1. *Let $Y \in L^p$ be given and $X = (X_1, \ldots, X_n) \in (L^p)^n$. The optimal replicating portfolio problem for X, $H(X)$, is defined as*

$$\min_{w\in\mathbb{R}^n} \rho\left(-S\left(Y, \mu + \sum_{i=1}^n w_i X_i\right)\right) \tag{1}$$

Proposition 1. *$H(X)$ has a solution for any $X \in (L^p)^n$.*

Proof. This problem has a guaranteed solution considering $f\colon \mathbb{R}^n \to \mathbb{R}$ as $f(w_1, \dots, w_n) = \rho\left(-S\left(Y, \sum_{i=1}^n w_i X_i\right)\right)$. ☐

Proposition 2. *$w^* = (w_1^*, \dots, w_n^*) \in \arg\min H(X)$ if and only if $R(Y|X) = -\left(\mu^* + \sum_{i=1}^n w_i^* X_i\right)$, where $\mu^* = -R(Y - \sum_{i=1}^n w_i^* X_i)$.*

Proof. We have by Definition that

$$\begin{aligned} R_{\rho,S}(Y|X_1, \dots, X_n) &= -\left(\mu^* + \sum_{i=1}^n w_i^* X_i\right) \\ \iff (\mu^*, w^*) &= \underset{\mu \in \mathbb{R}, w \in \mathbb{R}^n}{\arg\min}\, \rho\left(-S\left(Y, \mu + \sum_{i=1}^n w_i X_i\right)\right). \end{aligned}$$

Further, notice that

$$\underset{w \in \mathbb{R}^n}{\arg\min}\, \rho\left(-S\left(Y, \sum_{i=1}^n w_i X_i\right)\right) = \underset{w \in \mathbb{R}^n}{\arg\min}\, \rho\left(-S\left(Y, k + \sum_{i=1}^n w_i X_i\right)\right), \ \forall\, k \in \mathbb{R}.$$

We also have that

$$\min_{\mu \in \mathbb{R}, w \in \mathbb{R}^n} \rho\left(-S\left(Y, \sum_{i=1}^n w_i X_i\right)\right) = \min_{w \in \mathbb{R}^n} \rho\left(-f_S\left(Y - \sum_{i=1}^n w_i X_i + R\left(Y - \sum_{i=1}^n w_i X_i\right)\right)\right).$$

From these facts, we get the equivalence between both $w^* = (w_1^*, \dots, w_n^*) \in \arg\min H(X)$ and $R(Y|X) = -\left(\mu^* + \sum_{i=1}^n w_i^* X_i\right)$. ☐

3 Data and estimation procedures

Our empirical study examines the replication of the S&P 500 futures index[2] using a subset of S&P 100 stocks. The dataset includes daily adjusted closing prices, adjusted for dividends and stock splits, of 93 constituents[3] from the S&P 100 index. The stocks were obtained via the *BatchGetSymbols* `R` package (Perlin 2022). Our sample considers daily price observations

[2]We use the S&P 500 E-mini futures as our target index proxy. This data is downloaded from `https://br.investing.com/indices/us-spx-500-futures` (continuous contract, June 2023).

[3]Some stocks were excluded from the sample because they were delisted, merged, or ceased trading. The composition of the index considered refers to December 2022.

that span January 5, 2010, through December 30, 2022. This period covers multiple market cycles, including bullish phases and stress periods (e.g., the 2020 global market crash).

All price series are converted to logarithmic daily returns (log-returns) and scaled by 100, so that returns are expressed in percentage terms for ease of interpretation. The futures returns have a mean of 0.04% and a standard deviation of 1.1%, negative skewness, and high excess kurtosis, consistent with fat-tailed distributions. The S&P 100 constituent returns have an average standard deviation of approximately 1.7% and daily fluctuations that can reach gains of about +16% and losses of about -15%, reflecting idiosyncratic shocks absent at the aggregate index level. As expected, the futures index appears less risky than the average constituent.

The general formulation of our portfolio replication problem is given in Definition 1. In our empirical application, we consider stocks in the replicating portfolios with non-negative weights. Optimization is performed using the default `R` (version 4.3.2) optimization routines. To examine the impact of diversification, the replicating portfolio is further constrained to either 4 assets (concentrated) or 16 assets (diversified). We consider multiple risk measures—Expected Loss, Mean plus Semi-Deviation, Expectile, Expected Shortfall, and Maximum Loss (Example 1)—combined with alternative scoring functions such as SE, quantile, AE, Expectile loss, and LINEX (Example 2). This setup yields 25 distinct risk-and-score combinations (RASPs). For quantile and expectile scores, we test $\alpha \in 1\%, 2.5\%, 5\%$, values adopted in the risk forecasting literature (Kuester et al. 2006, Müller & Righi 2018) and aligned with Basel Committee recommendations (Basel Committee on Banking Supervision 2013).

We treat the returns X_t (for $t = 1, \ldots, T$) as realizations of a discrete probability space $\Omega = \omega_1, \ldots, \omega_T$ with uniform weights $\mathbb{P}(\omega_t) = 1/T$. Thus, the empirical distribution function and expectation are given by $F_X(x) = \frac{1}{T}\sum_{t=1}^{T} 1_{X_t \leq x}$, and $E[X] = \frac{1}{T}\sum_{t=1}^{T} X_i$, respectively. Based on this empirical distribution of the data, we adopt the historical simulation (HS) method for risk estimation. This method preserves the empirical characteristics of the returns and avoids restrictive distributional assumptions (Kuester et al. 2006, Righi et al. 2020).

In line with Sant'Anna et al. (2022), we consider a rolling-window estimation to estimate the replicating portfolio weights. We consider two lengths for the estimation window (in-sample period): a shorter window of 250 trading days (approximately 1 year) and a longer window of 1000 trading days (approximately 4 years). The short horizon allows adaptation to recent market dynamics, while the longer horizon improves statistical stability. In both cases, the out-of-sample test window, that is, rebalancing, is fixed at 120 trading days (around 6 months). Transaction costs were ignored in the baseline analysis, but robustness checks confirmed that including proportional transaction costs did not affect the results. The following steps can summarize the procedures: i) Initial training: Estimate the optimal replicating portfolio problem using the in-sample window and obtain the weights. With the weights, determine the returns of the replicating portfolio. ii) Out-of-sample test: Hold the replicating portfolio fixed for 120 trading days. For each out-of-the-sample window, we have 25 replicating portfolios. iii) Roll forward: Advance both windows by 120 days and re-estimate the optimal replicating portfolio. iv) Iterate: Continue until the end of the sample, generating a sequence of out-of-sample results.

This setup mimics an investor's semiannual rebalancing, balancing responsiveness and cost efficiency. Quarterly and annual rebalancing periods were also considered; however, because the results were similar, they were not presented. Furthermore, we considered a statistical portfolio. Thus, two strategies are compared: Rebalanced portfolios: weights updated every 120 trading days. Static portfolios: weights estimated once at the beginning and held fixed until 2022. For each strategy, the optimization selects a subset of assets and assigns weights.

To ensure robust evaluation, asset selection and portfolio construction are embedded in a Monte Carlo re-sampling procedure utilized in portfolio construction of the study of Righi & Borenstein (2018). For each scenario, we replicate the portfolio for each RASP and determine its performance. The process is repeated 1000 times, providing empirical distributions of replicating portfolio outcomes.

Given the returns of the replicating portfolio $X_{t,j,s}$ and the target asset y_t, we first evaluate descriptive statistics of the weights of the replicating portfolios. At this stage, we also quantify the concentration of the replicating portfolio ($Conc$), by $Conc_{j,s} = \sum_{i=1}^{n} (w_{i,j,s})^2$, where $w^*_{i,j,s}$ represents the weight of the asset i, given the risk measure j and the scoring function s. The $Conc$ metric varies only due to the Monte Carlo procedure for portfolios without replication (static portfolio). For portfolios with replication, *Conc* also varies in each rebalancing period (every 120 days) and in each Monte Carlo replication. The *Conc* metric is similar to the Herfindahl–Hirschman Index (HHI), discussed in Mainik et al. (2015), namely that higher values of *Conc* indicate greater concentration in the portfolio.

In line with previous studies such as Zhao et al. (2019) and Sant'Anna et al. (2022), we assess the different RASPs using: i) Average return (Av.ret): arithmetic mean of daily returns. ii) Cumulative return (Cum.ret): total of gains or losses of an investment in percentage over a period. iii) Standard deviation (SD): total volatility of daily returns. iv) Semi-deviation (SD^{-}): downside volatility of daily returns. v) Value-at-risk ($VaR_{5\%}$): tail risk of daily returns considering $\alpha = 5\%$. vi) Sharpe ratio (Sharpe): mean excess return per unit of volatility. vii) Treynor ratio (Treynor): mean excess return per unit of systemic risk (Beta). viii) Absolute error (Abs.E): average absolute difference between portfolio and target returns captures the tracking accuracy. All metrics are computed on a daily basis. To allow comparability with standard practice in the asset management literature, Av.ret, SD, SD^{-}, Sharpe, Treynor, and VaR are annualized.

To assess how closely the replicating portfolio matches the reference asset, we also apply these performance metrics to the S&P 500 futures index. We evaluated the results for the entire sample and for three sub-samples (2014-2016, 2017-2019, and 2020-2022) to consider different market moments and conditions. From 2014 to 2016, we experienced a period of recovery and economic stability after the American crisis. From 2017 to 2019, we experienced a period of increased volatility in the US market, particularly at the end of 2018, which coincided with the trade war between the United States and China. From 2020 to 2022, we experienced the economic crisis caused by the pandemic.

4 Results

In our framework, the $w_i, i = 1, \cdots, n$ estimates from regression models, are the weights of each stock i in the optimal replicating portfolio. Descriptive statistics of weights and concentration are available upon request. The average value of w_i estimates tends to be lower for more parsimonious risk measures, such as EL. On the other hand, w_i estimates average values are higher for more aggressive risk measures, such as ML, which reflects the maximum loss expected. This finding is reflected in the concentration of replicating portfolios. Portfolios constructed with EL and MSD have concentrations lower than those obtained with ML. For example, when considering portfolios with four assets and rebalanced every 120 days, with an in-sample size of 1000 observations and a α of 5%, we find that the average concentration of EL is approximately 0.4564, while for ML it is 1.063. This result occurs due to more restrictive measures, which generate a greater value of risk and consequently a greater determination of capital, concentrating their allocation in some assets with lower risk. It is also evident from our results that systematic rebalancing reduces concentration across almost all risk-score combinations. For example, in ML with the LINEX score, the average concentration drops from 1.7697 (without rebalancing) to 1.3204 (with rebalancing) (portfolios with four assets, an in-sample size of 1000 observations, and a α of 5%). This result suggests that rebalancing favors diversification and reduces exposure to idiosyncratic risks.

Results obtained with the loss functions SE and AE produce similar w_i values. Both functions exhibit a symmetric pattern of penalizing tracking errors. Regardless of the risk measure, portfolio concentration is moderate in these configurations. However, scoring functions that prioritize asymmetries and tail penalization, like Quantile and LINEX, invariably result in more concentrated portfolios. Greater allocation to a smaller subset of assets is the result of these scores' propensity to give larger weights to assets whose returns are more consistent with the behavior of the reference asset. It's interesting to note that the Expectile score behaves in an intermediate manner, with its corresponding concentration levels being higher than those of SE and AE but lower than those of Quantile.

In general, the results indicate that both the chosen risk measure and the scoring function significantly influence the allocation structure of the replicating portfolios. More conservative and asymmetric configurations tend to produce highly concentrated portfolios, whereas more parsimonious combinations lead to more balanced allocations. Periodic rebalancing reinforces this effect, smoothing portfolio weights over time and contributing to a more stable and robust replication strategy.

Table 2 reports the performance analysis of the S&P 500 index futures for the full sample and sub-samples. Table 1 describes the performance of replicating portfolios formed with 4 and 16 assets, considering an in-sample size of 1000 observations, $\alpha = 5\%$, full sample period. We present the results for portfolios without rebalancing and with rebalancing every 120 days. In line with Sant'Anna et al. (2022), the evaluation was carried out in two steps. First, we identified the risk measures ρ_S that generated the replicating portfolios with the best performance according to Av.ret, Cum.ret, SD, SD$^-$, VaR, Sharpe, and Treynor. In a second stage, the analysis emphasized portfolios whose results were closest to those of the underlying asset they replicated (lower Abs.E), namely the S&P 500 futures market index (Table 2), since the central objective of building a replicating portfolio is to mirror the behavior of the reference asset.

Most configurations exhibit substantially higher returns than the benchmark, whose average and cumulative returns were 0.4581 and 65.79, respectively. These findings corroborate the studies by Goel et al. (2018) and Sant'Anna et al. (2022), which obtained higher cumulative returns to track portfolios relative to the market benchmark. However, for the problem of replicating a reference asset, the desire is to have portfolios that do not differ from the asset they want to replicate, even if this difference is positive, i.e., the tracking portfolios outperform the index (Sant'Anna et al. 2022).

The highest Av.ret and Cum.ret are often obtained for ML$_{LINEX}$. In contrast, conservative configurations, including EL$_{AE}$, MSD$_{AE}$, and EVaR$_{AE}$, deliver lower performance returns. Sant'Anna et al. (2022) when considering a standard index tracking optimization model, which includes seven risk measures as constraints, and Ramos et al. (2023) when comparing risk measures for portfolio optimization, point to the EL with the best performance with respect to return metrics.

In addition to tracking the underlying asset, a replicating portfolio must, to be considered ideal, have a risk close to or lower than the asset it seeks to replicate. In general, EL$_{AE}$ resulted in portfolios that were less risky in terms of volatility (SD and semi-deviation) and tail risk (VaR). For portfolios with 16 assets, in some cases, the replicating portfolio MSD$_{AE}$ obtained lower SD and SD$^-$. Although the AE score, using the same risk measure, yields the lowest risk values among the score functions, the average risk estimates via SE are similar. The highest average risk values are observed for the LINEX function, being the highest values found for ML_{LINEX}. At the opposite extreme, ML_{LINEX} reported the highest levels of returns, reinforcing the classic trade-off between risk and return. Furthermore, ML_{LINEX} has the highest concentration among all the estimated replicating portfolios. Diversification helps investors reduce the risk of an investment (Shawky et al. 1997). However, a higher concentration of weights reduces the benefits generated by greater diversification. Concentrating most of the resources in one investment makes the portfolio more exposed to specific risks, which can contribute to greater portfolio risk.

We note a particularity with respect to the Sharpe ratio: portfolio size is the primary determinant of leadership. MSD_{SE} portfolios yield the highest Sharpe ratios for portfolios with four assets, while ML_{LINEX} portfolios yield the best results for portfolios with sixteen assets. The Sharpe ratio evaluates the excess return that an investor obtains per unit of total risk, measured by the portfolio's standard deviation. However, specific risks can be eliminated when building a portfolio, so some investors are only concerned with systemic risk. In this sense, the Treynor index becomes the most appropriate option to evaluate the efficiency of a replicating portfolio with respect to systemic risk. Regarding Treynor, we noticed a prevalence of portfolios obtained with the Quantile and Expectile score functions. While some configurations—especially in the non-rebalanced setting—achieve Treynor ratios close to or higher than the benchmark, several portfolios with rebalancing exhibit negative values. These negative Treynor ratios are primarily attributed to negative beta estimates, which occur when rebalancing alters portfolio composition and can induce net short market exposure. This inversion is a mechanical consequence of $\beta < 0$ and does not, by itself, indicate poor replication.

In terms of Abs.E, EL_{AE} and EL_{SE} present the lowest values. When these RASPs do not rank first, MSD_{AE} typically attains the lowest deviation from the reference asset (see Table 3). On the other hand, more aggressive scores - especially LINEX and, to a lesser extent, Quantile - exhibit the largest tracking errors, frequently $1.5/2\times$ higher than the best cases. Our result confirms an important trade-off: RASPs that generate higher returns tend to deviate further from the reference asset, whereas compositions that rely on parsimonious and symmetric loss functions (AE/SE) under EL/MSD yield tighter replication. In particular, replicating portfolios with the lowest absolute error coincide with those with lower volatility and tail risk, as well as with lower portfolio concentration, indicating that risk control and diversification are instrumental to faithful replication.

We also analyze the behavior of replicating portfolios in three sub–samples: 2014–2016, 2017–2019, and 2020–2022. Each of these periods captures distinct market regimes, ranging from stability and growth to increased volatility. When we examine the sub-samples, we observe that, in general, the patterns identified in the full sample are preserved. ML_{LINEX} delivers superior results in terms of Av.ret / Cum.ret (particularly for $\alpha \geq 2.5\%$), while EL with symmetric losses ($\text{EL}_{AE}/\text{EL}_{SE}$) yields lower risk (SD, SD^{-}, VaR) and a smaller absolute error relative to the reference asset. Regarding non-rebalanced replicating portfolios, we note that in more stable periods, such as 2017–2019 and $\alpha = 1\%$, $\text{EL}_{Quantile}$ achieves higher returns, while tracking performance remains superior under $\text{EL}_{AE}/\text{EL}_{SE}$ (and MSD_{AE}). For rebalanced portfolios during 2020–2022, RASPs with LINEX/Quantile present better results in terms of return/Sharpe for 16-asset portfolios, reflecting a more volatile regime, though without altering the leading models for index replication. It should be noted that Treynor can turn negative due to $\beta < 0$ in some windows; therefore, the replication assessment should primarily rely on Abs.E and SD/SD^{-}/VaR, where EL (AE/SE) and MSD_{AE} consistently outperform.

5 Conclusion

Risk-averse stochastic optimization based on models offers a simple formulation for replicating portfolios by connecting risk measures with score functions for tracking error evaluation between contingent claims and underlying assets, allowing control over different risk components, including asymmetry and tail behavior. Empirically analyzing 25 compositions of such functionals, we found that parsimonious, symmetric combinations have lower risk and tracking error than more aggressive portfolios, although the latter produce higher returns.

The results presented in this work establish a direct link between the allocation structure and the classical risk-return trade-off, as the choice of risk measures and score functions alters the balance between diversification and loss exposure. AE and SE, which penalize both sides of volatility, exhibit returns after minimization that are closer to the referenced benchmark and yield lower levels for the chosen risk metrics SD and SD^-. On the other hand, since LINEX and Quantile focus on extreme outcomes, they do not penalize positive deviations and thus achieve higher returns. Hence, these findings are reinforced by the fact that obtaining better rewards requires taking more risk, which, in our framework, is achieved by increasing portfolio concentration, as noted earlier. The Sharpe index further supports this behavior, indicating that for four-asset portfolios, MSD_{SE} provides the best equilibrium, while for sixteen-asset portfolios, ML_{LINEX} achieves the best balance.

The evidence is consistent across sample sizes and market conditions. Furthermore, this paper is the first to analyze the RASP framework, overcoming the limitations of related work Britten-Jones (1999), Li (2015), Shalit (2020). However, given that estimations rely on Historical Simulation, the findings are inherently tied to the empirical distribution of the observed data, which may limit their out-of-sample predictive power. In a literature review of energy markets, Halkos & Tsirivis (2019) highlights a series of studies that, unlike ours, focus on modeling dynamic dependencies and conditional risk, a reliable proxy for future development.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Anis, H. T., Costa, G. & Kwon, R. H. (2023), 'Risk-allocation-based index tracking', *Computers & Operations Research* **154**, 106219.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent measures of risk', *Mathematical Finance* **9**(3), 203–228.

Balter, A. G. & Pelsser, A. (2020), 'Pricing and hedging in incomplete markets with model uncertainty', *European Journal of Operational Research* **282**(3), 911–925.

Basel Committee on Banking Supervision (2013), 'Fundamental review of the trading book:

A revised market risk framework', *Consultative Document* . Available in: https://www.bis.org/publ/bcbs265.pdf.

Beasley, J. E., Meade, N. & Chang, T.-J. (2003), 'An evolutionary heuristic for the index tracking problem', *European Journal of Operational Research* **148**(3), 621–643.

Bellini, F., Laeven, R. J. A. & Rosazza Gianin, E. (2021), 'Dynamic robust Orlicz premia and Haezendonck–Goovaerts risk measures', *European Journal of Operational Research* **291**(2), 438–446.

Britten-Jones, M. (1999), 'The sampling error in estimates of mean-variance efficient portfolio weights', *The Journal of Finance* **54**(2), 655–671.

Gaivoronski, A. A., Krylov, S. & van der Wijst, N. (2005), 'Optimal portfolio selection and dynamic benchmark tracking', *European Journal of Operational Research* **163**, 115–131.

Goel, A., Sharma, A. & Mehra, A. (2018), 'Index tracking and enhanced indexing using mixed conditional value-at-risk', *Journal of Computational and Applied Mathematics* **335**, 361–380.

Guastaroba, G., Mansini, R., Ogryczak, W. & Speranza, M. G. (2020), 'Enhanced index tracking with CVaR-based ratio measures', *Annals of Operations Research* pp. 1–49.

Halkos, G. E. & Tsirivis, A. S. (2019), 'Energy commodities: A review of optimal hedging strategies', *Energies* **12**(20), 3979.

Huang, S.-F. & Guo, M. (2013), 'An optimal multi-step quadratic risk-adjusted hedging strategy', *Journal of the Korean Statistical Society* **42**(1), 37–49.

Kaina, M. & Rüschendorf, L. (2009), 'On convex risk measures on Lp-spaces', *Mathematical Methods of Operations Research* **69**(3), 475–495.

Kuester, K., Mittnik, S. & Paolella, M. S. (2006), 'Value-at-risk prediction: A comparison of alternative strategies', *Journal of Financial Econometrics* **4**(1), 53–89.

Li, J. (2015), 'Sparse and Stable Portfolio Selection With Parameter Uncertainty', *Journal of Business & Economic Statistics* **33**(3), 381–392.

Mainik, G., Mitov, G. & Rüschendorf, L. (2015), 'Portfolio optimization for heavy-tailed assets: Extreme Risk Index vs. Markowitz', *Journal of Empirical Finance* **32**, 115–134.

Meihua, X. C., Fengmin, X. & Hong, X. (2011), 'A mixed 0–1 LP for index tracking problem with CVaR risk constraints', *Annals of Operations Research* **196**(1).

Müller, F. M. & Righi, M. B. (2018), 'Numerical comparison of multivariate models to forecasting risk measures', *Risk Management* **20**(1), 29–5.

Perlin, M. (2022), 'Package 'BatchGetSymbols'.'. Available in: `https://cran.r-project.org/web/packages/BatchGetSymbols/BatchGetSymbols.pdf`.

Pflug, G. C., Pichler, A. & Wozabal, D. (2012), 'The 1/N investment strategy is optimal under high model ambiguity', *Journal of Banking & Finance* **36**(2), 410 – 417.

Ramos, H. P., Righi, M. B., Guedes, P. C. & Müller, F. M. (2023), 'A comparison of risk measures for portfolio optimization with cardinality constraints', *Expert Systems with Applications* p. 120412.

Righi, M. B. (2019), 'A composition between risk and deviation measures', *Annals of Operations Research* **282**(1), 299–313.

Righi, M. B. & Borenstein, D. (2018), 'A simulation comparison of risk measures for portfolio optimization', *Finance Research Letters* **24**, 105–112.

Righi, M. B., Müller, F. M. & Moresco, M. R. (2020), 'On a robust risk measurement approach for capital determination errors minimization', *Insurance: Mathematics and Economics* **95**, 199–211.

Righi, M. B., Müller, F. M. & Moresco, M. R. (2025), 'A risk measurement approach from risk-averse stochastic optimization of score functions', *Insurance: Mathematics and Economics* **120**, 42–50.

Rockafellar, R. T., Uryasev, S. & Zabarankin, M. (2006), 'Generalized deviations in risk analysis', *Finance and Stochastics* **10**(1), 51–74.

Sant'Anna, L. R., Filomena, T. P., Guedes, P. C. & Borenstein, D. (2017), 'Index tracking with controlled number of assets using a hybrid heuristic combining genetic algorithm and non-linear programming', *Annals of Operations Research* **258**(2), 849–867.

Sant'Anna, L. R., Righi, M. B., Müller, F. M. & Guedes, P. C. (2022), 'Risk measure index tracking model', *International Review of Economics & Finance* **80**, 361–383.

Shalit, H. (2020), 'The Shapley value of regression portfolios', *Journal of Asset Management* **21**(6), 506–512.

Shawky, H. A., Kuenzel, R. & Mikhail, A. D. (1997), 'International portfolio diversification: a synthesis and an update', *Journal of international financial markets, Institutions and Money* **7**(4), 303–327.

Zhao, L.-T., Meng, Y., Zhang, Y.-J. & Li, Y.-T. (2019), 'The optimal hedge strategy of crude oil spot and futures markets: Evidence from a novel method', *International Journal of Finance & Economics* **24**(1), 186–203.

Table 1: Metrics for optimal replicating portfolios with and without rebalancing. Portfolios are formed with 4 and 16 assets; in-sample size of 1000 observations; $\alpha = 5\%$. Stocks are S&P 100 constituents. Log-returns are $\times 100$. Period: Jan 5, 2010–Dec 30, 2022.

(a) Without rebalancing

RASP	Av.ret	Cum.ret	SD	SD^-	VaR	Sharpe	Treynor	Abs.E
4 assets								
EL_{SE}	0.7006	100.6299	16.8140	12.1198	24.6627	0.6126	0.8675	0.4771
EL_{AE}	0.6810	97.8138	16.5950	11.9603	24.3974	0.6043	0.8549	0.4754
EL_{LINEX}	0.9850	141.4803	25.0305	17.9619	36.5895	0.5911	0.8924	0.7934
$EL_{Quantile}$	0.7445	106.9311	18.1594	13.0797	26.5236	0.6057	0.8744	0.5197
$EL_{Expectile}$	0.7349	105.5519	17.6401	12.7145	25.8178	0.6149	0.8765	0.4945
MSD_{SE}	0.7369	105.8484	17.5877	12.6758	25.7025	0.6172	0.8809	0.4964
MSD_{AE}	0.7083	101.7327	16.9813	12.2406	24.8971	0.6127	0.8641	0.4816
MSD_{LINEX}	1.0440	149.9524	26.7972	19.2205	39.1678	0.5857	0.9325	0.8624
$MSD_{Quantile}$	0.8109	116.4721	19.4163	13.9827	28.3510	0.6199	0.9036	0.5494
$MSD_{Expectile}$	0.7910	113.6150	18.8990	13.6109	27.5687	0.6192	0.8966	0.5310
ES_{SE}	0.7383	106.0390	17.7225	12.7631	25.8906	0.6140	0.8849	0.5096
ES_{AE}	0.7414	106.4814	17.8204	12.8304	26.0363	0.6131	0.8848	0.5129
ES_{LINEX}	1.0010	143.7813	25.3671	18.2067	37.0442	0.5924	0.9078	0.8089
$ES_{Quantile}$	0.7868	113.0073	22.0034	15.7690	32.2276	0.5579	0.8521	0.7058
$ES_{Expectile}$	0.7811	112.1864	19.1571	13.7749	27.9779	0.6054	0.9623	0.5611
$EVaR_{SE}$	0.7327	105.2440	17.5266	12.6233	25.6111	0.6160	0.8836	0.5018
$EVaR_{AE}$	0.7182	103.1570	17.3161	12.4731	25.3651	0.6097	0.8733	0.4968
$EVaR_{LINEX}$	1.0147	145.7476	25.9654	18.6440	37.9676	0.5901	0.8955	0.8275
$EVaR_{Quantile}$	0.7861	112.9127	20.1543	14.4851	29.3794	0.5887	0.7936	0.6137
$EVaR_{Expectile}$	0.7738	111.1462	18.6770	13.4427	27.2419	0.6131	0.8931	0.5375
ML_{SE}	0.8666	124.4706	22.1213	15.8708	32.1622	0.5879	0.9219	0.6889
ML_{AE}	0.8666	124.4768	22.1220	15.8713	32.1636	0.5879	0.9219	0.6889
ML_{LINEX}	1.1522	165.4850	30.5443	21.8827	44.5860	0.5707	0.9206	1.0127
$ML_{Quantile}$	0.8685	124.7466	24.6173	17.6334	35.8527	0.5458	0.8940	0.8110
$ML_{Expectile}$	0.8612	123.6983	23.4587	16.8162	34.1342	0.5560	0.8851	0.7566
16 assets								
EL_{SE}	7.8851	1132.5460	182.7200	130.4337	266.1629	0.6739	1.1987	7.5578
EL_{AE}	8.0583	1157.4215	183.5571	131.1414	266.5220	0.6860	1.1927	7.5501
EL_{LINEX}	12.0093	1724.9004	245.8262	178.0501	362.4772	0.7670	1.0514	9.8581
$EL_{Quantile}$	10.7906	1549.8553	253.2556	180.7196	368.9463	0.6635	1.2181	10.6528
$EL_{Expectile}$	8.6055	1236.0177	201.2850	143.7263	292.7651	0.6687	1.1902	8.3538
MSD_{SE}	7.6705	1101.7228	185.8188	132.6446	271.8493	0.6465	1.1849	7.7509
MSD_{AE}	7.7566	1114.0814	183.5730	130.9666	267.8345	0.6605	1.1868	7.6099
MSD_{LINEX}	12.1158	1740.2004	246.9180	178.8987	363.9458	0.7707	1.0544	9.8978
$MSD_{Quantile}$	8.9784	1289.5733	207.4305	148.0896	301.0387	0.6746	1.2120	8.6174
$MSD_{Expectile}$	8.1298	1167.6832	194.2169	138.7098	283.4949	0.6549	1.1836	8.0806
ES_{SE}	7.7581	1114.3075	187.4829	133.7639	274.9573	0.6481	1.1951	7.8469
ES_{AE}	7.8584	1128.7040	188.2328	134.3339	275.3715	0.6525	1.1957	7.8658
ES_{LINEX}	12.0093	1724.9004	245.8262	178.0501	362.4772	0.7670	1.0514	9.8581
$ES_{Quantile}$	8.3448	1198.5735	193.9375	138.3795	282.3988	0.6725	1.2144	8.0644
$ES_{Expectile}$	8.0293	1153.2607	192.4095	137.4895	280.9359	0.6534	1.1853	8.0093
$EVaR_{SE}$	7.7452	1112.4481	187.3751	133.6448	274.8136	0.6461	1.1821	7.8241
$EVaR_{AE}$	7.9352	1139.7410	185.3633	132.3214	269.9691	0.6690	1.2112	7.7115
$EVaR_{LINEX}$	12.0388	1729.1376	246.2981	178.3970	363.1295	0.7673	1.0528	9.8819
$EVaR_{Quantile}$	8.3467	1198.8384	192.8672	137.7590	280.5391	0.6757	1.1974	7.9874
$EVaR_{Expectile}$	8.2718	1188.0812	193.0180	137.8417	281.2563	0.6703	1.2034	8.0165
ML_{SE}	9.7846	1405.3741	223.2746	160.0759	331.4830	0.6897	1.1248	9.3136
ML_{AE}	9.7846	1405.3741	223.2746	160.0759	331.4830	0.6897	1.1248	9.3136
ML_{LINEX}	12.1627	1746.9330	247.2335	179.0763	364.6151	0.7722	1.0611	9.9204
$ML_{Quantile}$	9.8553	1415.5215	222.4727	159.6667	328.4040	0.6921	1.1266	9.2006
$ML_{Expectile}$	9.8553	1415.5215	222.4727	159.6667	328.4040	0.6921	1.1266	9.2006

(b) Rebalancing every 120 days

RASP	Av.ret	Cum.ret	SD	SD^-	VaR	Sharpe	Treynor	Abs.E
4 assets								
EL_{SE}	0.6285	90.2659	15.1137	10.9288	22.6323	0.6055	-5.2146	1.0087
EL_{AE}	0.6034	86.6609	14.6026	10.5671	22.1035	0.6001	-5.3304	0.9978
EL_{LINEX}	0.8479	121.7800	30.7320	21.9754	45.2290	0.4207	-5.7051	1.4689
$EL_{Quantile}$	0.6693	96.1383	16.7441	12.0886	24.5389	0.5856	-5.0431	1.0511
$EL_{Expectile}$	0.6748	96.9198	16.3704	11.8301	24.0982	0.6028	-4.9910	1.0399
MSD_{SE}	0.6718	96.4949	16.1725	11.6745	23.7324	0.6086	-5.0565	1.0347
MSD_{AE}	0.6317	90.7345	15.2321	11.0151	22.7453	0.6041	-5.2078	1.0115
MSD_{LINEX}	0.9014	129.4648	32.3937	23.1585	47.6574	0.4242	-5.6461	1.5332
$MSD_{Quantile}$	0.7686	110.3978	19.0904	13.7429	27.4907	0.5966	-4.9952	1.1192
$MSD_{Expectile}$	0.7499	107.7157	18.4729	13.3068	26.7851	0.5993	-4.9942	1.1006
ES_{SE}	0.6550	94.0723	16.1546	11.6617	23.7177	0.5933	-5.0614	1.0335
ES_{AE}	0.6453	92.6787	16.1287	11.6447	23.7419	0.5852	-5.0908	1.0333
ES_{LINEX}	0.8647	124.2022	31.1830	22.3030	45.9313	0.4231	-5.6016	1.4878
$ES_{Quantile}$	0.7210	103.5640	21.9456	15.7207	30.5315	0.4947	-5.1754	1.1942
$ES_{Expectile}$	0.7078	101.6557	18.0008	12.9689	26.1183	0.5802	-5.6779	1.0870
$EVaR_{SE}$	0.6543	93.9736	15.9838	11.5407	23.5227	0.5989	-5.0718	1.0287
$EVaR_{AE}$	0.6349	91.1903	15.5219	11.2185	23.1280	0.5972	-5.1930	1.0188
$EVaR_{LINEX}$	0.8771	125.9735	31.4008	22.4455	46.1711	0.4256	-5.2774	1.4944
$EVaR_{Quantile}$	0.7187	103.2254	19.4939	14.0288	27.7500	0.5478	-4.9786	1.1257
$EVaR_{Expectile}$	0.7102	102.0091	17.7566	12.8033	25.8473	0.5894	-4.9983	1.0799
ML_{SE}	0.8066	115.8544	22.8501	16.3986	32.4089	0.5302	-4.8651	1.2352
ML_{AE}	0.8067	115.8739	22.8465	16.3956	32.4088	0.5304	-4.8581	1.2351
ML_{LINEX}	0.9812	140.9314	35.2246	25.1913	51.6857	0.4247	-5.0110	1.6446
$ML_{Quantile}$	0.8246	118.4437	25.8422	18.4768	35.3088	0.4913	-4.5973	1.3134
$ML_{Expectile}$	0.8210	117.9200	24.0162	17.1941	33.8693	0.5169	-5.5777	1.2723
16 assets								
EL_{SE}	6.3389	910.4593	177.1627	127.6835	249.1199	0.5650	-4.9129	7.5077
EL_{AE}	6.5285	937.6957	179.0635	129.1203	249.3460	0.5746	-4.8574	7.5123
EL_{LINEX}	10.6114	1524.1226	260.0728	188.7629	379.3809	0.6467	-5.6472	10.8851
$EL_{Quantile}$	9.1754	1317.8729	273.5005	197.2697	367.4553	0.5324	-4.4849	11.1671
$EL_{Expectile}$	6.9250	994.6468	201.6904	145.4959	277.4676	0.5432	-4.6741	8.4002
MSD_{SE}	6.0059	862.6287	183.4342	132.0677	264.9118	0.5182	-4.9546	7.8651
MSD_{AE}	6.0568	869.9410	176.6481	127.2390	249.2864	0.5416	-4.8080	7.5012
MSD_{LINEX}	10.7568	1545.0089	260.4498	189.0056	380.1729	0.6547	-5.7203	10.9075
$MSD_{Quantile}$	6.9507	998.3322	210.2355	151.7497	287.6614	0.5240	-4.5339	8.7216
$MSD_{Expectile}$	6.1537	883.8668	192.9685	139.1404	276.1914	0.5052	-4.8424	8.1982
ES_{SE}	5.9546	855.2623	179.8074	129.4568	259.1678	0.5227	-5.0442	7.7431
ES_{AE}	6.1587	884.5739	177.9936	128.1600	254.7483	0.5468	-5.1153	7.6358
ES_{LINEX}	10.6114	1524.1226	260.0728	188.7629	379.3809	0.6467	-5.6472	10.8851
$ES_{Quantile}$	6.3016	905.1077	184.7791	133.1586	257.9439	0.5395	-4.7632	7.7719
$ES_{Expectile}$	6.2276	894.4683	184.7499	133.1306	262.6130	0.5345	-5.0008	7.8595
$EVaR_{SE}$	6.1071	877.1647	180.8414	130.2362	260.3329	0.5341	-5.0992	7.7611
$EVaR_{AE}$	6.2247	894.0522	177.2957	127.7164	250.5899	0.5540	-4.8977	7.5391
$EVaR_{LINEX}$	10.6182	1525.0982	260.0760	188.7323	379.3183	0.6470	-5.6476	10.8862
$EVaR_{Quantile}$	6.8587	985.1259	195.3895	140.9618	267.0020	0.5559	-4.7159	8.0961
$EVaR_{Expectile}$	6.2130	892.3708	187.2741	135.0380	266.1642	0.5252	-4.9052	7.9503
ML_{SE}	8.0715	1159.3126	229.9737	165.9116	332.9278	0.5565	-5.3534	9.7658
ML_{AE}	8.0715	1159.3126	229.9737	165.9116	332.9278	0.5565	-5.3534	9.7658
ML_{LINEX}	10.8026	1551.5846	260.7892	189.2826	380.7106	0.6566	-5.7678	10.9254
$ML_{Quantile}$	7.6726	1102.0185	244.0544	176.5478	351.7651	0.5012	-4.7816	10.2360
$ML_{Expectile}$	7.6726	1102.0185	244.0544	176.5478	351.7651	0.5012	-4.7816	10.2360

Note: Av.ret = average portfolio returns; Cum.ret = cumulative portfolio return; SD = standard deviation; SD^- = semi-deviation; VaR = average value-at-risk at $\alpha = 5\%$; Sharpe = excess return/SD; Treynor = excess return/Beta (CAPM regression vs. market, r_f = 3m T-Bill). Av.ret, SD, SD^-, Sharpe, Treynor, VaR are annualized. Abs.E = absolute error between replicating and reference asset returns.

Table 2: Performance of the S&P 500 future market index. The log-returns are multiplied by 100. The analyzed period comprises January 5, 2010, to December 30, 2022.

Sample	Av.ret	Cum.ret	SD	SD^-	VaR	Sharpe	Treynor	Abs.E
Full sample	0.4581	65.7974	18.0998	13.6388	27.1741	0.3530	0.4257	0.0000
2014 - 2016	0.3723	17.9186	13.4009	9.9406	22.7975	0.4294	0.3772	0.0000
2017 - 2019	0.6601	31.4762	13.0821	10.0248	21.1566	0.6727	13.6038	0.0000
2020 - 2022	0.5268	25.1892	23.8831	18.1842	35.5147	0.3151	-15.7754	0.0000

Note: Av.ret is the average S&P 500 futures return; Cum.ret is cumulative return; SD is standard deviation; SD^- is semi-deviation; and VaR is the average 5% value-at-risk. Sharpe = expected excess return/SD; Treynor = expected excess return/β, with β from a single-factor regression on the market. Sharpe, Treynor, and β use the 3-month U.S. T-bill yield as r_f. For each scenario, we compute 1,000 portfolios and report the mean. Av.ret, SD, SD^-, Sharpe, Treynor, and VaR are annualized. Abs.E measures error relative to the futures index; for the S&P 500 futures benchmark, Abs.E = 0.

Table 3: Results for the Absolute Error (Abs.E) metric – comparison between 4-asset and 16-asset replicating portfolios, with (With reb.) and without rebalancing every 120 days (No reb.).

n	**Period**	α	**4 assets**		**16 assets**	
			No reb.	With reb.	No reb.	With reb.
Full sample 2014–2022						
250		1%	MSD_{AE}	EL_{SE}	EL_{SE}	EL_{SE}
		2.5%	EL_{SE}	EL_{SE}	EL_{SE}	EL_{SE}
		5%	EL_{SE}	EL_{AE}	EL_{AE}	EL_{SE}
1000		1%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
		2.5%	EL_{AE}	EL_{AE}	EL_{AE}	MSD_{AE}
		5%	EL_{AE}	EL_{AE}	EL_{AE}	MSD_{AE}
Sub-periods						
250	2014–2016	1%	MSD_{AE}	EL_{SE}	EL_{SE}	MSD_{AE}
	2017–2019	1%	EL_{SE}	EL_{AE}	EL_{SE}	EL_{SE}
	2020–2022	1%	MSD_{AE}	EL_{SE}	EL_{SE}	EL_{SE}
	2014–2016	2.5%	EL_{SE}	EL_{SE}	EL_{SE}	EL_{SE}
	2017–2019	2.5%	EL_{SE}	EL_{AE}	EL_{SE}	EL_{SE}
	2020–2022	2.5%	EL_{SE}	EL_{SE}	MSD_{SE}	MSD_{AE}
	2014–2016	5%	EL_{SE}	EL_{SE}	EL_{AE}	MSD_{AE}
	2017–2019	5%	EL_{SE}	EL_{AE}	EL_{AE}	EL_{SE}
	2020–2022	5%	EL_{SE}	EL_{SE}	EL_{AE}	EL_{SE}
1000	2014–2016	1%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
	2017–2019	1%	EL_{AE}	EL_{AE}	EL_{AE}	MSD_{AE}
	2020–2022	1%	EL_{AE}	EL_{AE}	EL_{AE}	MSD_{AE}
	2014–2016	2.5%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
	2017–2019	2.5%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
	2020–2022	2.5%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{SE}
	2014–2016	5%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
	2017–2019	5%	EL_{AE}	EL_{AE}	EL_{AE}	EL_{AE}
	2020–2022	5%	EL_{SE}	EL_{AE}	EL_{SE}	EL_{SE}