KATHOLISCHE UNIVERSITÄT EICHSTÄTT-INGOLSTADT

Three Essays on Asset Management and Capital Allocation

Kumulative Dissertation vorgelegt von

Christian Maximilian Cara

zur Erlangung des Grades

Doctor rerum politicarum
(Dr. rer. pol.)

im Rahmen des Dissertationsverfahrens an der
Wirtschaftswissenschaftlichen Fakultät Ingolstadt
der Katholischen Universität Eichstätt-Ingolstadt

Tag der mündlichen Prüfung:	16. Oktober 2025
Name des Erstgutachters:	Prof. Dr. Thomas Mählmann
Name des Zweitgutachters:	Prof. Dr. David Streich

Section A – Introduction and Synthesis

1 General introduction

Asset management is a cornerstone of modern finance, deeply intertwined with the fundamental challenge of allocating capital under uncertainty. As global financial markets become increasingly complex and interconnected, both institutional and individual investors seek not only to achieve superior returns but also to manage and mitigate risk through informed asset allocation. This quest takes place within a dynamic and often volatile environment where economic cycles, crisis events, and structural market changes continuously reshape the landscape of investment decision-making. Against this backdrop, the thesis *“Three Essays on Asset Management and Capital Allocation”* presents a trilogy of empirical investigations that explore, from different yet interconnected perspectives, key challenges and evolving strategies in contemporary portfolio management.

The three essays address prominent and long-standing questions within the field of asset management: How can investors best allocate their capital to manage risk, especially in turbulent markets? Is sophisticated portfolio optimization necessarily superior to simple heuristic-based strategies such as the equally-weighted (1/N) portfolio? And how do index-tracking strategies – often employed through exchange-traded funds (ETFs) – perform when markets experience systemic stress? These questions, while distinct in their empirical focus, converge on the core problem of how to construct robust, resilient, and efficient portfolios that perform well not just in average conditions but also in extreme scenarios.

The first essay, *“The performance of risk-based asset allocation in downward markets – an empirical examination,”* explores the relevance of risk-based portfolio construction methods during periods of heightened market volatility and drawdown. These methods, which deliberately de-emphasize return forecasts in favor of risk characteristics such as volatility, covariance, and correlation, have gained significant traction following the Global Financial Crisis. Their appeal lies in their robustness against the estimation errors that frequently compromise return-based optimization. The paper contributes to the literature by empirically examining how such risk-oriented strategies compare to the benchmark 1/N portfolio, particularly during market downturns when diversification and downside protection become paramount.

The second essay, *“A performance ‘horse race’: Does anything beat the 1/N portfolio?”* enters into the heart of a decades-long debate on the empirical and theoretical virtues of naïve diversification. Following the seminal work by DeMiguel, Garlappi, and Uppal (2009), the paper conducts a comparative performance analysis – often referred to as a "horse race – between the equally-weighted portfolio and a range of optimization-based portfolio strategies. With an expanded dataset and a longer out-of-sample period than prior studies, this paper revisits the robustness of the 1/N strategy, examining whether its simplicity continues to outperform or whether, under certain conditions, sophisticated models can offer a superior risk-adjusted return.

The third essay, *“Index tracking in crisis periods: An empirical investigation of the German DAX index,”* extends the discussion to the practical realm of passive investment strategies. Specifically, it evaluates the efficacy of index tracking – an investment technique central to ETF design – under conditions of market stress. With the growth of ETF investment accelerating in recent years, understanding how these strategies perform during crises is critical for both market participants and regulators. The paper examines whether various tracking methods, including constrained regression and optimization models, maintain their effectiveness during turbulent market periods, and how the granularity of asset selection (i.e., the number of stocks used) influences performance.

2 Relevance and contribution

Each paper contributes to an evolving understanding of portfolio construction under uncertainty. Collectively, they offer valuable insights into a set of practical challenges faced by investors, portfolio managers, and policy-makers alike.

First, the empirical evaluation of risk-based asset allocation strategies in the context of downward markets (Paper 1) addresses the increasingly recognized need for risk resilience in portfolio management. Traditional approaches, such as mean-variance optimization, often rely heavily on return forecasts that are prone to estimation error and structural instability. By focusing on volatility and correlation rather than expected return, risk-based strategies such as risk parity and minimum variance offer a compelling alternative. Yet their real-world performance, especially under stress conditions, remains under-examined. This essay fills that gap by providing a comparative analysis that evaluates their effectiveness not just in general, but specifically in times of crisis—when they are most needed.

Second, Paper 2 interrogates the practical superiority of optimized portfolios over the simple yet empirically effective 1/N strategy. It addresses a well-known paradox in asset management: the theoretically inferior naïve diversification approach often performs surprisingly well in empirical tests. This paradox challenges assumptions embedded in modern portfolio theory and calls for a deeper examination of the trade-off between model complexity and robustness. By using longer time periods and broader asset universes than earlier studies, this essay helps identify the circumstances under which optimization strategies might regain their theoretical advantage, if any, over simple heuristics.

Third, Paper 3 broadens the scope by analyzing index tracking – an increasingly dominant form of passive investment – within the context of crisis periods. While ETFs and index funds have democratized investing and lowered costs, their growing systemic importance raises new questions about their behavior during periods of stress. Using the German DAX as a case study, the paper tests the robustness of different index tracking models under real-world constraints. Importantly, it investigates how the dimensionality of the asset universe (number of stocks used) affects tracking performance, which has implications for fund construction, market liquidity, and investor outcomes.

3 Thematic cohesion and red thread

At first glance, each paper appears to tackle a separate aspect of asset management. However, a closer inspection reveals a unifying thread: all three papers critically assess portfolio construction methods in the presence of uncertainty, and particularly, under conditions where traditional models are likely to break down. Whether the uncertainty stems from estimation error (Paper 1), model misspecification (Paper 2), or systemic market crises (Paper 3), each paper addresses how investment strategies can be designed to remain effective, or at least resilient, in adverse environments.

The methodological connection lies in the shared empirical orientation and the common analytical framework of evaluating portfolio performance using out-of-sample testing, real-world data, and robustness checks. Each essay employs a rolling backtest methodology, comparing the ex-ante constructed portfolio with ex-post realized outcomes. The focus on empirical validation strengthens the thesis' contribution to the field of applied finance and ensures relevance to practitioners as well as scholars.

Another connecting dimension is the implicit tension between theory and practice. All three papers examine strategies that either originate from or challenge foundational portfolio theory. Whether it's Markowitz's mean-variance model, the risk parity paradigm, or passive replication of market indices, the thesis scrutinizes how these theoretical constructs fare when confronted with empirical realities such as estimation risk, market turmoil, and practical implementation constraints. In doing so, it provides not only a critique but also a constructive extension of classical finance models.

Moreover, the order of the essays reflects an intentional escalation in complexity and scope. Paper 1 begins with a critique of mean-variance optimization and a turn toward risk-based allocation. Paper 2 deepens the analysis by challenging the core assumption that optimization necessarily yields better results than simplicity. Finally, Paper 3 extends the conversation to market-wide replication strategies, bringing in the macroeconomic context of financial crises and institutional investment vehicles. This progression – from risk management within a portfolio to the performance of portfolios themselves, and ultimately to the systemic behavior of index replication – constructs a coherent narrative arc that speaks to the evolution of asset management thinking.

4 Deep synthesis and paper summary

This section includes a deep synthesis of each paper's motivation, content, and contribution.

Paper 1: The Performance of Risk-Based Asset Allocation in Downward Markets – An Empirical Examination

Motivation and Content

The first paper addresses a fundamental question in modern asset management: how can investors construct portfolios that maintain resilience during market downturns? This concern has grown increasingly important in the aftermath of the 2008 financial crisis, when the limitations of return-based optimization approaches became painfully evident. Classical mean-variance frameworks, although theoretically elegant, are highly sensitive to estimation errors in expected returns—a problem particularly exacerbated during volatile or crisis periods.

In response, risk-based portfolio strategies such as minimum variance, equal risk contribution, and risk parity have gained prominence. These approaches focus on the risk attributes of assets, largely avoiding the need for precise return forecasts. The motivation behind this shift is twofold: first, it aims to sidestep estimation error by relying primarily on covariance matrices, which are more stable over time; second, it seeks to construct more balanced portfolios that avoid concentration in high-volatility assets.

The paper empirically evaluates the performance of several risk-based strategies across multiple market regimes, with particular attention to downward trends and crisis periods. Using a dataset spanning multiple asset classes and historical market crashes, the study conducts rolling-window backtests to assess out-of-sample performance. Performance metrics include standard deviation, maximum drawdown, and risk-adjusted returns such as the Sharpe and Sortino ratios.

Contribution

The key contribution of this paper lies in its comprehensive comparison of risk-based strategies under adverse market conditions. The findings support the hypothesis that risk-oriented approaches tend to offer superior downside protection without significantly sacrificing long-term returns. Specifically, risk parity and minimum variance portfolios outperform naive and return-optimized portfolios in terms of maximum drawdown and volatility during crises. This suggests that risk-based frameworks not only offer a methodological alternative to mean-variance optimization but also provide practical utility for investors prioritizing capital preservation during downturns.

Paper 2: A Performance "Horse Race": Does Anything Beat the 1/N Portfolio?

Motivation and Content

Building on the insights from the first essay, the second paper turns its focus to one of the most debated topics in portfolio theory: the empirical performance of the naive 1/N strategy relative to optimized alternatives. The seminal work of DeMiguel, Garlappi, and Uppal (2009) challenged the assumption that optimized portfolios consistently outperform simple equal-weighting. Their results indicated that the 1/N rule often matched or exceeded the performance of more sophisticated models, particularly when transaction costs and estimation errors were accounted for.

This paper revisits the question using an extended dataset and alternative methodologies. The aim is to conduct a "horse race" among multiple portfolio construction techniques, including mean-variance optimization, robust optimization, and various shrinkage approaches, to determine whether any consistently outperform the naive benchmark over long investment horizons.

The study makes use of rolling-window backtests, similar to the first paper, but includes additional robustness checks such as varying the asset universe, investment horizons, and transaction cost assumptions. The performance metrics are expanded to include turnover rates, information ratios, and cumulative returns.

Contribution

This paper makes several important contributions to the literature. First, it confirms the robustness of the 1/N strategy across a wide range of market conditions, asset classes, and evaluation metrics. Second, it demonstrates that while some optimized strategies can outperform 1/N under certain conditions, their performance is often sensitive to parameter tuning and assumptions about return distributions. Third, it highlights the trade-offs between model complexity, estimation error, and practical implementability.

From a practitioner's standpoint, the results suggest that simple rules-based strategies like 1/N can offer a viable benchmark or fallback strategy when model inputs are uncertain or difficult to estimate. In turn, this raises important questions about the cost-effectiveness of sophisticated optimization frameworks in real-world asset management.

Paper 3: Index Tracking in Crisis Periods – An Empirical Investigation of the German DAX Index

Motivation and Content

The third paper explores the increasingly important domain of passive investing, with a particular emphasis on index tracking strategies during market crises. The explosive growth in assets under management in ETFs underscores the relevance of this topic. Yet, while ETFs are designed to replicate the performance of specific indices, their ability to do so accurately in turbulent times remains under-examined.

This paper investigates the efficacy of different index tracking methods—relative optimization, constrained regression, and classical mean-variance—using the German DAX index as a case study. The novelty lies in the segmentation of the dataset into crisis and non-crisis periods to assess how tracking accuracy and portfolio stability vary across regimes. Moreover, the study examines the role of portfolio size by varying the number of constituent stocks used in the tracking portfolio.

Key performance indicators include active return, tracking error, and the variance of tracking error. The methods are tested both in-sample and out-of-sample to simulate real-world application.

Contribution

The primary contribution of this paper is its empirical analysis of index tracking under stress conditions. The findings reveal that tracking performance tends to deteriorate during crisis periods, regardless of the method used. Among the techniques, constrained regression and relative optimization offer the most stable tracking during turbulence, while the Markowitz approach is more prone to large deviations due to its sensitivity to input errors.

The paper underscores the practical limitations of passive strategies in times of market stress and suggests that a more adaptive or hybrid approach might be necessary to maintain alignment with the benchmark. This has implications for both ETF design and the use of passive vehicles in institutional portfolio construction.

4 Conclusion and Outlook

The cumulative insight provided by these three papers is both critical and timely. In a world of increasingly frequent market disruptions – from the Global Financial Crisis to the COVID-19 pandemic, to geopolitical shocks and environmental risks – investors must continuously re-evaluate the robustness of their strategies. The thesis does not purport to provide a definitive ranking of all asset allocation methods, nor does it claim to offer a one-size-fits-all solution. Instead, it provides a data-driven, critical, and nuanced analysis of how various portfolio construction methods perform under realistic and often adverse conditions.

Ultimately, the thesis aims to contribute to a more grounded, empirically informed understanding of asset management. It encourages investors and academics alike to look beyond theoretical elegance and consider the messy, uncertain world in which portfolios are actually implemented. In doing so, it helps to bridge the gap between financial theory and practical asset management – an endeavor that is more important than ever in today's uncertain financial landscape.

Section B – Paper 1

The performance of risk-based asset allocation in downward markets – an empirical examination

Christian Cara

This version: March 2024

Abstract

This study seeks to explore whether risk-based asset allocation strategies offer improved safeguarding against price declines during periods of market volatility. Drawing upon insights gathered from a backtesting analysis spanning the daily returns of the 49 Fama and French (1997) industry portfolios as test assets from 1926 to 2022, this hypothesis finds partial support. Notably, the analysis reveals that the global minimum variance (GMVP) and maximum diversification (MD) portfolios exhibit a loss-mitigation attribute, as evidenced by the realized risk measures during the backtesting period. These investment strategies were able to achieve significantly better or equivalent results in terms of both return and risk compared to an equally-weighted industry benchmark. However, it's worth noting that the risk reduction achieved by the GMVP and MD portfolios primarily stems from extreme weighting of low-volatility industries. Implementing such extreme weights presents practical challenges, particularly for institutional investors.

Keywords: Risk parity; risk-based asset allocation; maximum diversification;

1 Introduction

Both institutional and individual investors endeavor to delineate the process of asset selection and investment decisions, leveraging portfolio optimization to navigate expected returns and associated risks. The overarching objective is to maximize expected utility while adhering to a specific risk tolerance level or targeted return (Lee, 2011; Qian, 2013). This entails amalgamating diverse investments with disparate price trajectories to harness the power of diversification. Diversification arises as companies, for instance, react differentially to economic or sectoral changes. Consequently, the risk-return profile within a portfolio doesn't merely result from a weighted average of individual asset risk profiles, but rather from an overarching risk mitigated by the diversification effect, influenced by the correlation of asset returns (Clarke et al., 2013).

Various portfolio optimization methodologies have emerged with the aim of effectively allocating investment capital among assets concerning return and risk. One such method is the classical mean-variance absolute optimization approach pioneered by Markowitz (1952). In absolute portfolio optimization, portfolios are chosen to exhibit the minimum risk, in the form of variance or standard deviation, given an expected return, or to yield the maximum expected return for a given level of risk.

A significant critique of classical absolute portfolio optimization revolves around its dependence on estimated expected returns. Expected returns are challenging to predict reliably (Merton, 1980), and estimation errors can significantly impact optimal weights and, consequently, portfolio performance. Remedies to this issue include approaches like the Black and Litterman (1992) model, which relies solely on a reference portfolio (e.g., an index) and subjective return expectations to ascertain an optimal allocation.

In the aftermath of the global financial crisis of 2008-2009 and the ensuing significant asset value losses, *risk-based approaches* gained traction. This shift is attributable, firstly, to these methodologies eschewing reliance on (imprecisely) estimated expected returns. Secondly, financial crises, characterized by unforeseen and simultaneous sharp declines across multiple asset markets, often prompt a heightened need for diversification and risk hedging among both institutional and individual investors. Particularly, *risk parity strategies* lean more towards risk allocation than capital allocation, making them appealing to risk-averse investors (Qian, 2013).

Against this backdrop, this research project endeavors to delve into and offer a nuanced response to the following inquiries:

- *Whether risk-oriented optimization strategies can effectively rival a naive equally-weighted portfolio over the long term.*
- *Whether, especially during volatile market phases, these strategies can deliver on the promise of enhanced diversification and consequently safeguard against losses.*

2 Background on risk-based asset allocation

2.1 Rationale for risk-based approaches among institutional investors

From the vantage point of enhanced loss mitigation, risk-based approaches offer potential advantages over traditional methodologies. Consequently, this research endeavor seeks to investigate the capacity for risk reduction through such approaches, particularly in specific macroeconomic scenarios like crises. To achieve this objective, diverse risk-based strategies will be juxtaposed against each other over an extensive observation period spanning from 1926 to 2022 and assessed relative to an equally-weighted benchmark.

Risk-based methodologies, designed to distribute and mitigate risk across available investments, resonate particularly well with the risk management ethos of institutional investors such as insurance companies and pension funds. These entities are bound by financial obligations to achieve certain investment outcomes. Especially during periods of crisis, marked by frequent capital market downturns, it becomes imperative for institutional investors to maintain adequate capital reserves and employ risk-hedging strategies to avert tangible damages stemming from failure to meet their commitments, irrespective of market conditions. Given their predictable and ongoing commitments, insurance companies and pension funds must allocate capital even amidst volatile market environments. In contrast, private investors, lacking specific expenditure plans, have the flexibility to endure market downturns and adjust their expenditure plans accordingly (Bikker et al., 2018). To mitigate default risks, asset-liability management necessitates an appropriate funding ratio, i.e., the ratio of assets to liabilities, and risk-mitigation strategies in asset allocation (Sun et al., 2006). Consequently, the prospect of achieving return objectives while maintaining lower risk exposure or robust downside protection may hold particular promise from the standpoint of institutional investors.

Thus, this research project is, in part, undertaken from the perspective of an institutional investor. As governmental regulations in many jurisdictions prohibit or severely restrict insurance companies and pension funds from engaging in risky short sales of uncovered securities, this constraint is duly accounted for in portfolio optimization by imposing a *non-negativity condition* on portfolio weights (Bikker et al., 2018; Molk and Partnoy, 2019). Institutional investors often employ leveraged risk-based strategies, such as a leveraged risk parity strategy. Leverage is primarily utilized to amplify less risky asset classes, such as bonds, within a mixed portfolio, with the aim of achieving enhanced returns while maintaining identical risk contributions from both asset classes. This is driven by the fact that the majority of risk in such a mixed portfolio is largely determined by the riskier equities, even with a 50/50% allocation (Benham et al., 2019). Since this research project does not incorporate bonds or less risky asset classes, a budget constraint is applied.

2.2 Foundations of risk-based approaches to asset allocation

A common feature among all risk-based methodologies is the omission of modeling expected returns to determine optimal weights. Therefore, to establish the optimal portfolio, only an estimation of the future portfolio risk is required, typically described by the variance-covariance matrix (Lee, 2011). By omitting the estimation of expected returns, the associated problem of estimation errors in traditional approaches is thus circumvented. Consequently, risk-based approaches are considered robust optimization methods (Maillard et al., 2010). Enhanced performance or improved risk reduction is expected to be achieved through better diversification according to these optimization approaches. Risk-based approaches can be grouped into two categories (see Figure 1).

On the one hand, the allocation strategies based on equal risk budget (ERB) and equal risk contribution (ERC) can be collectively referred to as risk parity approaches. These methodologies explicitly aim to achieve an equalization of risk contributions or proportions among individual assets or asset classes within an investment universe. On the other hand, alternative approaches to risk diversification, such as the naïve equally-weighted portfolio (EW), do not directly seek to establish risk contribution parity. Further alternative methods for risk-based portfolio optimization include global minimum variance portfolios (GMVP) and maximum diversification portfolios (MD). The individual techniques, along with their respective optimization functions and characteristics, will be discussed in the following sections (see also Lee, 2011).

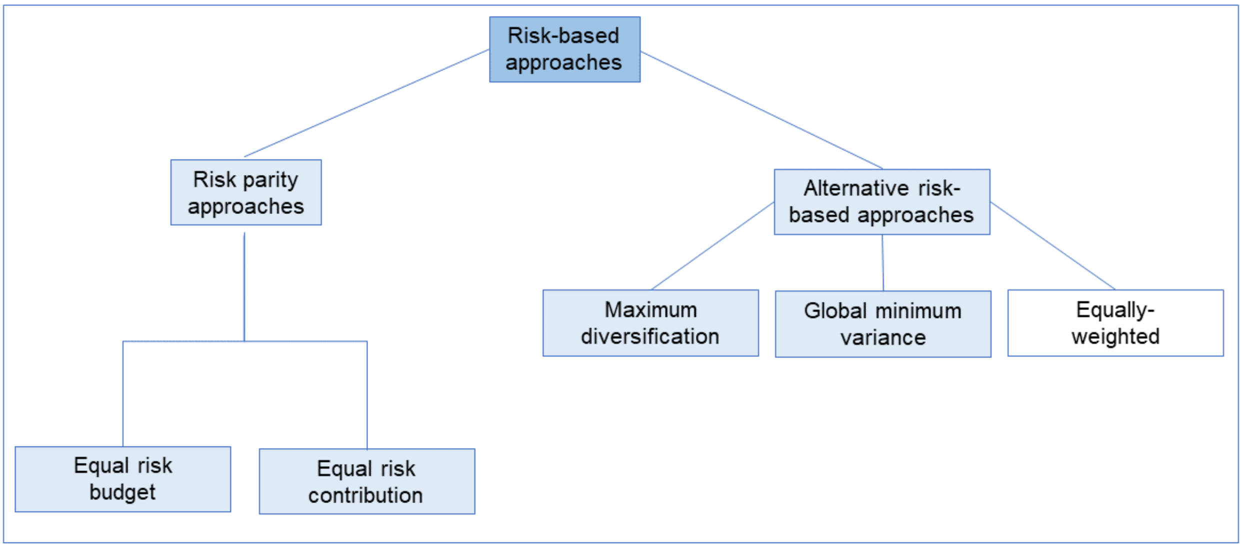


Figure 1: Categorization of risk-based asset allocation approaches (Source: own representation)

After this introduction, the next chapter delves first into the fundamentals of the risk parity approach, followed by an introduction to the equal risk budget strategy and the equal risk contribution strategy. This is followed by an exposition of alternative risk-balancing concepts, including the equally-weighted, global minimum variance, and the maximum diversification approaches.

2.3 The risk parity approach

2.3.1 Fundamentals

A traditional capital-based allocation, for example, of 50% equities and 50% bonds in the portfolio can result in the portfolio's return and risk being significantly influenced by equity returns, while bonds have little impact. Thus, there is actually no real diversification of the portfolio. In contrast, with the risk parity approach, compared to traditional asset allocation, significantly less is invested in equities and more in other asset classes. For a stock/bond portfolio to have both asset classes contribute equally to the overall risk (balanced risk portfolio), stocks would need to have a lower share in favor of bonds in the portfolio. Thus, the portfolio's risk budget can be spread more broadly across other investments. Accordingly, portfolio returns with lower volatilities can be expected.

Unlike, for example, asset allocation according to portfolio theory, the risk parity approach fundamentally does not require forecasts regarding expected returns. However, risk metrics and - depending on the chosen approach - correlations also need to be estimated. The latter is not the case for the equal risk budget (ERB) strategy as the first manifestation of the risk parity approach, as the weighting of individual investments in the portfolio is based on their inverse volatility. In contrast, correlation estimates are required for the second manifestation. This is the equal risk contribution (ERC) strategy.

We begin with the following definition.

Definition: A portfolio that fulfils the condition

$$w_1\partial_1\sigma_p=\ldots=w_N\partial_N\sigma_p, \quad \sigma_p=\sqrt{w\Sigma w^T}, \quad \partial_n\sigma_p=\frac{\partial\sigma_p}{\partial w_n}, \tag{1}$$

is called a **risk parity** portfolio.

The underlying idea is to achieve equal (or more generally: specific target) risk budgeting across all portfolio investments. Various risk measures can be used, but we will focus here on the standard deviation σ. Risk parity aims to ensure that each asset contributes equally to the overall volatility of the portfolio. This is achieved by demanding an identity of the **risk contribution (RC)** of all assets. The term $\frac{\partial\sigma_p}{\partial w_n}$ denotes the **marginal risk contribution (MRC)** of the n-th asset. The MRC indicates the additional portfolio risk generated by an infinitesimal increase in the proportion of the n-th asset in the portfolio and is formally defined as the derivative of the portfolio return standard deviation σ_p with respect to the proportion w_N. The RC, in turn, is the product of the proportion weight and MRC. Therefore, the following relationships hold:

The standard deviation of the portfolio return r_p is well known to be given by (Note: w denotes a row vector):

$$\sigma_p=\sqrt{w\Sigma w^T} \qquad (2)$$

The risk contribution of the n-th asset is calculated as:

$$RB_n=w_n\times MRB_n = w_n\times\partial_n\sigma_p \qquad (3)$$

For MRB follows:

$$MRB_n=\frac{(\Sigma w^T)_n}{\sigma_p} \qquad (4)$$

That is, the MRB is calculated by multiplying the variance-covariance matrix of asset returns with the allocation vector row-wise and summing. The n-th element of the (Nx1) vector Σw^T scaled by the portfolio standard deviation then yields the marginal risk contribution of the n-th asset.

$$MRB_n=\frac{\sum_{j=1}^{N} w_j\times Cov(r_n,\, r_j)}{\sigma_p}=\frac{Cov(r_n,\, r_p)}{\sigma_p} \qquad (5)$$

It can be shown that the sum of the N risk contributions equals the portfolio standard deviation:

$$\sigma_p=\sum_{n=1}^{N} RB_n \qquad (6)$$

If the (absolute) risk contribution RB is related to the total risk of the portfolio (σ_p), the relative contribution to the total risk (rRB) is obtained:

$$rRB_n = \frac{RB_n}{\sigma_p} = w_n \times \frac{Cov(r_n, r_p)}{\sigma_p^2} = w_n \times \beta_n \ , \qquad (7)$$

where β_n denotes the beta factor of the n-th asset with respect to the portfolio return.

2.3.2 Equal risk budget (ERB) strategy

The aim of this strategy is to make the **risk budgets** for all assets in the portfolio identical. This represents a simplification of the general definition of a risk parity strategy introduced above. Assuming that the return standard deviation represents the measure of risk, the risk budgets of the respective assets are defined as the product of standard deviation and portfolio share. Unlike the risk contributions, the correlations between individual assets are thus not taken into account, or it is assumed that the correlations of all asset returns with the portfolio return are identical. For the ERB approach, therefore, in a slight departure from the definition introduced above, the following objective applies to all assets 1 to N in a portfolio (Note: the risk budget - the product of standard deviation and portfolio share - can be interpreted as the risk contribution under perfect correlation):

$$w_1 \times \sigma_1 = w_2 \times \sigma_2 = \ldots = w_N \times \sigma_N \qquad (8)$$

For an asset n, the share w_N in the ERB portfolio is given in general form as:

$$w_n^{ERB} = \frac{1/\sigma_n}{\sum_{j=1}^{N} 1/\sigma_j} \qquad (9)$$

In the ERB strategy, the weights of the assets are distributed inversely proportional to their volatilities.

2.3.3 Equal risk contribution (ERC) strategy

The goal of the equal risk contribution (ERC) strategy is to ensure that the contributions of individual assets to the total portfolio risk (σ_p) are equal. This corresponds precisely to the definition of a risk parity strategy introduced in the fundamentals section 2.3.1. As shown above, the absolute risk contribution of an asset n to the total risk can be determined as follows:

$$MRB_n=\frac{(\Sigma w^T)_n}{\sigma_p} \tag{10}$$

The vector of marginal risk contributions $\partial_n\sigma_p$ across all assets is therefore given by:

$$c(w)=\frac{\Sigma w^T}{\sigma_p} \tag{11}$$

Furthermore, it has already been mentioned that the sum of RBs equals the total risk σ_p, and the RB of an asset n corresponds to the MRB weighted by the share w_N. Therefore, if all RBs are to be identical, it must consequently hold that:

$$\frac{\sigma_p}{N}=w_n\times MRB_n \quad \text{for } n=1, ..., N . \tag{12}$$

Thus, the optimal proportions w_n^{ERC} in the ERC portfolio can be determined by numerically minimizing the following objective function:

$$\min_w \sum_{n=1}^{N} \left[\frac{\sigma_p}{N}-w_n\times c(w)_n\right]^2 \tag{13}$$

subject to the constraints

$$\sum_{n=1}^{N} w_n=1 \quad \text{and} \quad 1 \geq w_n \geq 0, \tag{14}$$

where $c(w)_n$ denotes the n-th element of the vector c(w).

In the generalized case of a desired, not necessarily equal target risk allocation w_{target_n}, the function to be minimized is as follows:

$$\underset{w}{\operatorname{argmin}} \sum_{n=1}^{N} \left[w_{target_n} \times \sigma_p - w_n \times c(w)_n\right]^2 \tag{15}$$

Since in the ERC portfolio the absolute risk contributions are identical, this naturally applies to the relative risk contributions (RB/σ_p) of the N individual assets. It follows (see Equation (7) above):

Relative risk contribution in the ERC portfolio

$$= w_n^{ERC} \times \frac{Cov(r_n,\ r_p)}{\sigma_p^2} = w_n^{ERC} \times \beta_n^{ERC} = \frac{1}{N}\ , \tag{16}$$

where β_n^{ERC} denotes the beta factor of the n-th asset with respect to the ERC portfolio. The weight w_n^{ERC} can then also be expressed as follows:

$$w_n^{ERC} = \frac{1}{N \times \beta_n^{ERC}}\ , \tag{17}$$

where for an ERC portfolio holds: $\sum_{n=1}^{N} w_n^{ERC}=1$.

Thus, the allocation of an asset is inversely proportional to its beta factor. The higher the beta, the lower the weighting in the portfolio, and vice versa. Consequently, assets with a high (low) standard deviation or a high (low) correlation with the portfolio return are weighted lower (higher).

2.4 Alternative risk-based approaches

The risk parity approach is distinguished from other methods of asset allocation that also refrain from estimating expected returns. One such method is the global minimum variance approach, which determines the combination of available assets that minimizes risk. Additionally, the maximum diversification approach is considered among risk-based asset allocation concepts. Finally, the equally-weighted strategy, often referred to as a “naïve” approach, is used as a benchmark for comparison.

2.4.1 Equally-weighted (EW) approach

The equally-weighted approach can be considered the simplest form of portfolio diversification, as the weighting of each of the N assets in a portfolio is determined as follows:

$$w_n^{EW} = \frac{1}{N} \tag{18}$$

In this approach, the type of investment or which asset classes are included in the portfolio does not matter. The same applies to the return and risk metrics of the assets included. Therefore, it is advisable to use this strategy at most as a benchmark for assessing the performance of asset allocation strategies. As mentioned above, for example, the ERC approach would only lead to equally-weighted asset allocation in the extreme case of identical correlations between the assets included in the portfolio, as well as equal standard deviations of the assets. However, if the risks in a portfolio are distributed differently, the equally-weighted approach ultimately leads to a concentration of risk, as all assets receive the same weight, i.e., both the asset with the highest and the one with the lowest risk.

2.4.2 Global minimum variance (GMV) approach

The global minimum variance approach arises from portfolio theoretical considerations, aiming to find the efficient portfolio that exhibits the minimum risk. Accordingly, this so-called **global minimum variance portfolio** can be determined by taking the partial derivative of the portfolio variance function with respect to the weights of the individual assets n in the portfolio w_n and subsequently setting it to zero: $\frac{\partial \sigma_p^2}{\partial w_n}$.

- **Case 1: closed-form solution**

The weights of the GMVP can be determined analytically provided there are no constraints in the form of inequalities such as short selling prohibitions or asset limits. The weights of the GMVP then result as the solution to the following minimization problem:

$$\min_{w} w^T \Sigma w \tag{19}$$

subject to the constraints

$$w^T \imath = 1 \ , \tag{20}$$

where $\imath$ denotes the unit vector. The closed-form solution is given by:

$$w^{GMV} = \frac{1}{\imath^T \Sigma^{-1} \imath} \Sigma^{-1} \imath \ . \tag{21}$$

Note that w^{GMV} does not depend on the vector of expected returns μ, i.e., the GMVP can be determined without knowledge of μ.

- **Case 2: numerical solution**

In reality, many institutional investors are often unable to engage in short sales due to investment restrictions. To account for this, the *non-negativity constraint* on the weights must be added to the optimization problem. A closed-form solution is then no longer possible.

$$\min_{w} w^T \Sigma w \tag{22}$$

subject to the constraints

$$w^T \iota = 1 \quad \text{and} \quad 1 \geq w_n \geq 0 \tag{23}$$

From a risk and return perspective, portfolios on the Capital Market Line (CML) are more efficient than the GMVP. For example, if the total portfolio risk is to match that of the GMVP, this can also be achieved through a combination of the market portfolio with the risk-free asset, but here the expected return is higher.

2.4.3 Maximum diversification (MD) approach

To maximize the diversification effect in a portfolio, the so-called **maximum diversification approach** can be utilized, which can also be referred to as the "most-diversified approach". The diversification effect is determined using the diversification ratio (DR) (see Choueifaty et al., 2013):

$$DR = \frac{\sigma^T w}{\sqrt{w^T \Sigma w}} \ . \tag{24}$$

In the numerator of the diversification ratio, the weighted average of the standard deviations of individual assets in the portfolio is provided, while the denominator indicates the standard deviation of the portfolio. Considering a portfolio without the possibility of short sales ("long-only") and if at least one asset in the portfolio has $\sigma_n>0$, then: DR > 1. If all correlations between assets are equal to 1, the numerator and denominator of the diversification ratio would be equivalent. Otherwise, due to the diversification effect, the denominator yields a lower value than the numerator. Thus, the diversification ratio measures the diversification success of assets that are not perfectly correlated. Ultimately, this maximizes the distance between two volatility measures of the same portfolio. The numerator represents the portfolio risk in the absence of diversification, while the denominator indicates the (actual) risk with diversification.

The maximum diversification portfolio (MD) represents the portfolio that maximizes the diversification ratio:

$$\max_{w} \frac{\sigma^T w}{\sqrt{w^T \Sigma w}} . \tag{25}$$

With the budget constraint ($\sum_{n=1}^{N} w_n=1$) as the sole constraint, one obtains the closed-form solution:

$$w^{MD}=\frac{1}{\iota^T \Sigma^{-1} \sigma} \Sigma^{-1} \sigma . \tag{26}$$

Adding further constraints (e.g., position limits, short selling restrictions) requires numerical optimization.

For the resulting portfolio, it can be shown that all assets in the portfolio with a weight of $w_n > 0$ exhibit an identical (positive) correlation with the MD portfolio. In this case, the correlation for a portfolio P with the MD portfolio can be expressed as the quotient of their respective diversification ratios (Choueifaty and Coignard, 2008, p. 42):

$$\mathrm{Corr}(r_P, r_{MD})=\frac{DR_p}{DR_{MD}} , \tag{27}$$

where

DR_p = diversification ratio of portfolio P

DR_{MD} = diversification ratio of the MD portfolio

If a portfolio consists of only one asset n, its diversification ratio is 1. Therefore, in this case, for all portfolios (or assets n) in the maximum diversification portfolio, the following applies:

$$Corr(r_n, r_{MD}) = \frac{1}{DR_{MD}} \ , \quad \text{if } w_n > 0 \tag{28}$$

For w_n=0 follows

$$Corr(r_n, r_{MD}) \geq \frac{1}{DR_{MD}} \ . \tag{29}$$

As shown above, the correlation between an asset n and a portfolio P can also be expressed as follows:

$$MRB_n = Corr\left(r_n, r_p\right) \times \sigma_n \ \Leftrightarrow \ Corr\left(r_n, r_p\right) = \frac{MRB_n}{\sigma_n} = \frac{1}{\sigma_n} \times \frac{\partial \sigma_p}{\partial w_n} \tag{30}$$

Thus, for all assets i and j in the MD portfolio:

$$\frac{1}{\sigma_i} \times \frac{\partial \sigma_{MD}}{\partial w_i} = \frac{1}{\sigma_j} \times \frac{\partial \sigma_{MD}}{\partial w_j} \tag{31}$$

The quotient of MRB_n and σ_n can be interpreted as “relative marginal volatility” or “scaled marginal volatility” (see Demey et al., 2010, p. 14 and Roncalli, 2014, p. 173).

3 Empirical results

3.1 Data and methodology

The investment universe chosen for this study comprises the 49 industry portfolios formed by Fama and French (1997). The industry portfolios were constructed based on the industry classification of companies listed on the NYSE, AMEX, and NASDAQ stock markets using standard industrial classification (SIC) codes (Fama and French, 1997).[1]

The choice of industry portfolios is justified by several reasons. Firstly, these portfolios offer an extensive price history spanning over 96 years. This extended observation period allows for the examination of multiple macroeconomic phases and crises, enabling more robust conclusions about the efficacy of risk-based approaches, particularly during negative market phases, compared to a shorter time frame. Additionally, the selection of industry portfolios as the stock universe provides better direct and indirect diversification opportunities compared to a selection of individual stocks or indices. Manual selection of individual stocks post facto could be susceptible to survivorship bias, aside from the challenges in acquiring such a long time series of individual stock returns.

For assessing the potential for risk reduction, often quantified using symmetric risk measures, an approximately normal distribution of daily returns is required for the validity of such measures. The industry portfolios, comprising an average of 60 individual stocks per portfolio over the entire time period, align with the central limit theorem, allowing for an acceptable assumption of approximate normality for the sum (portfolio) of multiple return distributions, regardless of their individual distribution characteristics (Fama and French, 1997). The use of daily returns, which mostly cluster around zero on average, results in minimal differences between continuous and discrete returns, thus justifying the assumption of approximate time additivity for discrete returns and portfolio additivity for continuous returns.

Due to these advantages, the daily discrete returns of the 49 industry portfolios by Fama and French (1997) are utilized as the data basis for comparing the methods and assessing performance. A calendar year rebalancing frequency has been chosen for

[1] The data on the 49 time series of daily value-weighted industry portfolio returns is downloaded from the data library of Kenneth French (https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html#Research).

the respective optimization strategies. This selection is justified by the findings of previous studies, which demonstrate that pension funds, in practice, only partially engage in regular rebalancing or prefer ad-hoc methods (e.g., calendar-based or tolerance band-based) for rebalancing (e.g., Bikker et al., 2018; Sun et al., 2006).

To optimize the previously introduced risk-based approaches, it is necessary to select an appropriate risk measure, at least for a subset of them. Risk parity strategies, for instance, are typically optimized based on portfolio standard deviation but are fundamentally independent of it (Maillard et al., 2010). Therefore, it is essential to define the risk concept for this study. Risk concepts can be categorized into symmetric and asymmetric risk measures. Symmetric risk measures capture both negative and positive deviations as risk. Intuitively, this may seem contradictory as outperforming a mean return is seen as a desirable opportunity. However, in cases where the return distribution is symmetric, employing a symmetric risk measure can be justified. In such situations, increasing the probability of exceeding the mean return would correspond equally to increasing the probability of falling below it and thus represent risk equivalently (Lee, 2011).

Considering the daily return distributions of the individual industry portfolios in this study (see Figure A7 in the Appendix for industry-level return box plots), it is reasonable to assume an approximate normal distribution. Therefore, the optimizations will be performed based on the portfolio return standard deviation. For the sake of completeness, the evaluation will also include the use of competing asymmetric risk measures, aiming to gain a better understanding of the portfolio risk behavior, especially during negative market phases.

3.2 Identification of downward markets

To identify negative market phases, the (equally-weighted combination of the) returns of the industry portfolios themselves are used. This is justified by the aim of investigating whether risk-based approaches perform better than the benchmark during such phases. In assessing the risk and return behavior during volatile periods, it is crucial that the underlying return values used for determining portfolio weights experienced more pronounced downward deviations. Consequently, macroeconomic crises during which the market moved in a decoupled manner are excluded from consideration and remain unaccounted for. The relevant crises for the portfolios were delimited through the analysis of rolling annualized returns and rolling semi-annual annualized volatilities. See Figures A5 and A6 in the Appendix for a graphical depiction of the identified downward markets, together with the corresponding rolling mean returns and volatilities of the equally-weighted benchmark portfolio. The downward market episodes are labelled by their underlying characteristic event.

4 Empirical examination of the performance of risk-based asset allocation approaches

The following section will first discuss the performance of the previously introduced risk-based approaches over the entire observation period from July 1926 to October 2022. Following that, the performance of these methods will be specifically examined during negative market phases to determine whether risk-controlled optimization can more effectively reduce the increased market risks associated with these phases. A balanced weighting of individual industry portfolios is used as a benchmark for the ERB, ERC, GMVP, and MD portfolios.

4.1 Performance over the entire observation period

When comparing the annualized average daily returns over the entire 96-year period (07/1926-10/2022), presented in the first row of Table 1, it becomes evident that the average return of the equally-weighted benchmark portfolio significantly surpasses that of the GMVP. In contrast, the ERB and ERC portfolios exhibit returns only marginally below that of the benchmark. Only the maximum diversification portfolio managed to outperform the equally-weighted portfolio with a slightly higher absolute return. However, evaluating the strategies without considering the inherent risk provides only limited insights into their performance.

When considering volatility, which was used as the risk measure for portfolio optimization, it can be observed that all risk-based approaches exhibit a superior risk-return tradeoff, as expressed by the Sharpe ratio. This difference is particularly significant for the GMVP and MD portfolios. In the case of the MD portfolio, this is attributed to a slightly higher return accompanied by a slightly reduced volatility compared to the benchmark. For the GMVP, on the other hand, the better relative performance is mainly attributed to a substantial reduction in risk, with an almost 5 percentage point decrease in average annualized volatility, albeit at the expense of a slight reduction in return by 2.5 percentage points. As for the ERB and ERC portfolios, they lie within the range between the equally-weighted portfolio and the GMVP, a result that aligns with the findings of previous studies (Maillard et al., 2010).

Table 1: Full sample performance and risk statistics for five risk-based asset allocation strategies

This table contains full sample results of selected performance and risk measures for the five risk-based asset allocation approaches, 1/N, ERB, ERC, GMVP, and MD. The asset universe is composed of the 49 US industry portfolios of Fama and French (1997). The sample period comprises daily data points from July 1926 to October 2022. All five strategies are rebalanced yearly. All performance and risk measures are annualized.

Annualized full sample results (1926-2022)	1/n	ERB	ERC	GMVP	MD
Mean return	0.120	0.116	0.118	0.096	0.126
Volatility	0.171	0.161	0.157	0.118	0.158
Sharpe ratio	0.706	0.723	0.753	0.815	0.794
Semi standard deviation	0.141	0.134	0.131	0.100	0.126
Sortino ratio	0.851	0.866	0.900	0.963	0.997
Value at Risk	0.273	0.257	0.250	0.188	0.253

Assessing the portfolios based on the semi standard deviation, which considers only deviations in the negative domain, a more favorable picture emerges for the MD portfolio. While the evaluation of the GMVP, ERC, and ERB portfolios remains unchanged, the maximum diversification portfolio exhibits a significantly larger gap from the benchmark according to the Sortino ratio. This is due to achieving a 1.5-percentage points lower annualized semi standard deviation relative to a similar return, making the MD portfolio the best-performing strategy for the period based on the Sortino ratio. However, when considering the value at risk factor (at the 5% level), the GMVP is preferred in the overall assessment.

4.2 Return and risk behavior in negative market phases

In the following assessment of the risk and return behavior of risk-based approaches in negative market phases, both the annualized returns for the periods as well as the annualized value at risk and the maximum drawdown for the individual periods are taken into consideration. The use of performance measures, such as the Sharpe ratio previously used, is deliberately avoided here, as they typically result in a negative ratio due to predominantly negative returns during bearish phases. Negative Sharpe and Sortino ratios are then criticized for being no longer interpretable. In the case of identical negative returns, according to these measures, the investment with higher risk would always be preferred, which seems counterintuitive from a risk reduction perspective.

Table 2: Comparison of annualized discrete mean returns during negative market phases

This table reports mean returns for the five risk-based asset allocation approaches, 1/N, ERB, ERC, GMVP, and MD, during downward market periods. The asset universe is composed of the 49 US industry portfolios of Fama and French (1997). The negative market phases are described in Section 3.2. All five strategies are rebalanced yearly. Mean returns are annualized.

	1/n	ERB	ERC	GMVP	MD
Annualized mean return in downward markets					
Covid-19 Pandemic (Jan. 2020 - May 2020)	-0.363	-0.324	-0.235	-0.185	-0.117
Global Financial Crisis (Oct. 2007 - April 2009)	-0.358	-0.355	-0.343	-0.217	-0.192
Dotcom Bubble (April 2000 - Oct. 2002)	-0.023	-0.009	0.016	0.102	0.097
Black Monday (Oct. 1987 - April 1988)	-0.086	-0.081	-0.084	-0.063	-0.169
Recession (Jan. 1973 - Oct. 1974)	-0.367	-0.351	-0.335	-0.254	-0.101
Recession (Oct. 1968 - June 1970)	-0.247	-0.214	-0.205	-0.109	-0.161
Kennedy Slide (Dec. 1961 - Dec. 1962)	-0.154	-0.139	-0.143	-0.067	-0.162
Market Crash (May 1946 - June 1949)	-0.048	-0.039	-0.037	-0.038	-0.019
Recession (May 1937 - August 1938)	-0.083	-0.075	-0.063	-0.020	0.008
Black Friday (Oct. 1929 - July 1932)	-0.482	-0.475	-0.469	-0.388	-0.473

When considering the absolute annualized returns (Table 2) during negative market phases, a similar pattern emerges as in the overall analysis. Here too, all risk-based approaches were able to achieve significantly lower negative annualized returns. During the recent market crisis, the Covid-19 pandemic, both the GMVP and MD strategies achieved significantly lower negative annualized returns (17 and 24 percentage points, respectively) compared to an equally-weighted portfolio. This pattern repeats for the two portfolios for the majority of the included market phases. However, the ERC and ERB portfolios were unable to achieve significant deviations in returns from the benchmark, with the Covid-19 pandemic being the only exception.

When considering the annualized value at risk (at the 5% level) as a measure to quantify risk, the previous observations are also repeated here (see Table 3). Thus, the global minimal variance and maximum diversification portfolios can also exhibit significantly better (i.e., lower) values according to this risk measure. However, ERC and ERB achieved annualized value at risk figures comparable to the benchmark in individual periods. It is worth noting, however, that the ERC portfolio consistently outperforms an equally-weighted portfolio by 3 percentage points in the majority of phases.

Table 3: Comparison of annualized value at risk numbers during negative market phases

This table reports value at risk figures (i.e., the lower 5% percentile of the conditional strategy return distribution) for the five risk-based asset allocation approaches, 1/N, ERB, ERC, GMVP, and MD, during downward market periods. The asset universe is composed of the 49 US industry portfolios of Fama and French (1997). The negative market phases are described in Section 3.2. All five strategies are rebalanced yearly. VaR numbers are annualized.

Annualized VaR in downward markets	1/n	ERB	ERC	GMVP	MD
Covid-19 Pandemic (Jan. 2020 - May 2020)	0.931	0.915	0.874	0.829	0.797
Global Financial Crisis (Oct. 2007 - April 2009)	0.668	0.628	0.623	0.449	0.568
Dotcom Bubble (April 2000 - Oct. 2002)	0.305	0.288	0.269	0.198	0.241
Black Monday (Oct. 1987 - April 1988)	0.476	0.472	0.458	0.344	0.397
Recession (Jan. 1973 - Oct. 1974)	0.317	0.304	0.301	0.211	0.297
Recession (Oct. 1968 - June 1970)	0.265	0.242	0.238	0.166	0.243
Kennedy Slide (Dec. 1961 - Dec. 1962)	0.283	0.270	0.268	0.237	0.259
Market Crash (May 1946 - June 1949)	0.242	0.225	0.219	0.136	0.191
Recession (May 1937 - August 1938)	0.497	0.443	0.431	0.261	0.403
Black Friday (Oct. 1929 - July 1932)	0.553	0.531	0.502	0.410	0.461

The global minimum variance portfolio was able to achieve values lower by 10 percentage points in most phases, even when considering maximum losses (i.e., the maximum drawdown, see Table 4). While it could be argued here that a portfolio with low symmetric volatility tends to not only achieve lower lows but also lower highs, considering the results from Table 2, the assessment of the GMVP as less risky in terms of the risk-return ratio might not be unwarranted here as well. In contrast to the previous measures, the maximum diversification portfolio does not closely follow the GMVP and tends to perform worse. ERC and ERB achieve similar values to the equally-weighted portfolio.

Table 4: Comparison of maximum drawdowns of risk-based approaches during different market crises

This table reports the maximum drawdown for the five risk-based asset allocation approaches, 1/N, ERB, ERC, GMVP, and MD, during downward market periods. The asset universe is composed of the 49 US industry portfolios of Fama and French (1997). The negative market phases are described in Section 3.2. All five strategies are rebalanced yearly. Maximum drawdowns are annualized.

Maximum Drawdown in downward markets	1/n	ERB	ERC	GMVP	MD
Covid-19 Pandemic (Jan. 2020 - May 2020)	-0.383	-0.376	-0.358	-0.333	-0.307
Global Financial Crisis (Oct. 2007 - April 2009)	-0.566	-0.553	-0.545	-0.386	-0.489
Dotcom Bubble (April 2000 - Oct. 2002)	-0.252	-0.245	-0.239	-0.213	-0.244
Black Monday (Oct. 1987 - April 1988)	-0.340	-0.338	-0.336	-0.267	-0.318
Recession (Jan. 1973 - Oct. 1974)	-0.496	-0.481	-0.466	-0.400	-0.336
Recession (Oct. 1968 - June 1970)	-0.462	-0.427	-0.420	-0.281	-0.383
Kennedy Slide (Dec. 1961 - Dec. 1962)	-0.308	-0.297	-0.297	-0.254	-0.308
Market Crash (May 1946 - June 1949)	-0.324	-0.300	-0.297	-0.196	-0.263
Recession (May 1937 - August 1938)	-0.481	-0.453	-0.443	-0.311	-0.396
Black Friday (Oct. 1929 - July 1932)	-0.832	-0.826	-0.819	-0.761	-0.824

4.3 Time series of portfolio constituents' weights

The dynamic changes and differences in weighting of the 49 industry portfolios between the discussed methods are difficult to represent and describe punctually over such a long period as almost 97 years. Nevertheless, when examining Figures A1, A2, A3 and A4 in the Appendix, some clear differences in weighting can be observed. The figures depict, for each of the four strategies separately, the weights for the 49 industry portfolios using a stacked bar chart, which highlights rough differences.

On the one hand, in direct comparison between the GMVP and ERB portfolios, the theoretically possible extreme concentration is clearly evident in the analysis for GMVP. For example, the weights in the middle of the 20th century for the GMVP are mainly concentrated on commodity markets, which were likely less volatile at the time (oil, coal, mining, and gold), while in the first decade of this century, consumer goods were more heavily weighted. It is also noticeable that the weights of the GMVP show little consistency, with drastic corrections occurring frequently.

In comparison, Figure A4 shows that the MD portfolio is much better diversified in terms of weighting. However, looking at the weights of the ERB and ERC portfolios in Figures A1 and A2, the MD portfolio is poorly diversified in terms of portfolio concentration. Similarly, in terms of the consistency of weighting, the MD portfolio regularly makes extreme adjustments to the weights. For this reason, the GMVP and MD portfolios are seen as more critical in practical application, as they offer little weight-based diversification and consistency for institutional investors, which can make implementation and justification in practice more difficult.

4.4 Limits of the study

The increasing popularity of risk-based allocation strategies among institutional investors following the financial crisis is likely to continue in the future. Further research could shed light on the performance of the methods, particularly considering transaction costs. It should be noted that the approaches were evaluated only in relation to a US stock universe, limiting the scope of the presented analysis. Geographically diversified investment universes could further improve results in favor of risk-based approaches through greater diversification effects.

Furthermore, no weighting constraints were imposed in the analysis. The inclusion of such constraints in optimization can significantly alter the findings and is not uncommon in practice. As observed in the analysis of weights, these constraints could strongly affect the GMVP and MD methods, which turn out as the most effective in minimizing risk in this study. Weighting constraints could lead to a significant increase in risk or reduction in returns for GMVP and MD, aligning these approches more closely with ERC, ERB, and the benchmark. Therefore, analyzing the sensitivity of the methods and results to the setting of weighting constraints would be the next step to provide a more comprehensive answer to the posed hypotheses for practical application.

In addition, studying the performance of risk parity concepts using other risk concepts or rebalancing strategies could further differentiate any potential advantage.

5. Conclusion

This study aimed to investigate whether risk-based approaches to asset allocation could provide better protection against price losses during volatile market phases. Based on the findings obtained from an extensive backtest using the Fama and French (1997) industry portfolios and almost 97 years of daily return data, this hypothesis can be partially supported. Considering the risk measures determined during the backtesting period, a risk protection characteristic can be attributed especially to the GMVP and MD portfolios. These investment strategies were able to achieve significantly better or equivalent (to an equally-weighted industry benchmark) results in terms of both return and risk. However, the significant reduction in risk during market crises is not self-evident, as crises are mainly driven by systematic risks that are difficult to diversify.

When considering the market phases alongside a holistic view, a generally better performance compared to the benchmark can be observed for all portfolios. The ERB and ERC strategies cannot be attributed to significantly better protection against negative market phases. Only the ERC can be evaluated as the better investment strategy among the examined risk parity approaches, which is not surprising considering its conceptual consideration of correlations.

In terms of weight analysis, however, the ERB and ERC portfolios must be highlighted for implementing diversification in terms of portfolio concentration more effectively. On the other hand, the GMVP and MD portfolios achieved the mentioned risk reduction mainly through extreme weighting of low-volatile industries. However, such extreme weights are difficult to implement and justify in practice, especially for institutional investors, as mentioned earlier. Nevertheless, the results can be interpreted as an indication of the effectiveness of such methods in reducing market risks.

Literature

Benham, Frank, Roberto Obregon, and Timur Kaya Yantar. 2019. "Risk Parity." *Alternative Investment Analyst Review* 8 (2): 7-16. https://caia.org/aiar/3968.

Bikker, Jacob A., Dirk W. G. A. Broeders, and Dirk Jan de Dreu. 2018. "Stock Market Performance and Pension Fund Investment Policy: Rebalancing, Free Float, or Market Timing?" *21st Issue (June 2010) of the International Journal of Central Banking*, August. https://www.ijcb.org/journal/ijcb10q2a3.htm.

Black, Fischer, and Robert Litterman. 1992. "Global Portfolio Optimization." *Financial Analysts Journal* 48 (5): 28-43. https://doi.org/10.2469/faj.v48.n5.28.

Choueifaty, Yves, and Yves Coignard. 2008. "Toward Maximum Diversification." *Journal of Portfolio Management* (Fall 2008): 40-51. https://doi.org/10.3905/JPM.2008.35.1.40 .

Choueifaty, Yves, Tristan Froidure, and Julien Reynier. 2011. "Properties of the Most Diversified Portfolio." *Journal of Investment Strategies* 2 (Spring 2013): 49-70. https://doi.org/10.2139/ssrn.1895459.

Clarke, Roger, Harindra de Silva, and Steven Thorley. 2013. "Risk Parity, Maximum Diversification, and Minimum Variance: An Analytic Perspective." *The Journal of Portfolio Management* 39 (3): 39-53. https://doi.org/10.3905/jpm.2013.39.3.039.

Demey, Paul, Sebastien Maillard, and Thierry Roncalli. 2010. "Risk-based Indexation." Working paper, Lyxor Asset Management and Amundi Asset Management. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1582998

Fama, Eugene F., and Kenneth R. French. 1997. "Industry Costs of Equity." *Journal of Financial Economics* 43 (2): 153-93. https://doi.org/10.1016/S0304-405X(96)00896-3.

Lee, Wai. 2011. “Risk-Based Asset Allocation: A New Answer to an Old Question?” *The Journal of Portfolio Management* 37 (4): 11-28. https://doi.org/10.3905/jpm.2011.37.4.011.

Maillard, Sebastien, Thierry Roncalli, and Jerome Teiletche. 2010. “The Properties of Equally Weighted Risk Contribution Portfolios.” *The Journal of Portfolio Management* 36 (4): 60-70. https://doi.org/10.3905/jpm.2010.36.4.060.

Markowitz, Harry M. 1952. “Portfolio Selection.” *The Journal of Finance* 7 (2): 77-91. https://doi/abs/10.1111/j.1540-6261.1952.tb01525.x.

Merton, Robert C. 1980. “On Estimating the Expected Return on the Market: An Exploratory Investigation.” *Journal of Financial Economics* 8 (3): 323-361. https://doi/10.3386/w0444.

Molk, Peter, and Frank Partnoy. 2019. “Institutional Investors as Short Sellers?” SSRN Scholarly Paper No. 19-23. Rochester, NY: University of Florida Levin College of Law. https://papers.ssrn.com/abstract=3309731.

Qian, Edward. 2013. “Are Risk-Parity Managers at Risk Parity?” *The Journal of Portfolio Management* 40 (1): 20-26. https://doi.org/10.3905/jpm.2013.40.1.020.

Roncalli, T. 2014. “Introduction to Risk Parity and Budgeting.” Boca Ration, Florida. https://arxiv.org/abs/1403.1889.

Sun, Walter, Ayres Fan, Li-Wei Chen, Tom Schouwenaars, and Marius A. Albota. 2006. “Optimal Rebalancing for Institutional Portfolios.” *The Journal of Portfolio Management* 32 (2): 33-43. https://doi.org/10.3905/jpm.2006.611801.

A Appendix

A.1 Development of industry portfolio weights

Below, the annual weight distributions over the entire period are shown according to the individual optimization methods.

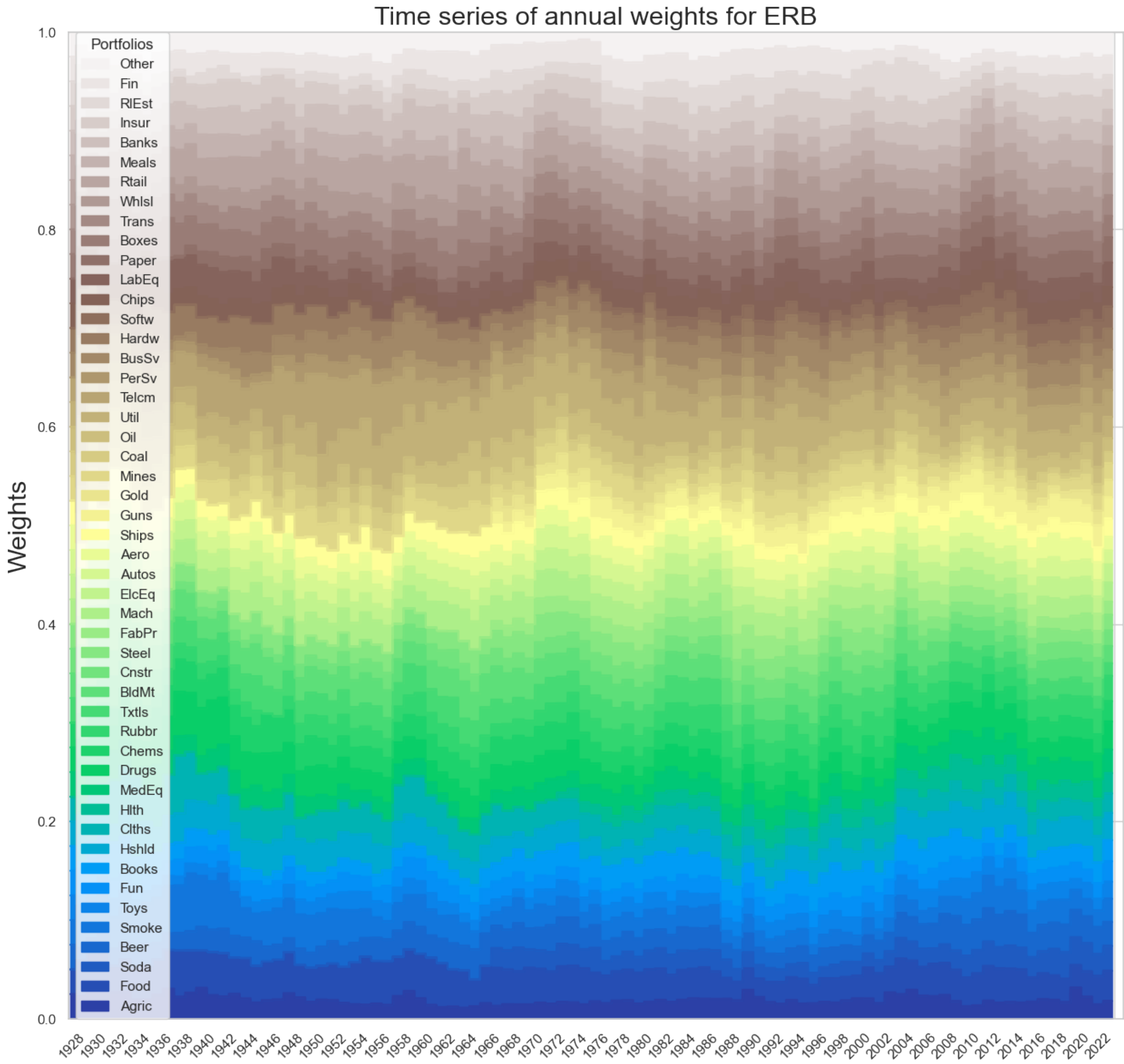
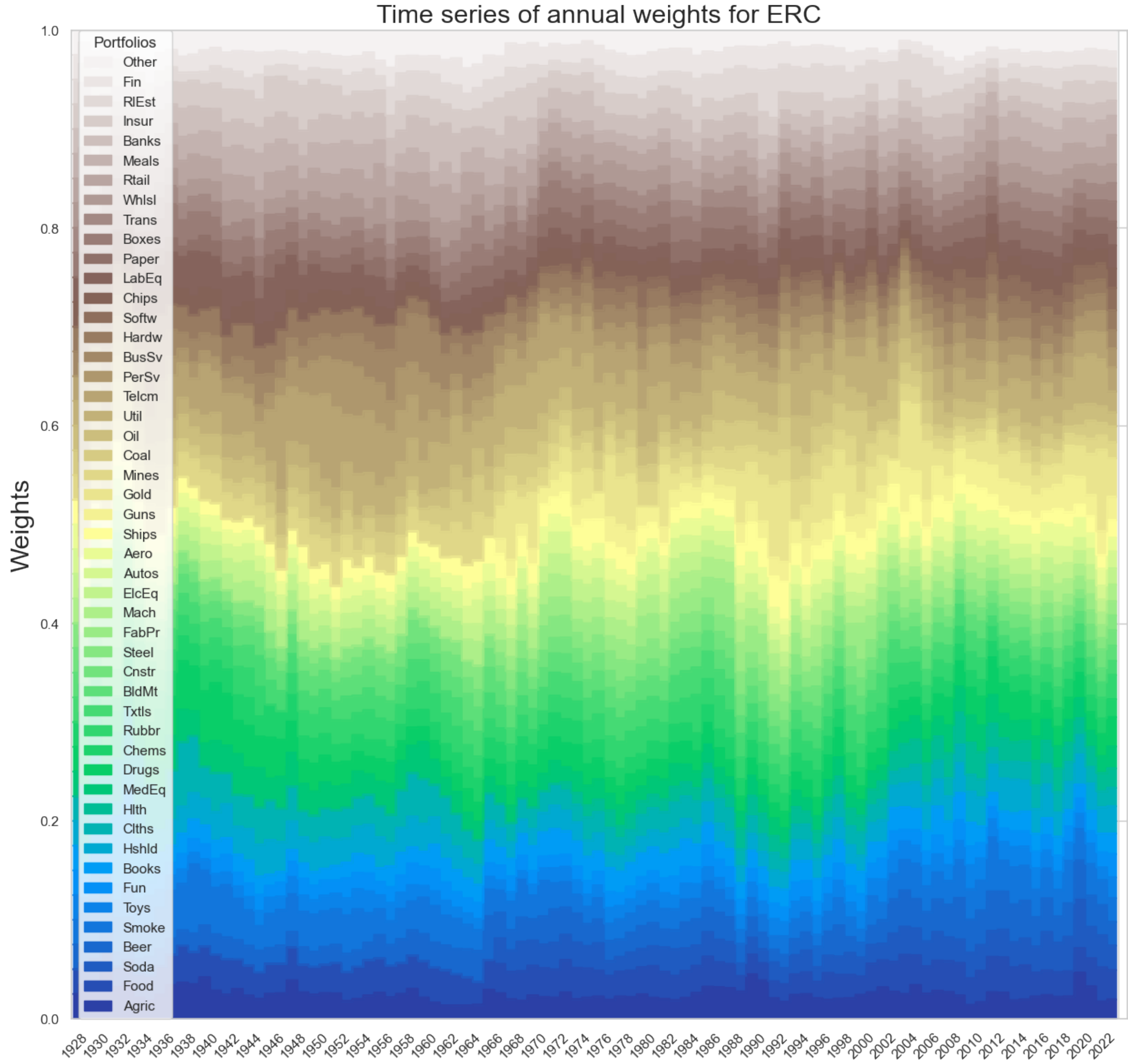


Figure A1



Time series of annual weights for ERC
Weights
Portfolios
Other
Fin
RlEst
Insur
Banks
Meals
Rtail
Whlsl
Trans
Boxes
Paper
LabEq
Chips
Softw
Hardw
BusSv
PerSv
Telcm
Util
Oil
Coal
Mines
Gold
Guns
Ships
Aero
Autos
ElcEq
Mach
FabPr
Steel
Cnstr
BldMt
Txtls
Rubbr
Chems
Drugs
MedEq
Hlth
Clths
Hshld
Books
Fun
Toys
Smoke
Beer
Soda
Food
Agric

Figure A2

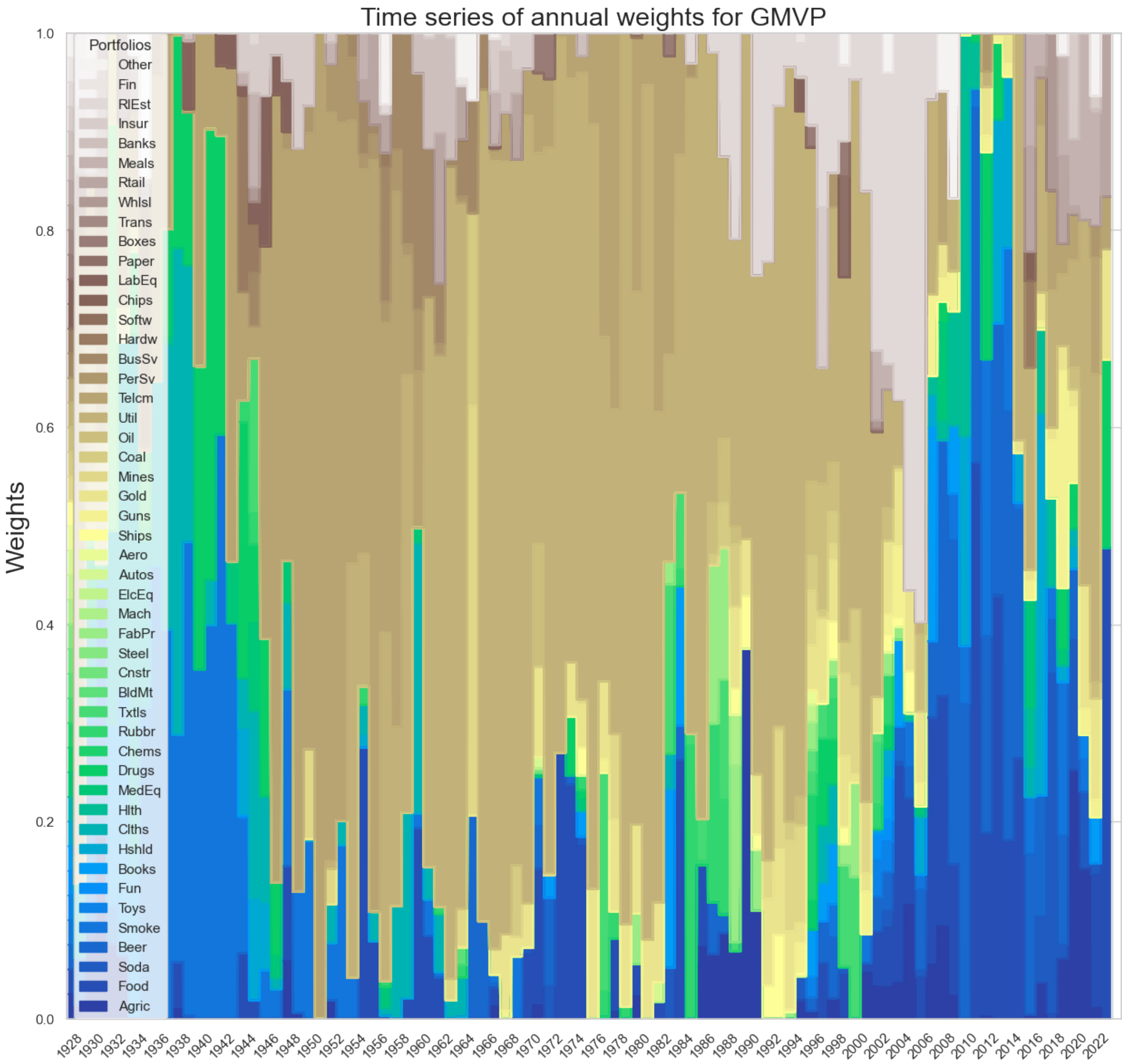
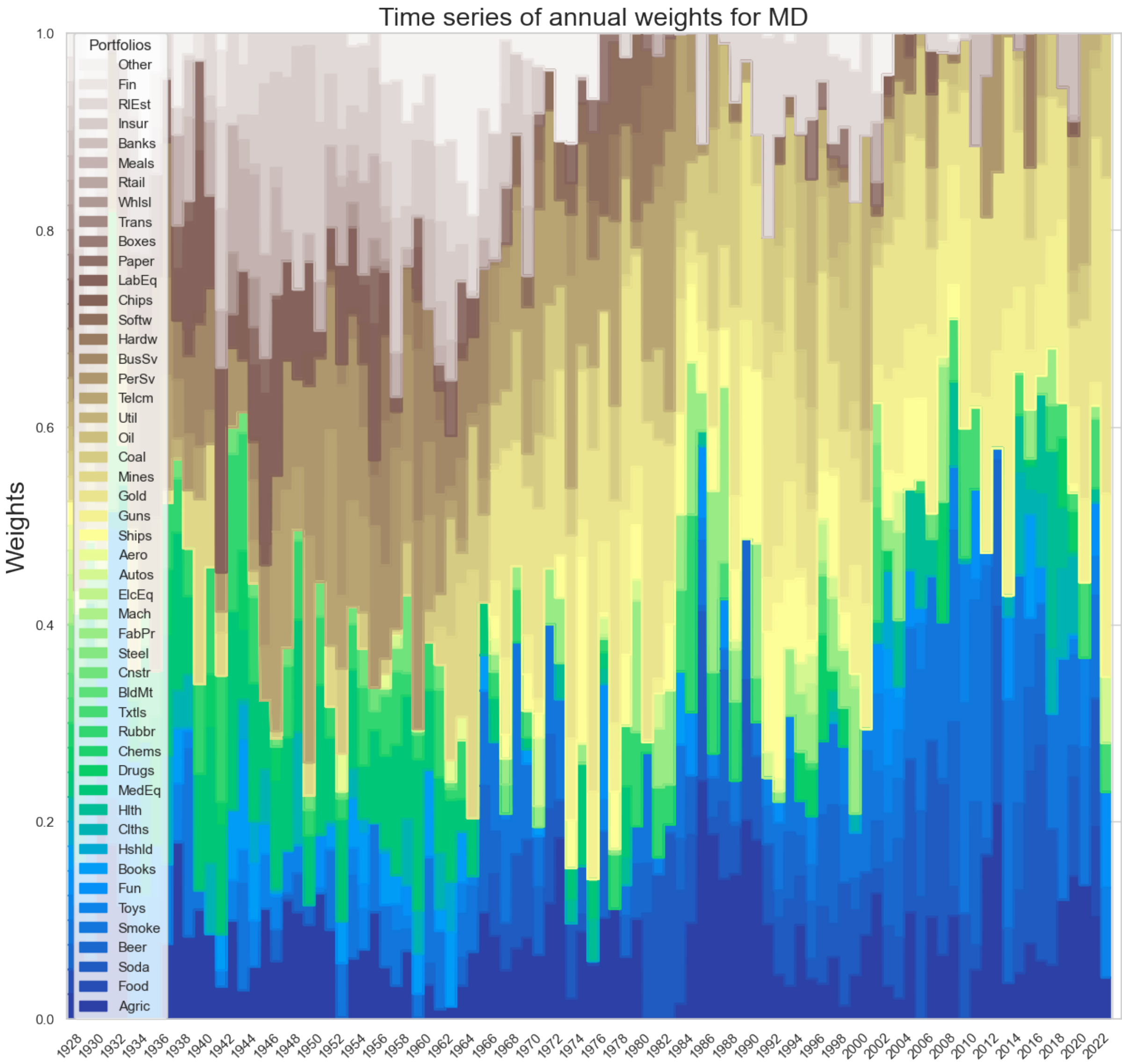

Time series of annual weights for GMVP
Weights
Portfolios
Other
Fin
RlEst
Insur
Banks
Meals
Rtail
Whlsl
Trans
Boxes
Paper
LabEq
Chips
Softw
Hardw
BusSv
PerSv
Telcm
Util
Oil
Coal
Mines
Gold
Guns
Ships
Aero
Autos
ElcEq
Mach
FabPr
Steel
Cnstr
BldMt
Txtls
Rubbr
Chems
Drugs
MedEq
Hlth
Clths
Hshld
Books
Fun
Toys
Smoke
Beer
Soda
Food
Agric
1.0
0.8
0.6
0.4
0.2
0.0
1928
1930
1932
1934
1936
1938
1940
1942
1944
1946
1948
1950
1952
1954
1956
1958
1960
1962
1964
1966
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022

Figure A3



Time series of annual weights for MD
Weights
Portfolios
Other
Fin
RlEst
Insur
Banks
Meals
Rtail
Whlsl
Trans
Boxes
Paper
LabEq
Chips
Softw
Hardw
BusSv
PerSv
Telcm
Util
Oil
Coal
Mines
Gold
Guns
Ships
Aero
Autos
ElcEq
Mach
FabPr
Steel
Cnstr
BldMt
Txtls
Rubbr
Chems
Drugs
MedEq
Hlth
Clths
Hshld
Books
Fun
Toys
Smoke
Beer
Soda
Food
Agric

Figure A4

A.2 Rolling annualized returns and volatility of the equally-weighted benchmark portfolio strategy

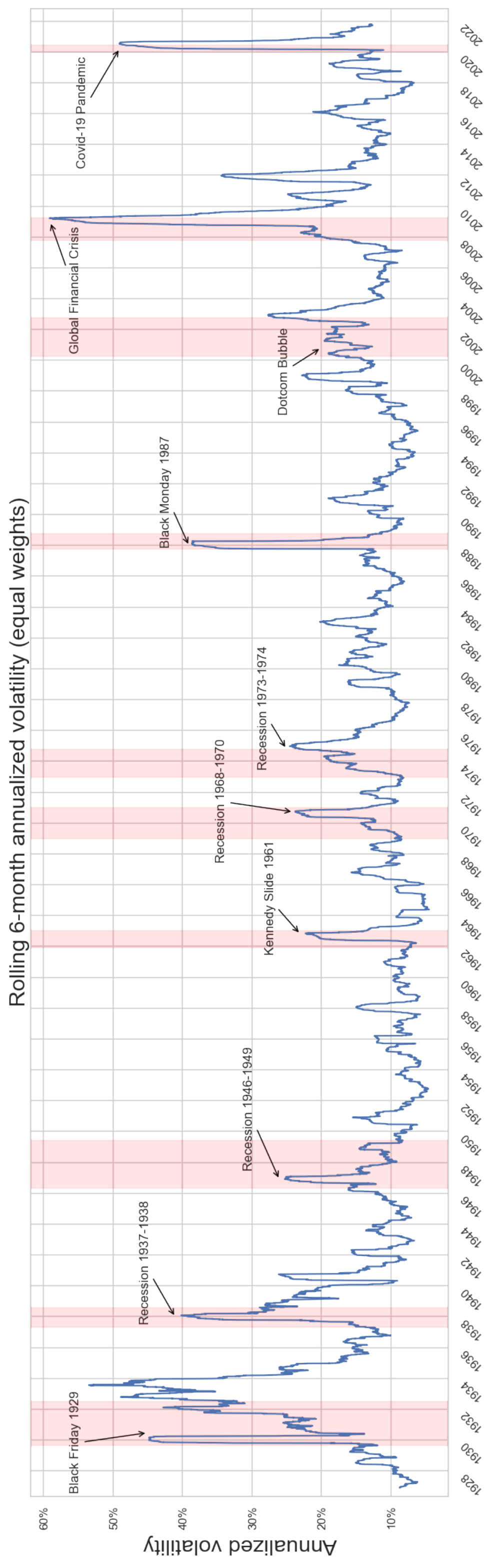
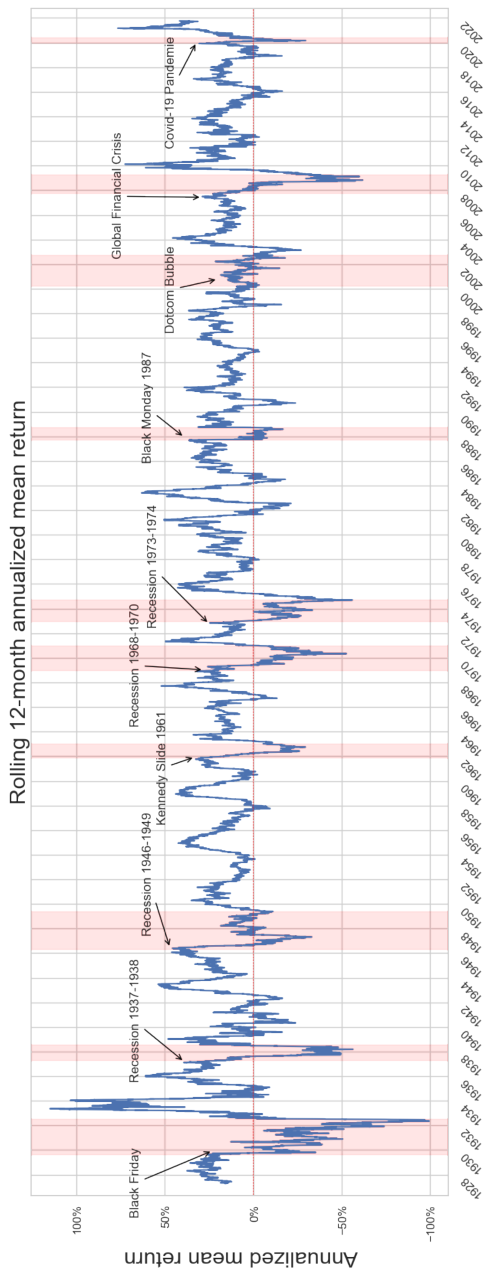


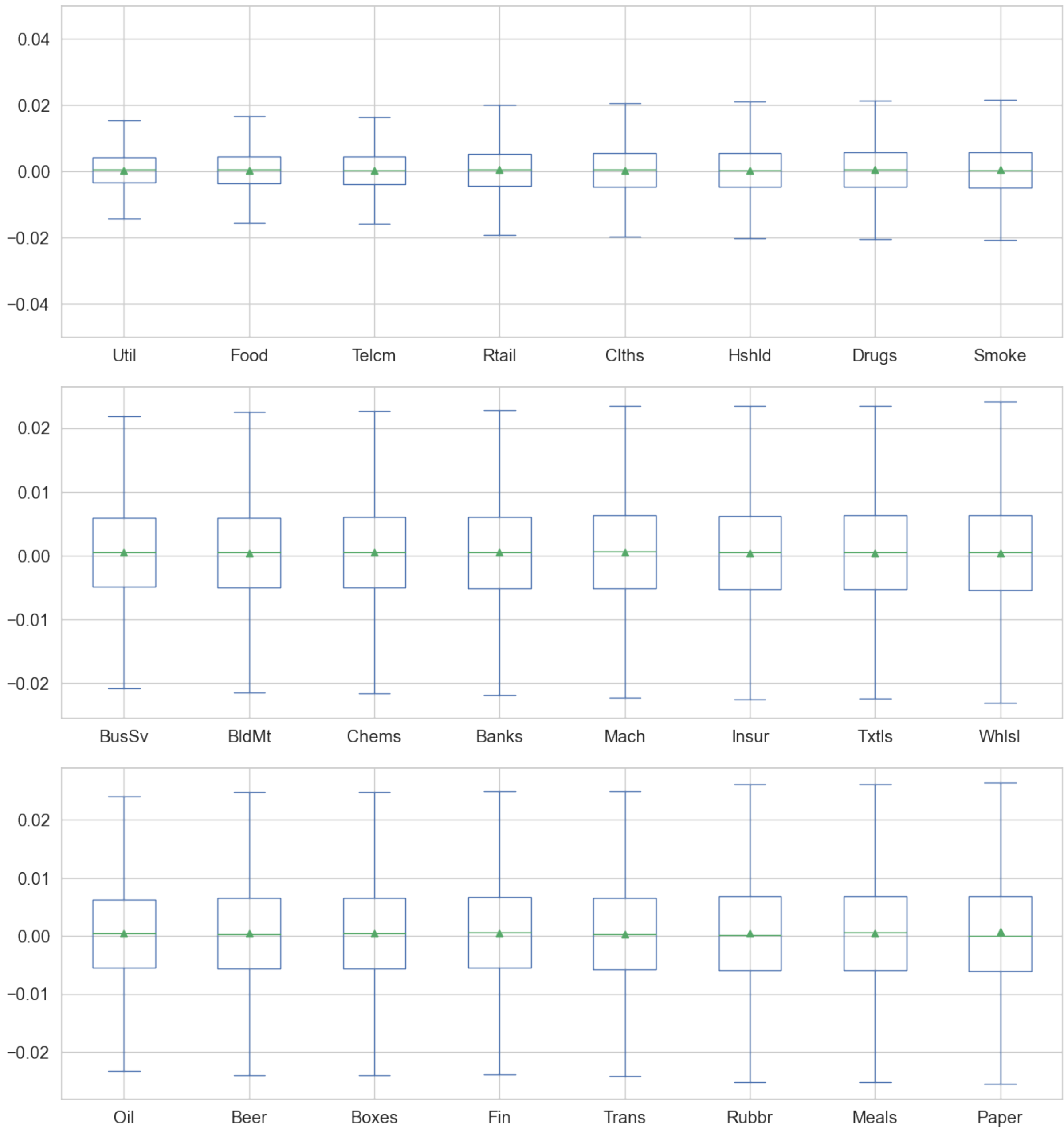
Figure A5



Rolling 12-month annualized mean return
Annualized mean return
100%
50%
0%
-50%
-100%
Black Friday
Recession 1937-1938
Recession 1946-1949
Kennedy Slide 1961
Recession 1968-1970
Recession 1973-1974
Black Monday 1987
Dotcom Bubble
Global Financial Crisis
Covid-19 Pandemie
1928
1930
1932
1934
1936
1938
1940
1942
1944
1946
1948
1950
1952
1954
1956
1958
1960
1962
1964
1966
1968
1970
1972
1974
1976
1978
1980
1982
1984
1986
1988
1990
1992
1994
1996
1998
2000
2002
2004
2006
2008
2010
2012
2014
2016
2018
2020
2022

Figure A6

A.3 Box plots of the daily return distributions of the 49 industry portfolios (excluding outliers)



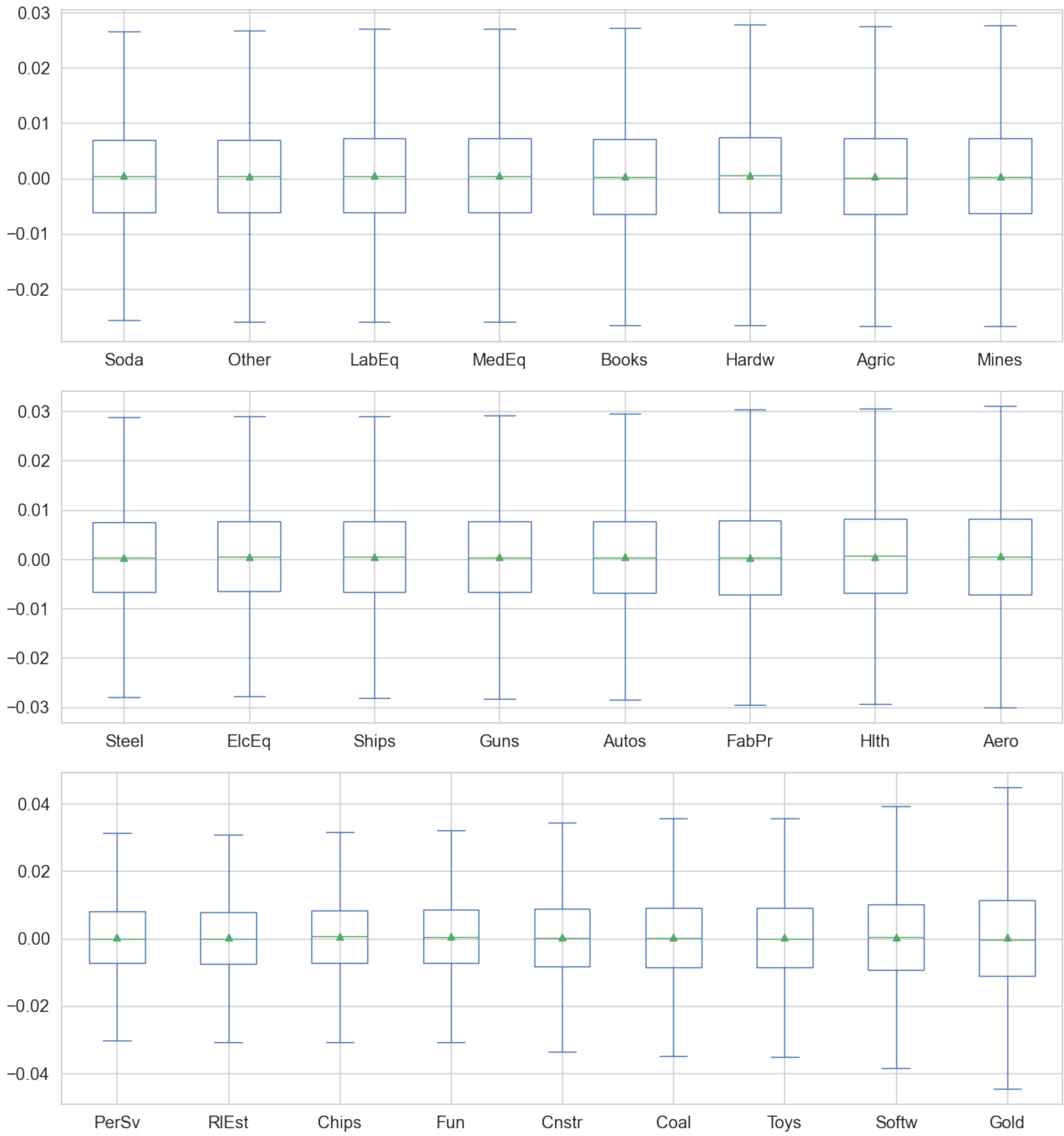

0.03
0.02
0.01
0.00
−0.01
−0.02
Soda
Other
LabEq
MedEq
Books
Hardw
Agric
Mines
0.03
0.02
0.01
0.00
−0.01
−0.02
−0.03
Steel
ElcEq
Ships
Guns
Autos
FabPr
Hlth
Aero
0.04
0.02
0.00
−0.02
−0.04
PerSv
RlEst
Chips
Fun
Cnstr
Coal
Toys
Softw
Gold

Figure A7

Section C – Paper 2

A performance "horse race": Does anything beat the 1/N portfolio?

Christian Cara

This version: August 2024

Abstract

In a widely cited paper, DeMiguel et al. (2009a) find that (μ-σ) optimized portfolios do not necessarily outperform simple portfolio construction approaches in out-of-sample tests. This paper conducts a backtest of the minimum variance portfolio, the maximum Sharpe portfolio, and a "naïve" (equally weighted) portfolio. The test assets consist of the monthly returns of the Fama and French 10, 30, and 48 industry portfolios. The Sharpe ratio and CAPM alpha are used to measure risk-adjusted performance. As its main result, the paper is able to confirm the key findings of DeMiguel et al. (2009a) using a more extensive and up-to-date dataset.

Keywords: naïve diversification; portfolio construction; (μ-σ) optimization;

1 Introduction

1.1 Problem statement

In their seminal paper from 2009, the authors DeMiguel, Garlappi, and Uppal (DGU in what follows) analyzed various approaches to the classical return-risk (μ-σ) portfolios, as influenced by Markowitz (1952), in comparison to a naïve portfolio.[2] In a naïve portfolio, the weights are not optimized but are instead equally distributed across all assets at each rebalancing date (1/N). The return-variance portfolios were consistently calculated based on data from the pre-evaluation period (in-sample) and then applied to the subsequent period (out-of-sample). The results of the study indicate that none of the various versions of the (μ-σ) optimized portfolios managed to consistently outperform a naïve equally-weighted portfolio across different datasets during the evaluation period.

Kritzman et al. (2010) criticize the DGU study, arguing that the in-sample period was not long enough, resulting in skewed outcomes, and that in general, a weight optimization is superior to naïve weight distribution. This study addresses the questions, using longer datasets than those employed by DGU, of whether a naïve portfolio allocation is risk-adjusted superior to various optimization methods over a long, non-simulated period, and whether increasing the available number of assets influences this relationship.

1.2 Structure of the paper

Chapter two provides an overview of Harry Markowitz's portfolio theory to establish a fundamental understanding of portfolio construction under the assumption of risk-averse investors. The following chapter three introduces the methods used in this study for optimizing portfolio weights. After presenting the datasets used, the approach of the rolling backtest is illustrated. Before the performance indicators used in this study are presented, chapter six explains how transaction costs are incorporated into the evaluation. This is followed by a presentation and interpretation of the results. The paper concludes with a critical assessment.

---

[2] See DeMiguel et al. (2009a).

2 Risk-return optimization according to Markowitz

Noble prize winner Harry Markowitz (1952) demonstrated in his famous study that if investors base their investment decisions on the expected future return ($\hat{\mu}$) and the expected variance ($\hat{\sigma}^2$) or standard deviation ($\hat{\sigma}$) of a portfolio return (as a proxy for investment risk), there are portfolios that are clearly inferior to others.[3] The expected return of a portfolio is the weighted sum of the expected returns of all securities. To determine the variance, the correlation coefficient between the asset returns and their standard deviations is also required. For a portfolio that, for simplicity, contains only two securities, the expected portfolio return ($\hat{\mu}_p$) and expected portfolio return variance ($\hat{\sigma}^2$) are calculated as follows:

$$\hat{\mu}_p = w_1\hat{\mu}_1 + w_2\hat{\mu}_2 \qquad (1)$$

$$\hat{\sigma}_p^2 = w_1^2\hat{\sigma}_1^2 + w_2^2\hat{\sigma}_2^2 + 2w_1w_2\hat{\rho}_{1,2}\hat{\sigma}_1\hat{\sigma}_2$$

where:

$\hat{\rho}_{1,2}$ = estimated correlation coefficient between the returns of asset 1 and asset 2

w = weighting of the asset return in the portfolio

$w_1+w_2=1$ (Full investment condition)

The formulas in equation (1) illustrate that the variance, as well as the expected return of a portfolio, is dependent, among other factors, on the weighting of the individual assets within the portfolio. Portfolio diversification occurs when the investable securities exhibit low covariances with each other. According to Markowitz (1952), this can already be achieved by investing in securities from different industries. The calculation of efficient portfolios is heavily reliant on the expected returns and the anticipated variance-covariance matrix. To estimate these, Markowitz suggests using a section of

[3] In what follows, the accent ^ indicates the estimated value of a specific moment of a given future return distribution.

historical time series data of the securities, while acknowledging that future multi-factor models could provide more accurate forecasts. The model assumes that investors base their investment decisions on maximizing the (μ-σ) ratio. Therefore, for a given variance, investors choose the portfolio that promises the highest return for that level of risk. Similarly, for a given expected return, the weights are chosen to minimize variance. Consequently, according to the model, investors are risk-averse, market securities are infinitely divisible, and transaction costs are not considered. The following section will discuss the methods used for determining portfolio weights and compare them with a naïve portfolio.

3 Methods for calculating portfolio weights

3.1 The naïve equal-weighted portfolio (EW)

As mentioned in the introduction, DeMiguel, Garlappi, and Uppal (DGU) highlighted that no (μ-σ) optimization method has consistently and significantly outperformed a naïve 1/N portfolio (abbreviated EW in the following) in the long term. This is partly because a naïve portfolio does not require estimates of expected returns, future variances, or correlations to determine the weights, thereby avoiding the impact of estimation errors on the allocation. Consequently, there is no need for extensive datasets of exceptional quality to calculate the optimal weights. Additionally, equal distribution of weights naturally leads to diversification, which helps to avoid concentration risks. Only in scenarios of extreme volatility did DGU's simulations demonstrate that (μ-σ) optimized weights can deliver better performance than a naïve allocation. The weight vector w is thus calculated as follows.

$w = \imath/N$, where $\imath$ is the unit vector (all elements of the vector are 1)

In addition to its inherent diversification, the naïve portfolio has the advantage of always investing in the best-performing security and based on market capitalization, overweighting small-cap stocks while underweighting large-cap stocks. This results in capturing the excess return associated with the small-cap factor premium (size-alpha).[4]

[4] See Kritzman et al. (2010).

3.2 The minimum-variance portfolio (MVP)

For the minimum-variance portfolio (MVP), if the efficient frontier is correctly specified and no risk-free asset is available, no other portfolio can offer a higher expected return for the given level of variance. This portfolio can be calculated without estimating expected returns, requiring only an estimation of the variance-covariance matrix.

$$\text{Objective function (w)} = w^T \widehat{\Sigma} w \rightarrow \frac{\text{min!}}{w} \qquad (2)$$

with $\widehat{\sigma}_p^2 = w^T \widehat{\Sigma} w$, where $\widehat{\Sigma}$ is the estimated future variance-covariance matrix. It holds that $w \geq 0$ (no short-selling constraint) and $w^T \imath = 1$ (the sum of the weights equals 1, i.e., 100%), where $\imath$ is a vector of ones.

Equation (2) illustrates that the weight vector (w) is selected to minimize the estimated expected portfolio variance $\widehat{\sigma}_p^2$ given an expected variance-covariance matrix ($\widehat{\Sigma}$). Furthermore, the non-negativity constraint (prohibiting short-selling) is always applied to the weights, alongside a budget constraint ensuring that all weights sum to 1 (or 100%). DGU also highlighted in their study that models prohibiting short-sells tend to yield slightly better Sharpe ratios than those without weight constraints. Additionally, Bielstein and Hanauer (2018) explained that the budget constraint and non-negativity condition (short-selling prohibition) are widely accepted restrictions in academic literature. The MVP has gained popularity in recent years because it does not require estimating expected returns. This approach benefits from the absence of estimation errors for expected portfolio returns since returns are excluded from the portfolio construction process.

3.3 The maximum Sharpe ratio portfolio (SR)

The maximum Sharpe ratio portfolio (SR) is the portfolio that maximizes the ratio of the expected excess return (return above the risk-free rate) to the expected standard deviation of returns. Consequently, in an investment universe without a risk-free rate and with correct estimation, it is considered the most efficient portfolio. It is calculated by maximizing the Sharpe ratio across all risky returns.

$$\widehat{SR}_p=\frac{\hat{\mu}_p-r_f}{\hat{\sigma}_p} \quad (3)$$

$$\text{Objective function }(w) = \frac{\hat{\mu}_p-r_f}{\hat{\sigma}_p} \rightarrow \frac{\max!}{w}$$

$$\hat{\sigma}_p=\sqrt{w^T\hat{\Sigma}w}$$

This equation emphasizes that the Sharpe ratio (SR) is maximized by choosing the appropriate weights (w) for the assets, thereby optimizing the portfolio for the highest possible return per unit of risk.

$$\hat{\mu}_p=w^T\hat{\mu}$$

Subject to:

$w \geq 0$ (no short-selling constraint)

$w^T\imath = 1$ (sum of weights equals 1 or 100%)

$\imath$ = all elements of the vector are 1

Instead of using a fixed risk-free rate, the study employs the rate of the one-month Treasury Bill (r_f). Simultaneously with the calculation of the MVP, the budget constraint and the non-negativity constraint are also applied.

3.4 The minimum value-at-risk portfolio (VaR)

A minimum value-at-risk portfolio (VaR) was not considered by DeMiguel, Garlappi, and Uppal. This portfolio optimizes weights to minimize the loss that, with a probability of α, will be exceeded over a given historical period and holding duration.

$$\text{Objective function }(w)=\hat{\mu}_p - \varrho_z(1\text{-}\alpha)\hat{\sigma}_p \rightarrow \frac{\min!}{w} \qquad (4)$$

$$\hat{\sigma}_p=\sqrt{w^T\hat{\Sigma}w}$$

$$\hat{\mu}_p=w^T\hat{\mu}$$

Subject to:

$w \geq 0$ (no short-selling constraint)

$w^T\imath = 1$ (sum of weights equals 1 or 100%)

$\varrho_z(1\text{-}\alpha)$ = (1-α)-quantile of the standard normal distribution

In this paper, α is set to 5% and consistently applied to the average monthly return and standard deviation (annualized) of the in-sample period. It is important to note that the value-at-risk does not indicate the magnitude of the loss that occurs with probability α. Additionally, this approach assumes a normal distribution of returns.

3.5 Summary of the applied approaches

To evaluate the advantage of the naïve portfolio relative to the other three approaches, each of the three optimization methods is applied a second time with a constraint on the maximum weight distribution, extending beyond the budget constraint and non-negativity condition. Bielstein and Hanauer (2018) used maximum weights of 5% in a universe of 1,000 stocks, and subsequent studies have even reduced this to 1.5%. Simulations and real data analyses by Tu and Zhou (2011) have shown that traditional (μ-σ) optimization approaches yield improved results compared to naïve optimization when incorporating the positive attributes of 1/N optimization. DGU combined the minimum-variance portfolio (MVP) with the 1/N portfolio. Additionally, the MVP was applied with a minimum weight constraint of half the naïve weight distribution (0.5*1/N), in addition to the non-negativity constraint. Both MVP approaches demonstrated a slightly higher Sharpe ratio than the naïve portfolio when applied to the 10-industry portfolios (adding the US market). Therefore, to mitigate the extreme weights resulting from optimization without weight constraints, the maximum weight for any single security is fixed at 20%. Table 1 provides a summary of the approaches used.

Table 1: Summary of optimization methods

Model	Abbreviation
Naïve Allocation	
0. Naïve Equal-Weighted Portfolio: w = 1/N	EW
Classical Approaches (with Budget Constraint & Non-Negativity Condition)	
1. Minimum-Variance Portfolio	MVP
2. Maximum-Sharpe-Ratio Portfolio	SR
3. Minimum-Value-at-Risk Portfolio (VaR)	VaR
Classical Approaches with Additional Weight Restrictions	
4. Minimum-Variance Portfolio with weight restriction: 0 ≥ w ≥ 0.20	MVP_C
5. Maximum-Sharpe-Ratio Portfolio with weight restriction: 0 ≥ w ≥ 0.20	SR_C
6. Minimum-Value-at-Risk Portfolio with weight restriction: 0 ≥ w ≥ 0.20	VaR_C

In the following section, the expected parameters, such as return and the variance-covariance matrix, are estimated based on historical data and assumed as expected for the out-of-sample period. Various approaches have been developed in academic research to make these estimates more accurate and robust for future predictions. DGU did not use shrinkage estimators like the Ledoit and Wolf (2008) estimator, arguing that a non-negativity constraint already yields similar results to shrinking the covariance matrix. Jagannathan and Ma (2003) mentioned that if the non-negativity constraint is applied in the weight determination process using the MVP method, it has a similar effect on the variance-covariance matrix as deriving it with shrinkage estimators. Due to the restrictions applied to the weights, the derivation and application of shrinkage estimators are omitted.

4 Datasets for the empirical analysis

DGU found through simulations that the likelihood of achieving better risk-adjusted returns through (μ-σ) optimization compared to naïve allocation increases when the in-sample period is extended, and fewer investable assets are available. Therefore, this analysis uses the 10-industry portfolios from Kenneth French's website[5] (as in the DGU study), as well as the 30- and 48-industry portfolios.

[5] https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html

In DGU's study, the 10-industry portfolios covered the period from 1963 to 2004 at a monthly frequency. For this study, a time series from July 1926 to October 2021 is available for both the 10- and 30-industry portfolios. The 48-industry portfolios do not have as long a history for each industry, but complete data is available starting from July 1969. Emphasis is placed on long historical data to ensure that the results obtained in this study are robust.

5 Rolling backtest

The months used to determine the weights based on the underlying strategy (in-sample) must always precede the period during which the weights are applied, and the performance is measured (out-of-sample). In the following, let T represent the total length of the return dataset, and M the number of months used in-sample to calculate the optimal strategy weights (window). Small t thus represents the holding period, which begins in the following month after the in-sample data period (i.e., t = M + 1).

The total available out-of-sample period is influenced by the length of the in-sample period and is therefore defined as (T-M). The weights used to multiply the returns at time t are calculated from the preceding (M) months. After the first calculation, for a holding period of one month, the first return of the dataset is removed from the in-sample period, and the previous month's return t is added. This process continues until the end of the dataset, resulting in T-M out-of-sample returns being calculated for a holding period of one month.

This approach is utilized by Tu and Zhou (2011) and DeMiguel et al. (2009b), among others. DGU also calculate T-M out-of-sample returns. One could also introduce a lag of one period between the calculation of the weights and the purchase of the securities. However, this is omitted here since monthly data is used. The procedure is illustrated in the following figure for an arbitrary dataset with 200 months, an in-sample period of 120 months for determining the optimal weights, and an out-of-sample period of 12 months.

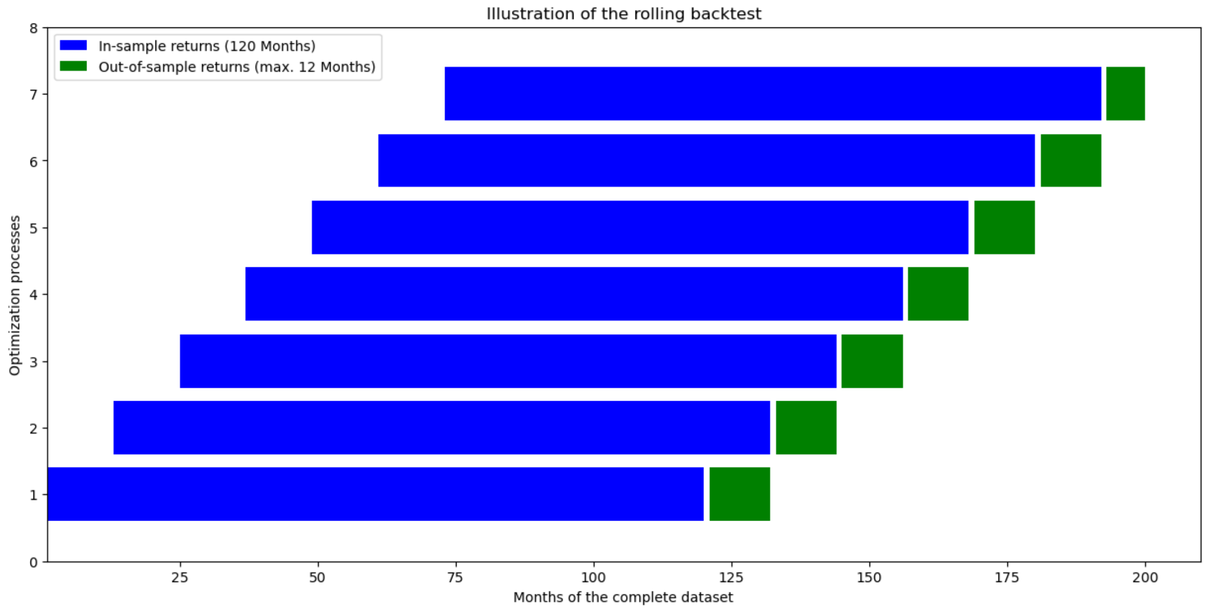


Figure 1: Illustration of the rolling backtest

Figure 1 illustrates that, by definition, the first M-months cannot be used as out-of-sample data. In practice, DGU suggest that in-sample periods of 60 or 120 months are commonly used for estimating the optimal asset weights of different strategies. The authors further demonstrated through simulations that when 25 investable assets are available, a dataset with monthly returns prior to the evaluation period and consisting of 3000 months is necessary to achieve more accurate estimates that can consistently outperform the naïve portfolio in (μ-σ) optimization (with 50 assets, 6000 monthly data points are required). However, this is not practical due to structural breaks in the data affecting estimation accuracy. The results in the DGU study showed no significant differences between an in-sample period of 60 or 120 months. If the weights are constrained to be positive, shorter in-sample periods may yield better results. Tu and Zhou (2011) also used an in-sample length of 120 data points in simulations and real data, with better results achieved at 240 months for industry portfolios. In a study by DeMiguel et al. (2009b), in-sample periods of 60, 120, and 240 months were employed, all yielding similar results, thereby motivating a focus on the 120-month window. A practitioner-focused study by Kritzman et al. (2010) considered a length of 5 years (60 months) insufficient and, therefore, included periods of 10 and 20 years. Conversely, Bielstein and Hanauer (2018) used an in-sample length of 5 years (60 months) when examining various forms of (μ-σ) optimization. The following study is based on in-sample periods of 120 data points (10 years) and, to mitigate potential effects from structural breaks, also on periods of 60 months (5 years).

DGU employed an out-of-sample time frame of one month. Similarly, in their subsequent study (DeMiguel et al., 2009b), the authors opted for a holding period of one month. A recalculation of weights after one month was also utilized by Hilliard and Hilliard (2018) when comparing the out-of-sample volatility and returns of optimized and naïve portfolios. Annual rebalancing, on the other hand, was used by Kritzman et al. (2010). Bielstein and Hanauer (2018) noted that rebalancing after one year is indeed practically relevant. Therefore, this study utilizes both a one-month and a one-year out-of-sample period. The returns achieved during the holding period are compounded. This means that when a holding period of 12 months is used, annual returns are evaluated.

6 Calculation of net returns

In the following, the performance indicators will consistently be calculated based on net returns. The approach of proportional transaction costs, as utilized by DGU, will be applied. Specifically, let $R_{k,p}$ represent the portfolio return after the weight calculations of optimization method k for each asset j, prior to rebalancing ($R_{k,p}=\sum_{j=1}^{N} R_{j,t}w_{k,j,M}$). However, the weights allocated to individual securities change during the holding period. The weights on the individual assets before rebalancing are denoted by $w_{k,j,M+}$, and the weights that are now implemented are denoted by $w_{k,j,t}$. The absolute amount that needs to be reallocated is therefore calculated as $\sum_{j}^{N}|w_{k,j,t}-w_{k,j,M+}|$. The net return is then calculated as follows:

$$\left(\left(1+R_{p,t}\right)\left(1-c\times\sum_{j=1}^{N}|w_{j,t}-w_{j,M+}|\right)\right)-1 \qquad (5)$$

where c represents the proportional transaction costs, which, as estimated by DGU, are set at 50 basis points. This means that if a gross return of 2% is generated in the first period, this corresponds to a net return of 1.49% (with $w_{j,M+}$ set to 0 for the first period).

7 Performance measurement

7.1 The Sharpe ratio

In their study, DGU evaluated the performance of out-of-sample portfolios using several metrics, including the Sharpe ratio. The Sharpe ratio is defined as the ratio of the average excess return (average return minus the risk-free rate) to the standard deviation of returns over the period.[6] This metric provides insight into the risk premium (return above the risk-free rate) relative to the risk taken (measured by the standard deviation) during the evaluation period.

To determine whether the difference between two Sharpe ratios is statistically significant, Jobson and Korkie (1981) developed a test. This test was adapted by Memmel (2003), with the test statistic for the null hypothesis (no significant difference between Sharpe ratios) for two portfolios calculated as follows:

$$\hat{Z}=\frac{SR_1-SR_2}{\sqrt{\hat{\theta}}} \tag{6}$$

where

$$\hat{\theta}=\frac{1}{T-M}\left(2\left(1-\hat{\rho}_{1,2}\right)+0.5\left(SR_1^2+SR_2^2-2SR_1SR_2\hat{\rho}_{1,2}^2\right)\right)$$

and $\hat{\rho}_{1,2}$ denotes the estimated correlation coefficient of the excess returns.

The resulting test statistic is compared against the critical value from the standard normal distribution, depending on the chosen significance level, to perform a one-sided test. As noted by DGU, this calculation assumes that portfolio returns are independent (i.e., returns are not dependent on previous returns) and normally distributed (i.e., no fat tails). However, this is often not the case in practice. Therefore, more robust methods for determining significance have been developed, though these are not addressed in this context.

[6] See Sharpe (1966).

7.2 The CAPM portfolio alpha

Another performance metric utilized in this analysis is the portfolio alpha based on the Capital Asset Pricing Model (CAPM). The model is expressed as follows:[7]

$$\acute{r}_p=\alpha_p+\beta_p\acute{r}_B+\varepsilon_p \qquad (7)$$

where:

- $\acute{r}_p$ denotes the portfolio excess return,
- $\acute{r}_B$ represents the benchmark excess return,
- β_p signifies the sensitivity of the portfolio return to the benchmark return (slope coefficient of a linear regression),
- α_p is the portfolio's alpha, indicating the benchmark adjusted average excess return of the portfolio,
- ε_p represents the error term in the regression with an expected value of zero.

The portfolio return based on the portfolio's exposure to the benchmark is calculated as $\beta_p\acute{r}_B$, thus representing the systematic risk associated with the benchmark through β_p. Accordingly, the return of an asset is not only dependent on the slope coefficient but also on the intercept of this linear regression. The intercept, α_p, indicates the extent to which the expected return of a portfolio exceeds (or falls short of) what would be predicted based on the systematic risk alone.

As Jensen (1968) later noted, a positive alpha suggests that the portfolio manager has made accurate predictions and achieved a higher return than would be expected given the level of systematic risk assumed. However, this is valid only if the portfolio manager does not exhibit market timing skills. Comparisons of portfolios based on alpha alone

[7] See, for example, Lintner (1965) and Sharpe (1966).

are less informative, as alpha is an absolute measure and does not provide insight into the systematic risk undertaken.

The benchmark for this analysis comprises the excess returns of the market, as used in the three-factor model on Kenneth French's website. This benchmark includes most stocks from the New York Stock Exchange, NASDAQ, and AMEX. This dataset is particularly suitable as its time series is consistent with that of the industry portfolios analyzed.

7.3 Maximum drawdown and variance

The drawdown DD of a portfolio or stock is calculated by determining the percentage difference between the current value and the previous highest value. Specifically:

$$DD_t = \frac{\max(P_{max} - P_t, 0)}{P_{max}}$$

This metric allows for the assessment of the maximum loss incurred at the worst investment point, providing insights into the robustness of the investment during periods of crisis. The maximum drawdown is defined as the largest loss observed over the entire period under consideration. When comparing portfolios, it is crucial to ensure that they cover the same time frame.

Additionally, variance is calculated out-of-sample to analyze whether the minimum variance portfolio (MVP) exhibits lower variance compared to the equal-weighted (EW) portfolio. Variance measures the squared deviations of returns from their mean. It is noteworthy that variance does not differentiate between positive and negative squared deviations from the mean.

8 Results from the backtest

8.1 Comparison based on the Sharpe ratio

Table 2: Sharpe ratio for an in-sample period of 120 months

The table displays the monthly (left) and annual (right) Sharpe ratios for each strategy, based on the dataset used and an in-sample period of 120 months. Returns during a holding period are compounded, so a 12-month holding period results in annual returns. The deviations of a strategy's Sharpe ratio from that of the naïve portfolio (EW) are shown in parentheses. The significance of these deviations is calculated using the procedure corrected by Memmel (2003) based on the method of Jobson and Korkie (1981). * p < 0.10, ** p < 0.05, *** p < 0.01.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.166958	0.156329	0.163937	EW	0.528819	0.504179	0.527521
MVP	0.166638	0.156642	0.17285	MVP	0.492803	0.482041	0.565094
	(-0.000319)	(0.000313)	(0.008913)		(-0.036016)	(-0.022138)	(0.037573)
SR	0.127272	0.120095	0.116954	SR	0.408526	0.390778	0.377198
	(-0.039686)**	(-0.036233)**	(-0.046984)**		(-0.120293)**	(-0.113401)*	(-0.150322)
VaR	0.151048	0.137065	0.139234	VaR	0.442215	0.431387	0.500062
	(-0.01591)	(-0.019264)	(-0.024703)		(-0.086604)**	(-0.072792)	(-0.027458)
MVP_C	0.186992	0.172777	0.179401	MVP_C	0.563463	0.53962	0.589582
	(0.020035)**	(0.016448)	(0.015463)		(0.034644)	(0.03544)	(0.062061)
SR_C	0.156502	0.14007	0.130458	SR_C	0.483416	0.453503	0.41408
	(-0.010455)*	(-0.016258)	(-0.033479)*		(-0.045403)*	(-0.050676)	(-0.11344)
VaR_C	0.174932	0.154036	0.149195	VaR_C	0.532429	0.462749	0.513897
	(0.007975)	(-0.002292)	(-0.014742)		(0.00361)	(-0.04143)	(-0.013624)

The Sharpe ratios were not annualized as in DGU but were maintained on a monthly basis for a holding period of one month. According to the results in Table 2, the naïve portfolio achieves higher Sharpe ratios compared to those calculated using the maximum Sharpe (SR & SR_C) or minimum value-at-risk methods. Specifically, the EW portfolio significantly outperforms the SR portfolio three times and the SR_C portfolio twice. Conversely, no significant difference is observed between the Sharpe ratios of the minimum variance portfolio and the naïve portfolio. It is noteworthy that for N=10, the naïve portfolio has a slightly higher Sharpe ratio. However, this difference initially reverses slightly in favor of the minimum variance portfolio when N=30, and the disparity increases for a spectrum of 48 industry portfolios as investable assets. This may suggest that the advantage of naïve portfolio allocation diminishes as the number of investable assets increases, a phenomenon also highlighted by DGU through simulations. With maximum weights of 20%, strategies incorporating this additional constraint show improved results. Despite this constraint, the VaR portfolio is still inferior (VaR_C) twice. However, as observed in DGU's study, the minimum variance portfolio with weight restrictions beyond non-negativity conditions (MV_C) is able to generate a higher out-of-sample Sharpe ratio compared to the naïve portfolio, with a significant difference noted.

The right side of Table 2 presents the results for a holding period of one year. It is evident that the naïve portfolio also outperforms the maximum Sharpe strategy (SR) significantly on two occasions. This superiority extends to comparisons between the naïve allocation and the minimum value-at-risk portfolio (VaR). Although not significant, the EW portfolio outperforms the VaR_C portfolio twice and the SR_C portfolio three times (once significantly). This suggests that even with a one-year holding period, the EW portfolio remains superior to these two optimization methods. Conversely, the results also indicate that, with a one-year holding period, the EW portfolio exceeds the MVP portfolio when N=10 and N=30, but this reverses for N=48. The MVP_C portfolio also demonstrates higher Sharpe ratios than the EW portfolio for a one-year holding period. Hence, the superiority of the EW portfolio diminishes slightly with a larger N.

The results for a reduced in-sample period of 60 months are available in Appendix Table A6, and they are similar: for a holding period of one month, the naïve portfolio allocation outperforms both SR and VaR optimization methods significantly on three occasions. However, the MVP portfolio shows higher Sharpe ratios twice, and the MVP_C portfolio does so three times (twice significantly) compared to the EW portfolio. The results for a one-year holding period are similarly aligned. The naïve portfolio remains superior to the classical approaches (SR and VaR). These findings are therefore comparable to those of DGU concerning the industry portfolios.

8.2 Comparison based on CAPM alpha

Table 3: CAPM alpha with an in-sample period of 120 months

The table displays annualized CAPM alpha values for each strategy, based on the dataset used and an in-sample period of 120 months. Returns during the holding period have been aggregated, resulting in annual returns for a holding period of 12 months.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.00865	0.004822	0.003237	EW	0.009578	0.008655	0.009415
MVP	0.018072	0.01381	0.01928	MVP	0.013289	0.012822	0.025093
SR	-0.00156	-0.001877	-0.011798	SR	-0.000997	0.000049	-0.000847
VaR	0.008319	0.002306	-0.001192	VaR	0.001494	0.001823	0.014569
MVP_C	0.019912	0.016753	0.017621	MVP_C	0.018875	0.019685	0.028735
SR_C	0.004825	0.001875	-0.007	SR_C	0.004531	0.008669	0.00449
VaR_C	0.013813	0.007536	0.001723	VaR_C	0.011942	0.006094	0.017529

Table 3 presents the CAPM alpha for each optimization approach across the three security universes and an in-sample period of 120 months. Initially, it is noteworthy that the portfolios optimized for maximum Sharpe ratio (SR and SR_C) exhibit the most negative alphas. This is interpreted as indicating that the excess returns of these strategies, relative to their beta (systematic risk), were often lower than the excess returns of the broad market. Consequently, these strategies did not demonstrate positive selection abilities. In contrast, the positive alphas of the EW, MVP, and VaR (except one) strategies can be observed. Comparing alphas among these strategies is not particularly meaningful, as it does not account for the differing levels of systematic risk. Nonetheless, it is worth noting that the MVP strategy, both with and without additional constraints, consistently yields the highest alpha, indicating a positive excess return relative to a benchmark with the same systematic risk. The alpha of the EW portfolio is always superior to that of the SR portfolios. For a one-month holding period, the alpha of the EW portfolio decreases with a larger N. The results for an in-sample period of 60 months differ in that the SR strategy shows positive alphas. This may be attributed to the shorter in-sample period, which, by avoiding larger structural breaks, produces better results for expected returns. The detailed results can be found in Appendix Table A7.

8.3 Comparison based on maximum drawdown and ex-post variance

Table 4: Maximum drawdown with an in-sample period of 120 months

The table presents the maximum drawdown for each strategy based on the dataset used and an in-sample period of 120 months. Returns during a holding period were compounded, so that with a holding period of 12 months, annual returns are obtained. The date in parentheses indicates the month in which the maximum drawdown occurred, compared to the previous monthly (left) or annual (right) all-time high.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.483182	0.53428	0.529187	EW	0.315634	0.334061	0.33059
	(2009-02-28)	(2009-02-28)	(2009-02-28)		(2009-06-30)	(2009-06-30)	(2009-06-30)
MVP	0.35858	0.396174	0.396778	MVP	0.228136	0.252093	0.218442
	(1938-03-31)	(2009-02-28)	(2009-02-28)		(1970-06-30)	(2009-06-30)	(2009-06-30)
SR	0.475504	0.609221	0.531378	SR	0.38822	0.452823	0.41585
	(1974-09-30)	(2009-02-28)	(2009-02-28)		(2009-06-30)	(2009-06-30)	(2009-06-30)
VaR	0.420572	0.444505	0.450376	VaR	0.283525	0.28098	0.317161
	(1974-09-30)	(1974-09-30)	(2009-02-28)		(1970-06-30)	(2009-06-30)	(2009-06-30)
SR_C	0.495598	0.515798	0.513321	SR_C	0.305561	0.376473	0.391238
	(1938-03-31)	(2009-02-28)	(2009-02-28)		(1942-06-30)	(2009-06-30)	(2009-06-30)
MVP_C	0.387956	0.397105	0.395777	MVP_C	0.23032	0.255322	0.218419
	(1974-09-30)	(2009-02-28)	(2009-02-28)		(1970-06-30)	(2009-06-30)	(2009-06-30)
VaR_C	0.442239	0.435458	0.446044	VaR_C	0.241373	0.287247	0.318443
	(1974-09-30)	(2009-02-28)	(2009-02-28)		(1970-06-30)	(1949-06-30)	(2009-06-30)

Table 4 displays the maximum drawdown for each strategy across the datasets, with holding periods of 12 months and one month. It is important to note that the right side of the table represents the maximum annual drawdown, while the left side shows the maximum monthly drawdown. Both holding periods indicate that the minimum-variance portfolio (MVP) exhibits lower maximum drawdowns compared to the equal-weighted (EW) strategy. Thus, a configuration where out-of-sample weights, which exhibit low in-sample variance, are used, results in a lower maximum drawdown than that of the EW strategy. The value-at-risk (VaR) strategy also achieves this outcome. The Sharpe ratio (SR) strategy, with one exception, shows higher losses, whereas the Sharpe ratio with constraints (SR_C) strategy, likely due to its broadly distributed weights, demonstrates significantly lower losses three times. When the in-sample period is reduced to 60 months, as shown in Appendix Table A8, the losses from the major crash of 1929, which reached its peak only in mid-1932, are included. However, this does not alter the fact that the MVP portfolio still exhibits lower maximum drawdowns compared to the VaR strategy. Therefore, the minimum value-at-risk (VaR) portfolio, in its applied form, cannot be credited with more effectively reducing maximum losses than the MVP portfolio.

Table 5: Variance for an in-sample period of 120 months

The table presents annualized variances for each strategy based on the respective dataset and an in-sample period of 120 months. Returns during the holding period were aggregated, so that with a holding period of 12 months, annual returns are generated.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.022342	0.027297	0.026745	EW	0.030519	0.036991	0.03462
MVP	0.014328	0.014626	0.015161	MVP	0.022237	0.022054	0.019818
SR	0.024448	0.031829	0.03167	SR	0.031478	0.047475	0.050618
VaR	0.016237	0.01778	0.019334	VaR	0.024873	0.028799	0.025337
SR_C	0.020528	0.024886	0.026796	SR_C	0.028823	0.039049	0.045099
MVP_C	0.016766	0.016033	0.016033	MVP_C	0.02465	0.023203	0.021332
VaR_C	0.017434	0.01886	0.019515	VaR_C	0.025238	0.030498	0.026563

Table 5 confirms the lower volatility of the MVP strategy compared to the EW strategy when considering out-of-sample variances. Interestingly, variances are slightly higher when the MVP strategy is constrained by maximum weights. This indicates that the minimum variance optimization process can deliver the desired results out-of-sample. The VaR portfolio ranks second in this regard. The naïve portfolio and the maximum Sharpe ratio portfolio, however, exhibit relatively high out-of-sample variances. This trend is also evident with an in-sample period of 60 months, which can be seen in Appendix Table A9. Consequently, a weakness of the naïve allocation seems to be its higher return volatility, regardless of the asset spectrum used.

9 Summary and critical appraisal

The success of the naïve portfolio in the datasets used here, as well as those employed by DGU, can partly be explained by the fact that the available securities always represent portfolios. This observation was also made by DGU. Consequently, the individual securities already exhibit lower variance. This could also be the reason why both the naïve portfolio and the minimum variance portfolio consistently show positive alphas. However, the results of this study also highlight that as the number of investable assets increases, the superiority of the EW portfolio is slightly diminished or disappears compared to the MVP. Nevertheless, the EW portfolio often outperforms the simple optimization methods of the maximum Sharpe ratio portfolio and the minimum value-at-risk portfolio, particularly in the comparison of Sharpe ratios. Specifically, the maximum Sharpe ratio portfolio was significantly inferior to the naïve portfolio.

When considering the out-of-sample maximum drawdown and variance, the results are less clear, with the MVP performing the best. The superior performance of the MVP strategy compared to the maximum Sharpe ratio strategy could primarily be attributed to the fact that the minimum variance portfolio does not require an estimate of expected returns, thus avoiding estimation errors that could negatively impact performance. With a holding period of one year and the consideration of the resulting annual returns, the advantage of the naïve portfolio allocation persisted.

Future studies should therefore employ larger asset universes (not just portfolios) across different holding periods when testing optimization methods against naïve portfolio allocation. The results here and in the DGU study suggest that naïve portfolio allocation should be used as a benchmark against optimization methods. It would also be promising to investigate whether the MVP strategy could yield even better results if, in addition to weight restrictions, the expected variance-covariance matrix was estimated using shrinkage estimators.

Literature

Bielstein, Patrick; Hanauer, Matthias X. (2018): Mean-variance optimization using forward-looking return estimates. Review of Quantitative Finance and Accounting Vol. 3, Nr. 52, S. 815-840. ISSN 0924-865X.

DeMiguel, Victor; Garlappi, Lorenzo; Uppal, Raman (2009a): Optimal versus Naïve Diversification: How Inefficient Is the 1/N Portfolio Strategy? The Review of Financial Studies Vol. 5, Nr. 22, S. 1915-1953. ISSN 0924-865X.

DeMiguel, Victor; Garlappi, Lorenzo; Nogales, Franciso J.; Uppal, Raman (2009b): A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms. Management Science Vol. 5, Nr. 55, S. 798-812. ISSN 0025-1909.

French, Kenneth R. (2021): Current Research Returns. Verfügbar unter: https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html (zuletzt geprüft am 28.12.2021).

Hilliard, Jimmy E.; Hilliard, Jitka (2018): Rebalancing versus buy and hold: theory, simulation and empirical analysis. Review of Quantitative Finance and Accounting Vol. 1, Nr. 50, S. 1-32. ISSN 0924-865X.

Jagannathan, Ravi; Ma, Tongshu (2003): Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps. The Journal of Finance Vol. 4, Nr. 58, S. 1651-1683. ISSN 0022-1082. 19

Jensen, Michael C. (1968): The Performance of Mutual Funds in the Period 1945-1964. The Journal of Finance Vol. 2, Nr. 23, S. 389-416. ISSN 0022-1082.

Kritzman, Mark; Page, Sébastien; Turkington, David (2010): In Defense of Optimization: The Fallacy of 1/ N. Financial Analysts Journal Vol. 2, Nr. 66, S. 31-39. ISSN 0015-198X.

Ledoit, Oliver; Wolf, Michael (2008): Robust performance hypothesis testing with the Sharpe ratio. Journal of Empirical Finance Vol. 5, Nr. 15, S. 850-859. ISSN 0927-5398.

Lintner, John (1965): Security Prices, Risk, and Maximal Gains From Diversification. The Journal of Finance Vol. 4, Nr. 20, S. 587-615. ISSN 0022-1082.

Markowitz, Harry (1952): Portfolio Selection. Journal of Finance Vol. 1, Nr. 7, S. 77-91. ISSN 1540-6261.

Sharpe, William F. (1964): Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. The Journal of Finance Vol. 3, Nr. 19, S. 425-442. ISSN 0022-1082.

Sharpe, William F. (1966): Mutual Fund Performance. The Journal of Business Vol. 1, Nr. 39, S. 119-138. ISSN 0021-9398.

Tu, Jun; Zhou, Guofu (2011): Markowitz meets Talmud: A combination of sophisticated and naïve diversification strategies. Journal of Financial Economics Vol. 1, Nr. 99, S. 204-215. ISSN 0304-405X.

Appendix

Table A6: Sharpe ratios for an in-sample period of 60 months

The table presents monthly (left) and annual (right) Sharpe ratios for each strategy based on the dataset used and a 60-month in-sample period. Returns during a holding period have been aggregated, so that a holding period of 12 months results in annual returns. In parentheses, the deviations of a strategy's Sharpe ratio from that of the naïve portfolio (EW) are shown. The significance of these deviations was calculated using the procedure corrected by Memmel (2003) based on the method of Jobson and Korkie (1981). * p < 0.10, ** p < 0.05, *** p < 0.01.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.146994	0.141307	0.159439	EW	0.414465	0.396541	0.541977
MVP	0.149814	0.151249	0.157461	MVP	0.406431	0.430083	0.53223
	(0.002819)	(0.009941)	(-0.001978)		(-0.008034)	(0.033542)	(-0.009747)
SR	0.122788	0.118022	0.134183	SR	0.385245	0.355126	0.40951
	(-0.024207)*	(-0.023285)*	(-0.025256)		(-0.029221)	(-0.041415)	(-0.132467)*
VaR	0.142019	0.129455	0.116403	VaR	0.425756	0.363934	0.46954
	(-0.004976)	(-0.011852)	(-0.043036)**		(0.011291)	(-0.032607)	(-0.072437)
MVP_C	0.160482	0.162559	0.166006	MVP_C	0.430024	0.467273	0.567459
	(0.013488)**	(0.021251)*	(0.006567)		(0.015559)	(0.070732)**	(0.025482)
SR_C	0.13887	0.140966	0.144548	SR_C	0.426548	0.386688	0.474166
	(-0.008125)	(-0.000341)	(-0.014891)		(0.012083)	(-0.009853)	(-0.067811)
VaR_C	0.14644	0.140501	0.12755	VaR_C	0.412914	0.409623	0.463297
	(-0.000554)	(-0.000806)	(-0.031889)*		(-0.001551)	(0.013082)	(-0.07868)

Table A7: CAPM alpha for a 60-month in-sample period

The table presents annualized CAPM alphas for each strategy based on the dataset used and a 60-month in-sample period. Returns during a holding period have been aggregated, so that a holding period of 12 months results in annual returns.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.007158	0.005103	0.005126	EW	0.006912	0.004463	0.012227
MVP	0.015572	0.019332	0.013895	MVP	0.009639	0.016251	0.017957
SR	0.002113	0.002837	0.005327	SR	0.01209	0.012208	0.001985
VaR	0.009855	0.00621	-0.011146	VaR	0.012375	0.00457	0.002316
MVP_C	0.015937	0.020289	0.014214	MVP_C	0.0123	0.021558	0.023855
SR_C	0.003793	0.010226	0.004682	SR_C	0.011145	0.008453	0.009018
VaR_C	0.008264	0.009305	-0.006321	VaR_C	0.008541	0.01111	0.003369

Table A8: Maximum drawdown for a 60-month in-sample period

The table shows the maximum drawdown for each strategy based on the dataset used and a 60-month in-sample period. Returns during a holding period have been aggregated, so that a holding period of 12 months results in annual returns. In parentheses, the date of the month in which the maximum drawdown occurred, compared to the preceding monthly (left) or annual (right) all-time high, is indicated.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.641275	0.645685	0.529187	EW	0.315634	0.334061	0.33059
	(1932-05-31)	(1932-06-30)	(2009-02-28)		(2009-06-30)	(2009-06-30)	(2009-06-30)
MVP	0.519015	0.445822	0.394075	MVP	0.244637	0.289335	0.308099
	(1932-06-30)	(1932-07-31)	(2009-02-28)		(1970-06-30)	(2009-06-30)	(2009-06-30)
SR	0.625983	0.721244	0.489015	SR	0.327491	0.518428	0.33219
	(1932-05-31)	(1932-05-31)	(2009-02-28)		(2009-06-30)	(1970-06-30)	(2003-06-30)
VaR	0.570282	0.5445	0.571273	VaR	0.284149	0.479422	0.378961
	(1932-06-30)	(2009-02-28)	(2009-02-28)		(1970-06-30)	(2009-06-30)	(2009-06-30)
SR_C	0.61601	0.630491	0.458785	SR_C	0.349893	0.365327	0.319238
	(1932-05-31)	(1932-06-30)	(2009-02-28)		(1942-06-30)	(1949-06-30)	(2009-06-30)
MVP_C	0.589487	0.591291	0.391811	MVP_C	0.259222	0.293722	0.307976
	(1932-05-31)	(1932-06-30)	(2009-02-28)		(2009-06-30)	(2009-06-30)	(2009-06-30)
VaR_C	0.575327	0.570614	0.546346	VaR_C	0.287628	0.419274	0.361322
	(1932-05-31)	(1932-06-30)	(2009-02-28)		(1942-06-30)	(2009-06-30)	(2009-06-30)

Table A9: Variance for a 60-month in-sample period

The table shows annualized variances for each strategy based on the dataset used and a 60-month in-sample period. Returns during a holding period have been aggregated, so that a holding period of 12 months results in annual returns.

	Holding period prior to rebalancing = 1 month				Holding period prior to rebalancing = 12 months		
	10 Industries	30 Industries	48 Industries		10 Industries	30 Industries	48 Industries
EW	0.031053	0.036724	0.028267	EW	0.062728	0.077817	0.031427
MVP	0.018537	0.018104	0.015755	MVP	0.038929	0.032843	0.019519
SR	0.032116	0.038297	0.035276	SR	0.058459	0.071611	0.075978
VaR	0.020987	0.023588	0.02351	VaR	0.038269	0.044951	0.028002
SR_C	0.027755	0.03069	0.026864	SR_C	0.04804	0.073258	0.037231
MVP_C	0.02287	0.019986	0.017028	MVP_C	0.048102	0.039408	0.020129
VaR_C	0.024224	0.023881	0.023305	VaR_C	0.042641	0.043884	0.031578

Section D – Paper 3

Index tracking in crisis periods:
An empirical investigation of the German DAX index

Christian Cara

This version: January 2025

Abstract

This paper examines the performance of index tracking strategies during periods of economic crisis, focusing on the DAX price index from 1997 to 2021. Four tracking approaches—relative optimization, the Markowitz (1987) approach, regression with constraints, and linear optimization—are introduced and analyzed. Crisis periods in the German economy were identified to separate the data into non-crisis and crisis subsets. Portfolio weights were fitted on non-crisis data and validated using crisis data, evaluating in-sample and out-of-sample performance metrics such as annualized mean active return and tracking error. Results show that while tracking errors increased during crises, mean active returns remained consistently positive across all approaches, indicating portfolio outperformance relative to the index. A critical assessment of assumptions and potential biases is included. Future research could explore dynamic portfolio strategies with regular rebalancing to enhance realism despite higher transaction costs.

Keywords: index tracking; economic crisis; tracking error; active return;

---

1 Introduction

1.1 ETFs being on the rise

Over the last years, Exchange traded funds (ETFs) have been on the rise as an investment vehicle. According to the financial services company Morningstar, the volume of global ETFs was about 7.2 trillion US-Dollars at the end of 2023. In 2023 alone, the increase in investments was more than 900 billion US-Dollars (see Jackson and Johnson, 2024).

These numbers show that ETFs are a common investment vehicle that is broadly used. As some ETFs are built to track certain indices, the question arises how they perform if a crisis like the recent COVID-pandemic hits the economy. This very question shall be the object of investigation in this paper.

1.2 Course of investigation

The aim of this paper is to analyze the strength of index tracking in crisis periods. Additionally, the impact of the number of stocks used to track the index will be examined. The stock index analyzed is the German index DAX in a period from the year 1997 up to and including August of the year 2021.[8]

To reach this objective, the theory of index tracking is introduced first. The following chapter starts with the general idea of index tracking and continues with four implementations of the sampling approach. Precisely, these are the a) relative optimization, the b) Markowitz approach, the c) regression with constraints approach and d) the linear optimization. The following section elucidates the periods of crisis for the German stock market. This method will be used to split the stock history into crisis and non-crisis periods. The fourth section of this paper contains the practical implementation of the approaches depicted before. After a brief analysis of the data used, the tracking portfolio will be fitted with non-crisis data using the different approaches. Furthermore, each model will be fitted with a varying number of stocks. In the further course of this investigation, the obtained model weights will be applied on crisis-data. The tracking performance will be measured by the active return and its variance. This section is followed by a critical evaluation to point out some possible methodological weaknesses

[8] In September 2021, the number of DAX index constituents was raised from 30 to 40. For reasons of better comparability, this study ends in August 2021.

of the analysis conducted. The paper ends with a short summary of the main findings and a brief outlook for future research.

2 Theory of index tracking

2.1 General idea

The general idea of index tracking is to replicate a given index as efficiently as possible. This replication could be used as an investment vehicle, like an ETF, that generates a return and risk profile equal to the index. Due to the fact that this approach does not need any active management, it is called a passive investment strategy.

In general, there are three approaches to track an index. The first one is called census approach. Within this method, the index is replicated using all its components and weights. This method tracks the index nearly perfectly but also generates high transaction costs due to portfolio adjustments. A second approach is called synthetic replication. In this case, derivatives like a total return swap are used to track the index's performance. The third concept, which is used in this paper, is called the sampling approach. In comparison to the census approach, this method only uses a subset of the index's components causing a tracking error but reducing trading costs dramatically (see Franzen and Schäfer, 2018).

In the following subsections, four possible implementations of the sampling approach are introduced. The first two ideas depicted, the relative optimization and the Markowitz approach of 1987, are examples of a quadratic optimization problem. A regression with constraints approach is introduced as a third method, the linear optimization as the final one. The description of these different approaches will be based mainly on Poddig et al. (2009).

2.2 Relative optimization

The main idea of relative optimization is to separate the benchmark risk and return profile from the portfolio. The aim is to create a portfolio that generates an optimal return for the additional risk taken by investing in the given portfolio. With some adjustments, this approach could be used for index tracking (see Poddig et al., 2009).

The tracking portfolio should have the following three properties. First, the portfolio should track the index as precisely as possible, therefore, the alpha of this portfolio against the index should be zero. Second, the portfolio beta (β_p) should be one to replicate the index. Finally, the residual risk of the portfolio (σ^2_{AP}), deducting the index risk, should be as low as possible. The target function of the weights (w) therefore is (see Poddig et al., 2009):

$$TF(w) = \sigma^2_{AP} \quad \rightarrow \quad \frac{\min!}{w}$$

The residual risk for the portfolio is also known as the tracking error, consequently, this approach minimizes the tracking error. The tracking error, also known as active risk, equals the variance of the portfolio’s active return (the difference between the portfolio and index return) (see Poddig et al., 2009).

To reach an optimal index tracking portfolio, the following constraints are necessary (see Poddig et al., 2009):

(1) $w^T\iota = 1$	Budget restriction: all weights sum up to one
(2) $w \geq 0$	Short selling restriction
(3) $\beta_p = 1$	No timing restriction
(4) $\alpha_p = 0$	No portfolio selection

The budget restriction prevents, that an invested amount unequal to 100% of the investment sum is used to track the index, whereas the short selling restriction avoids negative stock weights. This approach expands the relative optimization with constraints to generate a tracking portfolio with a minimum tracking error.

2.3 The Markowitz (1987) approach

The Markowitz (1987) approach uses the same target function as employed by the relative optimization: the minimization of the active risk or tracking error (see Poddig et al., 2009):

$$TF(w) = TE = \sigma_{AP}^2 \quad \rightarrow \quad \frac{\min!}{w}$$

The constraints of this approach equal the constraints (1) and (2) of the relative optimization setting in subsection 2.2. The main restriction of this approach is that the expected active return must be equal to a predefined return level (μ^*) (see Poddig et al., 2009):

(1) $\mu_A = \mu^*$ return level (generally 0)

Because the Markowitz approach abstains from the difficult task to estimate expected asset returns and focuses on the estimate of the variance-covariance matrix instead, which can be approximated by historical values, its application should be generally easier. The downside of this approach compared to the relative optimization is that it is not tied to a specific portfolio beta. This circumstance may lead to an unintentional timing effect. Furthermore, the portfolio's active return can have an expected value below zero (see Poddig et al., 2009).

2.4 Regression with constraints approach

The main idea of this approach is to explain the benchmark returns with the returns of a given number of stocks. The index or benchmark return is used as an endogenous variable in an ordinary least squares regression. The stock returns (r_{nt}) are the exogenous variables; the beta factors represent the weights. The residual of the benchmark return, and the weighted return of the portfolio, is the active return. The tracking error, that is, the target to minimize, is the square of these errors. This results in the following target function (see Poddig et al., 2009):

$$TF(w) = \sum_{t=1}^{T} \varepsilon_t^2 = \sum_{t=1}^{T} (r_{Bt}-(w_{P1}r_{1t}+w_{P2}r_{2t}+\ldots w_{PN}r_{Nt}))^2 \rightarrow \frac{\min!}{w}$$

The constraints for this regression are the budget restriction and the short selling restrictions (see subsection 2.2, equations (1) and (2)).

The target function is to minimize the mean squared error (tracking error) of the active return. The error term of a regression has an expected value of zero. The Markowitz approach minimizes the tracking error for a given return level. Since the return level for the active return is chosen to be zero (to track the index as close as possible), both methods have the same target function and inputs, therefore, they should lead to similar results (see Poddig et al., 2009).

In contrast to the relative optimization, for the regression with constraints approach, it is not necessary to estimate any input parameters, like the portfolio alpha and beta. This simplicity may also lead to problems, since the estimation only relies on historical values and may not be improved by superior estimation techniques (see Poddig et al., 2009).

2.5 Linear optimization

The last index tracking method introduced in this paper is linear optimization. The aim of this approach is to minimize negative active returns. Therefore, this method is not a pure index tracking approach. The rationale of this approach is that an investor wants to avoid a portfolio return (r_P) less than the benchmark's return (r_B). On the other hand, positive active returns will be accepted by the investor, leading to a one-sided risk definition. The target function representing this approach is shown below (see Poddig et al., 2009):

$$TF(w) = \sum_{\substack{t=1 \\ r_{Pt}<r_{Bt}}}^{T} |r_{Pt}-r_{Bt}| \rightarrow \frac{\min!}{w}$$

The residual returns can be represented by the auxiliary variables d_t^+ for positive active returns and d_t^- for negative active returns (see Poddig et al., 2009):

$$
\begin{aligned}
&d_t^+ = r_{Pt} - r_{Bt}, && \text{if } r_{Pt} > r_{Bt} \\
&d_t^+ = 0, && \text{else} \\
&d_t^- = r_{Bt} - r_{Pt}, && \text{if } r_{Pt} < r_{Bt} \\
&d_t^- = 0, && \text{else}
\end{aligned}
$$

Therefore, the target function, searching for values of the asset weights w_{Pi} and for the parameters d_t^+ and d_t^-, respectively, can be restated as follows (see Poddig et al., 2009):

$$TF = \sum_{t=1}^{T} d_t^- \rightarrow \min!$$

The constraints next to the budget restriction and the short selling restriction (see subsection 2.2, formulas (1) and (2)) are determined by the introduction of the two auxiliary variables and are shown in the formulars below (see Poddig et al., 2009):

$$\sum_{i=1}^{N} w_{Pi} r_{it} - d_t^+ + d_t^- = r_{Bt}$$

$$d_t^+ \geq 0 \quad \forall \quad t \in T$$

$$d_t^- \geq 0 \quad \forall \quad t \in T$$

Due to the availability of modern optimization algorithms and improved computing power, the large number of free parameters (w_{Pi}, d_t^+, d_t^-) in this approach is not a problem anymore (see Poddig et al., 2009).

3 Periods of crisis

To examine the efficiency of index tracking during crises, a definition of the term crisis and the determination of its duration is necessary. This paper grounds this definition on two sources. The first source is the European financial crises database operated by the European Systematic Risk Board (ESRB). The aim of this database is to provide definitions of crisis periods in the European Union for macroprudential oversight and policy creation. Crises are detected in a two-step approach. The first step is a quantitative analysis using a financial stress index based on the paper of Duprey et al. (2017) (see also European Systematic Risk Board, 2017a). The financial stress indicator index in this paper uses the domestic stock price index, 10-year government yields and the real effective exchange rate (see Duprey et al., 2017). The second step in the ESRB crisis detection is to separate between systematic crises and other events. A systematic crisis may be present if the funding supply of intermediaries contracts, bankruptcies of relevant financial institutions occur or policies are applied to preserve financial stability (see European Systematic Risk Board, 2017a).

The European financial crises database shows domestic crises up to and including the year 2017. In the relevant period, the database detects two crisis periods. The first one started in January 2001 and its management ended in November 2003. This crisis is the result of a stock market implosion of the 'new economy' and it mainly impacted on the financial sector. The second crisis shown in this database started in August 2007 and lasted for nearly six years until June 2013. This crisis started with the bankruptcy of Lehman Brothers and led to a drying up of the interbank market, causing severe fire sales and bankruptcies. In the later stage, the crisis turned into a sovereign debt crisis (see European Systematic Risk Board, 2017b).

Due to the fact that the ESRB financial crisis database does not cover the period from 2018 until 2021, the COVID crisis is not included. As described above, a systematic crisis is present if policies for financial stability are applied. These measures started in March 2020 and included grants, guarantees and tax reliefs. The financial sector was supported by the reduction of the countercyclical capital buffer to 0%. As parts of these measures were ongoing, the crisis lasted according to the definition mentioned before for some time. Therefore, in this paper, the period from March 2020 until December 2021 will be labeled as COVID crisis.

4 Index tracking during economic crisis

4.1 Data

The stock market index used for examining the power of index tracking in economic crisis is the DAX. This index represents the 40 largest German companies according to free float market capitalization. The DAX started 1987 with a value of 1.000 and was composed of 30 stocks until September 2021 (see STOXX Ltd., 2021a). There are several index versions of the DAX. The most popular one is a performance index, that takes dividend payments and subscription right issues into account. The index version used for this investigation is the price index, which solely represents the price changes of the contained stocks. The reason for this selection is that individual stock return time series without dividend payments will be used here for the index tracking task.

This paper analyzes a period of 25 years (from 1997 up to and including 2021). To track the DAX, all stocks used in the examination were part of this index for the whole period. The following table shows these stocks (see STOXX Ltd., 2021b)[9]:

Table 1: Stocks part of DAX from 1997-2021, used for index-tracking

adidas AG	Allianz SE	BASF SE
Bayer AG	Bayerische Motoren Werke AG	Deutsche Bank AG
Deutsche Lufthansa AG	Deutsche Telekom AG	Münchener Rückversich-e-rungs Gesellschaft AG
RWE AG	SAP SE	Siemens AG

The chart below shows the development of the DAX price index, and the periods of crisis as defined in section 3. The start of the crisis is the first day of the crisis's beginning month, whereas the last crisis day, as represented in the chart, is the last day of the crisis's ending month.

As the chart shows, all crises start with a steep decline in the index. Noteworthy is that the crisis's end is not defined as the lowest point of the index. A recovery phase, in

[9] All stocks are shown by their current stock name.

case of the COVID crisis even with an index increase above the initial level, is also part of the crisis period.

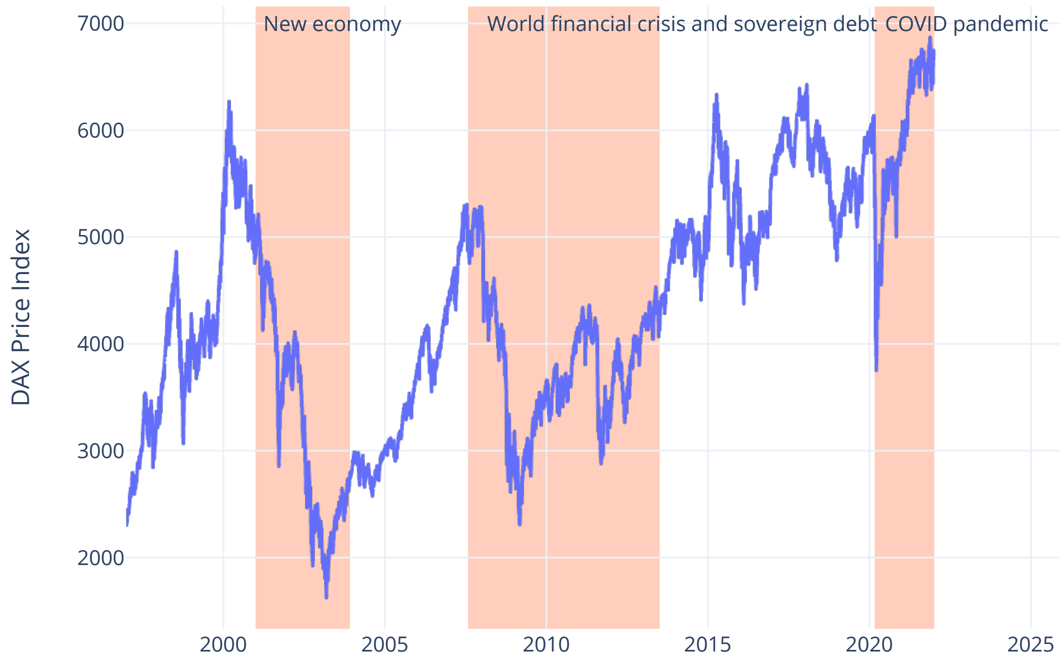


Figure 1: DAX price index and crises 1997-2021 (Source: Bloomberg, own calculation)

For the following analysis, the stocks and the index will be transformed into monthly excess returns, leading to a time series of 300 data points. For computational simplicity, the risk-free rate is assumed to be zero, therefore, the excess return equals the total return in this case. These time series are split into one training and one validation sample. The training sample is used to fit the index tracking weights utilizing the approaches introduced in section 2. This training sample contains the non-crisis data resulting in 172 data points, or 57% of all data. The validation sample contains crisis data to measure the performance of index tracking weights fitted in non-crisis periods. This sample covers the remaining 128 observations, 43% of the whole data set.

4.2 Fitting stock weights

In this section, the fitting process for the stock weights is described. The models used are the relative optimization, the Markowitz (1987) approach, the regression with constraints approach and the linear optimization as depicted in section 2. For each modeling approach, ten weight combinations are approximated, starting with three stocks to replicate the index, and ending with up to twelve stocks, resulting in 40 models.

To determine the extension order of the tracking portfolio, the in-sample (non-crisis period) correlation between each stock and the benchmark is used, as shown in the table below. The stock with the highest correlation is added to the portfolio first, followed by the stock with the second highest correlation and so on, ignoring the intercorrelation of the sample stocks. The reason for this is the basic idea that high correlations with the index may lead to a better tracking portfolio.

Table 2: Correlation of stocks vs. DAX price index for non-crisis periods (Data source: Bloomberg, own calculation)

Stock	Correlation with DAX price index
Siemens AG	0.77
Allianz SE	0.76
BASF SE	0.75
Bayer AG	0.71
Bayerische Motoren Werke AG	0.69
Deutsche Bank AG	0.68
Muenschener Rueckversicherungs-Gesellschaft AG	0.67
SAP SE	0.59
Deutsche Lufthansa AG	0.49
Deutsche Telekom AG	0.48
RWE AG	0.33
adidas AG	0.32

For all approaches, historic values are used to estimate the stock weights. For the relative optimization approach, the input parameters are the stock alphas and betas as well as the variance-covariance matrix. The starting weights are chosen in such a manner that all stocks are equally weighted in the portfolio. Furthermore, weight bounds are used. These bounds are set between 5% and 100% to ensure that all stocks used for the model will have a weight of at least 5% in the final tracking portfolio. These settings are also applied in the remaining approaches.

The figure below shows the weight distribution for the relative optimization using a varying number of stocks. It is noticeable that the stock with the highest correlation (Siemens AG, correlation of 0.77) has not obtained the highest weight in any scenario considered. For this approach, the correlation does not seem to be a good indicator for the weight level.

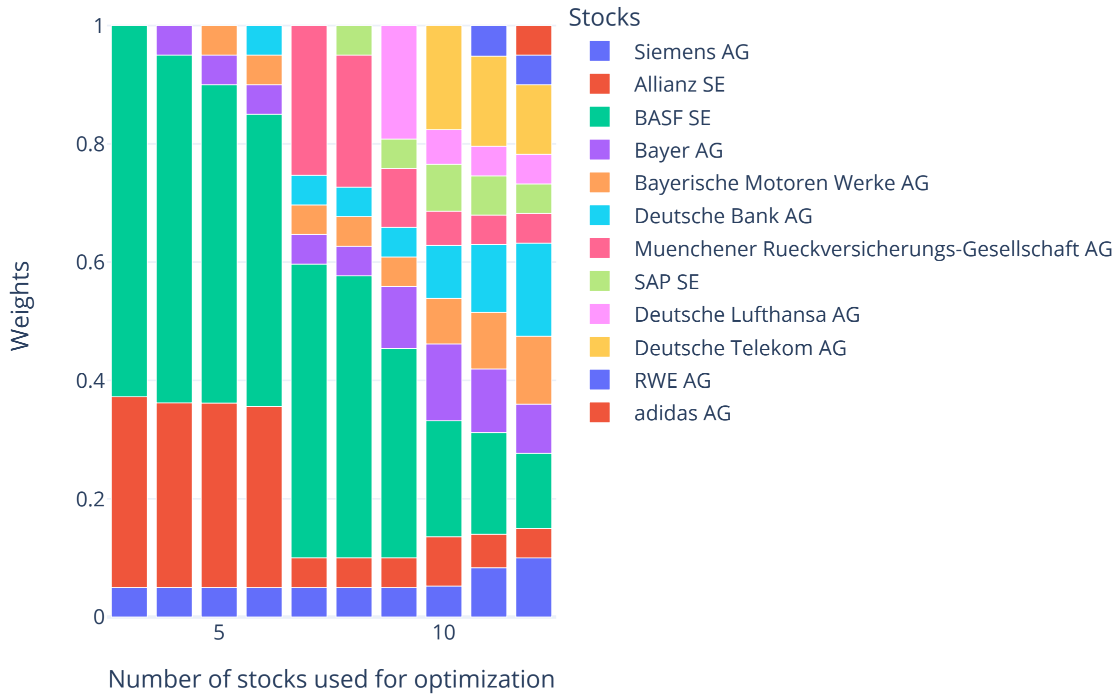


Figure 2: Stock weights of the relative optimization approach (Data source: Bloomberg, own calculation)[10]

To measure the in-sample goodness of fit of the stock weights, two statistics are used. The first one is the active return, the difference between benchmark and portfolio return, the second one is the active risk or tracking error, the variance of the active return. For a perfect fit, an active return and tracking error of zero are expected.

The Markowitz (1987) approach and the regression with constraints generate a similar result as expected. The tracking error for these approaches is near to zero (see Figure 4 below - the highest absolute tracking error occurs for 10 stocks with a value of -0.02%). For the relative optimization, the active return is very low as well, below 0.5% annualized for all portfolios (see Figure 3 below). Considerably different is the active return of the linear optimization. It is between 1.5% and 3% on a yearly basis. This circumstance is attributable to the methodology as shown in chapter 2.5. The linear

[10] Due to the fact that the weights resulting from the four approaches are not considerably different, only the weights for the relative optimization approach are shown exemplarily.

optimization minimizes negative active returns with no constraints on the positive ones, which are accepted in this approach.

The chart below shows the annualized mean active return for all four approaches for a varying number of stocks in the tracking portfolio.

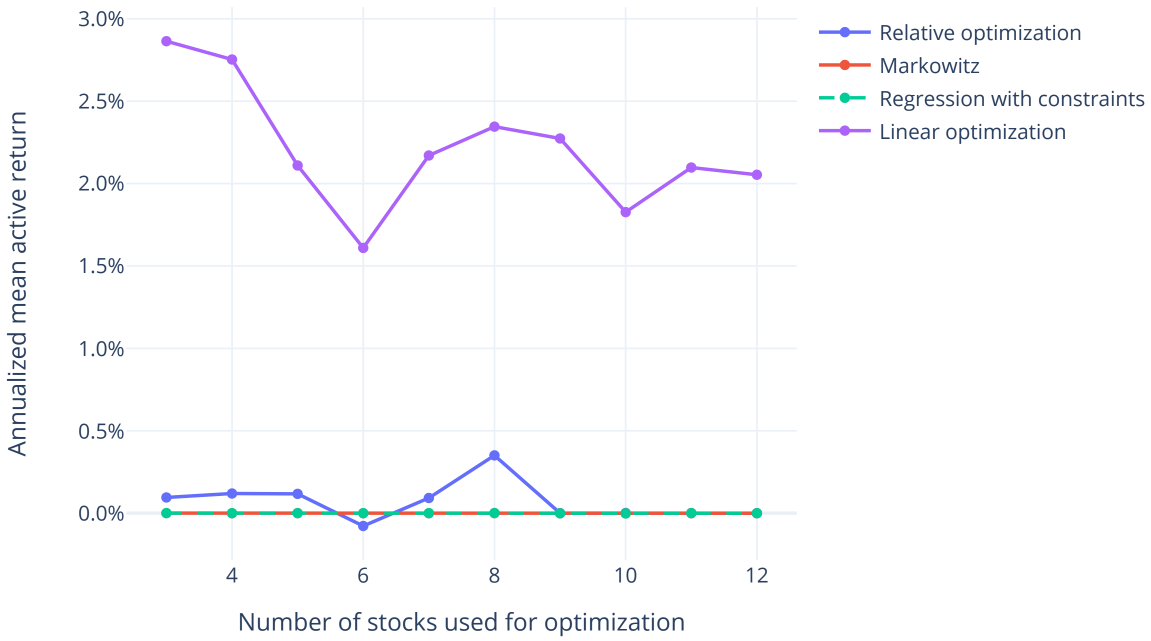


Figure 3: In-sample active return of the tracking portfolios (Data source: Bloomberg, own calculation)

As explained before, the Markowitz (1987) approach and the regression with constraints approach generate identical outcomes, resulting in overlapping lines in the chart. The tracking error path shown in Figure 4 below verifies that an increasing number of stocks leads to a decreasing tracking error. In other words, the more information (i.e., stocks) is added to the model, the more precisely the model can be fitted. There may be cases where the tracking error increases slightly with an additional stock, this is due to the bounds set for the stocks. Each stock has a minimum weight of 5% in the portfolio, which might potentially lead to suboptimal results.

The second statistic, the tracking error, based on the annualized returns for all approaches and portfolios, is shown in Figure 4.

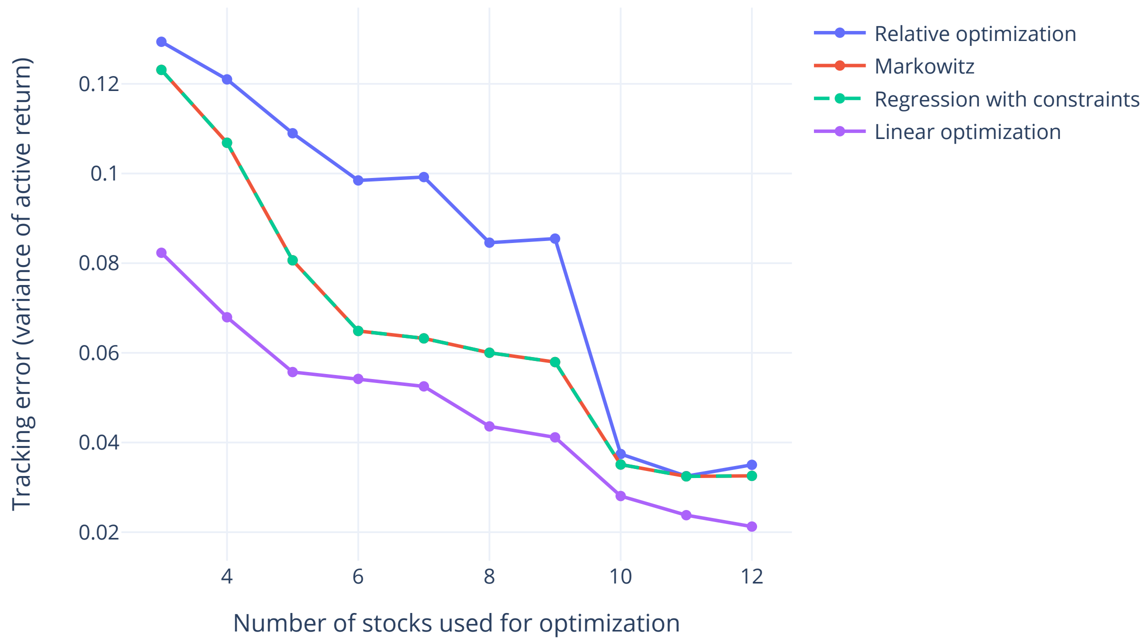


Figure 4: In-sample tracking error of the tracking portfolios (Data source: Bloomberg, own calculation)

4.3 Validation results

To examine the tracking strength of the fitted weights, they are tested out of sample. In this case, the validation period contains crisis periods as defined in section 3. The in-sample weights are multiplied with the monthly returns and the index return is subtracted, leading to the active return. The chart below shows the out of sample annualized mean active return for the four approaches tested. As mentioned in subsection 4.2, the Markowitz approach and the regression with constraints approach return the same weights, therefore their lines are overlapping.

Figure 5 shows a sharp increase in the annualized active return compared to the in-sample chart (see Figure 3), where all active returns except the one for the linear optimization were below 0.5%. Furthermore, the sign of all the active returns is striking: all models generate a positive active return in crisis periods. Finally, the active return shows a downward trend for an increasing number of stocks. This finding may lead one to conclude that the high returns, especially for portfolios with few stocks, result from 'lucky picks' that outperformed the index in the crisis periods.

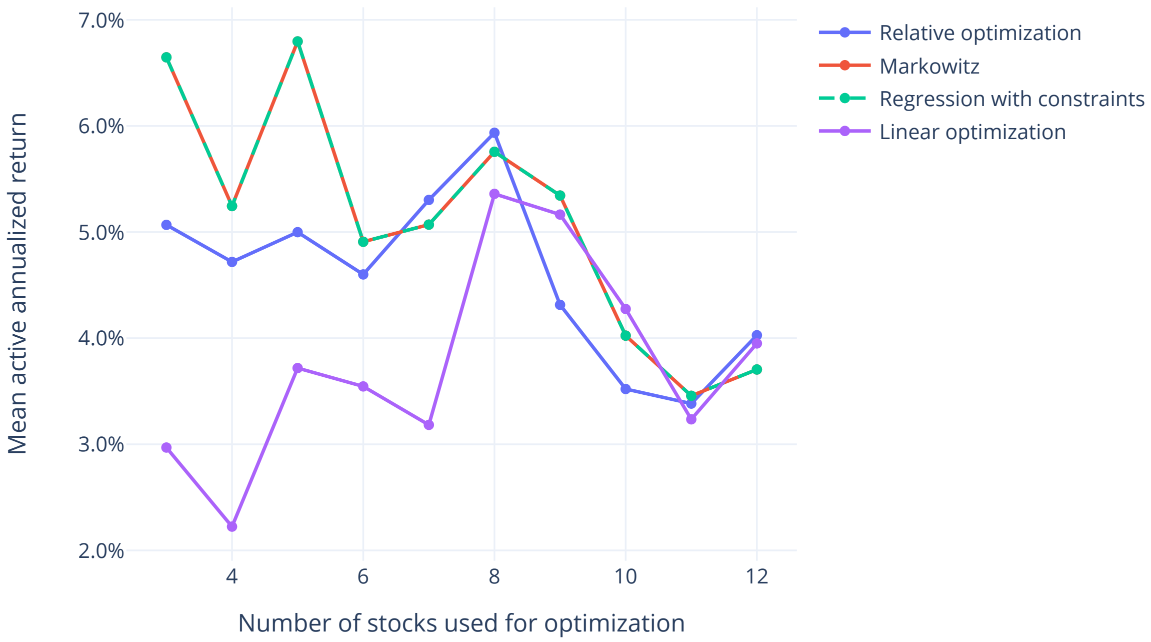


Figure 5: Out of sample active return of the tracking portfolios (Data source: Bloomberg, own calculation)

The second measure to assess the quality of the tracking portfolios is the active return variance, i.e., the tracking error. Figure 6 below shows the tracking error of the annualized active returns (see Figure 5):

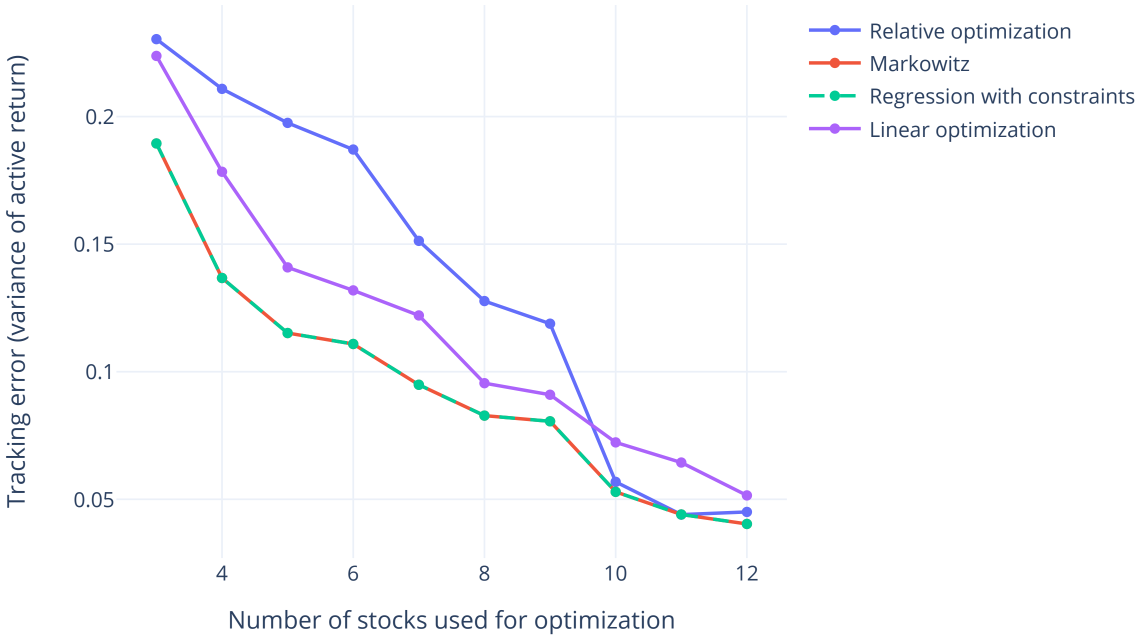


Figure 6: Out of sample tracking error of the tracking portfolios (Data source: Bloomberg, own calculation)

As in the case of the in-sample tracking error, the out of sample active return variance has a clear trend with a negative slope, a higher number of stocks in the portfolio reduce the volatility of the return around the benchmark. However, the scale has to be taken into account, e.g., the out of sample tracking error for portfolios with ten to twelve stocks is around 0.05, whereas the in-sample value is around 0.03, hence, the number increased approximately by 50%.

To sum up, all portfolios worked very well for the out of sample validation (i.e., the crisis) period. All of them could generate a positive mean active return, though the tracking error could increase by up to 100%.

4.4 Critical evaluation

Although the results of the analysis performed so far show clearly that index tracking performed well in crisis periods, methodological assumptions had to be made. This section tries to clarify these assumptions and to point out how they may bias the outcome.

The first assumption concerns the stock universe. The analysis only included stocks which were part of the DAX all through the period examined. This implies that the replication of the index may be optimized by adding stocks to the universe which have a higher explanatory power during certain time periods. Furthermore, the author predefined the order and number of stocks added to the tracking portfolio. This step could be optimized as well, e.g., by using a brute-force analysis and testing all possible portfolio combinations for a given number of stocks and picking the one performing best in-sample.

The second critical point that should be mentioned in this section is the partition of the time series in training and validation periods. The crisis periods were picked as defined by the ESRB and were focused on the German economy. That definition of crisis is just one of many and analyzing a more generally valid one would exceed this paper by far. Another effect of the time series partition is that the time series and their inherent structure is torn, and this may lead to biased results.

A final point to consider is the portfolio building approach. One set of weights is estimated for non-crisis periods and afterwards applied to crisis-periods. In practice, the portfolio would need a rebalancing period instead to ensure efficient index tracking. This procedure might possibly lead to lower tracking errors and active returns.

5 Conclusion

The aim of this paper was to examine the performance of index tracking in crisis periods for the use case of the German DAX index. After introducing the general idea of index tracking, a short overview of four tracking approaches was given. These approaches are the relative optimization, the Markowitz (1987) approach, the regression with constraints approach and the linear optimization. In the next section, periods of crisis for the German economy were defined to split the time series, covering the period from 1997 to 2021.

The following chapter contains the analysis performed, as well as a description of the results obtained. First, the data used was explained, including the tracked index (DAX price index) as well as the stock universe to represent it. In a first step, the stock weights were fitted for varying numbers of stocks using non-crisis data. All in all, 40 models were fitted, 10 for each approach. To measure the power of the tracking portfolio, the in-sample annualized mean return, and the tracking error were introduced. The second step of the analysis encompasses the out of sample validation. To do so, the in-sample weights were applied on crisis data. The resulting annualized mean active return and the tracking error were analyzed and compared to the non-crisis results. The final part of the analysis contained a critical evaluation of the assumptions used and the exploration of potential biases.

The main finding of this paper is the following: All four index tracking approaches worked decently on crisis data. Even though the tracking error doubled for certain portfolio combinations, the mean active return was constantly positive. Investors in these portfolios may suffer from weaker tracking power, but return-wise they outperformed the index.

For further research, a dynamic portfolio approach could be chosen. This may include choosing an optimal subset of all available index stocks and rebalancing it regularly leading to higher transaction costs but portraying the problem more realistically.

Literature

Duprey, Thibaut; Klaus, Benjamin; Peltonen, Tuomas (2017): Dating systemic financial stress episodes in the EU countries. Journal of Financial Stability, Vol. 32, pp. 30–56.

European Systematic Risk Board (2017a): A new database for financial crises in European countries, https://www.esrb.europa.eu/pub/pdf/occasional/esrb.op13.en.pdf?c79e7fcd59daca7c422b3e1cbcc01ec6.

European Systematic Risk Board (2017b): European financial crises database, https://www.esrb.europa.eu/pub/fcdb/esrb.fcdb20170731.en.xlsx?a0238978b006154b3fe798555c952e83

Franzen, Dietmar; Schäfer, Klaus (2018): Assetmanagement – Portfoliobewertung, Investmentstrategien und Risikoanalyse. 1st Edition, Stuttgart, Freiburg.

Markowitz, Harry M. (1987): Mean-Variance Analysis in Portfolio Choice and Capital Markets. New York, 1987.

Poddig, Thorsten; Brinkmann, Ulf; Seiler, Katharina (2009): Portfoliomanagement – Konzepte und Strategien. 2nd Edition, Bad Soden.

Ryan Jackson; Ben Johnson (2024): ETFs Cap Off Another Record Year of Flows with a Stellar December, https://www.wsj.com/articles/globalhttps://www.morningstar.com/articles/1073821/etfs-cap-off-another-record-year-of-flows-with-a-stellar-december-etf-assets-hit-9-trillion-11628769548.

STOXX Ltd. (2021a): Guide to the DAX Equity Indices, https://www.dax-indices.com/document/Resources/Guides/DAX_Equity_Indices.pdf

STOXX Ltd. (2021b): Historical Index Compositions of the Equity- and Strategy Indices, https://www.dax-indices.com/document/Resources/Guides/Historical_Index_Compositions_20.12.2021.pdf

Section E – General Summary and Outlook

This final section includes a brief overview of the paper's motivation, content, and contribution, as well as a more extensive discussion of future research directions.

1 Short summary

The past decades have seen an increasing complexity in financial markets, accompanied by heightened interest in asset management strategies that can reliably balance risk and return. Traditional approaches to portfolio management have evolved significantly since the foundational work by Markowitz (1952), driven by both theoretical refinements and empirical challenges. Among these challenges are estimation errors in expected returns, the structural changes in financial markets during crises, and the proliferation of passive investment instruments such as ETFs. This dissertation seeks to explore contemporary approaches in asset allocation through three self-contained but thematically interrelated essays. Each of the three papers delves into a specific yet complementary facet of portfolio management: the role of risk-based strategies in crisis periods, the empirical performance of naive versus optimized portfolios, and the efficacy of index tracking during turbulent markets.

Collectively, the papers offer a nuanced empirical and theoretical analysis of asset management strategies under varying market conditions, with a strong emphasis on robustness and real-world applicability. The overarching theme that connects these papers is the quest for practical, resilient portfolio strategies in the face of uncertainty – be it from model misspecification, estimation error, or macroeconomic shocks.
Next, I outline several promising directions for future research.

2 Future research directions

While the three papers make distinct contributions to the field of asset management, they also open several promising avenues for further inquiry. These directions span methodological refinements, data extensions, and conceptual frameworks.

Dynamic and regime-switching models

One of the limitations in traditional portfolio models is their static nature. Future research could explore regime-switching or dynamic models that adjust portfolio weights in response to changing market conditions. For instance, integrating macroeconomic indicators or market sentiment data into asset allocation frameworks may enhance performance during transitions between bull and bear markets.

Moreover, machine learning techniques such as reinforcement learning or recurrent neural networks could be leveraged to develop adaptive strategies that learn optimal asset allocations in real time.

Behavioral and ESG considerations

Recent years have seen growing interest in behavioral finance and environmental, social, and governance (ESG) criteria. Future studies could examine how behavioral biases affect investor preferences for different asset allocation strategies, especially under uncertainty. Similarly, ESG integration into portfolio optimization could be evaluated in terms of both financial performance and sustainability metrics.

For example, how do risk-based or naive portfolios perform when constrained by ESG score thresholds? Do ESG-enhanced index tracking strategies exhibit better resilience during crises?

International and emerging market contexts

All three papers primarily focus on developed market data, particularly the German DAX and diversified asset classes common to Western economies. An important extension would be to test these strategies in emerging markets, where volatility is higher, financial data is less stable, and market microstructures differ.

Moreover, cross-country comparisons could reveal structural differences in market behavior and portfolio performance, thereby refining the applicability of existing models.

Impact of high-frequency data and alternative datasets

With the rise of high-frequency trading and the availability of alternative datasets (e.g., satellite data, web traffic, social media sentiment), there is a growing need to understand how these new data sources can enhance portfolio decision-making. Research could investigate whether incorporating such data improves the estimation of risk parameters or the prediction of crisis periods, thereby benefiting both active and passive strategies.

Risk measures beyond variance

Traditional portfolio optimization relies heavily on variance or standard deviation as a proxy for risk. However, alternative risk measures such as conditional value-at-risk (CVaR), downside deviation, and drawdown duration offer potentially richer insights into portfolio performance during tail events. Future research could compare the performance of portfolios optimized under these alternative metrics.

Portfolio constraints and real-world frictions

Finally, incorporating realistic portfolio constraints – such as short-selling limits, transaction costs, taxes, and regulatory requirements – remains an underdeveloped area. Future studies could simulate investment scenarios with such constraints to evaluate the feasibility and performance of different strategies in institutional settings.

3 Concluding remarks

This thesis has explored some of the most pressing challenges and opportunities in contemporary asset management. Through three essays, it has provided empirical and theoretical insights into the comparative performance of risk-based strategies, naive versus optimized portfolios, and index tracking methods in crisis conditions. Taken together, the papers argue for a more cautious, empirically grounded approach to portfolio construction – one that respects the limits of model precision while striving for robustness and practical relevance.

While the field of asset management continues to evolve, the need for transparent, adaptive, and resilient portfolio strategies remains constant. It is the hope of this thesis that its findings contribute to this broader endeavor and pave the way for future research in building the next generation of investment frameworks.