mathematics

MDPI

Review

Quantitative Portfolio Management: Review and Outlook

Michael Senescall [1,*] and Rand Kwong Yew Low [1,2]

[1] Bond Business School, Bond University, Robina, QLD 4229, Australia; rlow@bond.edu.au or r.low@business.uq.edu.au
[2] UQ Business School, The University of Queensland, St Lucia, QLD 4072, Australia
* Correspondence: michael.senescall@bond.edu.au

Abstract: This survey aims to provide insightful and objective perspectives on the research history of quantitative portfolio management strategies with suggestions for the future of research. The relevant literature can be clustered into four broad themes: portfolio optimization, risk-parity, style integration, and machine learning. Portfolio optimization attempts to find the optimal trade-off of future returns per unit of risk. Risk-parity attempts to match the exposure of various asset classes such that no single asset class dominates portfolio risk. Style integration combines risk factors on a security level such that rebalancing differences cancel out. Finally, machine learning utilizes large arrays of tunable parameters to predict future asset behavior and solve non-convex optimization problems. We conclude that machine learning will likely be the focus of future research.

Keywords: portfolio management; asset management; portfolio optimization; diversification

MSC: 91-02

1. Introduction

At a small scale, early-life investors are more likely to engage in riskier portfolio strategies. However, as they accumulate more capital, the strategies of later-life investors become more conservative. While most personalized investment strategies focus on an individual's risk tolerance, this paper aims to identify the various quantitative investment strategies available to construct a portfolio with the goal of achieving a high return-to-risk ratio.

Portfolio management relies on the premise of obtaining the highest return (lowest risk) possible per unit of risk (for a given return). Markowitz [1] first approached this problem by treating asset returns as a stochastic process, defining risk as the variance in returns and creating Mean-Variance Optimization (MVO). The development of MVO is regarded as the catalyst of modern financial economics and continues to significantly impact portfolio management in the 21st century. However, the idea that an investor's utility can wholly be described as some function of the mean and variance of portfolio returns has been met with skepticism. Kahneman and Tversky [2] created prospect theory to address the discrepancy between MVO and actual investor behavior by treating the pain of capital loss more strongly than the reward of an equally large capital gain. There have since been many alternative utility functions developed to improve on MVO; however, none have been as influential. That said, practically implementing MVO has proved problematic, e.g., due to errors when estimating the forward-looking parameters and high transaction costs. Addressing these limitations has been at the core of relevant literature where quantitative portfolio management is fundamentally an optimization problem.

Changes within society and major economic/financial episodes have often led to broad-level changes in investment strategies. For example, Risk-Parity (RP) strategies gained traction due to their wealth-preserving effect during the 2008 Global Financial Crisis (GFC) [3]. These strategies involved the optimized diversification of a portfolio

check for updates

Citation: Senescall, M.; Low, R.K.Y. Quantitative Portfolio Management: Review and Outlook. *Mathematics* **2024**, *12*, 2897. https://doi.org/10.3390/math12182897

Academic Editor: Andrea Scozzari

Received: 17 July 2024
Revised: 13 September 2024
Accepted: 13 September 2024
Published: 17 September 2024

by allocating risk equally across asset classes, which ensured that part of the portfolio could benefit in the case of economic downturn. Much like MVO, however, the practical implementation of RP portfolios has many forms, each with their own limitations. More recently, technological advancements have seen the incorporation of novel data analytics into portfolio management.

With the goal of reducing the reliance on legacy financial theory—that is, to minimize model bias—Machine Learning (ML) has seen the birth of many new investment strategies. There are three key ML methods: Supervised Learning (SL), Unsupervised Learning (UL), and Reinforcement Learning (RL), all of which have shown potential for different research streams within portfolio management. Although the flexibility of ML can introduce new points of failure and complexity that may challenge traditional financial frameworks, there are well-established methods to address these issues [4]. For novel ML models which are yet to establish a handbook, risks are often mitigated by utilizing large datasets for training, and model performance is benchmarked against comparable contenders.

Effective implementation of future investment strategies requires accurate and reliable modeling of asset returns. The performance of a model, whether ML or statistical, is limited to the data it is provided. Published studies have attempted to make their results robust by training models on a diverse set of data, such as including different countries, stock exchanges, and asset classes. However, excess data can impact portfolio performance just as much as insufficient data. For example, models trained on a high-dimensional global dataset may not be able to detect peculiarities within a specific region.

While prior reviews evaluate the optimization of portfolios and the use of novel data analytics in financial markets individually, none comprehensively review quantitative portfolio management holistically over time (e.g., see Milhomem and Dantas [5], Goodell et al. [6], and Bartram et al. [7]). This review aims to investigate how strategies used in portfolio management have historically changed, in part, allowing us to identify possible future trends. By studying historical changes in portfolio management strategies, this review will facilitate effective future portfolio management and optimization. This investigation incorporates quantitative analysis of trends through data-driven approaches and identifies four core research themes from the literature that are intrinsically involved in the quantitative management of portfolios: portfolio optimization, risk-parity, style integration, and machine learning.

2. Research Methods

In reviewing the research on portfolio management, rather than using a more traditional narrative review [8,9], we followed a systematic process to collect and analyze citations relevant to a research question in an unbiased and replicable manner [10]. In the current context, this was achieved within the scope of quantitative portfolio management via a multi-step method depicted in Figure 1.

Table 1 summarizes the filtering method that we use. The initial step is to develop research questions, using key phrases and themes identified in a preliminary investigation in the study of portfolio management. Next, the key phrases and themes identified are used to create a search query within the Scopus comprehensive citation database. Lastly, matched citations are exported and manually processed via predefined inclusion and exclusion criteria.

Since it was unlikely that every relevant citation was indexed in the Scopus database, manual investigations and citation mapping tools (Connected Papers, Research Rabbit) were used as secondary sources (see https://www.connectedpapers.com/ (accessed on 30 June 2022) and https://www.researchrabbit.ai/ (accessed on 30 June 2022) to utilize these tools). Some innovative techniques analyzed in this review were sourced from suboptimal journals, and a more lenient selection criterion was used accordingly. A bibliometric analysis was conducted using the Anaconda development environment and the Python 3.10 and Pybliometrics 3.3.0 package. Visualizations of citations and keywords were generated using VOSviewer version 1.6.18.

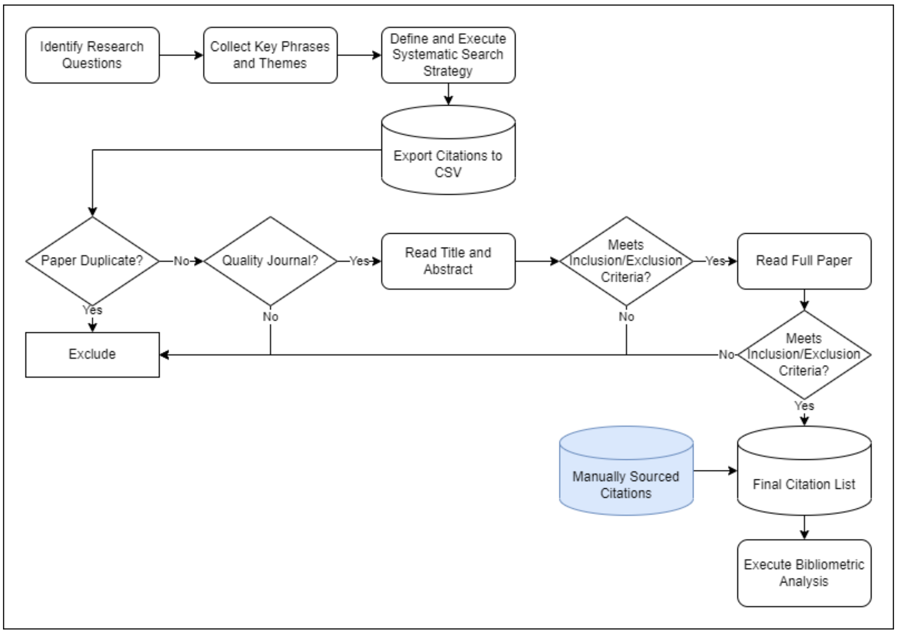

Figure 1. Research method flowchart. The start-to-finish process map used to generate the list of citations for bibliometric analysis, which follows the conceptual steps outlined in Linnenluecke et al. [10].

Table 1. Filtering method table.

Research Questions	**Question 1**	What is the history of quantitative portfolio management?
	Question 2	Given the results from Q1, what should be the focus of future quantitative portfolio management research?
Languages	All search queries and journal articles will be in English	
Research Quality	Articles must be published in a A/A*-quality journal (as per the ABDC database)	
Inclusion Criteria	Journal articles must be quantitative in nature and relevant to equity portfolios or their constituent parts	
Exclusion Criteria	Duplicate journal articles Review papers	
Search Query	TITLE-ABS-KEY("portfolio management" OR "portfolio optimization" OR "investment portfolio" OR "portfolio strateg*" OR "portfolio selection" OR "mean-variance" OR "portfolio choice" OR "optimal portfolio" OR "portfolio analysis" OR "modern portfolio theory" OR "diversification strategies" OR "markowitz" OR "risk parity" OR "volatility managed" OR "style integration") AND DOCTYPE(ar) AND SRCTYPE(j) AND LANGUAGE(English) AND SUBJAREA(ECON)	

3. Initial Results

A total of 473 papers on quantitative portfolio management were found when using the search terms and inclusion/exclusion criteria defined in Table 1. A mix of automated and manual systematic analysis was employed to identify a broad view of literature with respect to key dates, themes, citations, and phrases pertaining to quantitative portfolio management. Figure 2 shows that research on portfolio management has steadily increased since the 1970s, with a clear and strong upward trajectory since the turn of the century. We speculate that this trend occurred as a result of what is colloquially referred to as the technological boom of the 1990s, which caused an influx of investment in technological advancements [11]. This change in technology not only influenced financial markets but also created new directions for portfolio management research. For example, there appeared to be a significant increase in quantitative portfolio management research following the GFC in 2008 (Figure 2). This growth can be attributed to the increased demand for quantitative measures to improve the financial models, whose faults were deemed to be the reason for severe market failure [12]. The observed effects of the crisis were a revelation for the financial industry that more sophisticated modeling was required to account for the unexpectedly high asset correlations during the market downturn. The demand for improved statistical capabilities and advanced technological capabilities to run these models in financial markets has been growing since the GFC.

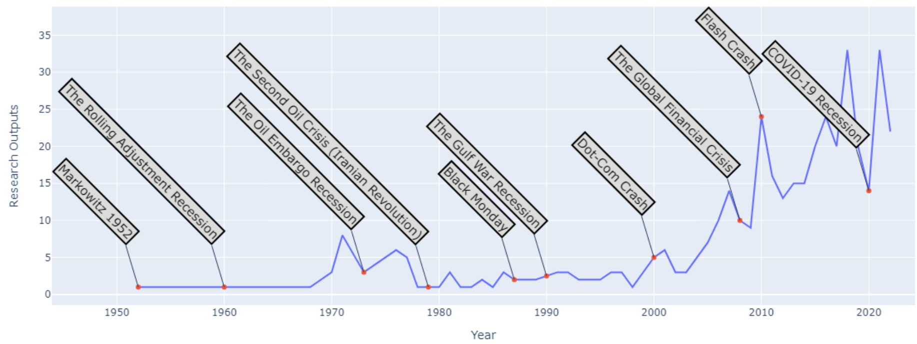

Figure 2. Portfolio management research outputs over time. From the final citation list, yearly quantitative portfolio management research outputs plotted alongside large market events to contextualize the academic progression of the field.

As investment strategies change, so too does portfolio management. The bibliometric analysis revealed four broad research themes within quantitative portfolio management over the time period studied (Figure 3). The themes are listed as they chronologically appeared in the literature, and they are not mutually exclusive:

1. Portfolio optimization
2. Risk-parity
3. Style integration (factor investing)
4. Machine learning

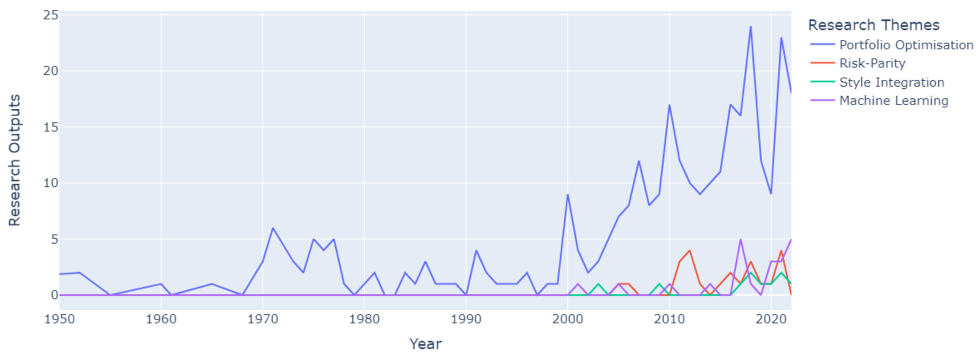

Figure 3. Portfolio management research outputs by theme over time. An adjustment of Figure 2 to contextualize the academic progression of each research theme that was identified.

Portfolio optimization has remained prevalent in the literature since its inception, marked by the seminal development of MVO by Markowitz [1], as seen in Figure 3. This endurance can be explained by the fundamental similarities between portfolio management and risk management, where the goal is to acquire an optimal risk-to-reward trade-off. While theoretically compelling, practical limitations mean that portfolios are almost never constructed without constraints or other preprocessing techniques.

The second theme, risk-parity strategies, has also been well-studied since its first mention in Qian [13] and has especially seen an influx in research activity following the GFC. Risk-parity portfolios attempt to balance risk contributions across asset classes by weighting them as inversely proportional to their risk. Interestingly, while risk-parity strategies have been used within the industry for a long time, the academic community only recently formalized the concept. For example, the first risk-parity fund, the All Weather asset allocation strategy, was launched by Bridgewater Associates in 1996 [14]. In contrast, the first academic analysis of the concept was only published by Qian [13] in 2005 and Maillard et al. [15] in 2010. Since the official recognition of risk-parity, it continues to be prevalent within the literature, as shown in Figure 3.

Moving past the GFC into the modern age of technology, there also grew a need to validate investment strategies with ever-growing datasets and investment opportunities. This was achieved by utilizing historical data for back-testing and less commonly simulated data as portfolio management shifted into the age of big data. Financial datasets grow in width faster than they grow in length, creating the need for practical dimensionality-reduction techniques that are capable of investing in tangible assets. To address this problem, investment funds harvest a unique source of returns packaged into a single asset, allowing portfolio managers greater risk management control. Style integration seeks to combine many styles into a single portfolio at a security level, which allows for trade differences to cancel out, in turn strengthening the performance. This intuitive concept has been met with skepticism, however, with some research suggesting there is no statistically significant difference in returns when compared to style mixing, which is the combination of preconstructed factor portfolios [16]. For example, consider a portfolio manager creating an equity investment portfolio that consists of exposures to two different styles/risk factors, there is "value" and "quality". With style integration, since implementation of the portfolio is performed at the stock level, the value style selects a long position in a specific stock such as Alphabet (GOOG), and the quality style selects a short position in the same stock (i.e., GOOG). As such, there will be no exposure in the investment portfolio to Alphabet, as trade differences cancel out, and no transaction costs are incurred, since the investment portfolio is constructed at the stock level. Alternatively, style mixing describes when the portfolio manager constructs an investment portfolio by purchasing a mix of readily available, pre-constructed risk factor or style portfolios such as a value ETF or a quality ETF rather than constructing the investment portfolio at the security level. Both of these ETFs would have incurred transaction costs, one in taking a long position in GOOG and the other in taking a short position in GOOG. As such, trade differences lead to increased transaction costs with no effective benefits, as the opposing trade positions cancel out.

Finally, ML is perhaps the most unique theme out of the four due to its ability to alter assumptions made by legacy financial theory and provide methods to quantify novel datasets. ML is a broad definition encompassing any statistical model capable of tuning its own parameters through learning processes without human intervention. While there is huge potential in utilizing ML for financial modeling, there has been minimal published work, likely due to the increased risk associated with unexplainable complexities (See Figure 3). ML, and its many subcategories discussed in detail later in this review, has impacted all other research themes within quantitative portfolio management.

Figure 4 displays a visual nodal mapping of the most-cited papers within the quantitative portfolio management literature. This figure shows a mapping of papers where the size of each node represents the number of citations, the connecting links represent where these citations are from, and the closeness of each node represents the degree of similarity. Despite changes in portfolio management research themes over time, the MVO model appears to still be highly topical across all fields of portfolio management, evidenced by Markowitz [1] being the most-cited author by a large margin. Another highly influential portfolio management paper was written by DeMiguel et al. [17], who empirically examined 14 portfolio optimization techniques to generate performance metrics for 8 different datasets. DeMiguel et al. [17] concluded that no optimization technique outperformed the 1/N portfolio, which raises the question: why spend so much time on developing these complex theoretical models if they do not work? Tu and Zhou [18] and Kirby and Ostdiek [19] contested DeMiguel's conclusion by asserting that his poor results were primarily due to research design, providing empirical evidence that, when relevant risks are considered, an investor can outperform 1/N. This contrary result brings hope to modern portfolio management and marks a new era.

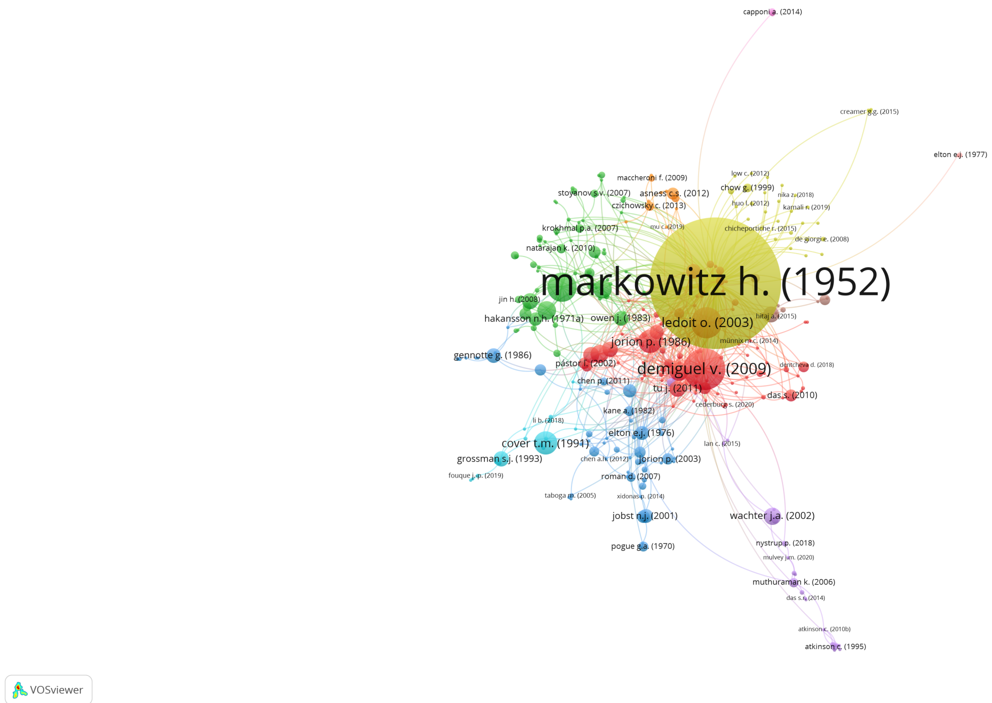

Figure 4. Mapping of the most-cited papers where arbitrary auto-generated clusters are highlighted with different colors [1,17,18,20–69].

Figure 5 shows a visual nodal mapping of key phrases used within the title or abstract of all papers examined, where the size of each node reflects the number of key phrase occurrences. The numbers of occurrences appear to be stable throughout the literature, with no nodes standing out. However, it is clear that "optimization" lies at the center of all portfolio management, and surrounding key phrases are used to guide the discussion, both in terms of content and terminology.

While this analysis is helpful, it does not provide the granularity required to meaningfully answer the two research questions outlined in Table 1. Accordingly, the following sections will extrapolate on these broad concepts to place them in the context of the current state of research.

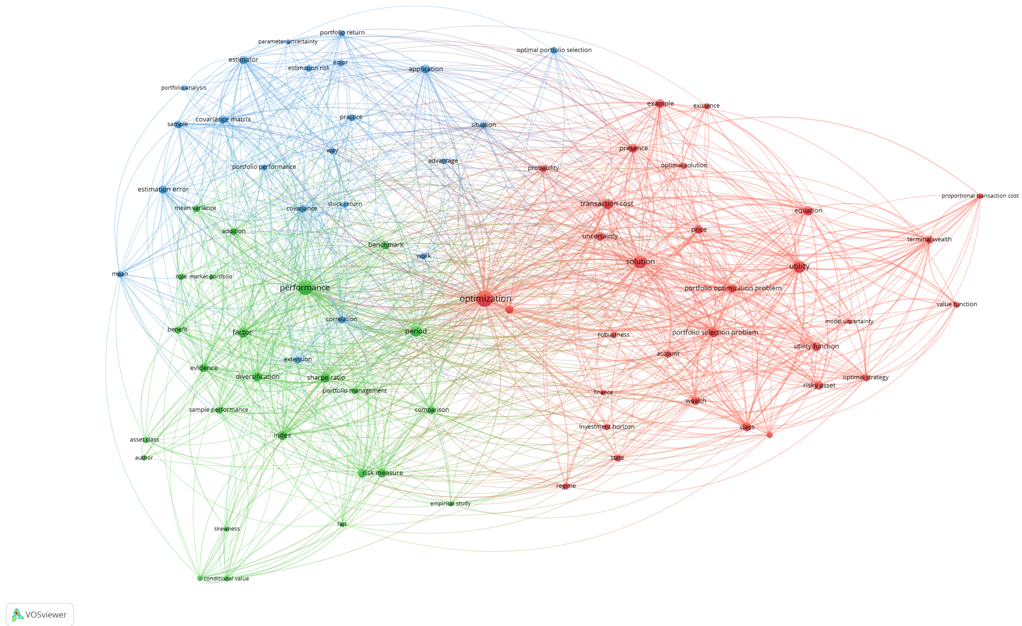

Figure 5. Mapping of key phrases in title and abstracts where arbitrary auto-generated clusters are highlighted with different colors.

4. Research Themes

4.1. Portfolio Optimization

Portfolio optimization is the process of defining a utility function to be used as a quantitative metric of investor satisfaction and then constructing a portfolio of assets such that this function is maximized. A loss function is the inverse of utility, where an investor wants to maximize their utility and minimize their loss. As expected, utility functions are unique to each investor to reflect their risk tolerance. While not utilized in this review, an array of literature explores human psychology within the context of financial markets, which can be used to refine investor utility. Although quantitative strategies can use individualized utility, the literature focuses on generalized utility estimates that can be applied across populations with only minor parameter changes. While a broad range of utility functions exist within the literature, the most notable from this review are outlined in Table 2.

Table 2. Utility function publications.

Study	Utility Function
Markowitz [1]	Mean-variance
Roy [70]	Safety first
Kelly [71]	Kelly criterion
Kahneman and Tversky [2]	Prospect theory (or loss aversion)
Gul [72]	Disappointment aversion
Sundaresan [73]	Habit utility
Abel [74]	Catching up with the Joneses
Knight [75]	Uncertainty aversion
Rockafellar and Uryasev [76]	CVaR optimisation

4.1.1. Mean-Variance Optimization

Mean-Variance Optimization (MVO), as established by Markowitz [1], assumes that investors are risk-averse, requiring greater reward for greater risks taken. Since the mid-20th century, portfolio managers and academics have used MVO to create optimized risk-adjusted portfolios, making it the most prolific utility function (Figure 2). The mean-variance function is defined as follows:

$$U = \mu - \frac{\gamma}{2}\sigma^2 \quad (1)$$

where μ and σ^2 are the mean and variance, respectively, and γ is a tuning parameter that represents an investor's risk aversion. This function incorporates the assumption that two assets with the same mean and variance of returns should be equally desirable to an investor. It is important to note that the function does not assume the normal distribution of asset returns, as is commonly mistaken. Mean-variance utility is arguably short-sighted in scope, lacking the ability to accurately estimate an investor's utility by failing to recognize that investors are likely interested in many other performance characteristics, e.g., skewness. The exploration of alternative utility functions has been prevalent within the literature to address these limitations.

4.1.2. Conditional Value-at-Risk Optimization

Value-at-Risk (VaR) is a downside risk measure quantifying the maximum amount of capital to be lost within a given timeframe with a certain degree of confidence. Under the condition that losses exceeding VaR signify an extreme downside event, Conditional Value-at-Risk (CVaR) is the expected shortfall defined by the mean of all conditional returns [76]. Notably, VaR has undesirable properties for optimization; e.g., it is not a coherent risk measure, so efficient linear programming techniques cannot be utilized [77]. Since the initial introduction of VaR as a risk measure, advancements in non-linear programming techniques have been made. However, they add unnecessary complexity. In contrast, CVaR satisfies all properties of a coherent risk measure and will always be equal to or greater than VaR, making it a reliable estimation of VaR. For loss function minimization, Rockafellar and Uryasev [76] suggest approximating CVaR using a Monte Carlo simulation from an estimated distribution:

$$\min_{(w,\alpha)} \alpha + \frac{1}{q(1-\beta)} \sum_{k=1}^{q} [-w^T r_k - \alpha]^+ \quad (2)$$

where

$$[t]^+ = \begin{cases} t \text{ when} & t > 0, \\ 0 \text{ when} & t \leq 0. \end{cases}$$

α represents VaR, β is the degree of confidence, q is the number of generated samples, w is portfolio weights, r_k is the kth vector of simulated returns, and 1 is a vector of ones. Low et al. [78] applies CVaR minimization to a portfolio of 12 US indices using a range of distribution estimation techniques to account for the higher asset correlation during bear markets. They find that CVaR optimization, complimented with a copula-based dependency structure, effectively targets down-side risk aversion and outperforms multivariate normal alternatives.

4.1.3. Kelly Portfolio Optimization

The Kelly criterion, first published by Kelly [71], formulates the optimal portion of wealth to place on a risky bet. When applied to asset allocation where there exists a universe of securities with a distribution of expected returns, the Kelly criterion can be rephrased as the optimization of long-term wealth. That is, a Kelly-optimal portfolio maximizes the expected log return. Applying this rule to a single-asset environment results in the following equation:

$$f^* = \frac{\mu - r}{\sigma^2} \quad (3)$$

where f^* is the fraction of wealth to invest, μ and σ^2 are the mean and variance of asset returns, respectively, and r is the risk-free rate of return.

Although the Kelly criterion aims to optimize an investor's utility, it conceptually departs from the traditional risk-to-reward trade-off that much of the literature focuses on. Although it may sound enticing to optimize long-term wealth, the fundamental lack of risk management in the unaltered Kelly criterion may lead to significant short-term drawbacks. It is unclear how long an investor must commit to such a strategy before it pays off, as the 'long-term' might be longer than the investor's lifespan. Despite these challenges, the Kelly criterion has been meaningfully implemented in a variety of different ways with added constraints [79].

4.1.4. Other Utility Functions

Safety-first utility [70] purports to invest only in safe assets until all liabilities are met, then significant risks can be taken for each additional dollar. Prospect theory [2] and disappointment aversion [72] are asymmetric utility functions that assume an investor will feel the pain of downside risk more than the pleasure of upside potential. Habit utility [73] assumes that once an investor's level of wealth, and therefore lifestyle, is ingrained as a habit, they become more risk-averse to avoid any lifestyle downgrades. The 'catching up with the Joneses' utility [74] defines risk relative to other investors; portfolio managers cannot control business cycles, but they can benchmark their performance to other market participants. Finally, uncertainty aversion [75] treats the accuracy of probability density function estimates as a risk metric.

Many utility functions beyond mean-variance are preferred frameworks for investor behavior. However, while these functions are often referenced as guiding concepts, they are rarely used directly for empirical validation. To bridge the gap between behavioral concepts and real-world strategies, research has been directed toward including higher-order moments in the portfolio construction procedure [2,78,80,81]. In particular, Adler [81] introduced full-scale optimization, a practical procedure designed to incorporate higher-order moments into the portfolio selection process. Adler [81] finds that full-scale optimization outperforms MVO, even out-of-sample, laying the foundation for practical optimization of higher moments.

Despite the compelling results of full-scale optimization, mean-variance utility has persisted as the dominant function in academia, primarily due to the existence of a closed-form solution and the convenience of implementing widely available linear programming methods. Mean-variance utility also provides a straightforward estimate of investor utility and can serve as the foundation of a broader strategy. As a result, the portfolio optimization literature has primarily focused on improving the limitations of MVO rather than on developing new utility functions.

The main downside of MVO is its underperformance in out-of-sample backtests, making it challenging to implement as a real-world strategy [17]. Primarily, this is due to the forward-looking nature of portfolio management. Future means and variances of asset returns are typically estimated using historical data, where the resulting estimation error leads to performance degradation. In addition, the transaction costs accrued from large swings in portfolio weights lead to poor out-of-sample performance. Improving these limitations has been a large focus of the MVO literature since its inception, where Table 3 identifies major contributions to the literature that seek to improve estimation error.

Many of the estimation improvements in Table 3 come from shrinking the optimization towards a trusted reference point. These methods are discussed below; however, a notable alternative approach developed by Demiguel [82] and Brodie [83] is to add a regularization penalty within the optimization procedure to stabilize the output weights. Regularization of MVO weights has resulted in lower transaction costs and better risk-adjusted performance that persists out-of-sample [82]. Due to its strong results, regularization is widely used in industry [84].

Table 3. Selected estimation error publications.

Study	Contribution
Black and Litterman [85]	Black–Litterman model
Jorion [86,87]	Bayes–Stein approach
Pastor [88], Pastor and Stambaugh [89]	Data-and-model
MacKinlay and L'uboš Pástor [90]	Missing factor model
Kan and Zhou [91]	Kan and Zhou three-fund portfolio
Tu and Zhou [18]	Combined portfolios
Demiguel [82], Brodie [83]	Regularization of portfolio weights

In addition to estimation error, there is the risk of over-fitting a portfolio strategy to historical data such that the backtested performance is an unreliable representation of likely future performance. Although only one historical sequence occurs in reality, numerous other probable historical scenarios could have unfolded but did not. Using Monte Carlo simulations to generate market data, a portfolio manager is able to evaluate the strategy in a wide array of 'what if' scenarios. Should the strategy perform poorly in the occurrence of a probable future market event, it can be modified to account for such scenarios. However, generating realistic market data has challenges of its own, as there will continue to be estimation errors when specifying the distribution of simulated data. Nonetheless, market simulations are a helpful tool utilized by many of the empirical studies referenced in this review.

4.1.5. Bayesian Estimation

Ever since the seminal work of Bayes [92], Bayesian estimation techniques have been utilized across virtually all research fields. The idea is to iteratively update an estimated Probability Density Function (PDF) as new information becomes available, shrinking the prior to the posterior by minimizing loss, which is usually a function of residuals. When implemented on large datasets, the posterior will likely shrink to a close approximation of the true PDF, regardless of the initial prior. However, financial time series datasets are notoriously short and high in dimensionality, causing the initial prior to impact the resulting posterior significantly [17]. Therefore, the financial literature has implemented Bayesian methods with this subjectivity in mind.

The Bayes–Stein approach developed by Jorion [86,87] is a portfolio application of the James–Stein estimator theorized by Stein [93] and James and Stein [94] which attempts to minimize the estimation error of mean asset returns. The sample mean is shrunk to the sample minimum-variance portfolio mean, effectively limiting the exposure to mean estimation risk by diversifying into the minimum-variance portfolio where no mean estimation is needed. Estimating mean return is notoriously more difficult than return variance, so the main contribution of the Bayes–Stein approach is to minimize the mean estimation risk. However, Jorion also uses traditional Bayesian estimation techniques to minimize the variance estimation risk.

The data-and-model approach developed by Pastor [88] and Pastor and Stambaugh [89] refines the Bayes–Stein approach by using an asset pricing model to estimate future asset returns, which are shrunk to the sample mean via a ratio of model mispricing variance and highest Sharpe ratio of factor portfolios. This refined methodology essentially allows an investor to select an asset pricing model and their prior belief of its performance, then these subjective inputs are fine-tuned via empirical observation.

The Black–Litterman model, developed by its namesake in [85], shrinks investor beliefs about future asset returns towards a benchmark portfolio. While older and less sophisticated than the data-and-model approach, the Black–Litterman model is practical for real-world applications and allows for greater flexibility, with subjective inputs applied directly to constituent assets. An investor can pick a benchmark portfolio, then input any divergent beliefs about future asset returns and corresponding uncertainty metrics into the model to generate posterior means and variances.

Although these Bayesian methods are conceptually compelling and improve sample-based MVO, they continue to underperform relative to the 1/N portfolio during out-of-sample backtests [17]. However, the Black–Litterman model is an exception, where the accuracy of investor beliefs primarily drives performance [95]. An alternative method of improving estimation risk is treating it no differently than any other risk factor, such as the diversification of uncorrelated assets. Kan and Zhou [91] formalize this idea by creating a three-fund portfolio where investments are divided across a risk-free asset, the tangency portfolio, and the minimum-variance portfolio. As a result, the estimation error for mean asset returns is diversified because the minimum-variance portfolio only needs to estimate variances. In addition, performance is significantly increased compared to sample-based MVO, providing evidence that estimation risk should be diversified.

Tu and Zhou [18] expands on this idea to combine other sophisticated Markowitz-related strategies with the 1/N portfolio. Although promising results are found, beating the 1/N portfolio remains an elusive task. Some have contested the idea that MVO is not practically useful by asserting that estimation error is largely caused by the over-reliance on rolling short-term samples [96]. However, large time frames are not always obtainable, so portfolios are often implemented with constraints, or unconstrained optimization is shrunk to a constrained portfolio, as is the case in Kan and Zhou [91] and Tu and Zhou [18].

4.1.6. Multi-Period Optimization

All the portfolio optimization techniques discussed thus far are fundamentally single-period methods. In single-period approaches, optimal portfolio weights are applied for the current rebalancing period only, disregarding inter-temporal dependencies. Although the greedy approach may still be optimal in cases with optimal substructure, path independence, or myopic investor utility, these assumptions often do not hold in real-world scenarios [97].

To solve for the multi-period portfolio, dynamic programming is utilized, working backwards from the portfolio's termination date to the present. A commonly used analogy is to imagine an agent randomly placed on a grid tasked with reaching a reward square by moving one square each period. Working backwards from the reward square, the utility of occupying each square is calculated iteratively by discounting the utility of previous adjacent squares. Once the utility of each square has been calculated, the agent can take one step to occupy the adjacent square with the highest utility, which is also one step closer to the reward square.

This concept, formalized with the Bellman equation from Bellman [98], becomes a bit more complex when the non-deterministic nature of financial markets is included. The stochastic Bellman equation is given as follows:

$$V(s) = \max_a \left[R(s,a) + \gamma \sum_{s'} P(s'|s,a) V(s') \right] \tag{4}$$

where $V(s)$ is the value of state s, $R(s,a)$ is the reward for taking action a in state s, γ is a discount factor, $P(s'|s,a)$ is the probability of arriving in state s' after taking action a from state s, and $V(s')$ is the value of state s'.

The dynamic programming solution is theoretically sound. It also presents new challenges to overcome. Notably, multi-period portfolio decisions are path-dependent, meaning that decisions made today will effect the available opportunities tomorrow. Therefore, if an unexpected structural market change occurs, a multi-period portfolio may be permanently impaired. Moreover, dynamic programming solutions often suffer from the "curse of dimensionality," rendering most multi-period solutions computationally intractable. Nonetheless, approximate dynamic programming and simplified multi-period approaches have shown promise in specific contexts where they are applied, especially goal-based robo-advisors [99,100].

4.2. Risk-Parity

Risk-parity is a constrained mean-variance portfolio where risk is equally distributed across asset classes, similar to how an investor might diversify capital equally across assets with the 1/N portfolio. In the same way that the minimum-variance portfolio is constrained to eliminate the estimation of mean returns, risk-parity is further constrained to eliminate the covariance estimates, leaving only the individual asset return variances. Although risk-parity is not a new concept, there is a scarcity of literature on the topic prior to its formalization by Qian [13]. Risk-parity has since been empirically tested. Maillard et al. [15] used global data across 13 asset classes and concluded that risk-parity outperforms the 1/N and minimum-variance portfolio on a risk-adjusted basis, highlighting the need for further research. Notable risk-parity contributions observed in this review are shown in Table 4 and explored below.

Table 4. Risk-parity publications.

Study	Contribution
Qian [13]	Formalisation of risk-parity
Bhansali [101]	Risk factors used instead of assets
Kirby and Ostdiek [19]	Volatility timing and reward-to-risk timing
Bhansali et al. [102]	Factor risk-parity
DeMiguel et al. [103]	Conditional mean-variance multifactor portfolio

4.2.1. Volatility Timing (VT)

Kirby and Ostdiek [19] developed a simple risk-parity framework by weighting assets inversely to their volatility on any given time step. Therefore, as asset variance increases, the weighting to that asset will shrink quadratically. A tuning parameter is added to modify how aggressively this strategy responds to volatility shocks, which indirectly tunes turnover and transaction cost tolerance. The formalization of VT can be represented as follows:

$$\hat{\omega}_{it} = \frac{(1/\hat{\sigma}_{it}^2)^\eta}{\sum_{i=1}^{N}(1/\hat{\sigma}_{it}^2)^\eta} \tag{5}$$

where $\hat{\omega}_{it}$ and $\hat{\sigma}_{it}^2$ are the estimated weight and volatility of an asset i at time t. N is the number of assets to be considered, and η is the volatility sensitivity parameter. Kirby and Ostdiek [19] implemented this strategy using US equities, and it outperformed the 1/N portfolio on a risk-adjusted basis. The simplicity of VT provides a promising starting point for further research.

4.2.2. Reward-to-Risk Timing (RRT)

Reward-to-risk timing (RRT), also developed by Kirby and Ostdiek [19], is an extension to the VT strategy. Under the RRT framework, asset weights are in a ratio to mean volatility and returns; an increase in volatility will not affect asset weights, provided there is a proportional increase in expected returns. Reintroducing mean asset returns creates extra estimation risk, potentially resulting in similar limitations identified in the portfolio optimization section above. To minimize estimation risk, mean asset returns are constrained to positive values, which is formalized as follows:

$$\hat{\omega}_{it} = \frac{(\hat{\mu}_{it}^{+}/\hat{\sigma}_{it}^2)^\eta}{\sum_{i=1}^{N}(\hat{\mu}_{it}^{+}/\hat{\sigma}_{it}^2)^\eta} \tag{6}$$

where $\hat{\mu}_{it} \geq 0$ is the mean returns of asset i at time t. This is a very intuitive strategy, where Kirby and Ostdiek [19] conclude that it shows more promise than VT on their dataset of US equities. The results persisted when implemented on UK equities when transaction costs were considered [104].

4.2.3. Factor Risk-Parity (FRP)

The original concept of spreading risk across asset classes makes little sense if these asset classes are strongly correlated to each other compared to alternative asset clustering techniques (Hierarchical Risk Parity (HRP) is also a prominent implementation of risk-parity which uses hierarchical clustering techniques in place of traditional risk factors; see the machine learning section below for more details). There is an abundance of research focusing on the discovery of uncorrelated risk factors that should increase diversification benefits compared to the naive assumption that separate asset classes are uncorrelated. Bhansali [101] explores this idea and concludes, among other things, that risk factor clustering is essential for a risk-parity portfolio to be practical.

Bhansali et al. [102] implemented FRP using nine asset classes and Principal Component Analysis (PCA) for factor identification. They found that asset-based risk parity results in portfolios concentrated on few risk factors that are not sufficiently diversified. Moreira and Muir [105] follows a similar method and applies FRP to a suite of traditional factor portfolios with promising results. However, further research by Cederburg et al. [69] and Barroso and Detzel [106] argues that Moreira's results underperform out-of-sample and do not survive friction costs. These are the same issues faced when implementing MVO portfolios, casting doubt on the efficacy of FRP.

4.2.4. Conditional Mean-Variance Multifactor Portfolio

To overcome the limitations associated with FRP, a Conditional Mean-Variance Multifactor (CMVM) portfolio was developed by DeMiguel et al. [103], where factor portfolio weights change via a parametric function of market volatility. The parameters are optimized using a modified mean-variance utility function adding a transaction cost value, which can be expressed as follows:

$$\min_{\eta \geq 0} \quad \frac{\gamma}{2}\eta^\top \hat{\Sigma}\eta - \hat{\mu}^\top \eta + \mathrm{TC}(\eta) \tag{7}$$

where γ is a risk-aversion parameter, $\hat{\mu}^\top \eta$ and $\eta^\top \hat{\Sigma}\eta$ are the extended mean and variance of portfolio returns, and $\mathrm{TC}(\eta)$ is a function for transaction costs (transaction costs are strictly defined in DeMiguel et al. [103], who borrow the bid–ask spread method from Abdi and Ranaldo [107]. However, the transaction cost method by itself is not a novel idea, and a portfolio manager can implement any function they please). To unpack this optimization into its different components, consider the weight of factor k at time t, such that

$$\theta_{kt} = a_k + b_k \frac{1}{\sigma_t} \tag{8}$$

where a_k and b_k are affine function parameters to be optimized, and σ_t is the variance of the market risk factor returns. Therefore, the portfolio returns can be expressed as follows:

$$r_{pt} = \sum_{k=1}^{K} r_{kt}\theta_{kt} = \sum_{k=1}^{K} r_{kt}\left(a_k + b_k \frac{1}{\sigma_t}\right) \tag{9}$$

where r_{kt} is the returns for factor k at time t. Finally, portfolio returns can be expressed in matrix form to demonstrate the usage of the extended factor-weight vector η seen in Equation (7). Portfolio return is derived in matrix form as follows:

$$r_{pt} = r_{ext,t}^\top \eta \tag{10}$$

where

$$r_{ext,t} = \begin{bmatrix} r_{1,t} \\ r_{2,t} \\ \vdots \\ r_{K,t} \\ r_{1,t}\dfrac{1}{\sigma_t} \\ r_{2,t}\dfrac{1}{\sigma_t} \\ \vdots \\ r_{K,t}\dfrac{1}{\sigma_t} \end{bmatrix}, \quad \text{and} \quad \eta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_K \\ b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix}. \tag{11}$$

The CMVM portfolio was implemented using nine risk factor portfolios constructed using US equities. Compared to other risk-parity strategies, the CMVM performed well, particularly in periods of high market volatility. This provides an optimistic outlook for the future of FRP strategies and outlines a potential future direction of research.

4.3. Style Integration

Investment styles and factors are often used interchangeably to denote unique sources of risk impacting any given asset. Many asset pricing models use this concept to calculate the price of an asset as the combined exposure to each risk factor (see CAPM [108,109], Fama-French Three-Factor [110], Arbitrage Pricing Theory [111], Carhart Four-Factor [112], Fama-French Five-Factor [113]). In addition, some investment funds specialize in a single investment style to harvest the returns of a given risk factor. However, if an investor wants diversified exposure to various investment styles, they have two options. First, they can apply optimal weights to many pre-constructed specialist investment funds, dubbed style mixing; or second, they can construct an array of style portfolios from a universe of assets and then integrate each portfolio such that trade differences of individual assets cancel out, dubbed style integration [114].

Style integration was first introduced by Frazzini et al. [115], who hypothesized its inherent benefit of lower transaction costs and therefore better performance compared to style mixing. Three US risk factor portfolios were individually analyzed and compared against an integrated portfolio of these same factors. They concluded that style integration provides excess risk adjusted returns, marking the inception of another potential portfolio strategy. Some doubt has since been cast on the claim that style integration provides lower transaction costs compared to style mixing; an intuitive concept that has been empirically challenged by Leippold and Rueegg [16].

Any portfolio construction strategy that combines many sub-portfolios can be implemented as a style integration approach, provided that the asset weights of each sub-portfolio are known. For example, an investment fund may not disclose its holdings, making style integration impossible. Many portfolio management studies are applied directly to a universe of assets or rely on preconstructed factor portfolios, making their method unsuited for style integration. However, existing methods often require minimal refinements to conform to the style integration framework. This means that many familiar strategies are yet to be tested and formalized in a style integration framework. For style integration strategies that have been formalized in the literature, Fernandez-Perez et al. [116] provide a comprehensive empirical analysis by applying five styles to a range of US futures contracts, currency futures, and US equities for robustness checks. Performance summaries of style integration strategies will be in reference to this study, with relevant contributions shown in Table 5.

Table 5. Showcased style integration studies.

Study	Contribution
Fitzgibbons et al. [114]	Equal Weight Integration (EWI)
Brandt et al. [117] Fischer and Gallmeyer [118] Ghysels et al. [119] DeMiguel et al. [120]	Optimal Integration (OI)
Barberis and Shleifer [121]	Rotation-of-Styles Integration (RSI)
Fernandez-Perez et al. [116]	Volatility Timing Integration (VTI) Cross-Sectional Pricing Integration (CSI) Style Momentum Integration (SMI) Principal Components Integration (PCI)

4.3.1. Equal Weight Integration

Fitzgibbons et al. [114] first introduced the Equal Weight Integration (EWI) to equally weight all styles, $w = 1/K$, where K is the number of investment styles being investigated. The intuition of this strategy is similar to the 1/N portfolio, where an investor attempts to diversify equally across the universe of assets. If the universe of assets to be explored does not have homogeneous risk factor exposure, the 1/N portfolio will overweight the more pervasive risk factors, as they will be prominent in a large percentage of the total assets. EWI solves this problem by equalizing the exposure of each risk factor, then it integrates all factors into a single portfolio constituting the universe of assets. Much like the benchmark usage of the 1/N portfolio for optimization, the EWI strategy is used as the benchmark for other style integration techniques.

4.3.2. Optimal Integration

The Optimal Integration (OI) strategy formalized in Fernandez-Perez et al. [116] optimizes investor utility by maximizing the function of expected excess style returns, $E_t[U(\sum_{k=1}^{K} \omega_k r_{k,t+1})]$, where ω_k is the weight of kth style portfolio and $r_{k,t+1}$ is the expected return at time $t+1$ (Brandt et al. [117], Fischer and Gallmeyer [118], Ghysels et al. [119], DeMiguel et al. [120]). The core idea is to optimize a set of style portfolio investments, so any optimization strategy discussed above could be adapted for use under an OI framework. While OI should theoretically be more robust than direct asset optimization, estimation risk still limits portfolio performance. Although OI significantly underperforms relative to the EWI benchmark, there exists room for improvement, as many optimization techniques are yet to be explored within the OI research stream.

4.3.3. Style Momentum Integration and Rotation-of-Styles Integration

The Style Momentum Integration (SMI) and Rotation-of-Styles Integration (RSI) strategies, introduced by Fernandez-Perez et al. [116] and Barberis and Shleifer [121], respectively, assume style performance momentum and allocate all available capital to the style with the highest excess return and Sharpe ratio, respectively, obtained from the previous time step. These strategies would likely be impractical if constructed using style mixing due to the necessity to liquidate the whole portfolio before rebalancing. However, due to the style integration framework, SMI and RSI have comparable turnover to other style integration strategies, making them viable. SMI performs marginally better than RSI, which both marginally underperform the EWI strategy. The core idea is to find the highest-performing style at the current time step, then allocate all capital to it. Although many variations of these strategies are yet to be tested, they will likely only be useful if future style utility can be estimated accurately.

4.3.4. Cross-Sectional Pricing Integration

The Cross-Sectional Pricing Integration (CSI) strategy [116] identifies the explanatory power of style portfolios on the cross-section of expected returns of a universe of assets, and styles with greater explanatory power are weighted more heavily. Fernandez-Perez et al. [116] implemented CSI by running a simple Ordinary Least Squares (OLS) regression over the previous 60 months of futures contracts with K style portfolios as independent variables, then the estimated style betas were used for another OLS regression where a function of the resulting R^2 was used to calculate style weights. CSI underperforms relative to EWI, SMI, and RSI, but due to there being many moving parts, CSI might be improved by minor methodology changes (e.g., non-linear regressions).

4.3.5. Principal Component Integration

Fernandez-Perez et al. [116] introduced the Principal Components Integration (PCI) strategy, which finds m principal components that explain at least τ percent of excess style portfolio returns then computes style weights as a function of the explanatory power of the eigenvectors. Reducing investment decisions into their constituent styles, as is the goal of style integration, is effectively a dimensionality reduction technique. However, this does not guarantee optimal reduction, so the core idea of PCI is to further reduce style return dimensionality by isolating the most important principal components. PCI has the highest turnover of style integration strategies reviewed and underperforms compared to EWI, SMI, RSI, and CSI. Only five styles were explored by Fernandez-Perez et al. [116], so it is hypothesized that performance may improve as the number of style portfolios increases due to the greater needs of dimensionality reduction.

4.3.6. Volatility Timing Integration

Fernandez-Perez et al. [116] introduced the Volatility Timing Integration (VTI) strategy, which weights styles inverse to their historic variance; a style integration approach to the Kirby and Ostdiek [19] VT model discussed above. Adopting a key benefit of a risk-parity approach, VTI has the lowest turnover of any style integration strategy explored. However, the overall performance is disappointing with negative Sharp, Sortino, and certainty equivalent ratios. VTI is the worst-performing style integration strategy explored by a significant margin. Style integration performance should roughly match the underlying model, so the robust multifactor risk-parity approaches explored above would likely improve performance.

4.4. Machine Learning

Machine Learning (ML) is a field of statistics that allows computers to learn patterns contained within data. For example, Ordinary Least Squares (OLS) is often considered a rudimentary machine learning model, as it learns the provided dataset by minimizing the sum of squared errors. More robust ML models exist, each with their contributions and drawbacks depending on the type of problem a practitioner wants to solve. The more complex models often have the benefit of less model bias; that is, less reliance on human-defined behavior. However, they suffer from high variance; that is, overfitting to noisy data due to overparameterization. Academia has been slow to adopt many novel ML techniques due to their unconstrained nature and perception of behaving as a black box. de Prado [4] challenges these perceptions and suggests that ML will be an invaluable tool to process complex datasets that have not been available in the past.

It would be impractical to document an exhaustive list of ML models used for portfolio management due to the large number of different models and variations thereof. Instead, an overview of topical models is provided, along with some examples of their usage. In brief, there are three types of ML models: supervised learning, unsupervised learning, and reinforcement learning. Supervised learning models, such as elastic-nets and random forests, require labeled data to generate regression or classification outputs. Unsupervised models, such as Convolutional Neural Network (CNN) and hierarchical clustering, can identify unlabeled patterns without intervention. Finally, Reinforcement Learning (RL) models are designed to interact with their environment, with positive (negative) reinforcements applied in response to desirable (undesirable) behaviors. Notable ML publications included in this review are outlined in Table 6 and discussed below.

Table 6. Showcased machine learning studies.

Study	Contribution
Zou and Hastie [122]	Elastic-net
Ho [123]	Random forests
Chen and Guestrin [124]	Extreme gradient boosting
Lundberg and Lee [125]	Shapley additive explanations
Jiang et al. [126]	Reinforcement learning portfolio optimization
Raffinot [127]	Hierarchical clustering
Jiang et al. [128]	Computer vision price signaling
Elkind et al. [129]	Panic selling behavioral analysis

4.4.1. Elastic-Net

The elastic-net model developed by Zou and Hastie [122] is a regularized OLS model that combines lasso and ridge regression penalties to desensitize the model to unimportant variables. The variable penalties λ_1 and λ_2 are often found by executing a k-fold cross-validation of data such that λ minimizes model variance. The primary difference between lasso and ridge regression is that lasso regression can shrink variable coefficients to 0, while ridge regression can only shrink variable coefficients asymptotically close to 0. Elastic-nets can replace OLS wherever linear or logistic regression is required; supervised learning methods are particularly useful to find accurate factor coefficients for asset pricing models and to calculate asset weights for hedge fund replication [130].

4.4.2. Random Forest (RF)

RFs, as introduced by Ho [123], attempt to predict a dependent variable via a flowchart-style decision tree, where each node is a binary selector [131]. The prediction power of a randomly selected subset of independent variables and an optimal binary split on the data is evaluated to generate a random forest node. Many different functions can be used to obtain variable predictive power, also depending on whether a regression or classification tree is used. Regardless of the function used, the concept is to select a variable and split it to maximize the information gain at each node. The decision tree is built in succession, repeating these steps to generate new nodes until the specified maximum length is reached. After one decision tree is generated, the process repeats itself an arbitrary number of times, and all models are combined into an ensemble. When data are passed to the final model, every decision tree will produce an output, and the average of each output is used as the final result, or for classification, the label with the highest number of votes is selected.

Decision trees, and therefore RFs, are non-linear, potentially providing higher-accuracy models compared to elastic-net and other linear models. However, in practice, RFs tend to overfit to training data because of the noisy financial data. Zhu et al. [132] suggest that a model should be selected based upon preliminary data exploration and domain knowledge, and they outline a new hybrid model that attempts to exploit the benefits of linear and non-linear models.

4.4.3. Extreme Gradient Boosting (XGB)

XGB builds upon decision trees to efficiently generate gradient-boosted models [124]. XGB also includes optimizations to further improve the performance of gradient boosting alone. For regression, an initial average value is generated from the training set, and residuals are recorded. Then, a decision tree is generated using the residuals as leaf nodes. The initial average is added to the predicted residual, multiplied by a learning rate. The residuals of these results are recorded, and the previous steps are repeated until the model's performance is satisfactory. XGB is a recently developed model, and while it does efficiently generate decision tree models, there is still room for improvement. Some literature has proposed that XGB should be adapted to include evolving data streams, the new adaptation being coined Adaptive XGBoost (AXGB), which poses an interesting alternative due to its efficiency in performance, training time, and memory usage [133].

4.4.4. Shapley Additive Explanations (SHAP)

Although sophisticated ML models can provide accuracy unmatched by traditional statistics, without an explanation of how the model works, it is unlikely to be practically implemented. Stakeholders may not want to take on the risk of an unexplained model, or regulators may simply ban the usage of unauditable development pipelines. The model explainability research stream addresses this problem, with SHAP being a particularly prolific factor importance metric.

Shapley [134] first introduced Shapley values as a game theory method to fairly distribute payoffs within a coalition. Fairness is derived by quantifying the impact of each actor on a final outcome, then gains and losses can be distributed as a ratio of individual contribution to all other actors. Particular subsets of actors may hold relationships such that different permutations of actors will have different results. Therefore, every permutation needs to be generated, then the mean of actor impact is taken, which can be formalized as follows:

$$\phi_j(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n-|S|-1)!}{n!}(v(S \cup \{i\}) - v(S)) \tag{12}$$

where N is a coalition of actors and S is a subset of those actors. i is an individual to identify their marginal contribution, n is the number of actors, and $v(S)$ represents the contribution of S.

Data scientists noticed that this game theory approach can be applied directly to statistical and AI models by replacing the coalition of actors with a vector of independent variables. However, the Shapley equation is computationally expensive in exponential time, making it unrealistic to implement for large datasets. To address this problem, Lundberg and Lee [125] proposed the SHAP method to approximate Shapley values in polynomial time. Although the exact implementation of Shapley value generation may differ depending on the software packages used, the core idea stays the same and has been a leap forward for explainable modeling.

4.4.5. Hierarchical Clustering

Hierarchical clustering is an unsupervised method of defining the cross-section of expected returns. The concept is to group similar assets together without explicitly defining the groups a priori. Prado [135] introduces hierarchical clustering within the context of asset management by developing the Hierarchical Risk Parity (HRP) portfolio, and Raffinot [127] explains that an accurate asset distance (dissimilarity) function must be derived for clustering to be effective. Although the distance function can be intuitive for broader data science contexts (e.g., simple Euclidean distance), financial applications may benefit from distance functions bespoke to investor preferences. After a distance function is defined, assets are either placed into a single cluster then iteratively split orthogonality from each other until the desired threshold of total clusters is found, or each asset starts with its own cluster that is then iteratively paired with others. Raffinot [127] builds a robust set of hierarchical clustering techniques along with popular portfolio management strategies and tests them empirically against US equities. They conclude that hierarchical clustering offers better diversification and risk-adjusted performance compared to traditional risk factor-based allocation.

4.4.6. Computer Vision Price Signaling

A CNN is a neural network architecture commonly used to classify images, which was first developed by Fukushima [136]. CNNs classify images by applying convolutions; that is, applying filters to an input image to find patterns. This is done in a computationally efficient manner by sampling a subset of pixels from the input image via a rolling window. Any number of filters can be stacked and applied to this subset of pixels. A filter is applied by multiplying an image pixel by the corresponding filter pixel, then the average value of all pixels within the rolling window is taken and mapped to a new image, which will be the output of the convolution layer. A non-linearity layer, usually a Rectified Linear Unit (ReLU) function, is then applied to normalize outputs to positive values. Next, the CNN applies pooling to shrink the size of convolution outputs to allow for efficient computing and increased generalization of filters [137]. This is done by applying a rolling window to the convolution output of a size greater than 1×1. The highest number within the window is selected, which generates another image to be used as the pooling output. The layers also do not necessarily need to be in the order outlined above; they can be mixed for different applications. For example, if processing a particularly low-resolution image, it may not be desirable to implement a pooling layer after every convolution and normalization.

CNNs have predominantly been used in finance as a method of time-series price signaling. This is achieved by passing historic asset chart images, along with known $t+1$ prices for data labeling, to the CNN, which can be interpreted in a 'human-like' manner. Jiang et al. [128] applied this concept to US stocks backtested from 1993–2019 and found that a CNN can provide more accurate pricing predictions than traditional risk factor analysis. These predictions also persisted through time scales and geolocations without the need to fine-tune the model, and pricing predictors were largely unique to traditional risk factors, adding portfolio diversification benefits. CNNs show large potential in discovering novel risk factors that can ultimately be used to develop efficient portfolios and better risk management.

4.4.7. Reinforcement Learning (RL)

RL models, such as Q-learning and Deep Reinforcement Learning (DRL), are designed to interact with their environment, with positive (negative) reinforcements provided in response to desirable (undesirable) behaviors. RL models are commonly used for the complete removal of financial models from the domain they operate. For example, to optimize a portfolio using an RL model, a simulation environment is provided without any prior knowledge of asset pricing or portfolio construction theory, and desirable outcomes are rewarded until the optimal strategy is found. Given a sufficiently sophisticated model with a well-designed utility function, RL can optimize any number of inputs to achieve an optimal outcome. Common limitations of RL include the difficulty of defining a utility function that advantageously incentivizes long- and short-term rewards and constructing an accurate simulation environment for training.

Much of the literature surrounding portfolio optimization using RL is not applied to traditional asset markets or published in top journals, likely because of its unorthodox nature and not due to a lack of potential. Jiang et al. [126] built a RL model for cryptocurrency markets using historical return data for available coins as a simulation environment. They defined utility as a function of expected returns and transaction costs. The RL portfolio outperformed all other benchmarks and achieved 400% returns within a 50-day period. However, it is difficult to analyze cryptocurrency risk due to the market's relative immaturity. Everyone stands to make money when the market goes up; however, this does not necessarily mean that good investment decisions were made.

Alternatively to optimizing a utility function directly, Cheng [138] uses RL to implement a style integration strategy for cryptocurrencies in which any number of constituent portfolios can be considered for selection. The optimal style weights are found via a series of sophisticated preprocessing steps, which are passed to an agent to make the final decision at each period. Similarly to Jiang et al. [126], Cheng [138] shows very promising results, with significant opportunities for future research. However, it remains difficult to compare legacy portfolio strategies on a risk-adjusted basis. Although RL should rightly be met with caution, it will likely strongly influence the future of portfolio optimization profoundly due to its unparalleled flexibility.

4.4.8. Panic Selling Behavioral Analysis

Panic selling is an intuitive behavioral concept whereby neurotic investors sell off large positions during times of poor market performance. Although there is a healthy pool of literature dedicated to behavioral finance, little has focused on quantitatively defining panic selling to obtain potentially useful insights for portfolio managers. Elkind et al. [129] fills this gap in the literature by analyzing individual brokerage accounts to predict investors who are likely to 'freak-out' in the near future. A novel dataset consisting of demographic attributes, such as marital status and self-declared investment experience, was used to train a Deep Neural Network (DNN) and a logistic classifier. The results of this study demonstrated that there is indeed a disproportionate number of investors who panic sell during market stress, and subtle indicators can identify these individuals ahead of time. Furthermore, the DNN outperformed linear classification across all provided metrics.

DNN models were first theorized by Ivakhnenko [139], where multiple layers of interconnected neurons are trained to minimize the output error by modifying the signal strength between a connected pair of neurons. DNN algorithms have significant potential to solve novel problems requiring non-linear relationships to be found due to their ability to approximate any continuous function [140]. However, a major limitation of DNNs is the need for large datasets to generalize to the given task appropriately. Although this problem is not unique to DNNs, its significant number of parameters and large degrees of freedom exacerbate the need for large and reliable datasets. However, lack of data availability is a shrinking problem, so DNNs will likely play an influential role in the future of quantitative portfolio management. They will likely be used primarily to analyze novel datasets to find approximate solutions to intuitive concepts that are hard to quantify, such as behavioral trends.

5. Conclusions

Effective portfolio management is essential for the long-term stability of any institution. Two key components drive performance: information and the processing of information. This review highlights the capabilities of quantitative strategies to collect and process information, suggesting the superiority of data-driven approaches over human intuition. Although the performance of a new investment strategy is difficult to quantify before it is implemented in the live market, many publications back-test their strategy on historical data and use the 1/N portfolio as a benchmark. Most investment strategies underperform relative to the 1/N portfolio on a risk-adjusted basis, and for the ones that outperform, it is unclear whether performance will persist when others replicate the strategy post-publication. However, optimists argue that underperformance is often due to research design and make the point that investors are happy to pay a premium for uncorrelated returns, making the 1/N portfolio benchmark gratuitous.

Keeping these arguments in mind, 473 quantitative portfolio management publications were analyzed using a combination of Scopus search queries and tool-assisted investigations, where four key research themes were identified: portfolio optimization, risk-parity, style integration, and machine learning. Many profitable and conceptually compelling models have been proposed, which provide valuable insights into the inner workings of financial markets. However, these models are limited by the assumptions made by their guiding economic theory, and they often oversimplify the statistical properties of asset returns, leading to parameter estimation errors. Although techniques to minimize estimation error can be utilized, the best-performing models treat estimation error as another risk to be diversified. Alternatively, model restrictions can be introduced, such that any parameters with large errors are simply removed.

Either way, model flexibility is sacrificed. Machine learning allows portfolio managers to loosen model assumptions or even do away with legacy financial theory completely. While black boxes hold greater governance risk, data peculiarities (e.g., non-linearity) are better modeled with machine learning and ultimately show promising results. Until recently, the literature has regarded machine learning with a justified degree of caution, often over-relying on legacy financial theory. The literature will continue to do so until novel techniques establish frameworks and garner trust within the broader investment community. As machine learning matures and begins to be trusted, journals will continue to loosen their requirements of strict theoretical backgrounds in favor of empirically driven approaches.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Markowitz, H. Portfolio Selection. *J. Financ.* **1952**, *7*, 77–91.
2. Kahneman, D.; Tversky, A. Prospect theory: An analysis of decision under risk. *Econometrica* **1979**, *47*, 263–292. [CrossRef]
3. Thiagarajan, S.R.; Schachter, B. Risk Parity: Rewards, Risks, and Research Opportunities. *J. Invest.* **2011**, *20*, 79–89. [CrossRef]
4. de Prado, M.L. Machine Learning for Econometricians: The Readme Manual. *J. Financ. Data Sci.* **2022**, *4*, 10–30. [CrossRef]
5. Milhomem, D.A.; Dantas, M.J.P. Analysis of new approaches used in portfolio optimization: A systematic literature review. *Production* **2020**, *30*, e20190144. [CrossRef]
6. Goodell, J.W.; Kumar, S.; Lim, W.M.; Pattnaik, D. Artificial intelligence and machine learning in finance: Identifying foundations, themes, and research clusters from bibliometric analysis. *J. Behav. Exp. Financ.* **2021**, *32*, 100577. [CrossRef]
7. Bartram, S.M.; Branke, J.; Rossi, G.D.; Motahari, M. Machine Learning for Active Portfolio Management. *J. Financ. Data Sci.* **2021**, *3*, 9–30. [CrossRef]
8. Krauss, C. Statistical Arbitrage Pairs Trading Strategies: Review and Outlook. *J. Econ. Surv.* **2017**, *31*, 513–545. [CrossRef]
9. Gillan, S.L.; Koch, A.; Starks, L.T. Firms and Social Responsibility: A Review of ESG and CSR Research in Corporate Finance. *J. Corp. Financ.* **2021**, *66*, 101889. [CrossRef]
10. Linnenluecke, M.K.; Marrone, M.; Singh, A.K. Conducting systematic literature reviews and bibliometric analyses. *Aust. J. Manag.* **2020**, *45*, 175–194. [CrossRef]
11. Weise, C.L.; Barbera, R.J. Minsky meets wicksell: Using the wicksellian model to understand the twenty-first century business cycle. In *Macroeconomic Theory and Macroeconomic Pedagogy*; Palgrave Macmillan: London, UK, 2016; pp. 214–233.
12. Lipton, A.; Rennie, A. *Credit Correlation: Life after Copulas*; World Scientific Publishing Company: Singapore, 2007.
13. Qian, E. *Risk Parity Portfolios: Efficient Portfolios through True Diversification*; Technical Report; PanAgora Asset Management: Boston, MA, USA, 2005.
14. Prince, B. *Risk Parity Is about Balance*; Bridgewater Associates: Westport, CT, USA, 2011.
15. Maillard, S.; Roncalli, T.; Teïletche, J. The Properties of Equally Weighted Risk Contribution Portfolios. *J. Portf. Manag.* **2010**, *36*, 60–70. [CrossRef]
16. Leippold, M.; Rueegg, R. The mixed vs the integrated approach to style investing: Much ado about nothing? *Eur. Financ. Manag.* **2018**, *24*, 829–855. [CrossRef]
17. DeMiguel, V.; Garlappi, L.; Uppal, R. Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy? *Rev. Financ. Stud.* **2009**, *22*, 1915–1953. [CrossRef]
18. Tu, J.; Zhou, G. Markowitz meets Talmud: A combination of sophisticated and naive diversification strategies. *J. Financ. Econ.* **2011**, *99*, 204–215. [CrossRef]
19. Kirby, C.; Ostdiek, B. It's All in the Timing: Simple Active Portfolio Strategies that Outperform Naïve Diversification. *J. Financ. Quant. Anal.* **2012**, *47*, 437–467. [CrossRef]
20. Capponi, A.; Figueroa-López, J.E. Dynamic portfolio optimization with a defaultable security and regime-switching. *Math. Financ.* **2014**, *24*, 207–249. [CrossRef]
21. Creamer, G.G. Can a corporate network and news sentiment improve portfolio optimization using the Black–Litterman model? *Quant. Financ.* **2015**, *15*, 1405–1416. [CrossRef]
22. Maccheroni, F.; Marinacci, M.; Rustichini, A.; Taboga, M. Portfolio selection with monotone mean-variance preferences. *Math. Financ.* **2009**, *19*, 487–521. [CrossRef]
23. Stoyanov, S.V.; Rachev, S.T.; Fabozzi, F.J. Optimal financial portfolios. *Appl. Math. Financ.* **2007**, *14*, 401–436. [CrossRef]
24. Asness, C.S.; Frazzini, A.; Pedersen, L.H. Leverage aversion and risk parity. *Financ. Anal. J.* **2012**, *68*, 47–59. [CrossRef]
25. Czichowsky, C. Time-consistent mean-variance portfolio selection in discrete and continuous time. *Financ. Stoch.* **2013**, *17*, 227–271. [CrossRef]
26. Krokhmal, P.A. Higher moment coherent risk measures. *Quant. Financ.* **2007**, *7*, 373–387. [CrossRef]
27. Natarajan, K.; Sim, M.; Uichanco, J. Tractable robust expected utility and risk models for portfolio optimization: Tractable robust expected utility and risk models. *Math. Financ.* **2010**, *20*, 695–731. [CrossRef]
28. Jin, H.; Xu, Z.Q.; Zhou, X.Y. A convex stochastic optimization problem arising from portfolio selection. *Math. Financ.* **2008**, *18*, 171–183. [CrossRef]

29. Mu, C.; Tian, W.; Yang, J. Portfolio choice with skewness preference and wealth-dependent risk aversion. *Quant. Financ.* **2019**, *19*, 1905–1919. [CrossRef]
30. Low, C.; Pachamanova, D.; Sim, M. Skewness-aware asset allocation: A new theoretical framework and empirical evidence. *Math. Financ.* **2012**, *22*, 379–410. [CrossRef]
31. Chow, G.; Jacquier, E.; Kritzman, M.; Lowry, K. Optimal portfolios in good times and bad. *Financ. Anal. J.* **1999**, *55*, 65–73. [CrossRef]
32. Huo, L.; Kim, T.-H.; Kim, Y. Robust estimation of covariance and its application to portfolio optimization. *Financ. Res. Lett.* **2012**, *9*, 121–134. [CrossRef]
33. Nika, Z.; Rásonyi, M. Log-optimal portfolios with memory effect. *Appl. Math. Financ.* **2018**, *25*, 557–585. [CrossRef]
34. Kamali, R.; Mahmoodi, S.; Jahandideh, M.-T. Optimization of multi-period portfolio model after fitting best distribution. *Financ. Res. Lett.* **2019**, *30*, 44–50. [CrossRef]
35. Chicheportiche, R.; Bouchaud, J.-P. A nested factor model for non-linear dependencies in stock returns. *Quant. Financ.* **2015**, *15*, 1789–1804. [CrossRef]
36. De Giorgi, E.; Post, T. Second-order stochastic dominance, reward-risk portfolio selection, and the capm. *J. Financ. Quant. Anal.* **2008**, *43*, 525–546. [CrossRef]
37. Hakansson, N.H. Multi-period mean-variance analysis: Toward a general theory of portfolio choice*. *J. Financ.* **1971**, *26*, 857–884.
38. Owen, J.; Rabinovitch, R. On the class of elliptical distributions and their applications to the theory of portfolio choice. *J. Financ.* **1983**, *38*, 745–752. [CrossRef]
39. Ledoit, O.; Wolf, M. Improved estimation of the covariance matrix of stock returns with an application to portfolio selection. *J. Empir. Financ.* **2003**, *10*, 603–621. [CrossRef]
40. Hitaj, A.; Mercuri, L.; Rroji, E. Portfolio selection with independent component analysis. *Financ. Res. Lett.* **2015**, *15*, 146–159. [CrossRef]
41. Münnix, M.C.; Schäfer, R.; Grothe, O. Estimating correlation and covariance matrices by weighting of market similarity. *Quant. Financ.* **2014**, *14*, 931–939. [CrossRef]
42. Dentcheva, D.; Stock, G.J. On the price of risk in a mean-risk optimization model. *Quant. Financ.* **2018**, *18*, 1699–1713. [CrossRef]
43. Gennotte, G. Optimal portfolio choice under incomplete information. *J. Financ.* **1986**, *41*, 733–746. [CrossRef]
44. Pástor, L'.; Stambaugh, R.F. Mutual fund performance and seemingly unrelated assets. *J. Financ. Econ.* **2002**, *63*, 315–349. [CrossRef]
45. Atkinson, C.; Wilmott, P. Portfolio management with transaction costs: An asymptotic analysis of the morton and pliska model. *Math. Financ.* **1995**, *5*, 357–367. [CrossRef]
46. Atkinson, C.; Ingpochai, P. Optimization of *N*-risky asset portfolios with stochastic variance and transaction costs. *Quant. Financ.* **2010**, *10*, 503–514. [CrossRef]
47. Chen, P.; Yang, H. Markowitz's mean-variance asset–liability management with regime switching: A multi-period model. *Appl. Math. Financ.* **2011**, *18*, 29–50. [CrossRef]
48. Chen, A.H.; Fabozzi, F.J.; Huang, D. Portfolio revision under mean-variance and mean-CVaR with transaction costs. *Rev. Quant. Financ. Account.* **2012**, *39*, 509–526. [CrossRef]
49. Cover, T.M. Universal portfolios. *Math. Financ.* **1991**, *1*, 1–29. [CrossRef]
50. Das, S.; Markowitz, H.; Scheid, J.; Statman, M. Portfolio optimization with mental accounts. *J. Financ. Quant. Anal.* **2010**, *45*, 311–334. [CrossRef]
51. Das, S.R.; Kaznachey, D.; Goyal, M. Computing optimal rebalance frequency for log-optimal portfolios. *Quant. Financ.* **2014**, *14*, 1489–1502. [CrossRef]
52. Elton, E.J.; Gruber, M.J.; Padberg, M.W. Simple criteria for optimal portfolio selection. *J. Financ.* **1976**, *31*, 1341–1357. [CrossRef]
53. Elton, E.J.; Gruber, M.J.; Padberg, M.W. Simple rules for optimal portfolio selection: The multi group case. *J. Financ. Quant. Anal.* **1977**, *12*, 329. [CrossRef]
54. Fouque, J.-P.; Hu, R. Optimal portfolio under fractional stochastic environment. *Math. Financ.* **2019**, *29*, 697–734. [CrossRef]
55. Grossman, S.J.; Zhou, Z. Optimal investment strategies for controlling drawdowns. *Math. Financ.* **1993**, *3*, 241–276. [CrossRef]
56. Jobst, N.J.; Horniman, M.D.; Lucas, C.A.; Mitra, G. Computational aspects of alternative portfolio selection models in the presence of discrete asset choice constraints. *Quant. Financ.* **2001**, *1*, 489–501. [CrossRef]
57. Jorion, P. Portfolio optimization with tracking-error constraints. *Financ. Anal. J.* **2003**, *59*, 70–82. [CrossRef]
58. Kane, A. Skewness preference and portfolio choice. *J. Financ. Quant. Anal.* **1982**, *17*, 15. [CrossRef]
59. Lan, C. An out-of-sample evaluation of dynamic portfolio strategies. *Rev. Financ.* **2015**, *19*, 2359–2399. [CrossRef]
60. Li, B.; Wang, J.; Huang, D.; Hoi, S.C.H. Transaction cost optimization for online portfolio selection. *Quant. Financ.* **2018**, *18*, 1411–1424. [CrossRef]
61. Mulvey, J.M.; Sun, Y.; Wang, M.; Ye, J. Optimizing a portfolio of mean-reverting assets with transaction costs via a feedforward neural network. *Quant. Financ.* **2020**, *20*, 1239–1261. [CrossRef]
62. Muthuraman, K.; Kumar, S. Multidimensional portfolio optimization with proportional transaction costs. *Math. Financ.* **2006**, *16*, 301–335. [CrossRef]
63. Nystrup, P.; Madsen, H.; Lindström, E. Dynamic portfolio optimization across hidden market regimes. *Quant. Financ.* **2018**, *18*, 83–95. [CrossRef]

64. Roman, D.; Darby-Dowman, K.; Mitra, G. Mean-risk models using two risk measures: A multi-objective approach. *Quant. Financ.* **2007**, *7*, 443–458. [CrossRef]
65. Taboga, M. Portfolio selection with two-stage preferences. *Financ. Res. Lett.* **2005**, *2*, 152–164. [CrossRef]
66. Wachter, J.A. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *J. Financ. Quant. Anal.* **2002**, *37*, 63. [CrossRef]
67. Xidonas, P.; Mavrotas, G. Comparative issues between linear and non-linear risk measures for non-convex portfolio optimization: Evidence from the S&P 500. *Quant. Financ.* **2014**, *14*, 1229–1242.
68. Pogue, G.A. An Extension of the Markowitz Portfolio Selection Model to Include Variable Transactions' Costs, Short Sales, Leverage Policies and Taxes. *J. Financ.* **1970**, *25*, 1005–1027. [CrossRef]
69. Cederburg, S.; O'Doherty, M.S.; Wang, F.; Yan, X.S. On the performance of volatility-managed portfolios. *J. Financ. Econ.* **2020**, *138*, 95–117. [CrossRef]
70. Roy, A.D. Safety First and the Holding of Assets. *Econometrica* **1952**, *20*, 431–449. [CrossRef]
71. Kelly, J.L. A New Interpretation of Information Rate. *Bell Syst. Tech. J.* **1956**, *35*, 917–926. [CrossRef]
72. Gul, F. A Theory of Disappointment Aversion. *Econometrica* **1991**, *59*, 667–686. [CrossRef]
73. Sundaresan, S.M. Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth. *Rev. Financ. Stud.* **1989**, *2*, 73–89. [CrossRef]
74. Abel, A.B. Asset Prices under Habit Formation and Catching up with the Joneses. *Am. Econ. Rev.* **1990**, *80*, 38–42.
75. Knight, F.H. *Risk, Uncertainty and Profit*; Houghton Mifflin Company: Boston, MA, USA, 1921.
76. Rockafellar, R.T.; Uryasev, S. Optimization of conditional value-at-risk. *J. Risk* **2000**, *2*, 21–41. [CrossRef]
77. Uryasev, S. Conditional value-at-risk: Optimization algorithms and applications. In Proceedings of the IEEE/IAFE/INFORMS 2000 Conference on Computational Intelligence for Financial Engineering (CIFEr) (Cat. No.00TH8520), New York, NY, USA, 28 March 2000; pp. 49–57.
78. Low, R.K.Y.; Alcock, J.; Faff, R.; Brailsford, T. Canonical vine copulas in the context of modern portfolio management: Are they worth it? *J. Bank. Financ.* **2013**, *37*, 3085–3099. [CrossRef]
79. Carta, A.; Conversano, C. Practical implementation of the kelly criterion: Optimal growth rate, number of trades, and rebalancing frequency for equity portfolios. *Front. Appl. Math. Stat.* **2020**, *6*, 577050. [CrossRef]
80. Low, R.K.Y. Vine copulas: Modelling systemic risk and enhancing higher-moment portfolio optimisation. *Account. Financ.* **2018**, *58*, 423–463. [CrossRef]
81. Adler, T.; Kritzman, M. Mean–variance versus full-scale optimisation: In and out of sample. *J. Asset Manag.* **2007**, *7*, 302–311. [CrossRef]
82. DeMiguel, V.; Garlappi, L.; Nogales, F.J.; Uppal, R. A generalized approach to portfolio optimization: Improving performance by constraining portfolio norms. *Manag. Sci.* **2009**, *55*, 798–812. [CrossRef]
83. Brodie, J.; Daubechies, I.; De Mol, C.; Giannone, D.; Loris, I. Sparse and stable Markowitz portfolios. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 12267–12272. [CrossRef]
84. Bruder, B.; Gaussel, N.; Richard, J.C.; Roncalli, T. Regularization of portfolio allocation. *SSRN Electron. J.* **2013**. [CrossRef]
85. Black, F.; Litterman, R.B. Asset Allocation: Combining Investor Views with Market Equilibrium. *J. Fixed Income* **1991**, *1*, 7–18. [CrossRef]
86. Jorion, P. Bayes-Stein Estimation for Portfolio Analysis. *J. Financ. Quant. Anal.* **1986**, *21*, 279–292. [CrossRef]
87. Jorion, P. Bayesian and CAPM estimators of the means: Implications for portfolio selection. *J. Bank. Financ.* **1991**, *15*, 717–727. [CrossRef]
88. Pastor, L. Portfolio Selection and Asset Pricing Models. *J. Financ.* **2000**, *55*, 179–223. [CrossRef]
89. Pastor, L.; Stambaugh, R.F. Comparing asset pricing models: An investment perspective. *J. Financ. Econ.* **2000**, *56*, 335–381. [CrossRef]
90. MacKinlay, A.C.; Pástor, L. Asset Pricing Models: Implications for Expected Returns and Portfolio Selection. *Rev. Financ. Stud.* **2000**, *13*, 883–916. [CrossRef]
91. Kan, R.; Zhou, G. Optimal Portfolio Choice with Parameter Uncertainty. *J. Financ. Quant. Anal.* **2007**, *42*, 621–656. [CrossRef]
92. Bayes, T.L. An essay towards solving a problem in the doctrine of chances. By the late Rev. Mr. Bayes, F.R.S. communicated by Mr. Price, in a letter to John Canton, A.M.F.R.S. *Philos. Trans. R. Soc. Lond.* **1763**, *53*, 370–418.
93. Stein, C.M. Inadmissibility of the Usual Estimator for the Mean of a Multivariate Normal Distribution. In Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Berkeley, CA, USA, 26–31 December 1954; University of California Press: Berkeley, CA, USA, 1955; Volume 1, pp. 197–206.
94. James, W.; Stein, C.M. Estimation with Quadratic Loss. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Los Angeles, CA, USA, 20–30 July 1961; pp. 361–379.
95. Allaj, E. The Black–Litterman model and views from a reverse optimization procedure: An out-of-sample performance evaluation. *Comput. Manag. Sci.* **2020**, *17*, 465–492. [CrossRef]
96. Kritzman, M.; Page, S.; Turkington, D. In Defense of Optimization: The Fallacy of 1/N. *Financ. Anal. J.* **2010**, *66*, 31–39. [CrossRef]
97. Balduzzi, P.; Lynch, A.W. Transaction costs and predictability: Some utility cost calculations. *J. Financ. Econ.* **1999**, *52*, 47–78. [CrossRef]
98. Bellman, R. On the theory of dynamic programming. *Proc. Natl. Acad. Sci. USA* **1952**, *38*, 613–640. [CrossRef]

99. Das, S.R.; Ostrov, D.; Radhakrishnan, A.; Srivastav, D. Dynamic portfolio allocation in goals-based wealth management. *Comput. Manag. Sci.* **2020**, *17*, 613–640. [CrossRef]
100. Wang, Y.; Aste, T. Dynamic portfolio optimization with inverse covariance clustering. *Expert Syst. Appl.* **2023**, *213*, 118739. [CrossRef]
101. Bhansali, V. Beyond Risk Parity. *J. Invest.* **2011**, *20*, 137–147. [CrossRef]
102. Bhansali, V.; Davis, J.; Rennison, G.; Hsu, J.; Li, F. The Risk in Risk Parity: A Factor-Based Analysis of Asset-Based Risk Parity. *J. Invest.* **2012**, *21*, 102–110. [CrossRef]
103. DeMiguel, V.; Martin-Utrera, A.; Uppal, R. A Multifactor Perspective on Volatility-Managed Portfolios. *SSRN Electron. J.* **2021**. [CrossRef]
104. Fletcher, J. Do optimal diversification strategies outperform the 1/N strategy in U.K. stock returns? *Int. Rev. Financ. Anal.* **2011**, *20*, 375–385. [CrossRef]
105. Moreira, A.; Muir, T. Volatility-Managed Portfolios. *J. Financ.* **2017**, *72*, 1611–1644. [CrossRef]
106. Barroso, P.; Detzel, A. Do limits to arbitrage explain the benefits of volatility-managed portfolios? *J. Financ. Econ.* **2021**, *140*, 744–767. [CrossRef]
107. Abdi, F.; Ranaldo, A. A Simple Estimation of Bid-Ask Spreads from Daily Close, High, and Low Prices. *Rev. Financ. Stud.* **2017**, *30*, 4437–4480. [CrossRef]
108. Sharpe, W.F. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Financ.* **1964**, *19*, 425–442.
109. Lintner, J. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *Rev. Econ. Stat.* **1965**, *47*, 37. [CrossRef]
110. Fama, E.F.; French, K.R. Common risk factors in the returns on stocks and bonds. *J. Financ. Econ.* **1993**, *33*, 3–56. [CrossRef]
111. Ross, S.A. The arbitrage theory of capital asset pricing. *J. Econ. Theory* **1976**, *13*, 341–360. [CrossRef]
112. Carhart, M.M. On Persistence in Mutual Fund Performance. *J. Financ.* **1997**, *52*, 57–82. [CrossRef]
113. Fama, E.F.; French, K.R. A five-factor asset pricing model. *J. Financ. Econ.* **2015**, *116*, 1–22. [CrossRef]
114. Fitzgibbons, S.; Friedman, J.; Pomorski, L.; Serban, L. Long-Only Style Investing: Don't Just Mix, Integrate. *J. Invest.* **2017**, *26*, 153–164. [CrossRef]
115. Frazzini, A.; Israel, R.; Moskowitz, T.J.; Novy-Marx, R. A New Core Equity Paradigm: Using Value, Momentum, and Quality to Outperform Markets. *AQR Cap. Manag. White Pap.* **2013**.
116. Fernandez-Perez, A.; Fuertes, A.M.; Miffre, J. A comprehensive appraisal of style-integration methods. *J. Bank. Financ.* **2019**, *105*, 134–150. [CrossRef]
117. Brandt, M.W.; Santa-Clara, P.; Valkanov, R. Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns. *Rev. Financ. Stud.* **2009**, *22*, 3411–3447. [CrossRef]
118. Fischer, M.; Gallmeyer, M.F. Heuristic portfolio trading rules with capital gain taxes. *J. Financ. Econ.* **2016**, *119*, 611–625. [CrossRef]
119. Ghysels, E.; Plazzi, A.; Valkanov, R. Why Invest in Emerging Markets? The Role of Conditional Return Asymmetry. *J. Financ.* **2016**, *71*, 2145–2192. [CrossRef]
120. DeMiguel, V.; Martín-Utrera, A.; Nogales, F.J.; Uppal, R. A Transaction-Cost Perspective on the Multitude of Firm Characteristics. *Rev. Financ. Stud.* **2020**, *33*, 2180–2222. [CrossRef]
121. Barberis, N.; Shleifer, A. Style investing. *J. Financ. Econ.* **2003**, *68*, 161–199. [CrossRef]
122. Zou, H.; Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc.* **2005**, *67*, 301–320. [CrossRef]
123. Ho, T.K. Random decision forests. In Proceedings of the 3rd International Conference on Document Analysis and Recognition, Montreal, QC, Canada, 14–16 August 1995; Volume 1, pp. 278–282.
124. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM, San Francisco, CA, USA, 13–17 August 2016; pp. 785–794.
125. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. *Adv. Neural Inf. Process. Syst.* **2017**, 4766–4775.
126. Jiang, Z.; Xu, D.; Liang, J. A Deep Reinforcement Learning Framework for the Financial Portfolio Management Problem. *arXiv* **2017**, arXiv:1706.10059.
127. Raffinot, T. Hierarchical Clustering-Based Asset Allocation. *J. Portf. Manag.* **2018**, *44*, 89–99. [CrossRef]
128. Jiang, J.; Kelly, B.T.; Xiu, D. (Re-)Imag(in)ing Price Trends. *SSRN Electron. J.* **2020**. [CrossRef]
129. Elkind, D.; Kaminski, K.; Lo, A.W.; Siah, K.W.; Wong, C.H. When Do Investors Freak Out? Machine Learning Predictions of Panic Selling. *J. Financ. Data Sci.* **2022**, *4*, 11–39. [CrossRef]
130. Chen, J.; Tindall, M.L. Hedge Fund Replication Using Shrinkage Methodologies. *J. Altern. Invest.* **2014**, *17*, 26–49. [CrossRef]
131. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
132. Zhu, M.; Philpotts, D.; Sparks, R.; Stevenson, M.J. A Hybrid Approach to Combining CART and Logistic Regression for Stock Ranking. *J. Portf. Manag.* **2011**, *38*, 100–109. [CrossRef]
133. Montiel, J.; Mitchell, R.; Frank, E.; Pfahringer, B.; Abdessalem, T.; Bifet, A. Adaptive XGBoost for Evolving Data Streams. In Proceedings of the 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 19–24 July 2020; pp. 1–8.
134. Shapley, L.S. A Value for N-Person Games. In *Contributions to the Theory of Games*; Princeton University Press: Princeton, NJ, USA, 1952.
135. Lopez de Prado, M. Building diversified portfolios that outperform out of sample. *J. Portf. Manag.* **2016**, *42*, 59–69. [CrossRef]

136. Fukushima, K. Neocognitron: A self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern.* **1980**, *36*, 193–202. [CrossRef] [PubMed]
137. Denker, J.S.; LeCun, Y. Transforming Neural-Net Output Levels to Probability Distributions. In Proceedings of the 3rd International Conference on Neural Information Processing Systems, Denver, CO, USA, 26–29 November 1990; Morgan Kaufmann Publishers Inc.: Burlington, UK, 1990; pp. 853–859.
138. Cheng, T.; Chen, K. A general framework for portfolio construction based on generative models of asset returns. *J. Financ. Data Sci.* **2023**, *9*, 100113. [CrossRef]
139. Ivakhnenko, A.G. Polynomial Theory of Complex Systems. *IEEE Trans. Syst. Man Cybern.* **1971**, *1*, 364–378. [CrossRef]
140. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control Signals Syst.* **1989**, *2*, 303–314. [CrossRef]