ARTICLE IN PRESS

European Journal of Operational Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Interfaces with Other Disciplines

Sustainable optimal stock portfolios: What relationship between sustainability and performance?☆

Beatrice Bertelli [a], Costanza Torricelli [b,*]

[a] University of Modena and Reggio Emilia, Italy
[b] University of Modena and Reggio Emilia and CEFIN - Centro Studi Banca e Finanza, Italy

ARTICLE INFO

JEL classification:
C61
G11
M14
Q01

Keywords:
Finance
Sustainable portfolio
Portfolio optimization
ESG score
Portfolio performance

ABSTRACT

The aim of this paper is to compare different strategies to combine sustainability and optimality in stock portfolios to assess whether there is an association between their average ESG (Environmental, Social, Governance) score and their financial performance and, if so, whether it depends on the specific strategy used. To this end, we confront the risk-adjusted performance of three ESG-compliant optimal portfolios resulting from: (i) optimizing on an ESG-screened sample, (ii) including a portfolio ESG-score constraint in the optimization on an unscreened sample, (iii) our original proposal of optimizing with an ESG-score constraint (so as to reach a target) over a slightly screened sample (so as to exclude companies with lowest sustainability). The optimization is implemented with Bloomberg ESG scores over a sample from the EURO STOXX Index in the period January 2007–August 2022 by minimizing portfolio residual risk. Two are the main conclusions from our results. First, we never find a significant negative association between portfolios' average ESG score and performance independently of the strategy used. Second, we find a positive association when the first and the third strategy are implemented with a high screening level. To be noted that the relationship between the ESG score and the risk-return ratio in the initial investment set plays a relevant role. If, as in our dataset, this relationship is essentially convex, with an appropriate level of screening portfolios are composed only by stocks whereby a higher ESG score is associated with a higher risk-return profile.

1. Introduction

In recent years, investors' attention towards environmental, social and governance (ESG) dimensions has significantly increased spurred by UN initiatives and programs (e.g. the 2030 Agenda and the Principles for Responsible Investment; PRI, 2017) and historical events. Crises played an important role in raising investors' awareness towards social responsibility and sustainability issues: the 2007-2008 global financial crisis highlighted the importance of corporate social responsibility (Cesarone et al., 2022), while Covid-19 pandemic transformed sustainability from a luxury good into a priority (Pástor and Vorsatz, 2020). Hence, sustainable investments have become central also in asset allocation and asset managers integrate these non-strictly financial aspects into their investment policies (van Duuren et al., 2016). According to the Global Sustainable Investment Review (GSIA, 2021) assets under management reached USD 35.3 trillion in 2020, (growing by 15% in two years) and they represent 36% of all professionally managed assets across the major markets (Europe, United States, Canada, Australasia and Japan). Primarily than obtaining a financial return, sustainable investors incorporate ESG assets in their portfolio to hedge specific risks such as climate risk (e.g. Engle et al., 2020; Alekseev et al., 2022) or simply to contribute to a better society and to promote good corporate

☆ We wish to thank Armin Varmaz for helpful clarification on the model setup, participants and discussants of DTX research talks (University of Bremen, January 2023), 95th International Atlantic Economic Conference (Rome, March 2023), Spring 2023 Semester Mathematical Finance Seminar Series (Texas Tech University, March 2023), Netspar International Pension Workshop '23 (Leiden, June 2023), International Finance and Banking Society Conference (Oxford, July 2023), Political Economy Seminar Series, (University of Modena and Reggio Emilia, September 2023), 17th International Conference on Computational and Financial Econometrics (Berlin, December 2023), European Financial Management Association 2024 Annual Meeting (Lisbon, June 2024). Costanza Torricelli kindly acknowledges support from the Department of Economics of the University of Modena and Reggio Emilia (FAR2022) and Italian Ministry of University (PRIN 2022ELYHCW). Usual caveat applies.

* Corresponding author.
E-mail address: costanza.torricelli@unimore.it (C. Torricelli).

https://doi.org/10.1016/j.ejor.2025.01.021
Received 23 June 2023; Accepted 16 January 2025
Available online 19 January 2025

Please cite this article as: Beatrice Bertelli, Costanza Torricelli, *European Journal of Operational Research*, https://doi.org/10.1016/j.ejor.2025.01.021

behaviour (Pedersen et al., 2021).

The literature has been giving increasing attention to sustainability with a focus on different issues such as the ranking of mutual funds (Cabello et al., 2014) or portfolio selection based on multiple criteria (Hallerbach et al., 2004; Ballestero et al., 2012). In particular, research studies have investigated several strategies according to which investors can set up ESG-compliant portfolios: from screening strategies and a combination of the latter with traditional portfolio theory to optimization problems that extend the mean-variance optimization model by considering a sustainability dimension beside risk and return. However, the existing literature is inconclusive about the relation between the ESG score and the financial performance of optimal portfolios and it has given little attention to the comparison of strategies with different outcomes in terms of sustainability. Moreover, most optimization models require a balanced dataset (i.e. all panel or cross-sectional data have measurements in all periods) for their estimation, which can be a relevant issue in general (e.g. firms may be delisted, firms may merge in the sample period) and it is even more so when a metric for sustainability is needed since the number of firms with ESG scores varies substantially over time.

Against this backdrop, the aim of this paper is to compare different strategies to combine sustainability and optimality in stock portfolios in order to assess whether there is, if any, a (positive or negative) association between portfolios' average ESG score and financial performance and whether such a relationship depends on the specific strategy used, i.e. the type of sustainability attained. In other words, we mean to answer the following questions: what type of relationship is there between financial performance and sustainability as measured by the portfolio average ESG score? Is it possible to do well by doing good and, if so, which is the best strategy to attain this?

To this end we provide comparative evidence on the performance of two philosophically different strategies to set up an ESG-compliant portfolio and we propose a third one resulting from a mixture of them. The first strategy, which is widely adopted in the industry, results from optimizing on an ESG-screened sample, the second is obtained by adding to the optimization problem a portfolio ESG-score constraint on an unscreened sample, while the third results from our original proposal of taking pros of both by optimizing with an ESG-score constraint (so as to reach a target) over a slightly screened sample (so as to eliminate only the worst companies in terms of ESG).

Our paper represents an empirical study contributing to the literature on optimal sustainable portfolios. Specifically, it is the second study after Varmaz et al. (2024) to implement the a priori approach as defined by Gasser et al. (2017) to solve a tri-criterion optimization model (risk-return-ESG), the first one to implement it over an EU stock sample and to compare the resulting financial performance with alternative strategies. Moreover, we depart from Varmaz et al. (2024) methodology because we combine their approach with the most widely used screening strategies.[1] Beside the relevant implication that our proposal has for the industry (i.e. attaining a portfolio ESG-target excluding the least sustainable stocks), by using different screening thresholds we are able to highlight the relevance of the relationship between the ESG score and the risk-return ratio in the initial investment set. Since in our dataset, this relationship is essentially convex (i.e. initially negative and then positive), by varying the level of screening we can test the association between portfolio average ESG scores and performances of sustainable optimal portfolios also in the presence of a monotone increasing investment set.[2]

The choice to follow Varmaz et al. (2024), which determines the optimal portfolio by minimizing residual risk for a given level of systemic risk and ESG score, has technical and practical advantages: it does not require the estimation of the covariance matrix, so it is suitable also when data are represented by an unbalanced panel (i.e. the time series of stocks returns have different numbers of observations over time since stocks can be listed/delisted, firms can merge, firms with ESG score vary), which is particularly relevant in empirical implementation since it allows considering also stocks with shorter time series. Moreover, this approach, by testing for different combinations of risk and sustainability, is particularly interesting in applications and for the financial industry also from a regulatory viewpoint. In fact, following the revision of the European Union's MiFID II directive (European Parliament, 2023) financial advisors are required to detect not only the market risk investors want to bear but also their ESG preferences.

Our dataset is based on the 586 stocks that composed the EURO STOXX Index in the period January 2007 – August 2022 and we approximate the ESG characteristic by means of the Bloomberg ESG disclosure score, which assesses firm's transparency on ESG issues.

Two are the main conclusions from our results. First, we never find a significant negative association between portfolio average ESG score and performance independently of the strategy used. Second, and by contrast, we find a positive association when the first and the third strategy are implemented with a high screening level. In fact, prior screening (Strategy 1) implies a superior risk-adjusted performance only with heavy screening, an ESG-score constraint in the optimization (Strategy 2) implies that the portfolio performance does not significantly worsen as the target ESG level increases; optimizing with an ESG-score constraint after a mild ESG screening (Strategy 3) does not significantly worsen performance as the target ESG level increases and, when the screening is higher, it obtains a superior performance. Furthermore, when testing for the ESG-compliant portfolio that better performs over time, we find that the comparative performance of the three strategies does not substantially vary over the financial cycle.

To be noted that the relationship between the ESG score and the risk-return ratio in the initial investment set plays a relevant role. If, as in our case, this relationship is essentially convex, with a specific level of screening portfolios are composed only by stocks whereby a higher ESG score is associated with a higher risk-return profile. In terms of policy implications for the financial industry, the third strategy we propose, allows sustainable investors to do well by doing good in the presence of both a convex and a monotonically increasing investment set, being thus able to attract a larger pool of investors towards ESG-compliant optimal portfolios.

The paper is organized as follows. Section 2 provides a critical review of the theoretical and empirical literature on ESG-compliant portfolio approaches. Section 3 illustrates the analytics of the optimization model, Section 4 the dataset, and Section 5 the empirical methodology. Section 6 presents results on the risk-adjusted performance of three different types of ESG-compliant portfolios and Section 7 compares portfolio performance over the whole period and over time. Section 8 presents some robustness tests. Last Section concludes.

[1] With respect to Varmaz et al. (2022, 2024) we implement the model over a different investment set whose relationship between ESG score and risk-return allows to highlight the role played by the opportunity set.

[2] In principle the relation between ESG score classes and the risk-reward ratio could be non-monotone (mainly convex or concave) or monotone (increasing or decreasing), whereby a mainly convex (concave) relation implies that the risk-reward ratio is lower (higher) for intermediate ESG scores, while a monotone relation implies that profitability steadily increases (or decreases) with sustainability. Given the market turn towards sustainability in the latter years, such a relation is expected to be mostly positive as witnessed by the empirical literature (for a survey see Friede et al. 2015; Whelan et al., 2021).

2. An overview of ESG-compliant portfolio approaches

The literature on ESG-related portfolios is rooted in the literature on socially responsible investing (SRI) and has been growing very fast in the latter years also spurred by the availability of scores and ratings useful to evaluate companies' nonfinancial performance in consideration of ESG factors.

In order to set up socially responsible and ESG portfolios, there are several strategies, characterized by different levels of complexity and sophistication and different outcomes in terms of sustainability: from screening strategies and a combination of the latter with traditional portfolio theory to optimization problems that consider the sustainability dimension beside risk and return. These various strategies can be essentially classified into three main strands as discussed in the following paragraphs.

2.1. Screening strategies

A first strand of literature proposes a strategy, widely used in practice because of its simplicity, that consists of implementing some sort of screening on the investment set. Negative screening excludes assets according to some socially responsible criteria such as the involvement in immoral activities or low ESG measures, while positive screening tilts portfolio towards companies outperforming in terms of social responsibility or sustainability. When positive screening implies the selection of most virtuous companies relative to industry peers, it is referred to as best-in-class.

As for the sort of screening, two are the main approaches taken. In the early literature, socially responsible investing consisted mainly in the exclusion of the so-called "sin stocks" i.e. stocks belonging to sectors considered unethical or immoral such as tobacco, alcohol, gambling and weapons (Blitz and Fabozzi, 2017; Hong and Kacperczyk, 2009). Later on, with the introduction of ESG ratings, screening is achieved by the exclusion (selection) of assets associated to low (high) scores. As for the portfolio composition, rather than resting on optimization, two main simple alternatives are followed: equal weights or weights resting on market capitalization.

Overall, the literature is inconclusive about the impact of ESG-related criteria on financial portfolio performance.[3] Although socially responsible firms could potentially benefit from higher profitability (Friedman, 1970; Bénabou and Tirole, 2010), empirical studies do not always find an overperformance associated to ESG portfolios with respect to a passive benchmark. For instance, Auer (2016) applies ESG screenings by using Sustainalytics scores over the components of the STOXX Europe 600 index in the period 2004-2012 and finds that only screenings based on the governance dimension realize a better performance with respect to the benchmark index. Bertelli and Torricelli (2024) implement negative and positive screening strategies using Bloomberg ESG scores and the EURO STOXX index. Overall, they prove overperformance of negative screening strategies over the long term (2007-2021) and non-overperformance of screened portfolios during periods of crisis such as the global recession and Covid-19 pandemic. Alessandrini and Jondeau (2020) show that negative screenings based on ESG scores on MSCI ACWI Index over the period 2007-2018 improve the overall ESG score of the resulting portfolios without reducing their risk-adjusted performance.

2.2. A two-step approach: traditional portfolio optimization over a screened sample

A second strand of literature takes a different approach resting on the idea of separating the consideration of the ESG dimension from the portfolio construction (Bender et al., 2017). The first step consists in screening based on ESG scores over the constituents of a diversified index, the second consists in setting up an optimal portfolio problem with the survived assets. Hence, the ESG issue is taken into consideration at the screening level of the investment set, over which a conventional optimal portfolio problem is solved based on some optimizing criterion (e.g. minimize portfolio risk or a tracking error).

With respect to pure-screening strategies, which use simple weighting techniques, these strategies allow the investor to meet financial objectives beside ESG ones, even if some trade-offs still emerge as demonstrated by Bohn et al. (2022). Starting from the MSCI ACWI Index, they implement negative screening and adopt two different strategies: Simple Exclusion by cap-weighting the survived stocks and Optimized Exclusion by weighting the survived securities to minimize the tracking error with respect to the benchmark. The Optimized Exclusion portfolio on one hand mimics the benchmark, but on the other it assigns higher weights to stocks correlated with the excluded ones and potentially just as undesired; however, it has a higher extra-return with respect to the benchmark than the Simple Exclusion portfolio. Liagkouras et al. (2020) first perform a screening over the constituents of FTSE-100 index to exclude assets that do not respect an ESG constraint, then they solve a mean-variance portfolio optimization model. They find that the optimal allocation of assets with high ESG score is characterized by a worse risk-return combination than optimized portfolios of the unscreened sample concluding that ESG investors must be willing to pay for sustainability. Similarly, Wang et al. (2022) show that screening, based on Bloomberg scores, reduces minimum variance portfolio performance in the Chinese stock market. In sum, the initial screening introduces constraints on the investment set that limit portfolio diversification and profitability according to traditional portfolio theory (Markowitz, 1952; Girard et al., 2007; Ortas et al., 2014).

2.3. Portfolio optimization including the ESG dimension

A third strand of literature, which aims to overcome the drawbacks of screening, proposes to address the optimal portfolio problem by including the ESG dimension beside risk and return over an unscreened sample. It results in an extension of the two-dimensional Markowitz optimization problem to a tri-criterion portfolio selection model that includes an additional linear objective (for instance an ESG target) to the portfolio mean and variance objectives (Hirschberger et al., 2013; Utz et al., 2014; Cesarone et al., 2022).[4]

Gasser et al. (2017) are the first to clearly argue that such a tri-criterion optimization model can be implemented either in an a posteriori or a priori fashion: the former identifies the set of efficient portfolios on an efficient surface (as opposed to the Markowitz efficient frontier) defined by feasible optimal combinations of return, risk and sustainability; the latter consists in finding the optimal portfolio by solving the optimization problem for a given set of investors' preferences about risk, return and ESG.

The a posteriori approach, adopted by most of the existing literature, requires the construction of the efficient (or nondominated) surface. For instance, Gasser et al. (2017) examine the relations between return, risk and social responsibility in a tri-dimensional space on a set of international stocks and find that mean-variance-ESG efficient portfolios tend to obtain a lower Sharpe ratio than mean-variance efficient portfolios. Further, Utz et al. (2014) use the approach by Hirschberger et al. (2013) to compute the mean-variance-ESG nondominated surface that

[3] In the literature prior to the diffusion of ESG rating, the measurement was often made in terms of Corporate Social Responsibility (CSR) whereby ESG can be thought of a metric for CSR (Kuzey et al., 2021).

[4] The socially responsible dimension can be represented by several measures: most studies use an aggregate ESG score or rating provided by different agencies (e.g. Refinitiv, Thomson Reuters, MSCI, Sustainalytics); but the focus could be also on a single dimension such as greenhouse gas (GHG) emission intensity (De Spiegeleer et al., 2021).

minimizes portfolio risk and maximizes portfolio expected return and ESG, and by means of an inverse optimization process they investigate whether US socially responsible mutual funds, that are based on negative screening, have a higher ESG risk tolerance than US conventional mutual funds in the period 2001-2010.[5] Their findings suggest that the initial screening is not sufficient to increase the sustainability of the fund, since conventional and socially responsible mutual funds do not show significant differences in terms of ESG risk tolerance and ESG score. Similarly, Utz et al. (2015) and Alessandrini and Jondeau (2021) prove that screening the initial investment set does not always guarantee a higher ESG quality and a superior financial performance with respect to efficient portfolios aiming to maximize portfolio ESG score. The former use an algorithm transformation method (the ε-constrained method) to convert the unconstrained problem into a constrained one and compute efficient portfolios on the nondominated surface that maximize portfolio sustainability by imposing conditions of risk and expected return that are no worse than those of screened US mutual fund portfolios.[6] The latter compare screened portfolios with optimal portfolios maximizing portfolio ESG score by imposing restrictions on the tracking error, transaction costs, and risk exposures over individual stocks from the MSCI ESG database in the period 2007-2018. The ε-constraint method is adopted also by Cesarone et al. (2022) in order to identify the risk-expected return-ESG relations embedded in the nondominated frontier. They minimize portfolio variance by imposing constraints on portfolio target return and target ESG for five different datasets representing indexes from major stock markets (Dow Jones Industrial, Euro Stoxx 50, FTSE100, NASDAQ100, S&P500) from 2006 to 2020. By varying such target levels they obtain efficient portfolios (i.e. on the efficient surface) and find that, over the full period 2006-2020, high-ESG portfolios show a better financial performance only in the US markets, whereas in the subperiod 2014-2020, after the Kyoto Protocol, a higher performance is recorded in four out of five datasets. Finally, another a posteriori application is proposed by Pedersen et al. (2021) who synthetize risk and return with the Sharpe ratio deriving a nondominated frontier, instead of a surface and thus restate the optimization problem across three dimensions (risk, return, ESG) in terms of a trade-off between ESG and Sharpe ratio. Boubaker et al. (2023) propose a multi-dimensional extension of such two-dimensional ESG--frontier for 334 energy firms in 2019 to examine the trade-off between ESG and Sharpe ratio. However, combining risk and return objectives into a Sharpe ratio objective turns the problem into a bi-criterion one and this may lead to a non-optimal solution in terms of risk and return, as pointed out by Steuer and Utz (2023). In the investigation of the risk-return-ESG efficient surfaces Steuer and Utz (2023) propose an innovative approach based on non-contour efficient fronts able to visualize the whole efficient surface and the optimal portfolio in absolute terms vs. in relative terms (e.g. Gardiner and Steuer, 1994; Figueira et al., 2010; Miettinen et al., 2010).

The a priori approach to solve the mean-variance-ESG optimization problem was first presented by Gasser et al. (2017) who illustrated the theoretical model without providing an empirical implementation. Only recently, Varmaz et al. (2024) reformulate the a priori approach by assuming the validity of a factor model for asset returns, thus overcoming some model-specific issues (i.e. the estimation of the covariance matrix and the identification of investors' return, risk and ESG preference parameters) and obtaining an analytical solution: the single optimal portfolio that minimizes residual risk and, by means of equality constraints, is consistent with investors' desired levels of systemic risk and sustainability. To be noted that the introduction of equality constraints may not guarantee Pareto optimality, however the approach does not require a balanced dataset to be implemented in contrast to optimization models based on an efficient mean-variance-ESG surface, whose implementation has been so far restricted to mutual funds, market indexes or stocks with long time series. The reformulation proposed by Varmaz et al. (2024) is flexible to accommodate the two competing interpretations of the ESG dimensions: ESG features have an effect on portfolio return because they modify portfolio exposure to systemic risks (e.g. Pedersen et al., 2021; Pástor et al., 2021) vs. ESG scores contribute to an additional expected return that is unrelated to assets' systemic risk (e.g. Bénabou and Tirole, 2010; Edmans, 2011; Friede et al., 2015; Hoepner et al., 2021).

Comparative inspection of the literature on the a posteriori and the a priori approaches shows that both have limitations. The a priori approach requires the problematic (and somewhat arbitrary) modelling of risk aversion and ESG attitude, making the selection of a single optimal portfolio strictly dependent on these pre-specified preference parameters. The a posteriori approach, which provides the entire Pareto optimal set, confronts the decision maker (DM) with two main issues: the cardinality of the solutions and the selection of an optimal portfolio comprehensible in terms of risk, return and ESG characteristics.

Summing up, the literature on portfolio optimization including the ESG dimension highlights a trade-off between Pareto optimality and the identification of an optimal portfolio that is comprehensible to the DM in terms of her/his preferences on risk, return and sustainability.

In this paper we follow the a priori approach by Varmaz et al. (2024), which, although not ensuring Pareto optimality, provides a closed-form optimal solution that is easily communicable to the investor and is aligned with her/his preferences, given that it simply requires investors to state their desired levels for portfolio return, risk, and ESG score (see Section 3.2).[7]

3. The analytics of the optimization model

All the strategies we compare produce an optimal portfolio and thus require an optimization framework. To this end, among alternatives described in Section 2.3, we follow the a priori approach reformulation by Varmaz et al. (2024) for three main reasons: it is intuitive and useful for the financial industry since it allows to set a priori desired levels of portfolio systemic risk and ESG score with no need of investors' preference parameters, it provides an analytical solution without the need of estimating the variance-covariance matrix, and it is thus suitable also in the presence of unbalanced panels. The latter feature is particularly relevant when accounting for the ESG dimension over time since the investment set consists of assets with different time of listing and availability of ESG scores, implying the dataset is an unbalanced panel.

In this Section we describe the optimization model that we use starting from a traditional mean-variance optimization framework extended to include an ESG constraint (Section 3.1) to end up with a reformulation that brings the above-cited technical advantages both in the presence and in the absence of an ESG constraint (Section 3.2).

3.1. Mean-variance optimization: the inclusion of an ESG objective

According to Markowitz (1952), risk-averse investors seek the portfolio that maximizes the expected return and minimizes the variance.

[5] Given that existing mutual fund portfolios are likely to be inefficient, Utz et al. (2104) implement the inverse optimization process over the closest matched portfolios that are on the nondominated surface.

[6] The ε-constraint method is a technique used to solve multicriteria optimization problems. It reformulates the original problem by optimizing only one of the original objectives while the others enter the optimization problem as constraints (Haimes et al., 1971) and uses the ε level comparison, which compares search points based on the pair of objective value and constraint violation of them.

[7] It is worth mentioning that real portfolios proposed by financial advisors often are not Pareto efficient, as witnessed by the reverse engineering process conducted by Utz et al. (2014) that have to proxy each analysed real portfolio through the closest efficient point on the non-dominated surface.

Hence, considering N risky assets, we recall the classical mean-variance portfolio optimization model:[8]

$$\max_{w} \alpha\mu_P + \lambda\sigma_P^2 \tag{1}$$

$$s.t. \quad w^T\mathbf{1} = 1$$

where:

$w \in \mathbb{R}^N$ is a vector of portfolio weights;
$\alpha \in \mathbb{R}_0^+$ represents investor's return preference;
$\mu_P \in \mathbb{R}^N$ is a vector of portfolio expected excess returns, defined as μw where $\mu \in \mathbb{R}^{1\times N}$ is a vector of expected asset excess returns
$\lambda \in \mathbb{R}^-$ represents investor's risk preference;
σ_P^2 is the portfolio return variance defined as w^TVw where V is an $N \times N$ positive semidefinite variance-covariance matrix of asset returns.

Because of investors' preferences for sustainable investments (Rossi et al., 2019; Hong and Kacperczyk, 2009) model (1) can be both modified to account for negative screening strategies based on ESG scores and extended by incorporating ESG beside market risk and return (Varmaz et al., 2024; Cesarone et al., 2022; Pedersen et al., 2021; Utz et al., 2014; Gasser et al., 2017). The threshold levels θ_s of ESG for screening are only used to shrink the set of feasible portfolio weights w.[9] Given θ_s, the set of feasible weight vectors w is

$$W_s = \{w \in \mathbb{R}^N \mid \mathbf{1}w = 1,\ w_i = 0\ \forall\ i \text{ such that } \theta_i < \theta_s\} \tag{2}$$

where $\mathbf{1} = (1, 1, \cdots, 1) \in \mathbb{R}^{1\times N}$.

Further, as previous studies we assume the additivity of ESG scores across assets in line with Drut (2010). It results in a tri-objective optimization problem that still represents a standard multi-criteria decision making (MCDM) model which with respect to equation (1) also maximizes the portfolio ESG score θ_P. Given a screening threshold $\theta_s \geq 0$ and the feasible set W_s in (2), the model (1) is restated as

$$\max_{w\in W_S} \alpha\mu_P + \epsilon\theta_P + \lambda\sigma_P^2 \tag{3}$$

where:

$\epsilon \in \mathbb{R}_0^+$ represents investor's ESG preference;
θ_P is the portfolio ESG score defined as θw where $\theta \in \mathbb{R}^{1\times N}$ is a vector of asset ESG scores.

The problem in (3) takes μ, V and θ as parameters that can be estimated from the data, whereas the parameters α, λ and ϵ must be specified a priori consistently with investors' preferences. However, investors might encounter some difficulties in quantifying their preferences with α, λ and ϵ because they are not directly observable. Rather, it is easier for investors to express their desired levels for portfolio expected excess return (μ_P^*) and ESG score (θ_P^*). Let w^* denote an optimal weight vector in (3) with optimal expected excess return $\mu_P^* = \mu w^*$ and ESG score $\theta_P^* = \theta w^*$. Then, omitting constant terms in the objective function, problem (3) with $\lambda = -1$ is equivalently stated as

$$\min_{w\in W_s} w^TVw \tag{4}$$

$$s.t.\ \mu w = \mu_P^*$$
$$\theta w = \theta_P^*$$

By setting investors' desired portfolio characteristics (μ_P^* and θ_P^*) consistently with their preferences α and ϵ, both programs in (3) and (4) bring to the same optimal portfolio weights.

3.2. A reformulation: with and without ESG constraint

Starting from problem (4), a reformulation that rests on Varmaz et al. (2024) brings both technical and practical advantages in the incorporation of ESG into mean-variance optimization. The model for asset returns is assumed to be a single-factor model in which the ESG dimension (e.g. the ESG score) is a characteristic that can affect the return without affecting the covariance structure among assets.[10] This view rests on the idea that the impact of ESG dimensions on stock returns can be driven by factors specific to individual firms rather than systematic market-wide effects and is supported by both the theoretical and the empirical literature (Bénabou and Tirole, 2010; Edmans, 2011; Friede et al., 2015). For instance, companies with higher ESG ratings may benefit from improved operational efficiency, reduced regulatory risks, or enhanced stakeholder trust, which can translate into higher profitability. Hence, an asset expected excess return can be described as a linear function of the market risk measure (beta) and of an ESG characteristic (e.g. ESG score or rating):

$$\mu_i = \beta_i(E(R_m)) + \theta_i c \tag{5}$$

where:

μ_i= expected excess return of asset i;
R_m = excess return of the market portfolio, i.e. the market factor;
β_i = sensitivity of asset i return to the market factor;
θ_i = ESG characteristic of asset i;
c = reward for the ESG characteristic.[11]

It follows that the portfolio expected excess return is

$$\mu_p = R_m\,\beta w + c\theta w \tag{6}$$

with $\beta w = \beta_P$ representing the portfolio beta and $\beta \in \mathbb{R}^{1\times N}$ is a vector of asset betas.

We have to recall that according to the CAPM (Sharpe, 1964; Lintner, 1965; Mossin, 1966), and more generally to a single-index model, portfolio variance can be rewritten as a function of the variance σ_{Rm}^2 of excess return of the market portfolio and residual risk (residual variance) σ_ϵ^2 :

[8] Markowitz's problem can be represented in a mean-variance plane because it assumes that investors select portfolios exclusively on the basis of the expected return and the expected variance of asset returns. This assumption is supported either by normally distributed returns (for any expected utility function) or by a quadratic utility function (for any return distribution), which represents risk-averse individuals.

[9] To be noted that the screening threshold θ_s is chosen by each DM, and there are two alternatives for interpretation: θ_s is a parameter associated with the DM similarly as the value function parameters; or θ_s is a decision variable of the DM (i.e. the value function is maximized over the weights vectors $w \in W_s$ as well as over the threshold θ_s which determines W_S). In the present paper we adopt the first interpretation.

[10] Varmaz et al. (2024) show that the model can be easily extended in order to accommodate more risk factors and also an ESG risk factor consistently with the theory that ESG can lead to a factor risk premium affecting returns (e.g. Pástor et al., 2021; Pedersen et al., 2021). At this stage of the analysis we consider a single-factor model and the theory according to which ESG can be seen as a characteristic affecting return without translating into more/less risk. Moreover, in the example proposed by Varmaz et al. (2024) there is a quite high correlation (-35%) between the market risk factor and the ESG risk factor.

[11] The variable c, representing the reward for the ESG characteristic, is independent of the specific stock, analogous to how the expected market return in the CAPM is independent of the individual asset. The extra return of a specific stock due to its ESG characteristic is given by the product $\theta_i c$ where θ_i denotes the ESG score of the stock, capturing the magnitude of its sustainability.

$$\sigma_P^2 = \sigma_{Rm}^2(\beta w)^2 + \sigma_\varepsilon^2\ w^T w \tag{7}$$

To be noted that the reformulation, in line with Varmaz et al.(2024), assumes i.i.d. residual returns (i.e. ε_i, residuals in the regression of an asset excess return over the risk factors and characteristics) implying that $Cov(\varepsilon_i,\ \varepsilon_j) = 0$ if $i \neq j$ and $Cov(\varepsilon_i,\ \varepsilon_j) = \sigma_\varepsilon^2$ if $i = j$.[12]

Substituting (6) and (7), problem (3) can be rewritten as

$$\max_{w\in W_S} \alpha \mathrm{R_m}\beta w + (\alpha c + \epsilon)\theta w + \lambda\left[\sigma_{Rm}^2(\beta w)^2 + \sigma_\varepsilon^2\ w^T w\ \right] \tag{8}$$

Let $w = w^*$ be an optimal solution to (8), then the optimal ESG is $\theta_P^* = \theta w^*$, the optimal beta is $\beta_P^* = \beta w^*$ and the optimal expected excess return $\mu_P^* = R_m\ \beta_P^* + c\theta_P^*$ is uniquely determined by (6). After omitting constant terms in the objective function and scaling by $\frac{1}{\sigma_\varepsilon^2}$, problem (8) with $\lambda = -1$ is equivalently stated as

$$\min_{w\in W_S} w^T w \tag{9}$$

$$s.t.\ \beta w = \beta_P^*$$
$$\theta w = \theta_P^*$$

The final optimization problem in (9) aims at minimizing residual risk by setting a desired level of portfolio beta and ESG score in line with investors' preferences. According to Varmaz et al. (2024) a more compact representation is

$$\begin{aligned}&\min_{w\in W_S} w^T w\\ &s.t.\ \ X^T w = b\end{aligned} \tag{10}$$

where:

$X = [1,\ \beta,\ \theta]$, a $N \times 3$ matrix that gathers the budget constraint (i.e. the sum of $w \in W_S$ is 1) and the variables on the left-hand side of the constraints of (9);

$b = \left[1,\ \beta_P^*,\ \theta_P^*\right]^T$, a vector that gathers the budget constraint and the variables on the right-hand side of the constraints of (9).

The solution differs across investors, because they have individual preferences for the desired values in vector b, and is represented as follows:

$$w^T = b^T\left(X^T X\right)^{-1} X^T \tag{11}$$

Problem in (10) is reported without specifying the subscript t, but it can be solved for each time t in our sample retrieving a vector of optimal weights. We are then able to compute the out-of-sample realized returns R_t of the portfolio at time t

$$R_t = w_{t-1}^T r_t = b^T\left(X_{t-1}^T X_{t-1}\right)^{-1} X_{t-1}^T r_t \tag{12}$$

where:

$r_t \in \mathbb{R}^N$ is a vector of assets returns at time t.

This model presents four main advantages with respect to the traditional mean-variance approach and its extension to incorporate ESG. First, by eliminating the portfolio variance from the objective function and by setting only equality constraints, it reduces the computational complexity of the problem and brings to an analytical solution.[13] Second, the model does not require the estimation of expected returns and the variance-covariance matrix. The latter has been criticized to be often a source of instability (López de Prado, 2020) or unreliable in the presence of a large number of assets (Shanken, 1992), and cannot be calculated when the panel is unbalanced, whereby panels of individual stocks are often unbalanced since stocks can be listed and delisted and firms can merge. Third, investors can more easily specify the desired level of risk, return and ESG for their portfolio, without having to set more abstract preference parameters. Fourth, the proposed model is flexible enough to accommodate different return-generating models from the simplest single-factor model to multi-factor models such as the Fama and French three-factor model (Fama and French, 1993).[14] Moreover, the ESG dimension can be included as a simple characteristic that affects stock returns only or it can be considered as a risk factor that affects stock returns by means of changing the risk.

This reformulation is useful also when optimizing without an ESG constraint, assuming that asset returns are described by a traditional CAPM model, hence driven by systemic risk only. The optimal portfolio can be found by minimizing residual risk as follows:

$$\min_{w\in W_S} w^T w \tag{13}$$

$$s.t.\ \beta w = \beta_P^*$$

The problem (13) is equivalent to (8) with $\lambda = -1$ if we impose the following assumption on the DM's preferences

$$\alpha c + \epsilon = 0 \tag{14}$$

Although the optimization model in (9), originally proposed by Varmaz et al. (2024), may be attractive to investors who have the opportunity to choose a portfolio with ESG and risk characteristics consistent with their preferences, such an approach is not without limitations. In particular, this approach imposes equality constraints and requires setting pre-specified target values for the constraints. In addition, a return generating model must be assumed, and following an a priori (vs. a posteriori) approach does not provide a complete representation of the efficient surface on which the optimal portfolios are situated. Hence, this approach may not guarantee Pareto optimality due to the presence of equality constraints, as discussed in Section 2.3.

4. Dataset and descriptive statistics

We can focus on single assets because the model in (9) is suitable also for unbalanced panels so that we do not have to assume investments in funds as most of the models in the literature (see Section 2.3). We start from all the stocks that were part of the EURO STOXX Index, a subset of the STOXX Europe 600 Index, from January 2007 to August 2022, so as to grant stocks' liquidity and market representativeness. All the index components belong to large, mid and small capitalization companies of 11 Eurozone countries therefore stock prices are expressed in the same currency (Euro) and are not affected by exchange rates.[15] The number of components in a given month is not fixed, but it is on average around 300 components every month. The final sample consists of 586 stocks

[12] The uncorrelation of residuals is one of the assumptions behind the factor pricing approach (see e.g. Cochrane, 2005), while equality of variances across assets requires assuming that risk factors explain a significant part of asset variances and it is often taken in the literature on this topic (see Varmaz et al., 2024; Daniel et al., 2020).

[13] The problem must be solved numerically in the case weight constraints are added (i.e. weights must not become negative).

[14] In a multi-factor framework, the approach by Varmaz et al. (2024) allows to set desired target levels for factors (beside beta and ESG score) as constraints, making the model appealing to many quantitative portfolio investment managers interested in ESG portfolio construction by means of factor investing (Melas, 2021).

[15] The index is very liquid, in fact it is frequently used as an underlying of both ETFs and derivatives. The 11 countries are: Austria, Belgium, Finland, France, Germany, Ireland, Italy, Luxembourg, the Netherlands, Portugal and Spain.

and their monthly total returns, which include also dividends beside capital gains, are retrieved from Bloomberg.

We assume that the expected stock excess returns are determined by a single-index model, as the one represented in equation (5), in which the only risk factor is the market factor and the ESG characteristic affects stock returns without modifying the risk profile.[16] In order to obtain the optimal weights solution in equation (11) we do not need to estimate the excess return of the market portfolio ($R_{m,t}$) and the reward for the ESG characteristic (c), because only the market beta (β_i) and the ESG score (θ_i) are required. Market betas are retrieved from Bloomberg and they are determined by comparing the price movements of the stock and the representative market for the past two years of weekly data: for example, for the Italian energy company Terna beta is calculated with respect to the FTSE MIB Index that is the primary benchmark index for the Italian equity market.

The ESG dimension in stocks is quantified by the Bloomberg ESG disclosure score that measures the amount of ESG data a company discloses based on public data (sustainability reports, annual reports, websites, publicly available resources and direct contact with the companies being assessed). The advantage of the Bloomberg ESG scores is that they are available for years further back with respect to other scores (e.g. Sustainalytics).[17] Moreover, this metric is supported by the literature whereby a higher ESG commitment is associated to a higher transparency in the disclosure of socially responsible information and it may have a positive outcome on corporate social responsibility (Chen and Xie, 2022). These scores range between 0, if there is no ESG data disclosure, and 100 when companies disclose every relevant data points. Bloomberg provides both individual scores on the three ESG pillars (environmental, social and governance) and an overall ESG score that equally weights the three individual scores.[18] In the present paper we focus on the aggregate measure of ESG to gauge overall sustainability and in line with most studies on optimal sustainable portfolios (Varmaz et al. 2024; Gasser et al., 2017; Cesarone et al., 2022; Pedersen et al., 2021).

Table 1 presents some descriptive statistics of key variables to solve the optimization problem: market betas and ESG scores. The average stock sensitivity (β) to the market portfolio is 0.964 suggesting that our sample on average well represents the reference market, although with a wide range (minimum and maximum betas are -0.997 and 2.558 respectively). Low and negative values for beta are mainly referred to stocks that have been listed towards the end of the analysed period thus showing little or negative co-movement with the market; while betas higher than 2 are mostly associated to aggressive stocks that have been delisted during the analysed time period. The average ESG score is 34.019 with great variability across stocks; the minimum value is 0, i.e. no Bloomberg score, while the maximum average score is 70.770. The correlation between market beta the ESG score for each stock is on average very low (0.043) indicating that on average the two variables are quite independent.

In Table 2 we report descriptive statistics focusing on different levels of ESG score (including the non-availability of the score itself) for the whole period and for two subperiods. Specifically, three different classes of ESG scores are defined (0-30, 30-50, 50-100) so as to have the same number of assets over the whole period, however the statistics obtained are robust to different ESG clustering (available upon request).

The statistics on different ESG-score classes provide more insight on the relation between ESG scores and the return-risk reward, while those on different subperiods provide information on the temporal distribution of ESG scores that has varied greatly over time. Descriptive statistics over the whole period (Panel A) and over subperiods (Panel B and C) show that the relation between return-risk reward (Mean/St. Dev.) and ESG score is substantially convex, being lower for intermediate ESG ratings (30-50) and higher for the lowest (0-30) and the highest rating (50-100) classes. By focusing on subperiods, the dynamic of the ESG market emerges. First, the number of stocks with a Bloomberg ESG score increases over time: the first subperiod (Panel B) from January 2007 to December 2014 represents an ESG market in its infancy, as demonstrated by the fact that most stocks (237) are in the first class with low (or no) scores, while the second subperiod (Panel C), from January 2015 to August 2022, refers to a more developed ESG market spurred, in 2015, by both the UN 2030 Agenda and the Paris Agreement on climate change and most stocks (233) are in the highest rating class (50-100).

5. Empirical strategy

We estimate three types of portfolios accounting for ESG dimensions in three different ways and thus corresponding to different ESG characterizations: the first uses only prior ESG screening (with different levels of tightness θ_s) to exclude stocks up to a certain level of ESG score with no guarantee on the resulting portfolio ESG score, the second uses a constraint in the optimization so as to reach a portfolio ESG-target θ_P with no guarantee on the ESG-score level of the stocks included in the optimal portfolio, the third combines features of both so as to obtain an optimal portfolio that on one hand reaches an ESG-target, on the other it does not include stocks whose ESG score is below a minimum level. Main features of the three strategies are summarized in Table 3.

Formally, the first results from prior screening of the sample based on a screening threshold $\theta_s \neq 0$ and optimization of residual risk with no ESG constraints θ_P based on (13); the second is obtained from an unscreened sample (i.e. $\theta_s = 0$) by minimizing residual risk with ESG-score constraint θ_P as in (9), while the third optimizes with an ESG constraint θ_P as in (9) yet over a slightly screened sample ($\theta_s \neq 0$).

All three strategies require specification of a desired level of systemic risk β, an ESG-score level θ_s for screening, and a desired portfolio average ESG-score θ_P. In the following we explain how we set these levels, robustness tests over these choices are presented in Section 8 and an analysis to evaluate how close the obtained optimal solutions w^* from problems (9) and (13) are to Pareto optimality is presented in the Appendix.

The different values for β, θ_s and θ_P are meant to represent portfolios attainable by investors, which can be used as the so-called Model Portfolios by the financial industry (Table 4).[19] Specifically, β can assume values 0.5, 1, and 1.5 so as to represent a defensive portfolio, a market tracking portfolio, and an aggressive portfolio respectively. As for the screening threshold θ_s, we consider it equal to zero when we consider an unscreened sample and equal to 20, 40 or 50 in the case of a screened sample. Hence in the latter case we identify three exclusion levels (ESG scores ≥ 20, ≥ 40, ≥ 50) that are consistent with the ESG-score distribution of the dataset (Table 2) and correspond to an increasing

[16] The three-factor Fama-French model is considered in the robustness in Section 8.

[17] Sustainalytics, for example, has a low coverage before 2014 and this is explained by the fact that before 2014, it was the needs of Sustainalytics clients that determined which companies received the ESG score (Auer, 2016).

[18] It has to be noted that the methodology for Bloomberg ESG Disclosure Scores was updated in early 2022, to account for the evolution of corporate ESG data reporting since the scores were originally created.

[19] On the other hand, each Model Portfolio is associated to a set G of individual DMs. Hence, the Strategies 1-3 can be associated with subsets G_1, G_2 and G_3 of the set G as follows. Strategy 1 is adopted by the subset G_1 of 3×3 individuals whose most preferred β_P^* is 0.5, 1 or 1.5; chosen ESG threshold θ_s is 20, 40 or 50; value function weights satisfy (14). Strategy 2 is adopted by the subset G_2 of 3×3 individuals whose most preferred β_P^* is 0.5, 1 or 1.5; chosen ESG threshold θ_s is 0; most preferred ESG score θ_P^* is 20, 40 or 50. Strategy 3 is adopted by the subset G_3 of 2×3 individuals whose most preferred β_P^* is 0.5, 1 or 1.5; chosen ESG threshold θ_s is 20; most preferred ESG score θ_P^* is 40 or 50. As a consequence, another interpretation of the paper is that it considers three groups of DMs (G_1, G_2, and G_3), each following a given Strategy (1, 2 or 3).

Table 1
Descriptive statistics of market betas and ESG scores Whole sample, whole period 2007–2022.

	Min.	Median	Mean	Max	St. Dev.	P(25)	P(75)
Mean β	-0.997	0.919	0.964	2.558	0.448	0.672	1.231
Mean ESG score (θ)	0.000	36.565	34.019	70.770	16.825	23.651	47.154
Corr (β, θ)	-0.860	0.048	0.043	0.897	0.381	-0.257	0.317

Notes: the table reports minimum, median, mean, maximum, standard deviation, 25th percentile and 75th percentile of the time series mean of market beta and ESG score. Corr (β, θ) indicates, for each of the 541 stocks Bloomberg assigns a score, the correlation between the time series of the score itself and the beta.

Table 2
Descriptive statistics of stock monthly returns for different ESG scores.

Panel A: Whole period: January 2007 – August 2022							
ESG score	Median (%)	Mean (%)	St. Dev. (%)	P(25) (%)	P(75) (%)	Mean/St. Dev.	N
0 - 30	1.369	0.756	5.409	-2.193	3.939	0.140	151
30 - 50	0.820	0.449	5.615	-2.502	3.856	0.080	161
50 - 100	0.973	0.759	5.186	-2.091	3.927	0.146	142
Panel B: First subperiod: January 2007– December 2014							
ESG score	Median (%)	Mean (%)	St. Dev. (%)	P(25) (%)	P(75) (%)	Mean/St. Dev.	N
0 - 30	1.349	0.569	5.761	-2.171	4.063	0.099	237
30 - 50	0.735	0.471	6.095	-2.668	4.736	0.077	163
50 - 100	0.865	0.846	5.125	-2.431	4.395	0.165	56
Panel C: Second subperiod: January 2015 – August 2022							
ESG score	Median (%)	Mean (%)	St. Dev. (%)	P(25) (%)	P(75) (%)	Mean/St. Dev.	N
0 - 30	1.393	0.953	5.036	-2.231	3.851	0.189	60
30 - 50	0.840	0.426	5.094	-2.113	3.459	0.084	158
50 - 100	0.973	0.666	5.276	-1.831	3.541	0.126	233

Notes: each panel of the table reports return statistics (median, mean, standard deviation, 25th percentile, 75th percentile and the ratio between mean and standard deviation) of equally weighted portfolios that are made up of stocks with an ESG score indicated in the first column. A score equal to zero refers to stocks for which Bloomberg does not assign a score. Portfolios are set up so that the 0–30 portfolio consists of all the stocks with an ESG score greater than or equal to zero and lower than 30. The composition of such portfolios can change over time because ESG scores are not constant over time. Hence, we calculate the return in month t for each portfolio and we calculate those statistics on the time series of monthly portfolio returns. "N" in the last column indicates the time series average number of assets in each portfolio.

Table 3
ESG-compliant portfolios: strategies at comparison.

	Prior ESG-screening	ESG-score constraint in optimization
Strategy 1	yes	no
Strategy 2	no	yes
Strategy 3	yes (mild)	yes

attention towards sustainable investments (weak, intermediate, high respectively), whereby we take 50 as representative of the upper class since higher scores (e.g. $\geq$ 60) would imply the impossibility to find, at many times t, optimal portfolios when implementing the screening. Similarly, as for the ESG score target θ_P, we identify preferences for increasing levels of average portfolio ESG score that we take as representative of the classes in Table 2 (20, 40, 50). As a result, we obtain 9 portfolios given by combinations of desired levels of beta and ESG screening threshold/target. As for the level of minimum screening to obtain the third type of portfolio, we take $\theta_s =$ 20.

We allow shorting stocks, for two main reasons: first, as with all assets (also non ESG ones), shorting is a way to display a view even more effective than simply not holding the asset; second ESG-investors look for an impact through their financial choices and this impact, which cannot be reached taking no position, can be attained by taking a short position even if not as much as a long one (where the engagement has more real implications, e.g. voting). Hence, in line with the literature (e.g. Pedersen et al., 2021; Fitzgibbons et al., 2018), optimal portfolios can be characterized by both long and short positions so as to improve investors' trade-off between risk and return and by shorting assets with a lower ESG score they can obtain a better overall portfolio score. Short positions can serve as a mechanism for investors to reach a specific ESG target and are thus fundamental also to avoid that such targets are reached by a reduction of diversification. On the other hand, by setting short sale constraints, investors avoid to have extreme long and short positions designed to exploit small differences in the structure of returns (Jacobs et al., 2014).

Once optimal weights are calculated according to equation (11), we compute out-of-sample realized returns with equation (12) for each period t and to do so we use beta and ESG score referred to period $t-1$. We have to recall that, differently from betas that are available monthly, Bloomberg provides ESG scores on an annual basis and are referred to a fiscal year, so in an out-of-sample perspective, the ESG score on December, 31 2006 impacts portfolio construction for the full fiscal year 2007. Then, starting from realized portfolio returns we measure portfolio risk-adjusted performance over the whole period (2007-2022) by means of the Sharpe ratio, since it is a widely used measure appropriate also for returns that deviate from a normal distribution (Auer, 2016).[20] To facilitate results interpretation, it is important to highlight the two distinct stages of the analysis: portfolio optimization with the comparison of alternative choices based on the linear value function in (8) and

[20] Studies by Schuhmacher and Eling (2011 and 2012) demonstrate that the conditions for the decision-theoretic foundation of the Sharpe ratio are the same of other admissible performance measures that are skewed and exhibit fat tails i.e. are more realistic. Further, also the resulting performance ranking is the same.

Table 4
Portfolios attainable: possible combinations of systemic risk and ESG-compliance.

Strategy 1			
Screening threshold (θ_s)	20	40	50
ESG target (θ_P)	-	-	-
β = 0.5 (defensive)	low exclusion	medium exclusion	high exclusion
β = 1 (mkt tracking)	low exclusion	medium exclusion	high exclusion
β = 1.5 (aggressive)	low exclusion	medium exclusion	high exclusion
Strategy 2			
Screening threshold (θ_s)	0	0	0
ESG target (θ_P)	20	40	50
β = 0.5 (defensive)	low ESG	medium ESG	high ESG
β = 1 (mkt tracking)	low ESG	medium ESG	high ESG
β = 1.5 (aggressive)	low ESG	medium ESG	high ESG
Strategy 3			
Screening threshold (θ_s)	-	20	20
ESG target (θ_P)	-	40	50
β = 0.5 (defensive)	-	low exclusion; medium ESG	low exclusion; high ESG
β = 1 (mkt tracking)	-	low exclusion; medium ESG	low exclusion; high ESG
β = 1.5 (aggressive)	-	low exclusion; medium ESG	low exclusion; high ESG

Notes: The table represents, for each strategy, features of portfolios resulting from combinations of systemic risk (β) and ESG screening threshold/target (θ_s, θ_P).

the quality judgment of a given portfolio based on out-of-sample performance measured by mean and standard deviation of excess return (defining the Sharpe ratio), over many monthly optimization results.

The Sharpe ratio for portfolio p is calculated as the ratio between the portfolio mean excess return μ and its standard deviation σ:[21]

$$SR_p = \frac{\mu}{\sigma} \tag{15}$$

We compare Sharpe ratios of different portfolios by means of a bootstrap test (Ledoit and Wolf, 2008). Such statistical test is robust to non-normality, correlation and errors due to small samples (Auer, 2016; Auer and Schuhmacher, 2013). The null hypothesis implies that the Sharpe ratios of two portfolios are the same:

$$H_0: \ \Delta = \eta_j - \eta_k = 0$$

where:

η_j = true Sharpe ratio of portfolio j
η_k = true Sharpe ratio of portfolio k.

Our final aim is to investigate the relation between portfolio Sharpe ratio and the desired level of portfolio systemic risk and ESG score and to compare this relation between the three optimal portfolio strategies.

6. Results

We start by comparing the risk-adjusted performance of the two philosophically farthest strategies to ESG-compliant portfolio, i.e. screening the investment set based on ESG scores before optimizing (Section 6.1) vs. constraining in terms of ESG while optimizing (Section 6.2). Finally, we propose a third strategy (Section 6.3) combining features of both by optimizing with an ESG constraint over a slightly screened sample.

For each strategy, we determine, for each month of the dataset, the optimal weights of portfolios: for the first two strategies we estimate 9 optimal portfolios characterized by the combination of different desired levels of portfolio β (0.5, 1, 1.5) and portfolio both θ_s and θ_P (20, 40, 50), while for the latter mixed strategy we estimate 6 optimal portfolios corresponding to a low screening threshold ($\theta_s = 20$), the desired levels of portfolio β (0.5, 1, 1.5) and two portfolio ESG score constraints θ_P (40, 50) reasonably consistent with a prior screening of 20. Then we calculate the out-of-sample realized return in the next month:[22] for each portfolio we obtain a time series of realized returns that we use to calculate portfolio performance by means of the Sharpe ratios. Robustness tests on the above-defined ESG scores, different performance measures and a different return-generating model are presented in Section 8.

6.1. Portfolio optimization over an ESG screened sample

The first strategy, which is widely used in the industry, consists in accounting for the ESG dimension by implementing first a negative screening strategy on the investment set and then using an optimization model. As for the choice of the latter, given the computational advantages stressed in Section 3.2, we adopt model (13) that minimizes residual risk with a specific level of systemic risk.

Table 5 reports results for different combinations of systemic risk and different levels of screening. Three main comments emerge. First, for each level of beta (i.e. systemic risk), the average portfolio ESG score θ_P (even if unconstrained) is always higher than the screening threshold and monotonically increases with the screening threshold, the mean (i.e. the portfolio return) increases with the screening threshold except for portfolio #2, the standard deviation (i.e. total risk) is quite stable: accordingly the Sharpe ratio (i.e. the risk-adjusted financial performance) tends to increase with the screening threshold. Second, for each level of screening, the mean increases with beta except for ESG $\geq$20 screening, standard deviation always increases with beta: accordingly the Sharpe ratio tends to increase with beta. Third and most interestingly, the last two columns highlight that, for portfolios mimicking the market (beta = 1) or aggressive portfolios (beta = 1.5), the better financial performance is statistically significant only when heavy screening is implemented (ESG $\geq$ 50) and θ_P results above 55.[23] Overall,

[21] The risk-free rate chosen to compute excess returns is the 1-month Euribor retrieved from the database of the German Central Bank (https://www.bundesbank.de/en/statistics/time-series-databases).

[22] Constraints on the level of β and θ_P, when it is the case, must be satisfied in each month.

[23] This result may also be interpreted for three DMs differentiated by β_P only. Then $\theta_S \in \{20, 40, 50\}$ is a choice of each DM, and the best choice based on out-of-sample results in Table 5 (with highest mean and lowest standard deviation) would have been $\theta_S = 50$ for all the three DMs.

Table 5
Optimal portfolio performance over a screened sample (no ESG target constraint) - Whole period 2007–2022.

#P	θ_S	β	θ_P	Mean	St.Dev.	SR	SR_i - SR_{i-1}	SR_i - SR_{i-2}
1	20	0.5	43.704	-0.013	4.772	-0.003		
2	40	0.5	51.417	-0.037	4.66	-0.008	-0.005	
3	50	0.5	56.589	0.041	4.537	0.009	0.017	0.012
4	20	1	44.098	-0.032	5.951	-0.005		
5	40	1	51.209	-0.006	5.877	-0.001	0.004	
6	50	1	56.241	0.200	5.643	0.036	0.037**	0.041*
7	20	1.5	44.494	-0.050	7.41	-0.007		
8	40	1.5	51.002	0.025	7.457	0.003	0.01	
9	50	1.5	55.894	0.360	7.386	0.049	0.046*	0.056*

Notes: #P indicates the progressive number of portfolios, θ_S the screening threshold implying exclusion of stocks with an ESG score lower than the threshold from the portfolio optimization. β indicates the average desired exposure to systemic risk and $\theta_P = \theta w^*$ the optimal portfolio ESG score, calculated as the time-average of the portfolio's weighted ESG scores at each month. Each combination of θ_S and β identifies an investor-specific portfolio. Mean, standard deviation and SR (Sharpe ratio) refer to the time series of out-of-sample portfolio realized returns. In each portfolio the budget constraint is respected, so the sum of portfolio weights is 1. The last two columns report, for each level of β, the approximated Sharpe ratio difference between portfolio with screening level *i* and portfolios with the one lower screening level *i-1* and two lower screening levels *i-2* respectively. ***, ** and * represent significance at 1, 5, 10% levels, respectively, in the Sharpe ratio difference.

Table 6
Optimal portfolio performance for different levels of desired beta and ESG score - Whole period 2007–2022, unscreened sample.

#P	θ_S	β	θ_P	Mean	St.Dev.	SR	SR_i - SR_{i-1}	SR_i - SR_{i-2}
1	0	0.5	20	0.089	4.928	0.018		
2	0	0.5	40	0.008	4.835	0.002	-0.016	
3	0	0.5	50	-0.033	4.874	-0.007	-0.009	-0.025
4	0	1	20	0.074	5.971	0.012		
5	0	1	40	-0.008	5.885	-0.001	-0.013	
6	0	1	50	-0.049	5.913	-0.008	-0.007	-0.020
7	0	1.5	20	0.058	7.271	0.008		
8	0	1.5	40	-0.024	7.194	-0.003	-0.011	
9	0	1.5	50	-0.064	7.213	-0.009	-0.006	-0.017

Notes: #P indicates the progressive number of portfolios; θ_S the screening threshold implying exclusion of stocks with an ESG score lower than the threshold from the portfolio optimization, β the average desired exposure to systemic risk and $\theta_P = \theta w^*$ the optimal portfolio ESG score, calculated as the time-average of the portfolio's weighted ESG scores at each month. Each combination of β and θ_P identifies an investor-specific portfolio. Mean, standard deviation and SR (Sharpe ratio) refer to the time series of out-of-sample portfolio realized returns. In each portfolio the budget constraint is respected, so the sum of portfolio weights is 1. The last two columns report, for each level of β, the approximated Sharpe ratio difference between portfolio with target ESG level *i* and portfolios with the one lower target ESG level *i-1* and two lower target ESG levels *i-2* respectively. ***, ** and * represent significance at 1, 5, 10% levels, respectively, in the Sharpe ratio difference.

Table 7
Optimal portfolio performance (with ESG score constraint) over a slightly screened sample - Whole period 2007–2022.

#P	θ_S	β	θ_P	Mean	St.Dev.	SR	SR_i - SR_{i-1}
1	20	0.5	40	-0.006	4.707	-0.001	
2	20	0.5	50	0.053	4.698	0.011	0.012
3	20	1	40	-0.028	5.856	-0.005	
4	20	1	50	0.031	5.791	0.005	0.010
5	20	1.5	40	-0.050	7.301	-0.007	
6	20	1.5	50	0.009	7.202	0.001	0.008

Notes: #P indicates the progressive number of portfolios, θ_S the screening threshold implying exclusion of stocks with an ESG score lower than the threshold from the portfolio optimization, β indicates the average desired exposure to systemic risk and $\theta_P = \theta w^*$ the optimal portfolio ESG score, calculated as the time-average of the portfolio's weighted ESG scores at each month. Each combination of θ_S, β and θ_P identifies an investor-specific portfolio. Mean, standard deviation and SR (Sharpe ratio) refer to the time series of out-of-sample portfolio realized returns. In each portfolio the budget constraint is respected, so the sum of portfolio weights is 1. The last column reports, for each level of β, the approximated Sharpe ratio difference between portfolio with target ESG level *i* and portfolio with the one lower target ESG level *i-1*. ***, ** and * represent significance at 1, 5, 10% levels, respectively, in the Sharpe ratio difference.

a high screening threshold enables obtaining optimal portfolios with both higher sustainability (θ_P) and higher performance.

From a sustainable finance perspective this strategy has, on one hand, the advantage of granting only the inclusion of stocks with the desired level of sustainability, on the other the disadvantage of an unknown (since unconstrained) ESG level in the final optimal portfolio (also because short positions are allowed).

6.2. Portfolio optimization over an unscreened sample with an ESG constraint

The second strategy, consists in accounting for the ESG dimensions by introducing an ESG average score constraint in the optimization model. Given the computational advantages stressed in Section 3.2, we adopt model (9) that minimizes residual risk with a desired level of portfolio beta and ESG score.

Table 6 reports results for different combinations of systemic risk and different levels of portfolio ESG target score. Three main comments are in order. First, for each level of beta (i.e. systemic risk), the mean (i.e. the portfolio return) decreases with the ESG target score, the standard deviation (i.e. total risk) is quite stable: accordingly the Sharpe ratio (i.e. risk-adjusted financial performance) tends to decrease with the ESG target score. Second, for each ESG score target, the mean decreases with beta, standard deviation always increases with beta: accordingly the Sharpe ratio tends to decrease with beta. Third and most interestingly, the last two columns highlight that, for each level of beta, the portfolio performance does not significantly worsen as the target ESG level increases. In sum, increasing the ESG portfolio target θ_P, i.e. portfolio sustainability, does not affect the optimal portfolio performance.

Comparison with Varmaz et al. (2022, 2024), which is the only one in the literature implementing the same strategy, highlights a main difference, since they find that the optimal portfolio performance increases along with the desired ESG score.[24] This different result can be explained by their dataset (stocks from the US S&P 500 Index and ESG scores from Refinitiv Datastream) whose descriptive statistics show a monotonically increasing relationship: higher ESG scores are associated

[24] It has to be noted that Varmaz et al. (2022), along with other research works on sustainable portfolios (e.g. Pedersen et al., 2021, Cesarone et al., 2022; Alessandrini and Jondeau, 2021) provide a comparative analysis of Sharpe ratios without testing statistical significance of differences.

to higher stock performance, whereas in our sample the relationship is convex (Table 2). A possible explanation is that the ESG target score constraint, in our investment set, forces the use of very heterogeneous stocks in terms of ESG and return/risk relationship.

From a sustainable finance perspective this strategy has, on one hand, the advantage of granting an optimal portfolio consistent with a target level of sustainability, on the other the disadvantage that such an average portfolio ESG score may be obtained by including also stocks with a very low level of sustainability.

6.3. Portfolio optimization with an ESG constraint over a marginally screened sample

Given the shortcomings of the previous two strategies, the third one consists in accounting for the ESG dimensions by both screening the sample and introducing an ESG average score constraint in the optimization model. The idea is of reaching an optimal portfolio with a target ESG score, including only stocks with a minimum level of sustainability, and for this reason the original investment set is only slightly screened (i.e. $\theta_S = 20$).

We screen each month the sample by excluding stocks with ESG score lower than 20 and we use model (9) with target scores of 40 and 50, obtaining 6 optimal portfolios. Results, for different combinations of systemic risk and different levels of portfolio ESG target score, are reported in Table 7. Three main comments emerge. First, for each level of beta (i.e. systemic risk), the mean (i.e. the portfolio return) increases with the ESG target score, the standard deviation (i.e. total risk) is quite stable: accordingly, the Sharpe ratio (i.e. the risk-adjusted financial performance) tends to increase with the ESG target score. Second, for each level of ESG target score, the mean decreases with beta, standard deviation always increases with beta, and, accordingly, the Sharpe ratio tends to decrease with beta. Third and most interestingly, the last column highlights that, for each level of beta, the performance does not significantly change as the target ESG level increases. In sum, also after a mild screening, increasing the ESG portfolio target θ_P, i.e. portfolio sustainability, does not affect the optimal portfolio performance.

In addition, Table 8 reports comparative results of the proposed third strategy with the previous two. Specifically, when comparing Strategy 3 with Strategy 1 (i.e. screening without any ESG target, Section 6.1), for comparability we can look at portfolios based on the same level of screening of 20, i.e. portfolios #1, 4, 7 in Table 5, characterized by ESG scores (θ_P) between 43.704 and 44.494 (Table 5), but always lower than the highest target in Strategy 3 of 50. It appears that, for each level of systemic risk, the absence of an ESG target constraint lowers SR, which are almost aligned to those of portfolios #1, 3, 5 in Table 7 that correspond to the lowest ESG target of 40. Hence, Strategy 3 brings the advantage of allowing investors to meet higher desired levels of portfolio average ESG score without sacrificing financial performance. When comparing Strategy 3 with Strategy 2 (i.e. optimization problem implemented with an ESG target score on an unscreened sample, Section 6.2), we can compare each portfolio in Table 7 with portfolios 2, 3, 5, 6, 8, 9 in Table 6 respectively. It appears that, for each level of beta, the sample screening, although mild, reverses the relationship: the higher the target ESG score, the higher the optimal portfolio performance. Moreover, portfolios #4 and #6 of Strategy 3 significantly outperform portfolios #6 and #9 of Strategy 2 highlighting that setting a high ESG target score is more favourable over a slightly screened sample (vs. an unscreened one).

These comparative results can be explained by the convex relationship observed in the investment set between ESG score and the ratio mean/standard deviation of returns. In fact, a mild negative screening excludes stocks characterized by a comparatively high ratio (Table 2, Panel A, first line) and, if intermediate ESG portfolio scores are targeted, portfolio will be mostly long on stocks with the lowest mean/standard deviation ratio (Table 2, Panel A, second line) whereas if highest ESG portfolio scores are targeted, portfolio will be mostly long on stocks with the highest mean/standard deviation ratio (Table 2, Panel A, third line).

Table 8
Sharpe ratio comparison of Strategy 3 vs. Strategy 1 and Strategy 2.

β	Strategy 1		Strategy 2		Strategy 3		Difference in SR	
	#P	SR	#P	SR	#P	SR	$SR_{s3} - SR_{s1}$	$SR_{s3} - SR_{s2}$
0.5	1	-0.003	2	0.002	1	-0.001	0.002	-0.003
0.5			3	-0.007	2	0.011	0.014	0.018
1	4	-0.005	5	-0.001	3	-0.005	0.000	-0.004
1			6	-0.008	4	0.005	0.010	0.013**
1.5	7	-0.007	8	-0.003	5	-0.007	0.000	-0.004
1.5			9	-0.009	6	0.001	0.008	0.010*

Notes: #P indicates portfolio number as enumerated in Tables 5–7 for Strategy 1, 2 and 3 respectively. The column "Difference in SR" shows, for each row, the difference in SR between Strategy 3 and Strategy 1 (on the left) and between Strategy 3 and Strategy 2 (on the right). Specifically, when Strategy 1 is considered, we subtract the SR of portfolio #1 to both portfolio #1 and #2 (of Strategy 3) as they share the same screening threshold (θ_S) and the same holds for higher level of systemic risk (β). ***, ** and * represent significance at 1, 5, 10% levels, respectively, in the Sharpe ratio difference.

7. Which is the best strategy for ESG-compliant portfolios on average and over time?

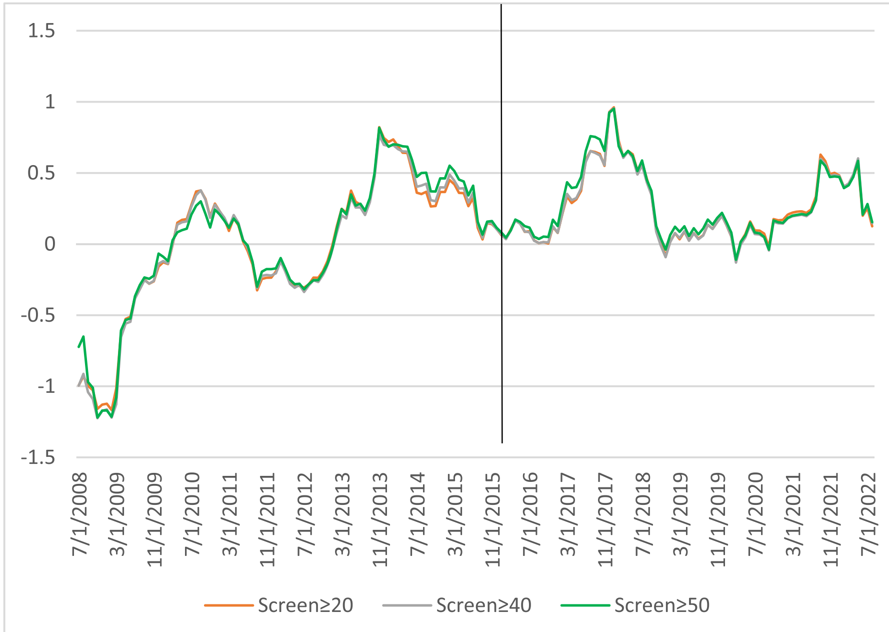
When comparing the three strategies over the whole period analysed, results in Section 6 can be summed up as follows. Optimization after screening (Strategy 1) implies a superior risk-adjusted performance only with heavy screening ($\theta_S = 50$). Accounting for ESG while optimizing (Strategy 2), returns portfolios with a performance that does not worsen as the target ESG level increases. Finally, optimizing with an ESG-score constraint after screening (Strategy 3) with a low threshold ($\theta_S = 20$) implies portfolios with a performance that does not worsen as the target ESG level increases, it is comparable to the one of Strategy 1 (with $\theta_S = 20$), it is higher than the performance of Strategy 2 (with $\theta_P = 50$). Comparatively, Strategy 3 does not worsen performance with respect to Strategy 1 (in which investors cannot control the average ESG score of the portfolio) and is able to overperform Strategy 2 (in which investors do not exclude stocks with a low ESG score).

Now the question is whether results change when focusing on shorter subperiods and, if so, whether they change over the financial cycle (e.g. in a procyclical vs anticyclical manner). To this end in this Section, we compute rolling Sharpe ratios, with window width equal to 18 months and different levels of systemic risk (β). Since they do not show relevant differences, for reasons of space in Fig. 1 we represent only the case of beta equal to 1 (other cases available upon request). In commenting results we will focus on the period after 2015, where the market is characterized by a greater ESG awareness following the publication of the Agenda 2030.

Panel a in Fig. 1 plots rolling Sharpe ratios for the first strategy: optimal portfolios based on heavy screening (ESG score greater or equal 50) tend to overperform other portfolios most of the time, although in the period 2020-2022, portfolios performances are rather aligned and do not seem to be driven by the screening threshold. The latter might be explained by the fact that both the Covid-19 pandemic and the greater maturity of the ESG market have led the market to a new normal, characterized by a convergence in the performance of portfolios of different shades of ESG.

Panel b in Fig. 1 plots rolling Sharpe ratios of the second strategy: rolling Sharpe ratios of portfolios with the highest ESG score (50) seem to reflect market phases being very close to/higher than Sharpe ratios of other portfolios in bullish market periods such as 2017 (when Sharpe ratio of all portfolios is increasing), and underperforming in periods of constant/bear market such as 2018-2019 (when Sharpe ratio of all portfolios is almost constant or decreasing). Finally, in the period 2020-

a) Portfolio optimization over an ESG screened sample



b) Portfolio optimization over an unscreened sample with ESG constraint

Fig. 1. Portfolios' rolling Sharpe ratio for different ESG strategies.
Notes: each figure represents the Sharpe ratio for β =1 and for different ESG levels (either screening thresholds θ_S or optimization constraints θ_P or both). Rolling window width = 18 months. The vertical line, corresponding to January 1st 2016, represents the coming into force of the UN 2030 Agenda.

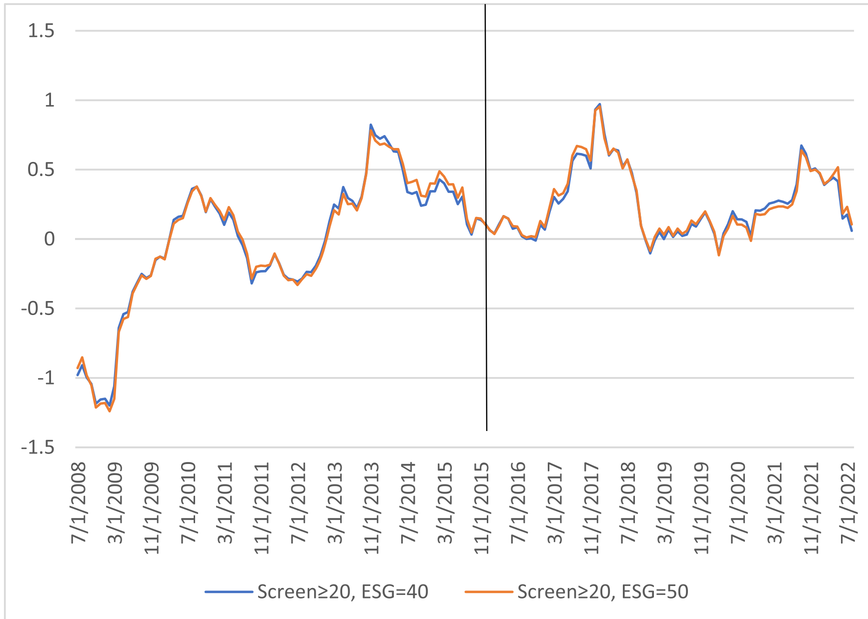
2022, when most stocks in the sample have a high ESG score, high-ESG portfolios overperform even when all Sharpe ratios are rapidly declining. Probably, in this latter period, a stronger awareness and demand for ESG assets is causing their performance to improve.

Panel c in Fig. 1 plots rolling Sharpe ratios of the latter strategy. Portfolios performance is almost overlapping regardless of the market phase; hence, relative portfolio performance does not show procyclicality features as in the screening before optimizing strategy (Fig. 1, Panel a). However, in 2021-2022 characterized by a fast-increasing awareness of ESG dimensions in financial decisions, the portfolio with the highest ESG target score slightly overperforms as in the optimizing without screening strategy (Fig. 1, Panel b).

Finally, Panel d in Fig. 1 plots rolling Sharpe ratios of Strategy 3 against Sharpe ratio of comparable portfolios in the first two strategies. Portfolios performance is almost overlapping over time since they share comparable sustainability features (screening threshold θ_S or target ESG score θ_P). The portfolio characterized by the highest level of sustainability (average ESG score of 50 over a mildly screened sample) tends to overperform other portfolios in both bullish market periods (2016-2017) and bearish market periods (2021-2022) after Covid-19 pandemic.

To sum up, when focusing on the period from 2015 following the

c) Portfolio optimization with ESG constraint over a marginally screened sample



d) Comparison of Strategy 3 vs. Strategy 1 and Strategy 2

Fig. 1. (*continued*).

publication of the UN 2030 Agenda, results show that the relative performance of optimal portfolios shows mild procyclicality only in strategies that account for ESG as a constraint in the optimization model, but such a relationship becomes countercyclical in the period 2020-2022 where the highest ESG portfolio overperforms even in a bearish market. By contrast, a strategy that accounts for ESG only by screening obtains quite similar performances across screening thresholds, even in the period following the Covid-19 pandemic outburst and does not show procyclicality features. Such a period is associated to a market in which investors are more aware and involved in sustainability issues, but the trend is not clear: it may result either in a superior performance of high ESG portfolios, or in new normal where there is a convergence in the performance of portfolios of different shades of ESG.

8. Robustness

This Section presents robustness tests to results presented in Section 6. We do so along three main lines: first we implement the three strategies described in Section 5 by including different levels of ESG screening and/or ESG target score; second we use alternative risk-adjusted measures with respect to the Sharpe ratio; third we consider the three-factor Fama-French model instead of the CAPM as a return generating process.

In terms of robustness w.r.t. the ESG scores, in Table 9 we report results from optimization of: the first strategy with different screening

Table 9
Optimal portfolio performance with more levels of ESG screenings and/or ESG target score - Whole period 2007–2022.

Panel A: optimization non-considering ESG over an ESG-screened sample

#P	θ_S	β	Mean	St.Dev.	SR	SR_i - SR_{i-1}	SR_i - SR_{i-2}
1	30	0.5	-0.024	4.686	-0.005		
2	40	0.5	-0.037	4.66	-0.008	-0.003	
3	50	0.5	0.041	4.537	0.009	0.017	0.014
4	30	1	-0.025	5.928	-0.004		
5	40	1	-0.006	5.877	-0.001	0.003	
6	50	1	0.2	5.643	0.036	0.037**	0.040*
7	30	1.5	-0.025	7.496	-0.003		
8	40	1.5	0.025	7.457	0.003	0.007	
9	50	1.5	0.36	7.386	0.049	0.046*	0.052*

Panel B: optimization with an ESG-score constraint θ over an unscreened sample

#P	θ_S	β	θ_P	Mean	St.Dev.	SR	SR_i - SR_{i-1}	SR_i - SR_{i-2}
1	0	0.5	30	0.048	4.853	0.010		
2	0	0.5	40	0.008	4.835	0.002	-0.008	
3	0	0.5	60	-0.074	4.969	-0.015	-0.016	-0.025
4	0	1	30	0.033	5.905	0.006		
5	0	1	40	-0.008	5.885	-0.001	-0.007	
6	0	1	60	-0.090	5.987	-0.015	-0.014	-0.021
7	0	1.5	30	0.017	7.213	0.002		
8	0	1.5	40	-0.024	7.194	-0.003	-0.006	
9	0	1.5	60	-0.105	7.269	-0.014	-0.011	-0.017

Panel C: optimization with an ESG-score constraint θ over a mildly ESG-screened sample

#P	θ_S	β	θ_P	Mean	St.Dev.	SR	SR_i - SR_{i-1}
1	30	0.5	40	-0.096	4.653	-0.021	
2	40	0.5	40	-0.223	4.954	-0.045	
3	30	0.5	50	0.018	4.565	0.004	0.025
4	40	0.5	50	-0.037	4.580	-0.008	0.037**
5	30	1	40	-0.099	5.912	-0.017	
6	40	1	40	-0.180	6.209	-0.029	
7	30	1	50	0.015	5.716	0.003	0.019
8	40	1	50	0.006	5.776	0.001	0.03**
9	30	1.5	40	-0.102	7.485	-0.014	
10	40	1.5	40	-0.137	7.795	-0.018	
11	30	1.5	50	0.012	7.23	0.002	0.015
12	40	1.5	50	0.048	7.346	0.007	0.024**

Notes: #P indicates the progressive number of portfolios, θ_S the screening threshold implying exclusion of stocks with an ESG score lower than the threshold from the portfolio optimization, β indicates the average desired exposure to systemic risk and $\theta_P = \theta w^*$ the optimal portfolio ESG score, calculated as the time-average of the portfolio's weighted ESG scores at each month. Each combination of θ_S, β (and θ_P in Strategy 2 and 3) identifies an investor-specific portfolio. Mean, standard deviation and SR (Sharpe ratio) refer to the time series of out-of-sample portfolio realized returns. In each portfolio the budget constraint is respected, so the sum of portfolio weights is 1. The last two columns (in Panel A and Panel B) report, for each level of β, the approximated Sharpe ratio difference between portfolio with screening/target ESG level *i* and portfolios with the one lower screening/target ESG level *i-1* and two lower screening/target ESG levels *i-2* respectively. The last column (in Panel C) reports, for each level of β, the approximated Sharpe ratio difference between portfolio with target ESG level *i* and portfolio with the one lower target ESG level *i-1*. ***, ** and * represent significance at 1, 5, 10% levels, respectively, in the Sharpe ratio difference.

thresholds θ_S (30, 40, 50), the second strategy with different levels of ESG target score θ_P (30, 40, 60), and the third strategy with different screening thresholds θ_S (30, 40) and different ESG targets θ_P (40, 50).[25] By comparative inspection of Table 5 with Table 9 – Panel A, Table 6 with Table 9 – Panel B, and Table 7 with Table 9 - Panel C, we can conclude that results are invariant, except for Strategy 3, for which we obtain a superior performance when the initial screening is not actually mild ($\theta_S = 40$), and the result is in fact aligned with Strategy 1 with heavy screening.

In terms of performance measures, against its popularity because of its straightforwardness and simplicity, Sharpe ratio suffers of two main shortcomings: it does not satisfy the monotonicity property (Aumann and Serrano, 2008; Cheridito and Kromer, 2013), it is not appropriate in the presence of non-symmetrical distributions. The violation of monotonicity can lead to situations in which an investor does not prefer a portfolio that produces a better result than another portfolio for every state of the world. To cope for this, we analyse Conditional Sharpe ratio which is based on a coherent risk measure quantified by the Conditional Value at Risk (CVaR), i.e. the expected loss that exceeds VaR. [26] As for alternative risk and performance measures that are useful in the presence of non-symmetrical distributions of returns, we compute the Calmar ratio and the Sortino ratio. As the (Conditional) Sharpe ratio, these measures consider the average excess return at the numerator of the ratio, however they differ for the risk measure used at the denominator: the Calmar ratio uses the maximum drawdown, i.e. the highest cumulated percentage loss incurred over the entire investment period; the Sortino ratio uses the square root of the lower partial moment of order two, i.e. an estimate of downside risk. Overall, results using different risk measures (available upon request) show that the performance of optimal stock portfolios resulting from the three strategies are qualitatively invariant.

In terms of robustness w.r.t. the return generating model, we assume expected returns to be generated by a multi-factor model. Specifically, we take a variation of the three-factor Fama-French model, whereby the factors are the three Fama-French factors (market, size, value) plus the ESG characteristic. We estimate betas from the three-factor Fama-French model where the European portfolios (i.e. market, small-minus-big and high-minus-low) are retrieved from Kenneth French's website and are converted into Euro by following Glück et al. (2020).[27] Factor betas are calculated in each month by means of time series regressions based on the previous 500 daily observations (almost two years). Results from beta estimates (available upon request) suggest that our sample well represents the European reference market and it is not over-exposed towards both small firms and companies with high book-to-market ratios. Average correlations between betas and the ESG characteristic are low, implying that the ESG dimension is quite uncorrelated with the financial risk profile. When implementing the three strategies to obtain ESG-compliant optimal portfolios, we consider a benchmark level equal to 1 for both factor loadings on the size and value factors and, as in the main analysis, we consider different levels of systemic risk and sustainability (measured either by the level of screening or by the target ESG score). Results (available upon request) prove that, also with a multi-factor return generating model, there is not a significant negative association between the portfolio average ESG score and performance.

9. Conclusions

The introduction of the ESG dimensions in setting up optimal portfolios has been becoming of uttermost relevance for the financial industry and, accordingly, the literature has been growing fast in recent years: theoretically it provides different approaches to the issue that are characterized by different level of sustainability, empirically it is still inconclusive about the comparative performance of the different approaches.

The aim of this paper is to assess whether there is, if any, a (positive or negative) association between portfolios' average ESG score and financial performance of sustainable stock portfolios and whether such a

[25] We could not use 60 as a screening threshold in Startegy 1 because in some months there are no stocks satisfying such a high the screening.

[26] We take the 95% VaR calculated on historical basis as the reference.

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

relationship depends on the specific strategy used, i.e. the type of sustainability attained. In other words, we mean to answer the following questions: what type of relationship is there between financial performance and sustainability as measured by portfolio average ESG score? Is it possible to do well by doing good and, if so, which is the best strategy to attain this?

To this end we compare the performance of two philosophically different strategies to set up an ESG-compliant optimal stock portfolio and we propose a third one resulting from a mixture of the two. The first type of strategy, which is widely adopted in the industry, results from optimization on an ESG-screened sample, the second one is based on adding a portfolio ESG-score constraint to the optimization problem on an unscreened sample, the third one that we propose in this paper takes pros of both, i.e. it optimizes with an ESG-score constraint (so as to reach a target) over a slightly screened sample (so as to eliminate only the worst companies in terms of ESG). It has to be stressed that our proposal, by using different screening thresholds, allows highlighting the relevance of the relationship between the ESG score and the risk-return ratio of the initial investment set in answering the paper's research questions.

As for the choice of the optimization approach we follow the reformulation by Varmaz et al. (2024) of the a priori tri-criterion optimization that minimizes portfolio residual risk by imposing a specific level of portfolio average systemic risk and ESG. The choice of this approach is made because it has two main methodological advantages: it reduces computational complexity and it is suitable in the presence of unbalanced panels, which are particularly common when ESG scores are used. Moreover, from an industry perspective, it is useful because it provides an optimal portfolio consistent with investor-specific desired levels of both systemic risk and sustainability.

Results are based on a sample starting from the 586 stocks that composed the EURO STOXX Index over the period January 2007 – August 2022, with Bloomberg ESG scores used to measure the ESG dimension of each stock. Two are the main conclusions from our results, which prove to be not far from Pareto optimality and robust with respect to the screening thresholds/targets taken, the risk-adjusted measures used and the return generating model (three-factor Fama-French model).

First, we never find a significant negative association between portfolio average ESG score and performance independently of the strategy used. Second, and by contrast, we find a positive association when the first and the third strategy are implemented with a high screening level. In fact, prior screening (Strategy 1) implies a superior risk-adjusted performance only with heavy screening (ESG score $\geq$ 50), an ESG-score constraint in optimization (Strategy 2) implies a portfolio performance not significantly worse as the target ESG level increases; optimizing with an ESG-score constraint after a mild ESG screening (Strategy 3) does not significantly impair performance as the target ESG level increases and robustness shows that, when the screening is higher (ESG score $\geq$ 40) it obtains a superior performance. To be noted that the relationship between the ESG score and the risk-return ratio in the initial investment set plays a relevant role. If, as in our case, this relationship is essentially convex, with an appropriate level of screening portfolios are composed only by stocks whereby a higher ESG score is associated with a higher risk-return profile.

Furthermore, when testing for the ESG-compliant portfolio over time by estimating rolling Sharpe ratios, we find that the comparative performance of the three strategies does not substantially vary over the financial cycle and we do not have clear evidence of performance cyclicality.

We can thus conclude that, independently of the ESG strategy adopted to set up a sustainable optimal portfolio, investors do not have to sacrifice performance for sustainability and thus they can do well by doing good. Even more, the adoption of an appropriate level of screening allows to increase profitability along with ESG score (Strategy 1 and 3). In particular, the strategy we propose (Strategy 3), which attains a specific ESG target and excludes less sustainable stocks, may be very attractive because it allows sustainable investors obtaining a portfolio in line with their sustainability preferences (as required by the MiFID II directive revision) while not changing or even improving the performance.

CRediT authorship contribution statement

Beatrice Bertelli: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Costanza Torricelli:** Writing – original draft, Supervision, Project administration, Methodology, Formal analysis, Conceptualization.

Appendix. Sustainable portfolios and Pareto optimality

Since the approach taken in Section 3, although granting alignment with preferred targets, does not guarantee Pareto optimality, in this Appendix we propose an analysis to evaluate how close to Pareto optimality are the sustainable optimal portfolios w^* from problems (9) and (13).

In the three-criteria problem (8), a feasible $w \in W_S$ defines expected excess return μ_P in (6), the ESG score $\theta_P = \theta w$ and the risk measure σ_P^2 in (7). Reversing the sign of the risk measure, an attainable criterion vector $g \in \mathbb{R}^3$ (for a vector maximization problem $v : max_{w \in W_S} g$) satisfies $g_1 \leq \mu_P, g_2 \leq \theta_P$ and $g_3 \leq -\sigma_P^2$, for some $w \in W_S$. Then the set G of attainable vectors g is convex, and for $\alpha > 0,\ \epsilon > 0$ and $\lambda = -1$, the optimal solution for (8) is Pareto optimal.

However, if the desired levels β_P^* and θ_P^* in (9) are not optimal for (8), then the optimal solution w^* for (9) is not Pareto optimal. Hence, we perform a test involving μ_P^*, θ_P^* and σ_P^{2*}, determined by the optimal solution $w^* \in W_S$ of problem (9) and by imposing the following equalities for the optimal $g^* \in G$: $g_1^* = \mu_P^*, g_2^* = \theta_P^*$ and $g_3^* = -\sigma_P^{2*}$.[28] Then we consider the reference point optimization problem (e.g. in Wierzbicki, 1979) of finding $w,\ g$ and a scalar ψ to

$$\begin{aligned} \min_{w \in W_S} \quad & \psi - \varepsilon \mathbf{1} g \\ s.t. \quad & \psi \geq s_k \left(g_k^* - g_k\right) \quad k = 1,\ 2,\ 3 \\ & g_1 \leq \mu_P \\ & g_2 \leq \theta_P \\ & g_3 \leq -\sigma_P^2 \end{aligned} \tag{A1}$$

where ε is a small positive number (such as $\varepsilon = 10^{-6}$), and figures $s_1,\ s_2,\ s_3$ must be chosen to scale the order of magnitude of $s_k g_k$ to the same for all k.

[28] To be noted that, for Strategy 1, the test is adjusted to account for the absence of the ESG-score constraint in line with problem (13).

The primary objective in (A1) is to minimize the largest scaled shortfall $g_k^* - g_k$ over the three criteria (at the optimum, ψ may be positive or negative) and if the optimal ψ is small in absolute value, then the optimal solution $w^* \in W_S$ of problem (9) is close to Pareto optimality.

To implement the test in (A1) g_k^* is required, so it is necessary to estimate μ_P^*, θ_P^* and σ_P^{2*}, determined by the optimal solution $w^* \in W_S$ of problem (9). By using equations (6) and (7) it is apparent that we need values for R_m (the excess return of the market portfolio), c (the reward for the ESG characteristic), σ_{Rm}^2 (the variance of excess return of the market portfolio) and σ_ε^2 (residual variance).[29] As for R_m and σ_{Rm}^2, consistently with our dataset, we use the monthly excess returns of the EURO STOXX Index and we estimate the variance of daily returns over a 30-day window of the EURO STOXX Index respectively. As for c, we choose the most plausible values and then we run robustness tests: initially we take it negative, given that in equilibrium sustainable assets have low expected returns because of the higher demand from investors (Pástor et al., 2021; Cornell, 2021), and small in absolute value, specifically -0.005 to scale it with the ESG score magnitude (tendentially greater than 50). Finally, as for σ_ε^2, according to the i.i.d. assumption we expect the risk factors explain a significant part of asset variances, hence we assume it sufficiently small, initially equal to 0.002 and then we run robustness.

Then in order to scale the order of magnitude of $s_k g_k$ across k, we consider all quantities to have an order of magnitude m such as $10^{-1} \leq m < 1$ and corresponding to the smallest order of magnitude over g_k.

Descriptive statistics for ψ are reported in Table A1 and are useful to investigate the distance between the optimal solutions w^* from problems (9) and (13) and Pareto optimal ones. For all strategies it emerges that average ψ has an order of magnitude which is lower or equal to m and the median is zero. This conclusion is confirmed by Fig. A1, which, for reasons of space is proposed only for portfolio #1 of Strategy 3, but it is very similar across portfolios and strategies (results available upon request).

In sum, we can conclude that our results based on optimization problems (9) and (13) are not far from Pareto optimality. Robustness (available upon request) over the necessary parameters, in particular c, σ_ε^2, and the estimation window for σ_{Rm}^2 provide results which are qualitatively the same.

Table A1
Descriptive statistics of the distance to Pareto optimality (ψ).

Panel A: optimization non-considering ESG over an ESG-screened sample

#P	θ_S	β	Min	Max	Median	Mean	St. Dev
1	20	0.5	0.000	1.854	0.000	0.147	0.265
2	40	0.5	0.000	1.854	0.000	0.147	0.265
3	50	0.5	0.000	1.854	0.000	0.145	0.264
4	20	1	0.000	1.996	0.000	0.105	0.235
5	40	1	0.000	1.996	0.000	0.105	0.235
6	50	1	0.000	1.996	0.000	0.105	0.235
7	20	1.5	0.000	4.062	0.000	0.219	0.476
8	40	1.5	0.000	4.062	0.000	0.219	0.476
9	50	1.5	0.000	4.062	0.000	0.218	0.475

Panel B: optimization with an ESG-score constraint θ over an unscreened sample

#P	θ_S	β	θ_P	Min	Max	Median	Mean	St. Dev
1	0	0.5	20	0.000	1.705	0.000	0.142	0.252
2	0	0.5	40	0.000	1.521	0.000	0.140	0.242
3	0	0.5	50	0.000	1.429	0.000	0.139	0.237
4	0	1	20	0.000	1.773	0.000	0.103	0.221
5	0	1	40	0.000	1.585	0.000	0.101	0.207
6	0	1	50	0.000	1.490	0.000	0.100	0.201
7	0	1.5	20	0.000	1.814	0.000	0.194	0.333
8	0	1.5	40	0.000	1.619	0.000	0.191	0.321
9	0	1.5	50	0.000	1.521	0.000	0.189	0.315

Panel C: optimization with an ESG-score constraint θ over a mildly ESG-screened sample

#P	θ_S	β	θ_P	Min	Max	Median	Mean	St. Dev
1	20	0.5	40	0.000	1.390	0.000	0.132	0.215
2	20	0.5	50	0.000	1.390	0.000	0.129	0.210
3	20	1	40	0.000	1.530	0.000	0.097	0.189
4	20	1	50	0.000	1.530	0.000	0.096	0.184
5	20	1.5	40	0.000	1.913	0.000	0.180	0.295
6	20	1.5	50	0.000	1.814	0.000	0.175	0.285

Notes: #P indicates the progressive number of portfolios, θ_S the screening threshold implying exclusion of stocks with an ESG score lower than the threshold from the portfolio optimization, β indicates the average desired exposure to systemic risk and $\theta_P = \theta w^*$ the optimal portfolio ESG score, calculated as the time-average of the portfolio's weighted ESG scores at each month. Each combination of θ_S, β (and θ_P in Strategy 2 and 3) identifies an investor-specific portfolio. Minimum, Maximum, Median, Mean, and Standard Deviation refer to the time series of ψ computed according to problem (A1). In each portfolio the budget constraint is respected, so the sum of portfolio weights is 1.

[29] This adds complexity to the problem and underscores the advantage of the methodology proposed by Varmaz et al. (2024), which does not require estimates over these variables that are difficult to obtain.

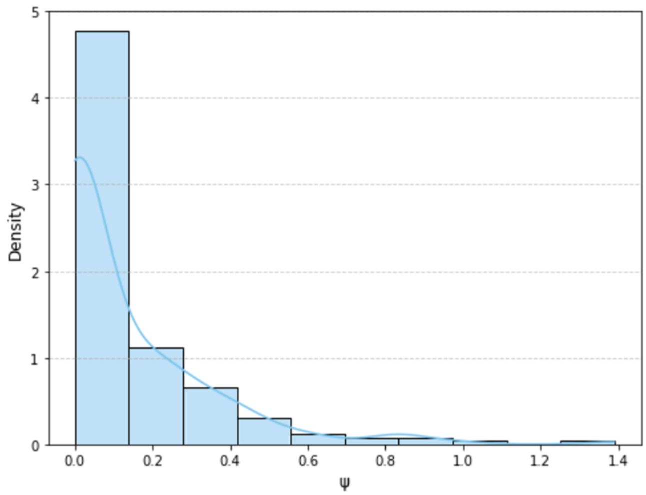


Fig. A1. Density plot of ψ.

References

Alekseev, G., Giglio, S., Maingi, Q., Selgrad, J., & Stroebel, J. (2022). A quantity-based approach to constructing climate risk hedge portfolios. https://pages.stern.nyu.edu/~jstroebe/PDF/AGMSS_QuantityClimateHedge.pdf.

Alessandrini, F., & Jondeau, E. (2020). ESG investing: From sin stocks to smart beta. *The Journal of Portfolio Management, 46*(3), 75–94.

Alessandrini, F., & Jondeau, E. (2021). Optimal strategies for ESG portfolios. *The Journal of Portfolio Management, 47*(6), 114–138.

Auer, B. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? *Journal of Business Ethics, 135*(2), 381–397.

Auer, B. R., & Schuhmacher, F. (2013). Performance hypothesis testing with the Sharpe ratio: The case of hedge funds. *Finance Research Letters, 10*(4), 196–208.

Aumann, R. J., & Serrano, R. (2008). An economic index of riskiness. *Journal of Political Economy, 116*(5), 810–836.

Ballestero, E., Bravo, M., Pérez-Gladish, B., Arenas-Parra, M., & Pla-Santamaria, D. (2012). Socially responsible investment: A multicriteria approach to portfolio selection combining ethical and financial objectives. *European Journal of Operational Research, 216*(2), 487–494.

Bénabou, R., & Tirole, J. (2010). Individual and corporate social responsibility. *Economica, 77*(305), 1–19.

Bender, J., Sun, X., & Wang, T. (2017). Thematic indexing, meet smart beta! Merging ESG into factor portfolios. *The Journal of Beta Investment Strategies, 8*(3), 89–101.

Bertelli, B., & Torricelli, C. (2024). The trade-off between ESG screening and portfolio diversification in the short and in the long run. *Journal of Economics and Finance*, 1–25. https://doi.org/10.1007/s12197-023-09652-9

Blitz, D., & Fabozzi, F. J. (2017). Sin stocks revisited: Resolving the sin stock anomaly. *The Journal of Portfolio Management, 44*(1), 105–111.

Bohn, J., Goldberg, L. R., & Ulucam, S. (2022). Sustainable investing from a practitioner's viewpoint: What's in your ESG portfolio? *Journal of Investment Management, 20*(2), 14–29.

Boubaker, S., Le, T. D., Manita, R., & Ngo, T. (2023). The trade-off frontier for ESG and Sharpe ratio: A bootstrapped double-frontier data envelopment analysis. *Annals of Operations Research*, 1–25.

Cabello, J. M., Ruiz, F., Pérez-Gladish, B., & Méndez-Rodríguez, P. (2014). Synthetic indicators of mutual funds' environmental responsibility: An application of the Reference Point Method. *European Journal of Operational Research, 236*(1), 313–325.

Cesarone, F., Martino, M. L., & Carleo, A. (2022). Does ESG impact really enhance portfolio profitability? *Sustainability, 14*(4), 2050.

Chen, Z., & Xie, G. (2022). ESG disclosure and financial performance: Moderating role of ESG investors. *International Review of Financial Analysis, 83*, Article 102291.

Cheridito, P., & Kromer, E. (2013). Reward-risk ratios. *Journal of Investment Strategies, 3*(1), 1–16.

Cochrane, J. H. (2005). *Asset pricing, revised edition*. Princeton, NJ: Princeton University Press.

Cornell, B. (2021). ESG preferences, risk and return. *European Financial Management, 27*(1), 12–19.

Daniel, K., Mota, L., Rottke, S., & Santos, T. (2020). The cross-section of risk and returns. *The Review of Financial Studies, 33*(5), 1927–1979.

De Spiegeleer, J., Höcht, S., Jakubowski, D., Reyners, S., & Schoutens, W. (2021). ESG: A new dimension in portfolio allocation. *Journal of Sustainable Finance & Investment*, 1–41.

Drut, B. (2010). Sovereign bonds and socially responsible investment. *Journal of Business Ethics, 92*(1), 131–145.

Edmans, A. (2011). Does the stock market fully value intangibles? Employee satisfaction and equity prices. *Journal of Financial Economics, 101*(3), 621–640.

Engle, R. F., Giglio, S., Kelly, B., Lee, H., & Stroebel, J. (2020). Hedging climate change news. *The Review of Financial Studies, 33*(3), 1184–1216.

European Parliament. (2023). Amendments to MiFID II and MiFIR The EU's markets in financial instruments. BRIEFING EU Legislation in Progress. EPRS_BRI(2022) 733546_EN.pdf (europa.eu).

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.

Figueira, J. R., Liefooghe, A., Talbi, E. G., & Wierzbicki, A. P. (2010). A parallel multiple reference point approach for multi-objective optimization. *European Journal of Operational Research, 205*(2), 390–400.

Fitzgibbons, S., Palazzolo, C., & Pomorski, L. (2018). Hit'em Where It Hurts: ESG Investing 2.0. Investments and Pensions Europe (November 1), AQR. https://www.aqr.com/Insights/Research/Trade-Publication/Hit-Em-Where-It-Hurts-ESG-Investing-2.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment, 5*(4), 210–233.

Friedman, M. (1970). The social responsibility of business is to increase its profits, New York Times, Section SM, September 13, 17.

Gardiner, L. R., & Steuer, R. E. (1994). Unified interactive multiple objective programming. *European Journal of Operational Research, 74*(3), 391–406.

Gasser, S. M., Rammerstorfer, M., & Weinmayer, K. (2017). Markowitz revisited: Social portfolio engineering. *European Journal of Operational Research, 258*(3), 1181–1190.

Girard, E. C., Rahman, H., & Stone, B. A. (2007). Socially responsible investments: Goody-two-shoes or bad to the bone? *The Journal of Investing, 16*(1), 96–110.

Glück, M., Hübel, B., & Scholz, H. (2020). Currency Conversion of Fama–French factors: How and why. *The Journal of Portfolio Management, 47*(2), 157–175.

GSIA. (2021). 2020 Global Sustainable Investment Review. http://www.gsi-alliance.org/wp-content/uploads/2021/08/GSIR-20201.pdf.

Haimes, Y. Y., Lasdon, L. S., & Wismer, D. A. (1971). On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man and Cybernetics, 1*(3), 296–297.

Hallerbach, W., Ning, H., Soppe, A., & Spronk, J. (2004). A framework for managing a portfolio of socially responsible investments. *European Journal of Operational Research, 153*(2), 517–529.

Hirschberger, M., Steuer, R. E., Utz, S., Wimmer, M., & Qi, Y. (2013). Computing the nondominated surface in tri-criterion portfolio selection. *Operations Research, 61*(1), 169–183.

Hoepner, A.G.F., Oikonomou, I., Sautner, Z., Starks, L.T., & Zhou, X. (2021). ESG shareholder engagement and downside risk. Technical report, SSRN.

Hong, H., & Kacperczyk, M. (2009). The price of sin: The effects of social norms on markets. *Journal of Financial Economics, 93*(1), 15–36.

Jacobs, H., Müller, S., & Weber, M. (2014). How should individual investors diversify? An empirical evaluation of alternative asset allocation policies. *Journal of Financial Markets, 19*(C), 62–85.

Kuzey, C., Uyar, A., Nizaeva, M., & Karaman, A. S. (2021). CSR performance and firm performance in the tourism, healthcare, and financial sectors: Do metrics and CSR committees matter? *Journal of Cleaner Production, 319*, Article 128802.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*(5), 850–859.

Liagkouras, K., Metaxiotis, K., & Tsihrintzis, G. (2020). Incorporating environmental and social considerations into the portfolio optimization process. *Annals of Operations Research*, 1–26.

Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in stock portfolios and capital budget. *Review of Economics and Statistics, 47*(1), 13–37.
López de Prado, M. M. (2020). *Machine learning for asset managers*. Cambridge University Press.
Markowitz, H. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.
Melas, D. (2021). The future of factor investing. *The Journal of Portfolio Management, 48*(2), 15–25.
Miettinen, K., Eskelinen, P., Ruiz, F., & Luque, M. (2010). NAUTILUS method: An interactive technique in multiobjective optimization based on the nadir point. *European Journal of Operational Research, 206*(2), 426–434.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*(4), 768–783.
Ortas, E., Moneva, J. M., & Salvador, M. (2014). Do social and environmental screens influence ethical portfolio performance? Evidence from Europe. *BRQ Business Research Quarterly, 17*(1), 11–21.
Pástor, Ľ., & Vorsatz, M. B. (2020). Mutual fund performance and flows during the COVID-19 crisis. *The Review of Asset Pricing Studies, 10*(4), 791–833.
Pástor, L., Stambaugh, R. F., & Taylor, L. A. (2021). Sustainable investing in equilibrium. *Journal of Financial Economics, 142*(2), 550–571.
Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics, 142*(2), 572–597.
Principles for Responsible Investment. (2017). The SDG investment case.
Rossi, M., Sansone, D., Van Soest, A., & Torricelli, C. (2019). Household preferences for socially responsible investments. *Journal of Banking & Finance, 105*(C), 107–120.
Schuhmacher, F., & Eling, M. (2011). Sufficient conditions for expected utility to imply drawdown-based performance rankings. *Journal of Banking & Finance, 35*(9), 2311–2318.
Schuhmacher, F., & Eling, M. (2012). A decision-theoretic foundation for reward-to-risk performance measures. *Journal of Banking & Finance, 36*(7), 2077–2082.
Shanken, J. (1992). On the estimation of beta-pricing models. *The Review of Financial Studies, 5*(1), 1–33.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance, 19*(3), 425–442.
Steuer, R. E., & Utz, S. (2023). Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing. *European Journal of Operational Research, 306*(2), 742–753.
Utz, S., Wimmer, M., Hirschberger, M., & Steuer, R. E. (2014). Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *European Journal of Operational Research, 234*(2), 491–498.
Utz, S., Wimmer, M., & Steuer, R. E. (2015). Tri-criterion modeling for constructing more-sustainable mutual funds. *European Journal of Operational Research, 246*(1), 331–338.
Van Duuren, E., Plantinga, A., & Scholtens, B. (2016). ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics, 138*(3), 525–533.
Varmaz, A., Fieberg, C., & Poddig, T. (2022). Portfolio optimization for sustainable investments. Available at SSRN 3859616.
Varmaz, A., Fieberg, C., & Poddig, T. (2024). Portfolio optimization for sustainable investments. *Annals of Operations Research, 341*(2), 1151–1176.
Wang, Z., Liao, K., & Zhang, Y. (2022). Does ESG screening enhance or destroy stock portfolio value? Evidence from China. *Emerging Markets Finance and Trade, 58*(10), 2927–2941.
Whelan, T., Atz, U., Van Holt, T., & Clark, C. (2021). *ESG and financial performance: Uncovering the relationship by aggregating evidence from 1000 plus studies published between 2015-2020* (pp. 520–536). New York: NYU STERN Center for sustainable business.
Wierzbicki, A.P. (1979). The use of reference objectives in multiobjective optimization-theoretical implications and practical experience. IIASA Working Paper WP-79-066.