DOING WELL, WHILE DOING GOOD

ESG-SCREENING AND THE MITIGATION OF UNFAVORABLE EXPOSURES FROM THE PERSPECTIVE OF AN INSTITUTIONAL INVESTOR

Master Thesis

October 2022

Author: **A.H.M. Evers**

Company Supervisor: **M. Maradona**
Investment Strategy
Fiduciary Advice
MN Services

Academic Supervisors: **Prof. Dr. L. Spierdijk**
Financial Engineering & Management
Faculty of Behavioural, Management and Social Sciences (BMS)
University of Twente

Dr. B. Roorda
Financial Engineering & Management
Faculty of Behavioural, Management and Social Sciences (BMS)
University of Twente

"Be who you are and say what you feel, because those who mind don't matter, and those who matter don't mind."

Bernard M. Baruch

Acknowledgements

Working on this thesis for the past year was the final part before completing my master's degree in Financial Engineering and Management. With that, this work represents the capstone of all my student endeavors.

In the past year, I had the opportunity to perform my research within the Investment Strategy team of MN Services, as part of the Academic Excellence Program together with PGGM. One of the greatest insights that this experience has given me, is that recognizing and admitting your own ignorance is the first step towards knowledge. For my time here, I want to express credits and appreciation:

- to Marissa Maradona, my supervisor at MN, for her brilliant expertise, significant contribution, and endless patience.
- to the other members of the Investment Strategy team – you know who you are – for being there and helping me whenever needed.
- to the AXP team, in particular Liakos, the BA team and my fellow interns, for all the learning, help and wonderful times.

I am also very grateful for my academic supervisors. I thank my first supervisor Laura Spierdijk for all her expertise and adjustments but above all for her valuable support, both on and off paper. And I thank Berend Roorda: for his thorough feedback on my draft, his approachability during the entire master's program, and the exciting education.

For the memories past, present and future and the unforgettable years at university, I owe my friends.

In closing, I feel a deep sense of gratitude for the opportunities that my parents have given me.
For their unconditional support, advice and love, I cannot thank them and my siblings enough.

Teun Evers
The Hague / Hertme, October 2022

Executive Summary

Investors are increasingly looking to incorporate ESG preferences into their investment strategies. When (institutional) investors hold large and passively mandated portfolios, a popular approach to sustainable investing is the exclusion of securities via a screening policy. A well-known policy is screening via so-called ESG (Environmental, Social, Governance) ratings, yet the implications of screening on portfolio performance and efficiency are not always clear.

Problem context

Institutional investors hold large and passive portfolios to gain advantage from the diversification benefits and equity premium that a broad market index has to offer. From a theoretical perspective, exclusion from a portfolio restricts the investment opportunity set and therefore decreases diversification opportunities, resulting in increased specific risk for which an investor is not rewarded. Exclusion based on ESG scores should therefore -in theory- lead to worsened risk/return characteristics. Moreover, so-called 'sin-stocks' (securities with low ESG-ratings) have exhibited superior returns in the past (see 2.1.2), so excluding these from a portfolio would result in missed opportunities.

Empirical research however shows diverging results as to what extent exclusion decreases performance. Results vary along three main items: the performance measures chosen to assess portfolio performance, the timeframe of the data used, and the original index, i.e. the index from which exclusion is performed (a regional index, large cap index, world index etc.).

Research goal

The goal of this research is two-fold. First, we want to clarify the ambiguous implications of ESG-screening on portfolio performance, using the perspective of an institutional investor. Second, we explore how rebalancing the weights of the screened portfolio through optimization can help to mitigate unfavorable exposures that result from screening (optimized exclusion).

Achieving the first goal requires assessing the performance of screened portfolios through theoretically and practically relevant performance measures (see 3.1). The effects on each performance measure are graphed against the different levels of exclusion (5% exclusion, 10% exclusion, ... , 95% exclusion), which offers a pragmatic insight for institutional investors who do not immediately seek to optimize their screening policy, but rather want to objectively observe the effects of exclusion on each level. Furthermore, the timeframe used (2013-2021) provides recent market insights while accounting for an ESG-score coverage above 95%. Finally, the original index from which exclusion is performed, is the broadest index available, recognized by institutional investor to represent the market index: the MSCI ACWI.

Optimization

The results of exclusion, based on an exclusion policy using MSCI EVA *industry adjusted* ESG ratings (industry adjusted ratings have the best coverage and prevent us from screening out entire industries) over the period 2013-2021, are positive (section 3.5). We show that the aggregate ESG score of a portfolio can be improved without deterioration of risk-adjusted performance (risk in terms of overall volatility). In fact, we observe *improved* risk-adjusted performance through the annualized Sharpe ratio increasing from 0.29 to 0.35, and to 0.5 for respectively the MSCI ACWI, 50% screened portfolio and 90% screened portfolio. The aggregated ESG-scores for these exclusion levels are 5.28 (ACWI), 7.01 and 9.06 (out of 10). However, exclusion leads to undesirable ex-post tracking error (1.22% and 3.44% for 50% and 90% exclusion, where 2% is the maximum desirable for enhanced indices), which is an important measure for institutional investors. See Table 1 on the next page for a performance summary.

Furthermore, undesirable sectoral and regional exposures are observed. This increases a portfolios risk towards sector-and regional-specific shocks. We therefore aim to achieve the second goal through an optimization approach that minimizes the ex-ante tracking error of a portfolio, while controlling for these exposures. Essentially, we look for a way to replicate the benchmark index performance (as institutional investors), while maintaining a high aggregate ESG-score through screening (see 4.1: Rationale). We apply the minimization of the tracking error through quadratic optimization, whereby we add some linear constraints in order to maintain sectoral and regional exposures, as well as portfolio turnover that is an outcome of monthly exclusion and rebalancing (see 4.3: Optimization Design). Note that minimizing the ex-ante tracking error requires the covariance matrix of asset returns which, depending on the dimension of the problem, has to be estimated. The suboptimal estimation of our covariance matrix could lead to skewed optimization results. Therefore, caution should be considered when interpreting these results.

Results

We ultimately show that screening and optimization go relatively hand-in hand up until the 80% mark of exclusion. Up until here, there is a tradeoff between absolute performance on the one hand (where screened portfolios score better) and relative performance on the other (in favor of the optimized portfolios). Relative performance is enforced by minimization of the tracking error and limitations of exposures (see 5.1: Overall Performance). Moreover, optimization reduces the turnover compared to the screened portfolios. We believe that a large part of the absolute performance decrease (compared to the plain screened portfolios) is caused by restricting exposures of higher-return sectors and regions that would otherwise (unrestricted) be represented more. If an investor prefers absolute performance regardless of tracking error or regional and sector exposures that are different from the benchmark, then optimization beyond the 80% mark is more devastating to your performance than before the 80% mark. When an investor wants to keep performance close to the benchmark performance, which was our intention in the first place, optimization does help from 80% onwards, and maybe even from 65% onwards where the screened and optimized tracking error results invert.

Overall, this research shows that it is possible to construct a quality ESG portfolio with a significantly smaller subset of constituents than the original benchmark, while maintaining exposure close to the benchmark and not underperforming in terms of risk and return (see table 1 below). At a 90% exclusion level, the optimization manages to construct a portfolio with only 14% of the original market capitalization (248 from 2642 constituents) while increasing the aggregate ESG score from 5.28 to 9.13 out of 10.

	MSCI ACWI	**Excl 50%**	**Opt 50%**	**Excl 90%**	**Opt 90%**
Average Constituents #	2642	1238	1238	248	619
Average Marketcap %	100	58.5	58.5	14.04	14.04
Annual. Return %	11.11	11.93	11.47	14.29	11.51
Annual. Volatility %	13.13	13.04	12.89	13.21	12.67
Annual. Sharpe Ratio	0.29	0.35	0.32	0.5	0.32
Annual. Tracking Error (Ex-Post) %	-	1.22	1.41	3.44	2.69
Annual. Information Ratio	-	0.17	0.06	0.24	0.03
Annual. Turnover %	10.87	59.04	49.73	120.07	109.78
Max Drawdown %	-21.25	-19.82	-20.7	-16.83	-20.29
Average Portfolio IAA Score	5.28	7.01	7.17	9.06	9.13

Table 1: Portfolio performance for the benchmark (MSCI ACWI), screened portfolios (Excl) and their subsequent optimized (Opt) portfolios for the 50% and 90% exclusion levels. For the full tables including regional and sector exposures, see Appendix D.

Glossary

- **Constituent:** a security or stock that is part of an index.

- **Developed Market:** a country that is developed in terms of its economy and capital markets, and has an elevated level of regulation and oversight, a market exchange, and good liquidity in debt and equity markets.

- **Efficient frontier:** a line in the space of the investment opportunity set representing the set of optimal portfolios offering the highest expected return given a defined level of risk or vice versa.

- **Emerging Market:** a country that is becoming more engaged with global markets. It has some, but not all of the characteristics of a developed market.

- **ESG**: Environmental, Social and Governance.

- **IAA Score**: Industry Adjusted Average, or Industry Adjusted Score, is an ESG rating attributed to a company through the weighted average of all 3 ESG pillar scores, normalized for the range of scores set by industry peers.

- **Index Tracking Problem (ITP)**: the problem of reproducing the performance of an index by using a portfolio of assets that is a subset of the index, with the goal of minimizing the tracking error.

- **Information ratio (IR)**: measurement of a portfolios return beyond the returns of its benchmark, compared to the volatility of those returns. Calculated as the difference in portfolio and benchmark returns, divided by the tracking error.

- **Institutional Investors:** legal entities that pool large funds of various investors to purchase securities. Examples are pension funds, commercial banks, insurers, hedge funds and sovereign wealth funds.

- **Large-cap:** refers to a company with a market capitalization value of more than $10 billion.

- **Market capitalization:** the total value of all the shares of a company's stock, calculated by multiplying the stock price with the number of outstanding shares.

- **Maximum drawdown (MDD):** a measure of an asset's largest price drop from a peak to a through. An indicator for downside risk over a specified time period.

- **Mid-cap:** refers to a company with a market capitalization value between $2 billion and $10 billion.

- **Modern Portfolio Theory (MPT)**: an investment theory developed by Harry Markowitz in 1952. The theory allows investors to construct an asset portfolio that maximizes expected return for a given level of risk.

- **MSCI**: Morgan Stanley Capital International, a global investment data and index provider.

- **MSCI ACWI**: MSCI's All Country World Index, is MSCI's flagship global equity index, capturing the full opportunity set of large- and mid-cap stocks across 23 developed and 24 emerging markets. The index covers approximately 85% of the free float-adjusted market capitalization in each market.

- **Negative/exclusionary screening:** the exclusion of a stock, sector, country or other issuer from a fund or portfolio based on a strategic policy.

- **Risk factors, factor characteristics:** factors are statistical determinants of expected stock returns. Identified factors include growth versus value, size, credit rating, volatility, momentum, investment etc.

- **Sharpe ratio (SR)**: a measure to compare the return of a stock or portfolio with its risk, calculated as the return in excess of the risk-free rate, divided by the assets volatility.

- **Sin-stocks / sin-industries:** companies or industries that are engaged in controversial products or services (weapons, alcohol, gambling, tobacco etc.).

- **Specific, diversifiable risk:** risk that applies only to a particular company, industry, sector, or geographical region.

- **Sustainable Investing, SI, or SRI**: Sustainable and Responsible Investments, Socially Responsible Investment, or Sustainable and Responsible Investing, is an investment approach that considers environmental, social and governance (ESG) factor in portfolio selection and management.

- **Systematic risk:** risk inherent to the entire market.

- **Tracking error (TE):** a measure of divergence of the price behavior of a portfolio and the price behavior of a benchmark, calculated as the standard deviation of the difference between the portfolio and the benchmark return.

- **Turnover (two-way)**: a measure of how many securities in a fund or portfolio are bought or sold over a given period of time.

- **Volatility:** risk of a stock or portfolio, calculated as the standard deviation of returns over a given period.

Table of Contents

1 Introduction

The topic of sustainable investing has become increasingly popular to investors who wish to do more than just multiplying their capital. Making a positive impact on the world through investing in companies aligned with their own and society's values, as well as mitigating climate risk are some drivers for investors to engage in sustainable investing.

1.1 Sustainable Investing

1.1.1 Definition

We start by defining sustainable investing, also named SRI (Sustainable and Responsible Investments, or Socially Responsible Investment) according to the Global Sustainable Investment Alliance (Global Sustainable Investment Alliance, 2021, p. 7): *"Sustainable investment is an investment approach that considers environmental, social and governance (ESG) factors in portfolio selection and management."* The GSIA predicts that by 2025, ESG-mandated assets are to make up $96 trillion of the total $165 trillion of global assets under professional management, up from a $19 trillion of a total $64 trillion in 2014, thus representing an increase from 30% to 58%.

There exist multiple definitions of sustainable investing, or SRI, but one common similarity between the definitions is the mentioning of integrating ESG factors in analysis, research, and selection. ESG criteria can be used in terms of company-level scores and ratings (or portfolio level using an aggregate of company-level scores) by means of separate scores for each pillar (E, S and G), or a weighted average thereof. These scores and ratings are published by several rating agencies. Each pillar's score is determined by various sub-metrics. In the case of ESG scores from MSCI, that are used within this research, the E, S and G scores are generally set according to respectively 13, 16 and 6 subcategories. The use of these scores or ratings is a result of the acknowledgement of investors that issues on either 3 pillars can affect the performance of companies or portfolios across regions, sectors, and asset classes as well as through time.

1.1.2 Approaches to Sustainable Investing

There exist several approaches for investors to commit to sustainable investing. It is common that a single strategy or product adopts a combination of these approaches. Both the GSIA and Eurosif (Eurosif, 2018) discern the following approaches:

ESG Integration	The systematic and explicit inclusion by investment managers of ESG factors into financial analysis.
Corporate Engagement & shareholder action	Employing shareholder power to influence corporate behavior, including through direct corporate engagement, filing or co-filing shareholder proposals, and proxy voting that is guided by comprehensive ESG guidelines.
Norms-based screening	Screening of investments against minimum standards of business or issuer practice based on international norms.
Negative/exclusionary screening	The exclusion from a fund or portfolio of certain sectors, individual companies, countries, or other issuers based on activities considered not investable. Exclusion criteria can refer, for example, to product categories (e.g., weapons, tobacco, gambling) company practices (e.g., animal testing, violation of human rights, corruption) or controversies.
Best-in-class/positive screening	Investing in sectors, companies, or projects selected for positive ESG performance relative to industry peers, and that achieve a rating above a defined threshold.
Sustainability themed/thematic investing	Investing in themes or assets specifically contributing to sustainable solutions – environmental and social – (e.g., sustainable agriculture, green buildings, lower carbon tilted portfolio, gender equity, diversity).
Impact investing and community investing	Impact investing concerns investing to achieve positive, social, and environmental impacts – requires measuring and reporting against these impacts, demonstrating the intentionality of investor and underlying asset/investee, and demonstrating the investor contribution. Community investing is where capital is specifically directed to traditionally underserved individuals or companies. Some community investing is impact investing, but community investing is broader and considers other forms of investing and targeted lending activities.

Table 2: Description of the various approaches to Sustainable Investing. Source: Global Sustainable Investment Alliance (2021).

Of these strategies, ESG integration is the most popular (worldwide) when measured in assets dedicated, followed by negative/exclusionary screening and corporate engagement and shareholder action. Strategy preferences however differ per region, with negative/exclusionary screening being the most used approach in Europe, followed by corporate engagement. In the US, ESG integration as a strategy is larger than all other strategies combined (see Appendix A: Sustainable Investing Assets by Strategy & Region 2020). When looking at the definitions, another view is that ESG integration can even be regarded as an umbrella term for some subsequent approaches, as these all imply some form of inclusion of ESG factors into the buy/sell and inclusion/exclusion decision.

The least used strategies are impact/community investing and positive/best-in-class screening. One could argue however that the latter is an inverse methodology of negative/exclusionary screening when taking ESG performance on a company or sector level as a threshold below which companies or sectors are excluded from a portfolio. This research indeed focuses mainly on these two similar approaches. The other approaches are not discussed in the remainder of this report, but merely serve to the readers comprehension of the different approaches to sustainable investing.

Considerations in Screening Approaches.

The pros and cons of screening, specifically exclusion (negative screening), have been widely documented in academic research. For example, Blitz and Swinkels (2020) question the effectiveness of exclusion by arguing that it merely leads to a transfer of ownership of a stock from a concerned to a less-concerned investor, and that investors could achieve more by engagement as an active shareholder (see corporate engagement & shareholder action). Broccardo, Hart and Zingales (2021) substantiate this view by stating that exclusion is less effective than 'voice' in a competitive world, while acknowledging the limitations of social responsible investors that cannot attain a majority of vote without reducing the diversification of their portfolio. Investors that do hold a significant amount of the outstanding shares of a company, often institutional investors, can significantly contribute to the company's sustainability policy. Institutional investors are legal entities that pool large funds of various investors to purchase securities. Pension funds, commercial banks, insurers, hedge funds and sovereign wealth funds are examples of institutional investors.

In some cases, engagement can be less effective, e.g. when the core business of a company happens to be inherent to unsustainable activities (one cannot persuade an oil refiner to fully convert to solar panel production). In this respect, large investors often resort to screening, in particular by excluding so-called 'sin-industries' or 'sin-stocks', which are industries engaged in controversial products or services (weapons, tobacco, gambling, or adult entertainment) or malpractices (see Table 2). A second reason why institutional investors resort more to top-down approaches like screening (which includes negative/exclusionary, best-in-class/positive and norms-based screening) is that it can be systematically applied on a portfolio containing hundreds or even thousands of securities. These larger portfolios are often traded as part of a passive benchmark mandate. As such, any form of screening as an SI approach can be more effective than engaging on a company-level.

1.2 Organizational Background

This research is conducted within the Investment Strategy team of MN Services ('MN'). MN is the third largest pension investor in the Netherlands with a total of €178 billion assets under management (2021), of which €100 billion is managed on behalf of the PMT pension fund (Pensioenfonds Metaal & Techniek). MN is considered an institutional investor, which is the perspective we use throughout this research.

The listed equity portfolio of PMT constitutes approximately 30% of the total PMT assets under management and includes 3 mandates of developed market regions (US, Europe, Pacific) and a listed emerging market mandate. The four mandates of listed equities are complemented by a private equity mandate and together form the Equity Cluster of PMT. The objective of the Equity Cluster of PMT is to achieve 3% return in excess of the pension liabilities and is managed to follow the total return of the MSCI All Country World Index (MSCI ACWI). This is the flagship index constructed by MSCI and is regarded a proxy for the equity risk premium in the market on the long term. The MSCI ACWI is capitalization-weighted, meaning that constituents with a higher market capitalization will receive a higher weighting in the index.

4 years ago, PMT introduced the 'Bewuste Selectie' strategy in the 3 developed market mandates as part of their sustainable investing strategy. Currently, the strategy also includes the emerging market mandate and therefore the listed equity part of the Equity Cluster.

Conforming to the Bewuste Selectie strategy, PMT imposes a screening approach on the index universe, which is defined as the MSCI ACWI, according to several layers of exclusion rules, or 'legs.' The first leg concerns exclusion based on international treaties. The second leg comprises exclusions based on fundamental ethical responsibility, resulting in the removal of securities directly involved in nuclear weapons, controversial weapons, civil-use weapons, fur, tobacco, gambling, and adult entertainment. These securities were earlier referred to as sin-stocks.

The third leg of the Bewuste Selectie strategy involves the screening based on ESG factors. Specifically, companies with the 3 lowest ESG ratings (CCC, B and BB) according to MSCI IVA ESG ratings are excluded from the total portfolio. Note that we earlier mentioned that negative screening more often involves sin-stocks or exclusions that are regulations based, while screening using ESG ratings is better associated with best-in-class/positive screening (as exhibited in Table 2, securities are included in a portfolio for their particular good ESG rating). ESG ratings can however also be used as an exclusion policy.

For the sake of clarity and consistency, we will further on only use *exclusion* (negative screening) based on ESG ratings as our screening policy.

After exclusions, the remaining securities are reweighted based on their original market capitalization proportions. The third leg (ESG exclusion) currently accounts for the exclusion of approximately 26% of securities from the benchmark portfolio, i.e. the original MSCI ACWI index, equal to 15% of the initial market capitalization. The entire Bewuste Selectie strategy (including legs 1, 2 and 3) consequently leads to the exclusion of 54% of the securities amount, equal to 30% of market capitalization (see Table 3 below).

		MSCI ACWI (benchmark portfolio)	*Bewuste Selectie*	*Exclusion %*
Total Filter	#Securities	2,985	1,385	54%
	%Market cap	100%	70%	30%
Leg 3 (ESG)	#Securities	2,985	2,196	26%
	%Market cap	100%	85%	15%

Table 3: Amount of securities and market capitalization excluded for the total Bewuste Selectie filter, and leg 3 separate.

1.3 Problem Context

1.3.1 The Bottom Line

In theory, the systematic exclusion of securities from the chosen index universe is detrimental for two reasons. First, the exclusion policy restricts the investment opportunity set, thus reducing diversification efficiencies that could possibly impact the risk and return characteristics of a portfolio. Second, exclusions from the original benchmark lead to deviations in returns relative to this benchmark (Blitz & Swinkels, 2021). This translates to *tracking error*, which is a measure of divergence of the price behavior of a portfolio and the price behavior of a benchmark and is considered an important measurement for institutional investors that wish to capture the benefits of following an index.

Both reasons are derived from financial theory, on which we will elaborate more in the next sections. *Empirical research*, however, is less conclusive about the effects of exclusion, or ESG-focused portfolios. On the one hand for example, individual high-quality ESG firms tend to perform financially well, so focusing a portfolio towards quality ESG-stocks, or therefore excluding bad-quality ESG stocks, should be beneficial and is shown to be so. Indeed, exclusion leads to the increase of diversifiable risk or at least less efficient diversification compared to an unbiased portfolio, but in some instances, it can be offset by the lower volatility that quality ESG-stocks exhibit. On the other hand, securities that are excluded from portfolios exhibit higher expected returns from investors that demand a compensation for their ‘sin exposure.’ So sin-stocks *also* perform well on an individual basis. Screening out sin-stocks could then lead to missed opportunities.

We argue that whether exclusion or ESG investing is beneficial or not comes down to a matter of perspective. Empirical conclusions in literature vary along many variables such as the original benchmark index used (from which the screening takes place), but also the data timeframe, the ESG ratings provider, and most importantly: the performance measures along which the portfolio performance is assessed. This variation and subsequent divergence in empirical results makes it difficult for MN and other institutional investors to make a solid assumption about what effect the step-by-step exclusion has on their portfolios.

We therefore aim to clarify the ambiguous implications of ESG-screening on portfolio performance and explore the possibilities of maintaining the benchmark performance post-exclusion. We will discuss this more in detail in section 1.4.

The next paragraphs of this section will provide more background on the bottom-line statement of the problem context. This background is needed to attain a proper understanding of the research context and the formulation of the research goal and research questions in the next section. Paragraph 1.3.2 discusses Modern Portfolio Theory, which is the theory behind the argument that restricting the investment universe leads to an increase in risk that can otherwise be diversified by combining assets that are not or weakly correlated. It also implies that if a portfolio is optimal, excluding securities from it can only lead to worse risk/return characteristics. Paragraph 1.3.3 then discusses active and passive investing, and the Index Tracking Problem. This is relevant information, because the Index Tracking Problem describes the issue of replicating an index with only a subset of the original amount of securities and is therefore parallel with our own exclusion problem. Finally, we provide some background information on literature describing the minimum amount of securities needed to maintain proper diversification and performance.

1.3.2 Modern Portfolio Theory

Modern Portfolio Theory (MPT) was developed by Harry Markowitz in 1952. MPT is an important financial concept that describes ways of diversifying and allocating assets within a portfolio to maximize the portfolio’s expected return, given a specified level of risk (volatility) (Markowitz, 1952). Specifically, the primary objective of MPT is to maximize expected return while reducing diversifiable, or *specific* risk. This is risk that can be reduced by diversifying assets. On the opposite is *systematic* risk, that cannot be diversified away as it affects the entire economy and most investments. Efficient capital markets reward investors that bear systematic, or market risk; but because diversification can be obtained, those investors are not rewarded for bearing specific, diversifiable risk (Barnett & Salomon, 2006). Under MPT, an investor can reduce the diversifiable risk of a portfolio by combining securities that are weakly correlated or not correlated at all. This is referred to as diversification. The relation holds that for each level of risk, there is a portfolio of assets that maximizes the expected return. These combinations can be plotted on a graph with the portfolio risk on the x-axis and the portfolio expected return on the y-axis:

Figure 1: A schematic illustration of the Efficient Frontier and the investment opportunity set.

Here, the green area represents the investment opportunity set that we mentioned earlier, which represents all feasible portfolios. Within this set of feasible portfolios there exists a subset of portfolios that have the highest expected return for each level of risk. On the most left point of the efficient frontier exists a portfolio which bears the minimum risk. The portfolio risk is measured by the variance of its expected returns; therefore, it is called the minimum variance portfolio. Portfolios that are still on the 'border' of the investment opportunity set but below the minimum variance point (the part of the line that is not green thick) are unwise to invest in, since they provide less expected return, but an equal amount of risk compared to a portfolio that is on the efficient frontier.

Screening & Diversification

Under Modern Portfolio Theory (and mathematically in general), reducing the scope of the initial investment set through any kind of screening whilst achieving a more diversified portfolio compared to the initial investment set, is impossible. Because screening leads to the exclusion of certain firms, industries, and sectors, screened portfolios are more inclined to carry a considerable amount of specific risk (Barnett & Salomon, 2006; Rudd, 1981). This additionally shifts the mean-variance frontier towards less beneficial risk-return characteristics (Renneboog, Ter Horst, & Zhang, 2008). In studying the restriction of the investment universe, Pedersen, Fitzgibbons and Pomorski (2021) show that the ESG-SR frontier for investors who apply screening (thus the frontier from a subset of constituents exhibiting a certain ESG score) is strictly dominated by the unconstrainted efficient frontier, as illustrated in the figure below.

They show that *"minimizing the variance among all portfolios must provide a result that is at least as small as minimizing over the subset with a given ESG score".*

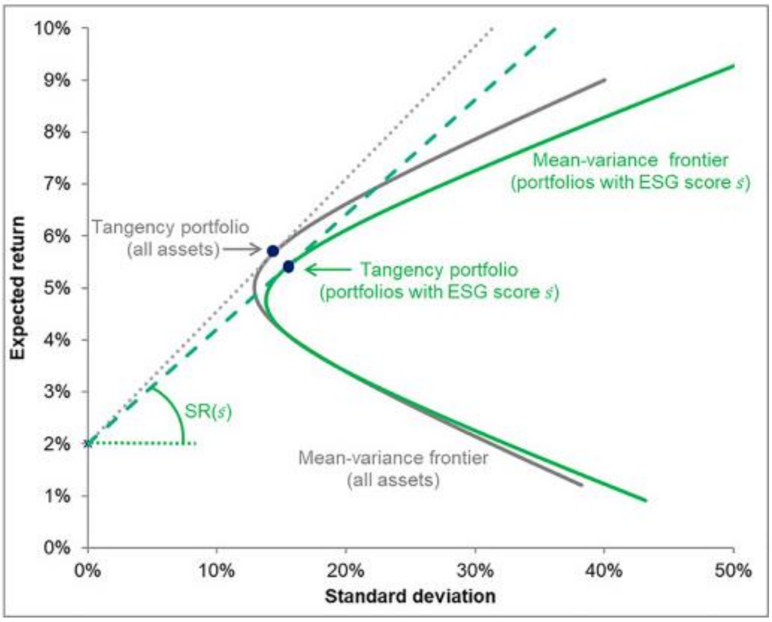


Figure 2: Schematic illustration of the ESG Efficient Frontier and the unconstrained Efficient Frontier (Pedersen, Fitzgibbons, & Pomorski, 2021).

On the contrary, they also find that investors without ESG-related restrictions may find a higher aggregate ESG score of their optimal portfolio than ESG-restricted investors, because unrestricted investors can take short positions in poor ESG assets resulting in hedged ESG risks or financing for leveraged positions in quality ESG assets.

There also exist contrarian viewpoints to the one that ESG-screening limits the investment opportunity set (Figure 2) and thereby decreasing diversification and increasing risk. Hoepner (2010) acknowledges that screening leads to worsened portfolio diversification through the number of securities and the correlation of securities. But he argues that another important diversification driver is the *average specific risk* of stocks, which can be reduced through screening as there seems to be a negative relationship between a firm's ESG rating and its specific risk. Similarly, Verheyden and Feiner (2016) find that ESG screening reduces tail risks, and find a positive effect originating from ESG screenings with an increase in annual return performance and a decrease in volatility, drawdowns and CVaR (Conditional Value at Risk) for screened universes.

1.3.3 Index Tracking and Tracking Error

Within the institutional investors' context, there are typically 2 general investment approaches: *active investing* and *passive investing*. With active investing, managers use information and forecasting models to let their portfolio *outperform* a benchmark portfolio that is of a similar asset class and/or focus as their active portfolio mandate. With passive investing, a given benchmark is replicated by a *tracking portfolio* that matches the performance of a benchmark. The latter strategy of course demands less input and work and is most often the choice of investors believing that the marketplace reflects all available information in the price paid for securities, and that it is impossible to consistently outperform the broad market benchmark. Often, institutional investors allocate the majority of their equity investments towards a passive mandate.

The most logical choice for a passive fund manager to construct a tracking portfolio is to simply buy all the stocks (constituents) that make up the benchmark index. This is known as *full replication*. However, often portfolio managers are imposed constraints that restrains them from full replication. On the practical side, certain stocks in the index could carry such a small weight that administrative and liquidity issues arise. Second, full replication demands almost continuous trading to rebalance the portfolio holdings. When transaction costs are high and tied to trading frequency, full replication can be costly. On the strategy side, portfolio managers might be given constraints on their portfolio such as a maximum weight for all constituents or screening constraints when adopting an exclusion policy (such as ESG-screening). For these reasons, many passively managed portfolios hold fewer stocks than the amount of constituents in the benchmark index. The problem of reproducing the performance of an equity index without buying all the underlying constituents is referred to as the Index Tracking Problem (Beasley, Meade, & Chang, 2003).

Whilst there is an abundant amount of strategies and quantitative methods for constructing a portfolio to replicate an index, it all comes down to designing a tracking portfolio whose *tracking error* relative to the benchmark is as small as possible (Rachev, Stoyanov, & Fabozzi, 2007). As stated earlier in this section*, excluding securities from the capitalization-weighted index leads to deviations in returns, or tracking error* (Blitz & Swinkels, 2021). In our case, the capitalization-weighted index is the MSCI ACWI, which is designed to represent the performance of the full opportunity set of large- and mid-cap stocks across developed and emerging markets and is therefore the most geographically diversified index.

Tracking error can also be seen as a measure of dispersion of specific risk as compared to its benchmark. This can be explained as follows: the residual return of a portfolio is the share of the return that is not explained by the benchmark, i.e. results from overweighting or underweighting securities as compared to the benchmark index. The residual risk, which is the diversifiable (specific) risk, measures the variations in residual return (Le Sourd, 2007). Blitz and Swinkels (2021), who examined the impact of excluding sin industries on expected portfolio risk and return, found indeed that exclusion leads to under-diversification and exposure to unwanted and diversifiable risk (or tracking error). They also show that this tracking error can be translated into an equivalent loss in expected return.

Because tracking error is a symmetric phenomenon, a larger tracking error does not directly imply worse performance. As Blitz and Swinkels (2021) argue, *"the outperformance of investors who exclude sin stocks is equal to the underperformance of investors who end up owning these stocks instead – and vice versa"*. However, we know that individuals assess their loss and gain perspective in an asymmetric manner (loss aversion). The researchers consequently state that the degree of tracking error can be limited by reweighing the remaining constituents (i.e. the stocks that are *not* screened out but are in the 'new' portfolio) through any optimization technique that selectively changes the weights of the remaining constituents so that they provide a hedge for the excludes stocks, rather than naively reweighting the remaining stocks based on market capitalization. Methods mentioned include increasing weights of stocks from the same industry, or stocks that offer similar factor characteristics. The former is only possible when one is not excluding entire industries.

1.3.4 The Minimum Needed

While it may be clear that screening can reduce specific risk, but not necessarily the *total* risk of a portfolio, another question that naturally arises is how far one can go with exclusion. That is, what is the bare minimum amount of stocks that a portfolio needs to eliminate most specific risk, in order to maintain its diversification to a fair degree?

The mythical legend within financial literature is that *"95% of the benefit of diversification is captured with a 30-stock portfolio"* (Fisher & Lorie, 1970). A stock portfolio holding 128 securities comes even more close to the diversification benefits of the benchmark index. Although the research and the data (NYSE-traded stocks during 1926-1965) are anything but recent, it is still a prevailing sentiment of 30 stocks being enough to hold a properly diversified portfolio, also supported by Statman (1987). More recent research nuances this view. The diversification benefits in the mentioned paper are measured in the reduction of total volatility, which includes both specific and systematic risk. As we mentioned in paragraph 1.3.2, MPT assumes that proper diversification cannot prevent systematic risk but only specific risk. In a more recent paper, Surz and Price (2000) state that R-squared and tracking error should be used as measures, rather than overall risk, to illustrate improvement in diversification as they are measures of diversification. R-squared measures the proportion of the variance that is explained by the market, or a benchmark. It can thus be used to measure *undiversifiable risk*. Tracking error on the other hand measures diversifiable risk, as explained in the previous section. The results of Surz and Price showcase that a 15-stock portfolio can only achieve 75-80% of available diversification (related to the market portfolio) and even a 60-stock portfolio achieves less than 90% of available diversification. Studying individual firm data over 1962 – 1997 in the US, Campbell et al. (2001) report an increase in specific (unsystematic) risk relative to the overall variability of the stock market, advocating for larger portfolios to minimize diversifiable risk. Accordingly, Statman (2004) states that the optimal number of stocks within a portfolio is at least 300, but this also depends on the cost of increasing diversification and the expense ratio of the portfolio concerned. Similarly, Haensly (2020) achieves a number of at least 300 stocks through decomposing total portfolio risk into systematic and specific risk and performing a simulation analysis.

Following the conclusions of a literature review on the topic (Zaimovic, Omanovic, & Arnaut-Berilo, 2021), it suffices to say that evaluating the number of assets needed for diversification is impacted by an enormous amount of different factors: ones definition of diversification, investor preferences, the change over time of the assets features etc. Nevertheless, the size of a well-diversified portfolio today is higher than in the past, primarily caused by increased unsystematic risk and decreased correlations between stock returns.

Regardless, it is clear that in order to be sufficiently diversified, it is not necessary to hold the entire universe of stocks available. However, portfolios that are screened in accordance with specific policies, such as ESG, are not randomly selected, let alone selected because of their negative correlation. In fact, it could be argued that remaining, high-quality ESG stocks are in some way correlated, thus bearing specific risk (Kurtz, 1997).

1.4 Research Design

Viewing this problem through the lens of an institutional investor prioritizes three things. First, regardless of the outcome of exclusion on the portfolio's performance, we should aim to replicate a similar performance by limiting residual risk relative to the benchmark. Continuously aiming to improve performance relative to the benchmark is time and resource consuming and not sustainable over time, as we assume that on the long term, the MSCI ACWI is a proxy for equity risk premium. Second, it means that we should aim to capture the diversification benefits of the MSCI ACWI, as this is one of the reasons to invest in such a broad index in the first place. Finally, our problem context is bound to tracking and maintaining the original index, because of the fact *that we use an original index*. If we were to outperform the MSCI ACWI but with a better ESG score, we might as well achieve that by constructing our own ESG portfolio bottom-up. But that is not the nature of the problem. Our concern is this: given the fact that we have the MSCI ACWI index universe, and given the fact that we perform exclusion, what happens and how can we maintain performance.

As stated earlier at the beginning of section 1.3, financial literature presents diverging results on the effectiveness and effects of screening policies, be it a general screening policy or specifically ESG-focused; while screening seems unfavorable, proper risk/return characteristics (in terms of Sharpe-ratio) and diversification can still be maintained with fewer securities. However what this exact number of securities is, remains unclear. Especially given the screening procedure is not random or optimized, but rather based on a rule-based screening policy. In other words: there is not a clear 'point of no return' when applying an ESG screening policy on a market benchmark. The effects of screening also differ for each original investment universe, with most research performed on US markets. These results, along with often occurring practical issues concerning portfolio management of institutional investors, are the core reasons for MN to perform a study on this subject themselves.

1.4.1 Research Objectives

In addition to maintaining a certain return profile, risk characteristics and allocation are equally important to maintain a consistent performance and (region-or sector related) shock protection. The first research goal is therefore to determine what performance measures (portfolio KPI's) capture relevant information for an institutional investor to track, given our problem context. That is, maintaining performance measures relative to the benchmark consistently over time, while accounting for allocation and practical constraints. Consequently, these performance measures should be calculated for our benchmark in order to have a baseline measurement. We aim to provide an objective view given our specific index used, while using performance measures that are both relevant in theory and practice.

The second goal is to apply a pre-determined exclusion policy on the benchmark portfolio, which will result in screened portfolios for each percentage amount of excluded securities (5% exclusion, 10%, etc.). Each screened portfolio will be measured on the performance measures that we defined earlier, in order to compare the effects of ESG-screening to the benchmark.

Paragraph 1.3.3 mentioned research that the degree of tracking error can be limited by reweighing the remaining securities (that are by default reweighted on their market capitalization) through any optimization technique. This is called *rebalancing*. The third research goal is therefore to develop an optimization program that rebalances the remaining securities in such a way that the deviations from the original benchmark performance measures are limited.

Consequently, evaluating the performance measures of these 'optimized' portfolios allows us to provide an insight into the dynamics that several constraints or optimization techniques impose on screened portfolios, and how they can be applied to sustain original benchmark performance.

In summary, we now have the following research objectives:

1. **Understand** the important drivers behind the broadly diversified equity benchmark and **identify** the performance measures related to risk, return and diversification that are relevant for institutional investors to maintain;

2. **Apply** a single pre-set screening policy accounting for ESG-scores and **analyze** the performance measures of the benchmark portfolio and the screened portfolio;

3. **Develop** a model that reweighs the remaining securities in the screened portfolios, leading to an optimized portfolio for each screened portfolio that accounts for the performance measures being within bounds of the original benchmark portfolio;

4. **Summarize** practical insights on the dynamics of screening and optimization on consistent portfolio performance.

Note that we now speak of 3 different portfolios. At first, we have the benchmark portfolio, which is the original MSCI ACWI index from which we perform the screening procedure. Resulting from the screening procedure, we have multiple screened portfolios for each level of exclusion. Then for each screened portfolio, we develop an optimized portfolio resulting from the optimization model. This results in 1 benchmark portfolio, and one screened and optimized portfolio for each level of exclusion.

1.4.2 Research Questions

The main goal of this research is to quantify the effects of exclusion on a benchmark portfolio and to propose a method that can generate an optimal portfolio by rebalancing the remaining constituents of a screened portfolio in such a way that the performance measures and diversification benefits are consistently analogous towards the benchmark portfolio, while improving the aggregated ESG score of the portfolio.

In order to achieve this, and the research objectives in the previous section, we formulate the following main research question:

Does there exist a trade-off between portfolio performance and ESG-screening, and can rebalancing the weights of a screened portfolio through optimization help to suppress deviations in performance from the benchmark?

The main research question can be answered by breaking down to the following sub questions, related to the research objectives discussed in the previous paragraph. Behind each sub question, the method that will be used to solve the question is written within brackets.

1. What performance metrics are most relevant and practical to quantify risk/return and diversification of a portfolio in the context of institutional investing? *(Method: literature study)*

2. How do the performance metrics of the screened portfolios differ from the benchmark portfolio? *(Method: data analysis)*

3. What methods can be used to optimize our screening target while accounting for the performance measures? *(Method: literature study)*

4. Can we reweigh the screened portfolios in such a way that we maintain performance while improving the aggregated ESG score through exclusion? *(Method: modelling)*

5. At what stage of exclusion can diversification or proper performance no longer be maintained, for the screened and the resulting optimized portfolios? *(Method: data visualization / data analysis)*

Note that at *question 2*, only the benchmark and screened portfolios are subjected to a data analysis on the performance characteristics. Eventually, it will also be necessary to perform this analysis on the rebalanced version of the screened portfolios i.e., the optimized portfolios in order to consistently report the differences. This will inherently be solved through *question 4*.

1.4.3 Research Outline & Methodology

Now that we have formed the research goals and corresponding questions, let us clarify and highlight the problem statement and how each chapter and sub question will aid in our problem solution. Figure 3 represents a schematic view of the problem context of the previous paragraph. If theoretical evidence were sound and applicable enough for an institutional investor to assess the impact of exclusions on its performance, this research would not have been viable. Although there is a vast amount of research available, both theoretical and empirical, about the effects of exclusion and a shrinking investment set, the conclusions do not satisfy the needs of an institutional investor. Those needs are simply the assessment against multiple practical performance measures and exclusion applied on a widely adopted and diversified index such as the MSCI ACWI. Although often mentioned as the most diversified equity portfolio, or 'market portfolio,' we do not assume ACWI to be the optimal portfolio according to MPT. Doing so would imply that any deviation from it would lead to worsened risk/return characteristics. What we do assume is that the geographical and sectoral diversification benefits of ACWI should be maintained (as pointed out earlier). Furthermore, selecting the ACWI as a benchmark brings practical advantages such as the availability of market data and corresponding ESG scores.

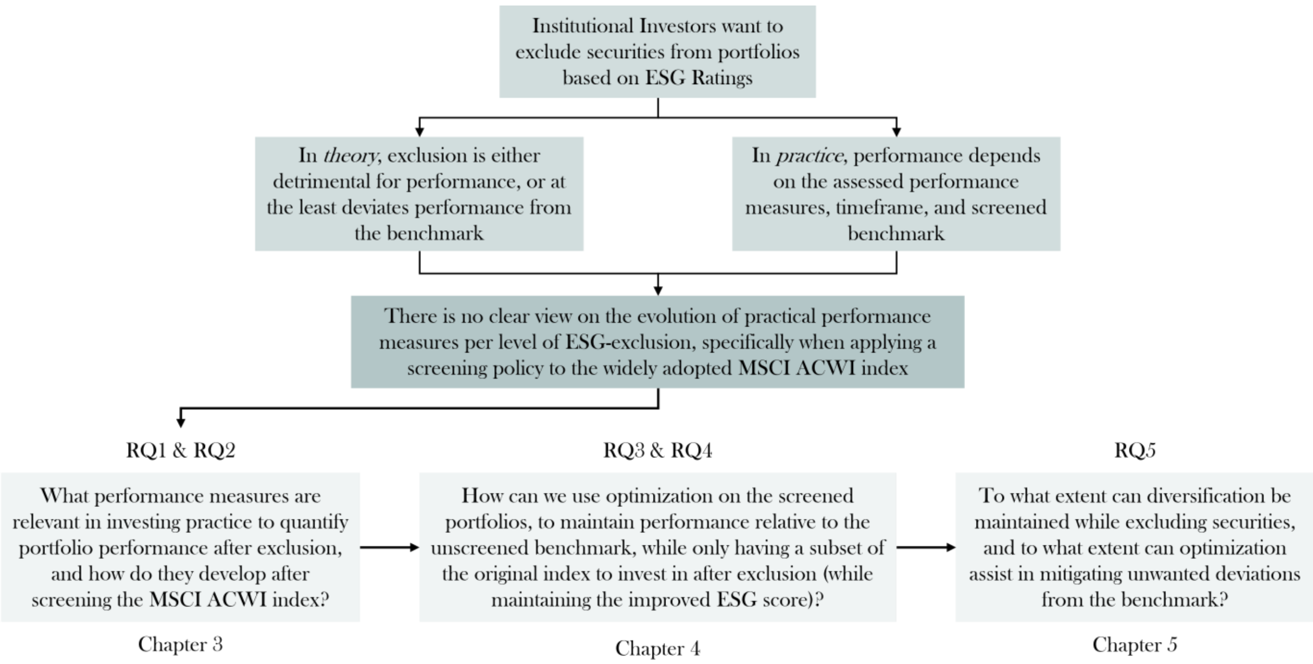


Figure 3: Schematic view of the problem context.

The figure also shows the, sometimes combined, research questions and in what chapter they will be answered.

Before we answer research question 1 and 2 in Chapter 3, we first perform a brief literature review that expands the views on literature already mentioned in the problem context section 1.3. First, we will consult and summarize the extensive amount of research regarding ESG and financial performance, broken down in company-level performance to prime the reader on this topic, and portfolio-level performance where we find out what performance measures have been used as gauges in previous studies. Performance measures found in earlier literature studies will eventually be complemented by measures deemed relevant from practical experience within MN. Then, we dive into earlier works on portfolio optimization, which will also help us in the selection of relevant performance measures for Chapter 3, as well as a good understanding of earlier used optimization methods in preparation for Chapter 4.

Following Chapter 2, the remainder of the thesis contains two major parts: screening and optimization. Part one is where we apply and discuss screening in Chapter 3, and part two is where we apply necessary optimization on this screening in Chapter 4. The results and comparison of screening versus optimization will subsequently be discussed in Chapter 5.

In detail: Chapter 3 will start with summarizing theoretical and practical knowledge necessary to select the performance measures that we will use throughout the research. These performance measures will then all be addressed and explained. By an exploratory data analysis, we will calculate some of these performance measures on the MSCI ACWI, which results we will use later to explain some of the effects of screening. Then, the screening procedure is introduced and executed whereafter the effects of screening on the selected performance measures is assessed. Chapter 3 thus answers the first and second research questions.

Chapter 4 answers the third and fourth research question by circling back to the theoretical context in Chapter 2 where we discussed similar and useful optimization methods. Together with the practical needs for optimization we encountered through Chapter 3, we will now design an optimization and discuss its practical implementation.

Chapter 5 then discusses the results of the optimization method along the same performance measures that were used to assess the screened portfolios. This enables us to answer our final research question, that is to what extent diversification can be maintained when systematically excluding securities, and to what extent can optimization assist in mitigating unwanted exposures. Chapter 6 consequently summarizes all earlier findings from the research questions, which allows us to answer the main question of this research.

Note that the appendix will also contain the complete result tables that are used throughout this thesis, for a printed overview and easy comparison.

1.4.4 Data

The ESG and Equity data used for the purpose of this research is provided by MSCI. For the ESG data, we use the MSCI ESG Ratings – Equities and MSCI ESG Ratings Time-Series – Equities. This dataset contains the ESG ratings and scores for constituents of multiple MSCI indices, including the ACWI, from 2007 up until 2021. We also use ACWI time series data for all securities comprising ACWI starting from 2001 up until 2021, with separate closing weights. All mentioned data has a monthly frequency. The ACWI time series data is consequently merged with the closing weights data and the ESG data. Except from some outlier cleaning in stock returns as a result of currency changes with the introduction of the Euro, and the handling of a small amount of missing values in sector designations, no large data cleaning operations have been performed.

After merging the ESG dataset with the Equities dataset, each constituent gets assigned a monthly risk-free rate value for which we use the 1-month US Treasury Yield (see 3.1 - Sharpe Ratio). Finally, we manually designate region codes (US, EMEA, APAC, EM) to each constituent based on the securities' country code and the MSCI country-region mapping, which is provided in Appendix E.

1.5 Research Contribution

The novelty of this research as compared to published academic research comes through a variety of factors. First, we show the effects of systematic screening based on MSCI ESG-Scores on very recent data (2021), whereby the coverage ratio of ESG scores is above 95% for all markets. Any research that examined this as well used less recent data or data from earlier years and as a result less coverage. Furthermore, we also examine the effects of screening along a wide variety of performance measures beyond the risk-adjusted return (Sharpe-ratio), which is something that we have only encountered occasionally (see paragraph 2.1.2 – Related Work). Some extra portfolio performance measures are also selected to provide a practical view for the specific use for institutional investors.

The main novelty of this research is the pragmatic and systematic exclusion in small steps that show results for each level of percentage screening. We do not solely use the best or most efficient exclusion approach, but rather provide a complete view on the effects of exclusion as for each investor, the 'point of no return' is subjective and based on his risk appetite.

Finally, the optimization method that is proposed is a combination of an index tracking problem that minimizes the tracking error and tries to replicate the benchmark index with fewer securities, while being constrained in its allocation (regions & sectors) and turnover as to retain diversification and keep optimization realistic for large investors. This method is a combination of the work of Alessandrini and Jondeau (2021) and Branch et al. (2019). Branch et al. minimizes tracking error via optimized exclusion, but does not impose constraints on other variables and neither reports results along different exclusion levels. Alessandrini and Jondeau do control for these exposures, but perform optimization only on the 200 firms with the largest market capitalization in the MSCI ACWI, and only on one level of exclusion, which is not explicitly mentioned. We will further discuss these papers in section 2.2 (literature study) and section 4.1 on the rationale of optimization design.

To the best of our knowledge, an optimized exclusion approach like ours has not yet been researched in financial literature. We do not try to find an optimal solution, but observe for each level of exclusion what the subsequent effect of this optimization is. The research could therefore serve as a simple but practical guide for investors and researchers to examine the effects of exclusion, and assess whether this form of optimization works for them, or use it as an inspiration or baseline to perform more advanced optimization on.

2 Theoretical Context

This chapter aims to provide the reader a deeper understanding of some of the theoretical concepts already touched upon in section 1.3. Note that the goal of this chapter is not to provide an exhaustive overview of all published work on sustainable equity investing, as this would be a study at itself. Rather, the goal is to familiarize the reader with the current perception on ESG investing and (corporate) performance, and the different considerations of implementing ESG into portfolio management. At the end of this chapter, we also want to have a more detailed view on the effects of screening on a portfolio so that we can compare these views with our own screening results. Furthermore, we want to have a good perspective on the variety of performance measures used throughout similar studies. Finally, we want to familiarize ourselves with similar work regarding optimization and ESG portfolios so that we have a solid information base for the answering of research questions 3 and 4. We will discuss these findings in section 2.3.

2.1 ESG & Financial Performance

In parallel with the appetite of investors towards sustainable investing, academic publications surrounding sustainable equity investing have been soaring in the past years. In a recently published book, Coqueret (2022) estimates the pace of development of SRI literature to be roughly two serious papers per day.

The relationship between ESG and financial performance, i.e. the returns of sustainable investments against the returns of non-sustainable investments, is one of the central questions in sustainable investing, to which not a single answer can be given. In a meta-analysis on 2200 primary studies, Friede, Busch and Bassen (2015) found 90% of the papers reporting a nonnegative relation between ESG and Corporate Financial Performance (CFP), thus on a *company-focused level*.

They suggest that the common mixed perception of investors on ESG investing is biased by results on *portfolio-focused* studies, which are overlaid by various idiosyncratic and systematic risks, as well as implementation costs. Because aggregated individual firm performance within portfolios may be different than primary firm data, we make the important distinction between studies explicitly focusing on company-level performance; and studies on the fund/portfolio level. Studying the former focuses on the impact that ESG has on operational metrics such as ROE, ROA, the cost of capital or the stock price. Studies on the fund level typically focus on risk-adjusted attributes such as the Sharpe ratio (a portfolios performance in excess of the risk-free rate, adjusted for return volatility, or risk) or other performance measures.

2.1.1 Company-level

Whelan and Atz (2021) performed a meta-analysis on 1000 research papers between 2015 and 2020 and found a positive relationship between ESG and corporate financial performance in 58% of studies. 13% of studies showed neutral impact, 21% showed mixed results and 8% depicting a negative relationship. This percentage is less than the earlier mentioned meta-analysis from Friede et al. but provides a more recent view. Evans and Peiris (2010) present results that find a positive relationship between ESG ratings and stock returns, valuations, and a company's operating performance. Eccles, Ioannou and Serafeim (2014) have tracked the financial performance of 180 companies within an 18 year-period and conclude that high sustainability companies outperform low sustainability companies, in terms of stock market and accounting performance. Similarly, Geczy and Guerard (2021) find that high ESG stocks earn higher returns than low ESG stocks.

On the other hand, regarding company-level performance, Chava (2014) finds that stocks that are excluded on an environmental basis exhibit higher expected returns from investors (as they demand compensation for exposure to risks associated with sin stocks) thus bearing a higher cost of capital. Similarly, El Ghoul et al. (2011) conclude that 'greener' firms have a lower implied cost of capital. Further reinforcing the existence of a 'sin premium,' Hong and Kacperczyk (2009) report significant positive abnormal returns for sin stocks, and Bolton and Kacperczyk (2021) find greater average stock returns for companies bearing a larger carbon footprint. Above findings seem to predict that ESG tastes from investors reduce the cost of capital and thus lowering a company's attractiveness.

However, Blitz and Fabozzi (2017) find that this sin premium becomes economically small and statistically insignificant when additionally controlling for more recently established asset pricing factors, in particular profitability, investment, and low risk. In other words, the strong historical returns of sin stocks observed in the earlier studies can be fully explained after all, when accounting for all relevant factor characteristics.

The mixed convictions on individual company-level performance and ESG score and the lack of consistent differences in performance between SRI and conventional strategies can perhaps be explained by Pedersen, Fitzgibbons, and Pomorski (2021). They explain market drivers by differentiating between 3 types of investors. On the one hand, ESG-unaware investors are not aware of ESG scores and seek to maximize their unconditional mean-variance utility. Second, ESG-aware investors similarly have mean-variance preferences, but they use ESG scores as an indicator for high future performance. Finally, ESG-motivated investors use ESG information and prefer high ESG scores as they seek an optimal trade-off between high expected return, low risk, and a high average ESG score.

Consequently, a market driven by ESG-aware investors will lead to higher expected returns of high scoring ESG stocks. On the contrary, a market driven by ESG-motivated investors yields lower expected returns for high scoring ESG stocks.

2.1.2 Portfolio-level

A positive relationship between ESG and company-level performance sets the expectation that a portfolio consisting of only quality ESG stock also performs well. Indeed, De and Clayman (2015) report a strong association between ESG ratings and stock returns, researching stock data from 2007 up until 2012. They conclude that within a portfolio, the highest return stocks -in terms of both simple return and risk(volatility)-adjusted return- always had superior ESG profiles. Moreover, portfolio performance could be improved by eliminating lower-tail ESG stocks (screening). This would also reduce overall portfolio volatility. Adjusting the returns for risk would yield equivalent results. More interesting, a statistically significant positive correlation between ESG and stock returns was not found, with the exception of during the peak of the 2007/2008 financial crisis. On ESG ratings and stock volatility, there was a strong statistically significant negative correlation, especially during periods of high market volatility. The authors argue that this implies portfolio diversification opportunities through the reduction of the average stock-specific risk, an argument which we earlier saw in research of Hoepner (2010) (paragraph 1.3.1). The finding that ESG ratings bring lower total volatility and consequently higher risk-adjusted returns (in terms of Sharpe ratio) is supported by Verheyden et al. (2016) and Kumar et al. (2016), who furthermore state that ESG factors bring higher returns, and that each industry is affected different by ESG factors.

The same meta-analysis that we mentioned earlier (Whelan & Atz, 2021) reports that 59% of investment studies, typically focused on attributes such as the Sharpe ratio and Alpha (excess return earned on an investment above the benchmark return), reveal a similar or better performance against conventional investment approaches.

Opponents of ESG tilted investing, stating that SRI is financially detrimental, not necessarily conclude that quality ESG-stocks are bad for performance but rather that excluded 'sin' stocks outperform benchmarks (Fabozzi, Ma, & Oliphant, 2008; Hong & Kacperczyk, 2009). Excluding these stocks from a broad benchmark would therefore result in missed out opportunities, as shown by Trinks and Scholtens (2017). Similarly, Dimson, Marsh, and Staunton (2020) report superior returns for specifically the tobacco and alcoholic beverages industries, but at the same time state that while large-scale exclusions could be costly and may exert temporary downward pressure on stocks, impact on portfolio returns on the long term is small.

Generally speaking, balancing the outperformance of the excluded securities with the advantageous risk-adjusted characteristics of quality ESG stocks results in a view that sustainable investing on a portfolio level does not really improve performance, but also doesn't hurt (Coqueret, 2022). Bauer, Koedijk and Otten (2005) for example studied ethical funds against conventional funds and find no evidence of differing risk-adjusted returns. While focusing only on the period 1990-2001, also more recent studies support this view (Hornuf & Yüksel, 2022; Blankenberg & Gottschalk, 2018).

Specifically focusing on screening, Cai, Cooper and He (2022) report that positive screening (opposite from negative screening) drastically reduces the universe size, thus reducing diversification, increasing volatility and cutting performance on risk adjusted-measures, while also incurring serious transaction cost penalties. Negative screening however does not result in performance degradation nor in extra transaction costs.

Related Work

The closest work yet to our own research setup is the work by Alessandrini and Jondeau (2020). On a data set from 2007 to 2017, they perform ESG-score based exclusions on the MSCI ACWI universe on the 10%, 25% and 50% exclusion levels. The ESG scores used here are industry-adjusted. Where a weighted average score combined with a high exclusion percentage at some point leads to the exclusion of entire industries that inherently perform poor on ESG (e.g. Energy), an industry adjusted ESG score measures the performance of each stock relative to its industry peers and thus prevents the exclusion of entire industries. The findings record improved ESG scores for each of the portfolios without deterioration of the risk adjusted performance (Sharpe ratio), and tracking error remaining relatively low. However, the screening also results in significant regional bets in favor of Europe and against US and emerging markets constituents. The screening also results in sectoral bets, in favor of information technology stocks and against financial and energy stocks. In a later research, Alessandrini and Jondeau also apply an optimization program accounting for these deviations. In the following section, this optimization and other research integrating ESG into portfolio optimization is discussed.

2.2 ESG & Portfolio Optimization

Portfolio optimization takes place in two stages. The first stage is optimizing between weights of different asset classes to hold i.e. asset allocation. Here, you attribute weights to the different asset classes in your overall portfolio (e.g., 60% equities/stocks, 40% debt/bonds). The second stage is optimizing the weights of assets within a single asset class, for instance, optimizing the weights of the constituents in an equity portfolio. This section will focus on the latter. The weights of the constituents in portfolio optimization are determined according to a given objective function. Examples are tracking error or mean-variance optimization which is most well-known as a component of Modern Portfolio Theory (Markowitz, 1952).

One approach that researchers have taken in previous years is the direct integration of ESG scores in the objective function of a portfolio optimization. This makes sense when investors value improved ESG scores more than financial gains. In case of the mean-variance optimization, a variety of extensions to the traditional mean-variance analysis have been proposed. Rather than maximizing expected returns given a specific level of risk, one could seek to maximize a combination between

ESG scores and returns (Baracchini & Addessi, 2012; Gasser, Rammerstorfer, & Weinmayer, 2017). Fish, Kim, & Venkatraman (2019) also use mean-variance analysis but adjust the returns with ESG metrics rather than directly involving them in the utility function. And while Drut (2010) presents a maximization of a mean-variance utility while subjecting the optimization to a minimum level of portfolio ESG score, Schmidt (2020) on the other hand minimizes the risk minus the ESG score while constraining expected return.

Further tailored towards an optimized exclusion problem, where – like in this research - a limited portfolio universe is optimized, Alessandrini and Jondeau (2021) propose an optimization whereby the aggregate ESG score of a portfolio is maximized, while constraining tracking error, turnover, regional and sectoral exposures, as well as risk factor exposures. The inspiration for the constraint settings come from an earlier paper published by these authors (see previous section). Their optimization is performed on the 4 separate regions US, Europe, Pacific and Emerging Markets whereby they select the 200 firms with the biggest market capitalization for that given region to ease computational burden. That selection for each region roughly corresponds to 'screening out' 50% of the firms, thereby matching the amount of stocks used in their previous analysis (see Alessandrini & Jondeau, 2020). However for the worldwide ACWI portfolio their set only covers a small amount of stocks. Nevertheless, they show that maximizing a portfolio ESG score (8,97/10) while keeping the tracking error (2,63%), turnover (99,56%), regional and sectoral exposures, and risk factors within stated limits can be achieved through optimization.

In other optimized exclusion methods, Branch & Cai (2012) combine the earlier described Index Tracking Problem (paragraph 1.3.3) with ESG-screened portfolios and show that a portfolio constructed of only socially responsible stocks can deliver market performance and statistically insignificant tracking error. However, thereby using data from 1996 to 2008 and only applied to the S&P500 as the benchmark index. Branch, Goldberg, & Hand (2019) also come close to our research question given the fact that they first perform binary exclusion (as we do with screening) where they experience tracking error, and then optimize by rebalancing the remaining securities so that they minimize tracking error with the MSCI ACWI. They argue that there is a clear trade-off between excess tracking error when performing cap-weighted exclusion (i.e. screening a portfolio and reweighting the remaining constituents on market capitalization) on the one hand, and risk of unwanted exposures when optimizing tracking error on the other hand. The trade-off results from correlation, as excluding unwanted constituents in combination with risk minimization leads to overweights in constituents that are correlated with the excluded constituents.

2.3 Conclusions on theory

So far, we have studied the financial literature on the relationship between ESG and financial performance, both on the company level and the portfolio level (paragraphs 2.1.1 and 2.1.2), as an extension of the literature already mentioned within the problem context (section 1.3). Furthermore, we shortly discussed related work and other examples of ESG-linked optimization designs.

2.3.1 General

It is hard to form a singular view on ESG and company-level performance. While there seems to be a positive relationship between ESG and stock price and accounting measures, looking at the relationship from an asset pricing perspective might imply less favorable characteristics for ESG stocks. Furthermore, the relationships also depend on one's definition of ESG or sustainability, and market dynamics. This makes it challenging to form a consistent viewpoint about what the exclusion of sin stocks, and thus the inclusion of quality ESG stocks would do on an aggregate portfolio level. In other words, research on individual assets performance provides insufficient information to form assumptions about portfolio performance post-screening.

Following Pedersen, Fitzgibbons and Pomorski (2021) as explained by Figure 2 in paragraph 1.3.2 we would expect screening processes to reduce the investment set and therefore moving the efficient frontier to the right thereby decreasing the Sharpe ratio, which is by definition suboptimal, at least in theory and in-sample. But when we look at empirical results on the screening of the MSCI ACWI index for instance, we see no deterioration in Sharpe ratio over time, but all the more in sector, regional, and risk factor exposures.

Although research on the fund level also depends highly on the choice of the rating agency (De Spiegeleer et al. (2020), see discussion in 6.2), the studied markets and the choice of performance measures, it shows clearer and more practical implications and an overarching conclusion can still be made on the effects of screening on a whole portfolio. While some studies report strong associations between stock returns and ESG ratings, most can be said about the negative relationship between ESG ratings and stock volatility. Although all agree that general screening limits the diversification opportunities, some of the diversification could be maintained because of the stock-specific risk of quality ESG stocks. Screening out poor performing ESG stocks could result in missed out opportunities, but still achieves the goal of the fund bearing a higher aggregate ESG score, while maintaining risk-adjusted returns, in part due to the decreased volatility.

2.3.2 Performance Measures & Optimizations

Next to the obvious stock return and volatility, by far the most used measure of portfolio performance is the Sharpe ratio as a measure of risk-adjusted return, as well as Alpha. When scholars specifically research screening, tracking error from the original benchmark is also used. Studies on ESG criteria in optimizations primarily focus on integral optimization. What we mean with that is that a portfolio is used as starting point, and the optimization program chooses which stocks to include or exclude. This is inherently different from our view that we want to optimize a fixed set of securities on which screening has already been performed, and we want to maintain the size of that set so that we can compare the optimized set against the screened set on an equal basis. Although the research of Alessandrini and Jondeau (2021) comes close to our intentions as to the maintaining of benchmark characteristics, they do not perform optimized exclusion but rather end up with a portfolio where an unknown amount of securities is screened out. We do see use in the performance measures that Alessandrini & Jondeau report and optimize on. These are (in addition to return, volatility, Sharpe ratio and tracking error) the exposures per sector and region, aggregate ESG score, and portfolio turnover (representing the amount of transactions).

The core of the research performed by Branch et al. comes even closer, as they minimize the tracking error on already screened portfolios (optimized exclusion), also using MSCI ACWI as a benchmark index. Their conclusion that there seems to be a tradeoff between tracking error and unwanted exposures (that form when minimizing the tracking error) is also seen in the optimization results of Alessandrini and Jondeau (2021). Although measured on a different time period and no similar exclusion percentages, they report an increase in tracking error (when optimizing for ESG scores and restricting regional and sectoral exposures) from 1.43 to 2.82 for respectively the screened MSCI ACWI universe and the optimized MSCI ACWI universe.

3 Effects of Exclusion on Portfolio Performance

This chapter starts with summarizing and explaining the performance measures that we deem important to assess post-exclusion portfolio performance in a practical and pragmatic way. The choice of these performance measures in part comes from the conclusions we made on the literature search in the previous chapter, where we looked at frequently used performance measures in related works. Discussions with investment strategists at MN yielded some additional performance measures, which we will also discuss. Hereby we propose an answer to the first research question, namely: *what performance metrics are most relevant and practical to quantify risk/return and diversification of a portfolio in the context of institutional investing?* In the second section, we will perform an exploratory data analysis on the MSCI ACWI index against some of the performance measures used. These results can later be used to explain effects of screening.

The third section introduces the part of this chapter that concerns the exclusion procedure. We discuss the ESG ratings and their availability and what effect it has on the results. Thereafter, we will describe the screening procedure, and in the subsequent section we discuss the screening results in comparison with the default benchmark performance. The latter section thus gives us an answer to the second research question: *how do the performance metrics of the screened portfolios differ from the benchmark portfolio?* Finally, we also discuss our own screening results against the results that we summarized in the literature review. As we earlier mentioned, one of the goals is to provide a practical overview on the effects of exclusion. This is done by plotting the performance measures against the sequential levels of exclusion, so that it becomes clear from what exclusion mark performance could be disproportionally decreasing with respect to the increased ESG score.

3.1 Performance Measures

The measures that were derived from the reviewed literature in the previous chapter primarily capture *returns*, *risks,* and the *trade-off* between the two. They are derived into two categories: *absolute* measures (stand-alone measures based on the portfolio characteristics) and *relative* measures (performance relative to the market capitalization weighted equity benchmark portfolio, here the MSCI ACWI). Absolute measures that we will use are return, risk and the Sharpe ratio as the risk-adjusted return. Alpha, as the active return over the benchmark, was also mentioned in multiple studied papers, but in accordance with MN, we find the information ratio to be a better relative measure to gauge benchmark outperformance. The information ratio namely adjusts the active return over the benchmark (alpha) for the volatility in dispersion between the benchmark and the tracking (screened) portfolios, also known as tracking error.

Tracking error is one of the most important measures that we will use. In addition to being an often-used measure by institutional investors, it also is the main optimization objective in the Index Tracking problem, in which we recognize a parallel with our own problem context.

Furthermore, we found that tracking error can be used as a measure of diversifiable risk. As we do not measure the beta of the portfolio but only the *total risk* in terms of returns in volatility, adding the tracking error as measurement will provide insights that help us dissect the different types of risk.

We choose to further replicate the performance measures mentioned by Alessandrini and Jondeau in their optimization framework, namely turnover, aggregate ESG score, and sectoral and regional exposures. The latter two are shown to be an important driver of risk and return, and not monitoring or constraining them could result in unwanted exposure to industry or region-related shocks. Furthermore, industry and region factors are useful in explaining variability in global equity returns (Menchero & Morozov, 2011; Norges Bank Investment Management, 2019). Although they also measure and constrain a portfolio's risk factors, we refrain from measuring them as we do not have enough data to calculate the factor loadings for all relevant factors, and because it would increase complexity. In addition to above performance measures, we add some simple measures like the average amount of constituents that are in the portfolio, and their corresponding total market capitalization compared to the MSCI ACWI benchmark. Finally, corresponding to practical relevance as discussed with MN, we add the maximum drawdown (MDD) measure, which serves as an indicator to identify a portfolios performance during market downturns. This is especially important as two portfolios could have the same tracking error, volatility, and information ratio, but their maximum drawdowns as compared to the benchmark could differ significantly.

Each performance measure is calculated on a monthly basis, and then (when applicable) annualized for the overall reporting. IAA score, sectoral exposures and regional exposures are reported overall as averages. Below, we illustrate the calculation method per performance measure:

Return

Calculated by multiplying the monthly closing weights with the total percentage return for each constituent. The total percentage return is determined by first calculating the percentage price return per constituent denominated in USD. Consequently, the dividend yield is determined and added to the percentage price return, resulting in the total return per constituent per month. The annualized return is calculated using the geometric mean of returns. The reason we choose the geometric mean rather than the arithmetic mean, is that the latter fails to account for compounding, or the compound annual growth rate of a portfolio. Moreover, when the variance in return is large from year to year, calculating an arithmetic average will overstate the actual average annual return.

Standard Deviation (volatility)

The monthly volatility is calculated as the standard deviation of the monthly returns over a 3-year period. The monthly volatility is annualized by multiplying with the square root of 12.

Sharpe Ratio

The Sharpe ratio is calculated by dividing the arithmetic mean of the monthly excess return by the monthly standard deviation of excess returns. The excess return is calculated by subtracting the risk-free rate from the portfolio's monthly return. For the overall calculation in the performance measures tables, the measures are calculated using a 3-year window. As the risk-free rate, we use the 1-month US Treasury Yield since our investments are denominated to Dollar values and the majority of securities is US designated. The monthly Sharpe ratio is annualized by multiplying with the square root of 12. As stated above, when the volatility of returns is higher, the geometric mean is lower. Calculating the Sharpe ratio by then dividing by the standard deviation would essentially imply that you discount your result for volatility twice, which should be avoided.

Tracking Error

The tracking error (TE) is a measure of divergence of the price behavior of a (tracking) portfolio and the price behavior of its benchmark. It is measured as the standard deviation of the difference between the portfolio and the benchmark return:

$$TE = \sigma(r_p - r_b)$$

Where r_p and r_b are respectively the portfolio return and the benchmark return. Above tracking error is the *ex-post*, or the backward-looking tracking error using historical returns. This version is most often used in reporting historical portfolio performance, as are for example the portfolio returns and historical volatility. *Ex-ante* performance measures on the other hand, are estimations of future performance observations and are generally used to control implied risk. Specifically, the formula of the *ex-ante* tracking error is given by the standard deviation of the active return:

$$TE = \sqrt{Var(r_p - r_b)} = \sqrt{E\left[\left(r_p - r_b\right)^2\right] - (E[r_p - r_b])^2} = \sqrt{(w_p - w_b)^{\mathsf{T}} \Sigma (w_p - w_b)}$$

where Σ is the forecast covariance matrix, and w_p and w_b are the constituent weights of the portfolio and the benchmark portfolio, respectively.

The monthly ex-post tracking error is calculated by taking the 3-year standard deviation of the monthly difference between the portfolio return and the benchmark return. It is annualized by multiplying with the square root of 12.

Information Ratio

The monthly information ratio is calculated by taking the 3-year arithmetic mean monthly return of the portfolio and the benchmark and subtracting these, and then divide them by the 3-year tracking error of the portfolio respective to the benchmark. The monthly information ratio is annualized by multiplying with the square root of 12.

Turnover

For the turnover (two-way), we need both the initial weights of each month and the closing weights. As we only have the closing weights of the index constituents, we compute the initial weights by dividing the closing weights by their total percentage return that month. Next, the turnover of month t is calculated by taking the sum of the absolute difference between the closing weights of month t-1, and the initial weights of month t. The turnover is annualized by multiplying with 12. The turnover is essential in portfolio management, especially when handling substantial amounts of money, as it is a measure of trading frequency or rebalancing. A higher turnover (high amount of buying and selling the stocks) results in high transaction costs. While the turnover from screening out stocks should be relatively limited, it is important to control when optimizing over the weights over the portfolio.

Maximum Drawdown

The maximum drawdown of a portfolio is calculated as the peak value of a portfolio before a largest drop, minus the lowest value before a new high is established, and that divided again by the peak value before the largest drop (to take a percentage of the peak value). The maximum drawdown is a one-of value over the entire measurement period.

IAA Score

The IAA score, or Industry Adjusted Average Score, is the ESG rating of use. It is calculated by taking the weighted average of a company's E, S and G rating and normalized for peer ratings (see 3.3.). The IAA score is calculated by aggregating the portfolio constituents' scores according to the weight of each constituent

Sectoral & Regional Exposures

Monthly sector and region exposures are calculated by summing the weight of each constituent belonging to the sector or region in question.

3.2 Exploratory Analysis of the Benchmark

Table 4 illustrates the annualized performance per sector, region and ESG rating as compared to the MSCI benchmark. Immediately some remarkable results can be seen. For instance, we see that the energy sector has delivered bad returns over the measurement period (2013 – 2021) with a high volatility and MDD compared to the benchmark. On the other hand, the information technology sector has experienced returns twice as high as the benchmark, while also bearing the highest aggregate ESG score of all sectors. In fact, 60% of all constituents of the AAA-bucket are within the information technology sector.

Furthermore, the outperformance of the US region as compared to the EU region is hard to oversee, especially taking the even lower volatility into account, also resulting in a much higher Sharpe ratio for US stocks. European and APAC constituents however see a far higher ESG score than their US and Emerging Markets counterparts, with especially the latter region exhibiting significantly poor ESG performance.

Our own results incline towards the literature in favor of a positive relation between ESG and financial performance. We can clearly see that sectors that have an overall higher aggregate ESG score have performed better over the period 2013-2021, while industries that exhibit significantly lower aggregate ESG scores than the overall benchmark, such as energy and real estate, also display weakened risk-return characteristics.

The literary conclusions on portfolio-level performance also correspond to our findings. On portfolio level, strong relationships were found between ESG ratings and stock returns while a negative relation was often found between ESG ratings and volatility. We can see this through the higher Sharpe ratio in the AAA rating bucket that decreases along with the ESG rating. Although this decrease in Sharpe ratio can be primarily attributed to the decrease in return (as there is no clear decreasing pattern in volatility through the buckets), we do see a significant increase in maximum drawdown when the rating of a bucket decreases. We will be able to form more conclusions on our own results with respect to academic literature when we discuss the screening results in section 3.5.

	MSCI ACWI	Energy	Materials	Industrials	Consumer Discretiona	Consumer Staples	Health Care	Financials	Information Technology	Comm. Services	Utilities	Real Estate
Average Constituents #	2642	143	240	395	341	216	192	503	258	119	135	96
Average Marketcap %	100	6.57	5.25	10.38	11.98	8.99	11.45	18.41	16.52	5.5	3.18	1.77
Annual. Return	11.11	0.41	6.05	10.22	13.05	8.4	13.69	7.75	22.35	8.55	8.36	4.15
Annual. Volatility	13.13	22.81	17.5	15.03	15.31	11.09	12.55	16.91	15.39	13.11	12.47	11.07
Annual. Sharpe Ratio	0.29	-0.19	0	0.22	0.38	0.11	0.48	0.09	0.89	0.12	0.11	-0.24
Max Drawdown %	-21.25	-53.01	-32.56	-26.21	-21.22	-13.69	-15.02	-33.97	-17.48	-16.11	-18.74	-23.63
Average Portfolio IAA Score	5.28	4.54	5.32	6.04	4.78	5.59	4.95	4.96	6.17	4.98	6.06	2.35

	MSCI ACWI	US	EMEA	APAC	EM	AAA	AA	A	BBB	BB	B	CCC
Average Constituents #	2642	716	450	459	1016	122	316	458	538	490	397	153
Average Marketcap %	100	56.61	20.77	11.29	11.34	7.86	14.84	21.52	22.09	16.33	9.63	3.31
Annual. Return	11.11	15.49	7.89	4.91	1.92	16.38	11.11	12.29	10.03	10.5	10.52	4.79
Annual. Volatility	13.13	13.49	15.08	13.16	15.92	13.59	12.69	13.74	13.2	14.06	13.47	16.08
Annual. Sharpe Ratio	0.29	0.57	0.08	-0.13	-0.26	0.62	0.3	0.36	0.22	0.25	0.25	-0.09
Max Drawdown %	-21.25	-20.13	-23.76	-23.41	-31.47	-15.05	-19.1	-20.84	-23.09	-24.44	-21.34	-39.66
Average Portfolio IAA Score	5.28	5.06	6.52	5.71	3.68	9.61	7.95	6.53	5.07	3.6	2.16	0.6

Table 4: Annualized performance measures per sector, region, and ESG-rating.

3.3 ESG Scores & Coverage

For this research, we make use of MSCI ESG data that keeps separate Environmental, Social and Governance scores for worldwide securities. For each company, MSCI calculates an Industry Adjusted Score, which is defined by taking the weighted average of all 3 ESG pillars and normalizing the score based on the score ranges set by the benchmark scores within the industry (GICS classification) set. This score then corresponds to the ESG Rating, ranging between AAA (best) and CCC (worst). Table 5 below illustrates the mapping of the scores towards the ratings.

LETTER RATING	*LEADER/LAGGARD*	*FINAL IAA SCORE*
AAA	Leader	8.571 – 10.0
AA	Leader	7.143 – 8.571
A	Average	5.714 – 7.143
BBB	Average	4.286 – 5.714
BB	Average	2.857 – 4.286
B	Laggard	1.429 – 2.857
CCC	Laggard	0.0 – 1.429

Table 5: Mapping of the Industry Adjusted Score to letter ESG Ratings.

In the remaining of this thesis, we will refer to the Industry Adjusted Score as the Industry Adjusted Average (IAA score). We will use the IAA score as our exclusion policy, as it has a greater coverage than the unadjusted weighted average of the 3 pillars, and excluding on IAA score avoids excluding entire industries when reaching higher percentages of exclusion.

Figure 4 and Figure 5 on the next page report the coverage of ESG scores for each defined region of the MSCI ACWI in both numbers of firms as well as market capitalization. The coverage for IAA scores and ESG ratings is equal, as ESG ratings are based on the IAA score. The 'World' Index is here the MSCI ACWI index. It can be seen that because of the drastic increase of ESG coverage in the Emerging Markets region, the overall ESG coverage in terms of both market capitalization and number of stocks reaches above 95%. For this reason, we only consider the data starting from 2013.

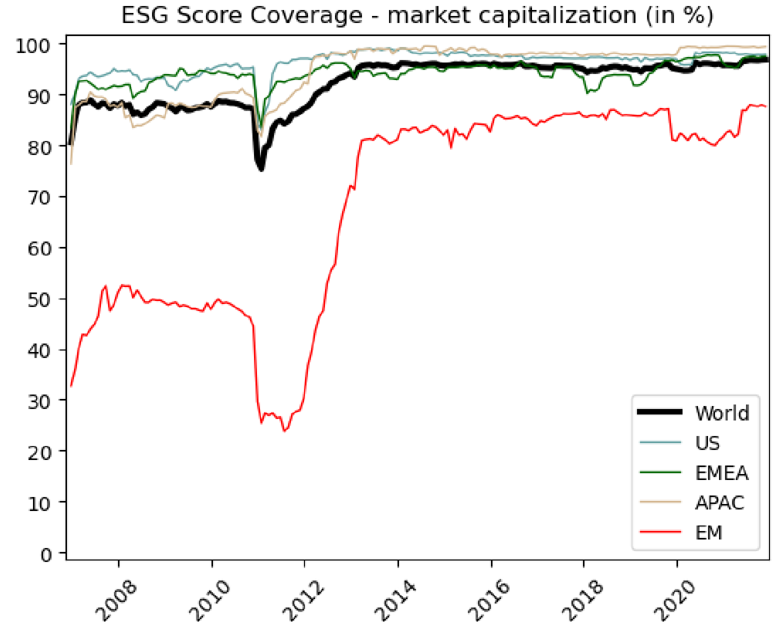


Figure 4: Coverage of IAA Scores per sub-region of the MSCI ACWI index, in terms of numbers of firms (%).

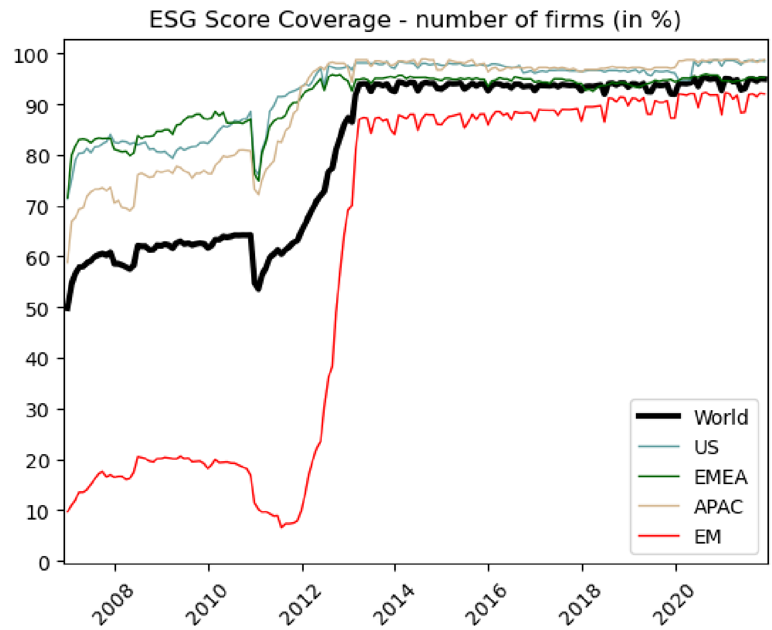


Figure 5: Coverage of IAA Scores per sub-region of the MSCI ACWI index, in terms of market capitalization.

3.4 Screening Procedure

As mentioned earlier, the screening procedure is carried out using the Industry Adjusted Score for each constituent. The starting point is not the MSCI ACWI universe, as not all constituents have an ESG score (see 3.3). All constituents without an ESG score are filtered out, resulting in the Pure portfolio, which represents around 95% of the market capitalization of MSCI ACWI. The performance of Pure is also incorporated in future performance evaluations, as this is a fairer benchmark to measure optimization performance against rather than MSCI ACWI. Note that all relative performance measures such as tracking error are still calculated with the MSCI ACWI as a benchmark.

Per month, constituents are ranked on their IAA score, where the total number of constituents is divided into 5% bins and for each exclusion step of 5% an extra bin is removed for that month. Whenever it occurs that the bin split is made in a range where the ESG scores are the same, the policy ranks the concerning constituents in the order they appear in the constituent list for consistency purposes. The remaining securities are then reweighted on market-capitalization. Figure 6 below illustrates that (logically) the evolution of the constituent amount per exclusion level is constant, and that the exclusion of market capitalization tends to move slower in the early stages of exclusion, but then accelerates in later stages. The screening results will be discussed in detail in the next section. The reason that we choose for monthly exclusion is that we have monthly price data.

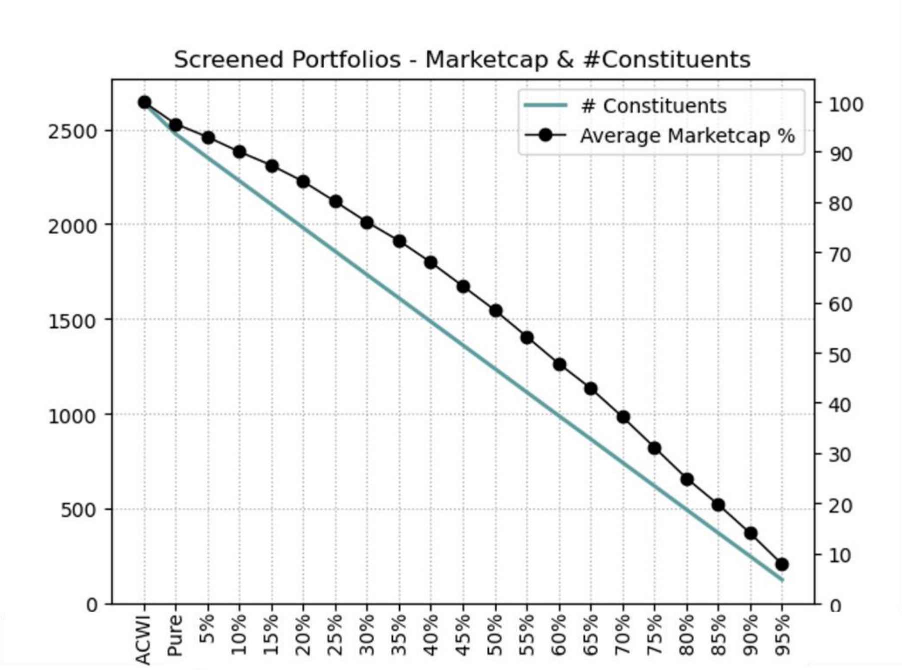


Figure 6: Evolution of number of constituents (left y-axis) and percentage market cap (right y-axis) per level of exclusion. Starting from 5% exclusion on the left up until 95% exclusion on the right.

3.5 Screening Results

On the next page, Figures 7-11 display the evolution of performance measurements per exclusion level. Further, Table 6 shows for each exclusion level the portfolio performance according to the performance characteristics introduced in section 3.1. First of all, as expected, excluding more securities results in an improvement of the aggregated ESG (IAA) score, as seen in Figure 9. This is inherent to the screening design because the screening excludes the worst ESG-scores for every step of exclusion.

Risk-adjusted returns

Because we see a significant increase in annualized return, with similar levels of volatility, the result is an increasing Sharpe ratio the more securities are excluded. In section 1.3 and specifically 1.3.2 on Modern Portfolio Theory, we mentioned that any shrinking of the investment set would possibly lead to an increase in specific risk and thus less favorable risk/return characteristics. Although annualized volatility does not fully capture specific risk, it is at least interesting to see that risk/return characteristics improve for each level of exclusion. As mentioned earlier, empirical evidence on the effects of screening is also dependent on the index from which screening is performed, as well as the timeframe of the dataset. Especially timeframe is an important aspect here, since we think that a large amount of the improved returns can be attributed to the fact that the information technology sector gains significant exposure the more securities are excluded. In section 3.2, we saw that indeed the information technology sector exhibits the highest aggregated ESG rating (thus it makes sense that its exposure grows, the more low ESG rated stocks are excluded), but also annualized returns that are double the returns of the MSCI ACWI. We therefore think that the improved returns are mostly owed to an increase in IT exposure, which is a sector that has happened to perform extraordinary well in the sample space (2013-2021). This observation is also in accordance with empirical evidence discussed in paragraph 2.1.2 on ESG and portfolio-level performance; namely that the highest return stocks had superior ESG profiles and that portfolio performance could be improved by eliminating lower-tail ESG stocks.

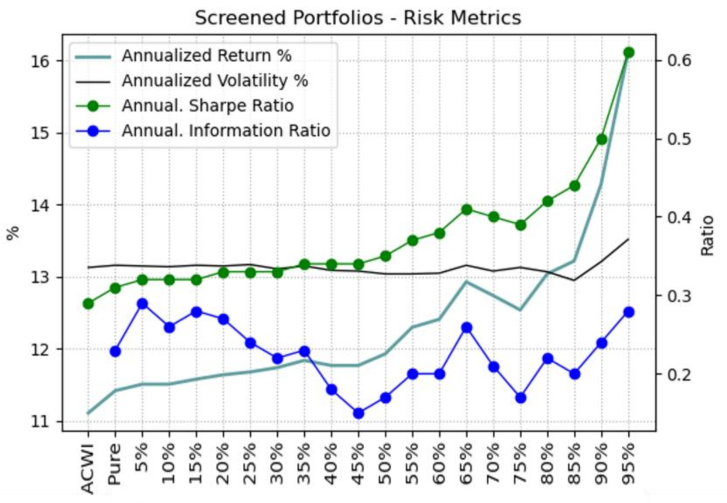


Figure 7

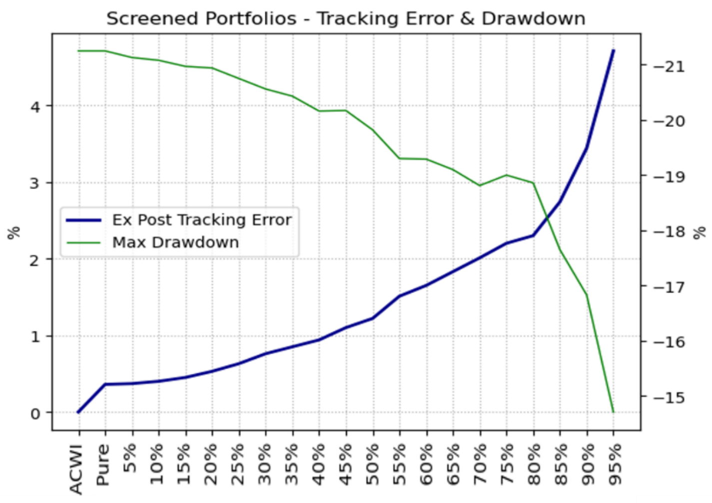


Figure 8

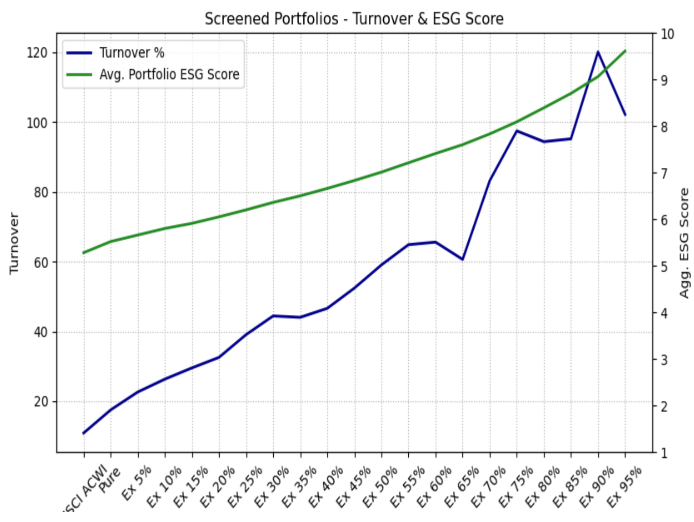


Figure 9

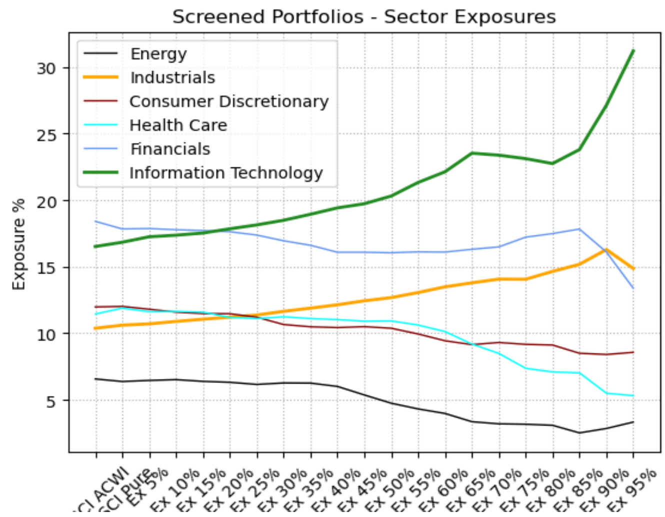


Figure 10

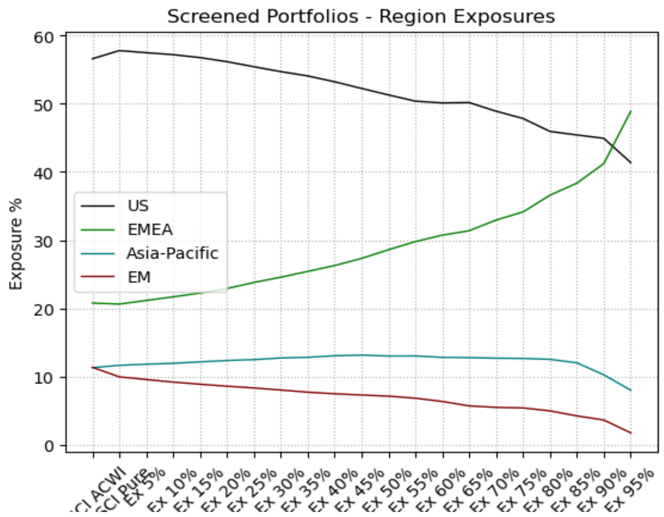


Figure 11

Figures 7-11: Evolution of performance measures per level of exclusion

	MSCI ACWI	Pure	Ex 5%	Ex 10%	Ex 15%	Ex 20%	Ex 25%	Ex 30%	Ex 35%	Ex 40%	Ex 45%	Ex 50%	Ex 55%	Ex 60%	Ex 65%	Ex 70%	Ex 75%	Ex 80%	Ex 85%	Ex 90%	Ex 95%
Average Constituents #	2642	2477	2352	2229	2105	1981	1857	1733	1610	1486	1362	1238	1114	990	867	743	619	495	371	248	124
Average Marketcap %	100	95.58	92.97	90.03	87.34	84.11	80.12	76.02	72.3	67.99	63.26	58.5	53.21	47.85	42.94	37.22	31.15	25.02	19.75	14.04	7.85
Annual. Return %	11.11	11.42	11.51	11.51	11.58	11.64	11.68	11.74	11.84	11.77	11.77	11.93	12.3	12.41	12.93	12.74	12.54	13.04	13.22	14.29	16.12
Annual. Volatility %	13.13	13.16	13.15	13.14	13.16	13.15	13.17	13.11	13.15	13.09	13.08	13.04	13.04	13.05	13.16	13.08	13.13	13.07	12.95	13.21	13.52
Annual. Sharpe Ratio	0.29	0.31	0.32	0.32	0.32	0.33	0.33	0.33	0.34	0.34	0.34	0.35	0.37	0.38	0.41	0.4	0.39	0.42	0.44	0.5	0.61
Annual. Tracking Error (Ex-Post) %	-	0.36	0.37	0.4	0.45	0.53	0.63	0.76	0.85	0.94	1.1	1.22	1.51	1.65	1.83	2.01	2.2	2.3	2.74	3.44	4.71
Annual. Tracking Error (Ex-Ante) %	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Annual. Information Ratio	-	0.23	0.29	0.26	0.28	0.27	0.24	0.22	0.23	0.18	0.15	0.17	0.2	0.2	0.26	0.21	0.17	0.22	0.2	0.24	0.28
Annual. Turnover %	10.87	17.5	22.63	26.31	29.53	32.55	39.06	44.42	44.02	46.6	52.39	59.04	64.84	65.57	60.61	83.16	97.43	94.35	95.16	120.07	102.16
Max Drawdown %	-21.25	-21.25	-21.13	-21.08	-20.97	-20.94	-20.75	-20.56	-20.43	-20.16	-20.17	-19.82	-19.3	-19.29	-19.1	-18.81	-19	-18.86	-17.65	-16.83	-14.71
Average Portfolio IAA Score	5.28	5.52	5.66	5.8	5.91	6.05	6.2	6.36	6.5	6.66	6.83	7.01	7.21	7.41	7.6	7.83	8.09	8.39	8.7	9.06	9.61
Regional Exposures %																					
US	56.61	57.81	57.5	57.21	56.78	56.18	55.44	54.72	54.09	53.23	52.26	51.3	50.39	50.13	50.18	48.91	47.85	45.95	45.43	44.94	41.38
EMEA	20.77	20.62	21.15	21.68	22.23	22.9	23.79	24.56	25.41	26.27	27.32	28.59	29.79	30.73	31.38	32.95	34.14	36.59	38.36	41.22	48.9
Asia-Pacific	11.29	11.63	11.8	11.93	12.14	12.34	12.47	12.71	12.81	13.04	13.13	13	13.01	12.8	12.76	12.68	12.64	12.51	12	10.23	7.99
Emerging Markets	11.34	9.94	9.55	9.17	8.86	8.57	8.31	8.01	7.69	7.46	7.28	7.11	6.81	6.33	5.68	5.46	5.38	4.95	4.21	3.61	1.73
Sector Exposures %																					
Energy	6.57	6.38	6.46	6.52	6.39	6.32	6.16	6.27	6.26	6.01	5.37	4.75	4.32	3.98	3.36	3.2	3.17	3.09	2.52	2.85	3.33
Industrials	10.38	10.61	10.71	10.89	11.06	11.21	11.37	11.65	11.89	12.14	12.44	12.68	13.06	13.49	13.79	14.07	14.06	14.65	15.18	16.29	14.87
Consumer Discretionary	11.98	12.02	11.81	11.59	11.48	11.47	11.21	10.66	10.49	10.43	10.5	10.38	9.95	9.44	9.15	9.31	9.17	9.12	8.5	8.41	8.57
Health Care	11.45	11.89	11.62	11.65	11.59	11.19	11.11	11.24	11.11	11.03	10.91	10.93	10.62	10.13	9.19	8.49	7.37	7.1	7.03	5.5	5.32
Financials	18.41	17.84	17.87	17.78	17.72	17.63	17.38	16.95	16.61	16.09	16.09	16.05	16.12	16.1	16.31	16.49	17.22	17.49	17.83	16.11	13.39
Information Technology	16.52	16.84	17.25	17.37	17.53	17.85	18.14	18.49	18.94	19.42	19.73	20.31	21.33	22.13	23.53	23.38	23.12	22.75	23.8	27.12	31.21
Materials	5.25	5.34	5.25	5.24	5.23	5.26	5.32	5.45	5.5	5.54	5.62	5.53	5.44	5.22	5.05	5.01	5.21	5.09	5.12	5.09	5.16
Consumer Staples	8.99	9.11	9.03	9.05	9.09	9.17	9.42	9.49	9.56	9.75	9.86	10.02	10.2	10.43	10.7	11.02	11.06	10.46	10.03	8.75	6.85
Communication Services	5.5	5.3	5.33	5.23	5.25	5.22	5.14	4.98	4.81	4.75	4.58	4.34	3.8	3.69	3.71	3.8	4.16	4.41	4.21	4.07	5.19
Utilities	3.18	3.27	3.3	3.33	3.35	3.38	3.46	3.55	3.57	3.61	3.7	3.8	3.96	4.1	4.05	4.19	4.44	4.74	4.88	4.96	5.54
Real Estate	1.77	1.4	1.39	1.34	1.31	1.29	1.28	1.28	1.25	1.23	1.2	1.2	1.21	1.28	1.15	1.05	1.02	1.1	0.9	0.84	0.58

Table 6: Portfolio performance for each level of exclusion.

After exclusion, remaining constituents are reweighted on market capitalization.

Volatility

A second view at volatility development provides us with a new insight. Academic research that we previously discussed, mentioned that although diversification efficiencies shrink and specific risk increases, this could be compensated through the reduction of the average stock-specific risk that high-quality ESG stocks exhibited. Some research also found a strong statistically significant negative correlation between ESG ratings and stock volatility, especially during periods of high market volatility (see 2.1.2). When we take a look at Figure 7 on the upper left side, we see that the *total* annualized volatility stays virtually equal throughout exclusion levels. This is not an unforeseen event, since we saw in section 3.2, Table 4 that volatility along the rating buckets and regions do not exhibit large outliers as compared to the benchmark index. Only some individual sectors show larger differences in volatility, so a larger difference in sector exposure post-exclusion would hence result in a larger volatility change. However, Figure 10 shows that the sector exposure of the 5 largest industries plus the energy sector remains relatively equal in the first stages of exclusions. In the final stages, from 85% exclusion onwards, larger changes are seen in the sector exposures, and hence also in the annualized volatility of the screened portfolio (see Figure 7).

However, when we take a look at the time-series development of volatility along different exclusion levels, we see indeed what previous academic research has also found: that high-ESG portfolios exhibit lower volatility during periods of high market volatility, whereas they show higher volatility during relatively calm periods.

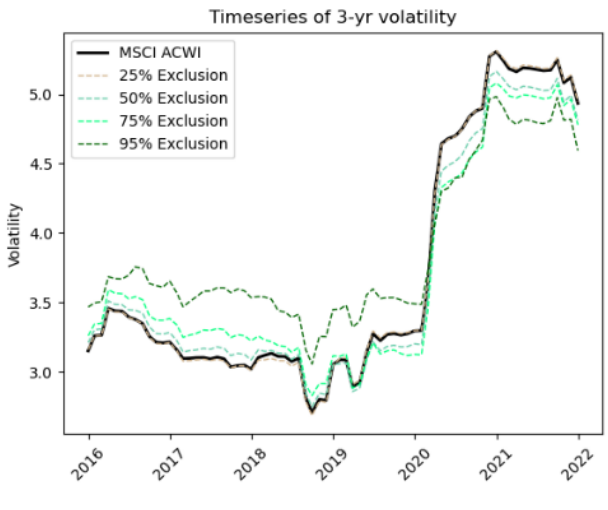


Figure 12: Timeseries of the 3-year standard deviation of returns (volatility - %) for various exclusion portfolios.

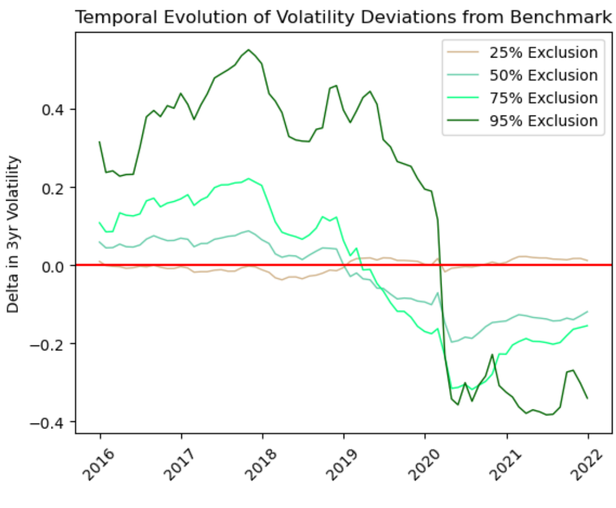


Figure 13: Timeseries of the difference in screened portfolio volatility (%) with the MSCI ACWI benchmark (red line). A positive difference implies a higher volatility for the screened portfolio.

Here, Figure 12 shows the time-series of the 3-year volatility for 4 different screened portfolios and the MSCI ACWI, where we clearly see that the more exclusion has taken place, the more volatility the portfolio carries.

When market volatility increases, for instance during the Covid pandemic volatility in 2020, this inverts. Figure 13 shows the delta in volatility from the benchmark (red line). Volatility levels above the red line means that the portfolio carried a higher volatility than the benchmark, and below the red line that it carried a lower volatility.

Regional Exposure

In addition to changes in the largest sectors, we also observe a change in region exposures. These observations are in line with Alessandrini & Jondeau (2020). As mentioned in the literature study, they observed no deteriorated risk-return performance, but a significant bias towards region, sector, and risk factor exposures. The shift in region exposures makes sense, because EMEA-linked firms tend to have a higher ESG score (avg. 6.52) in favor of the US (5.06), APAC (5.71) and Emerging Markets (3.68). Excluding the lower tail of ESG ratings thus reveals a decrease of EM exposure from 11.34% (ACWI) to 1.73% (95% excluded). Similarly, EMEA exposure increases from 20.77% to 48.9% in respectively the benchmark and our 95% screened portfolio. APAC stocks show less of a deviation, while US stocks also showcase a major decrease in favor of higher ESG-quality EMEA stocks.

Tracking Error & Information Ratio

As earlier observed in the problem context of section 1.3, scholars demonstrate that exclusions from the benchmark lead to tracking error. As Figure 8 exhibits, we indeed see an increasing exponential pattern of ex-post tracking error with the increasing of exclusion. A good reference to a 'high' tracking error is according to the typical levels of tracking error defined by Vardharaj et al. (2004). An index fund should theoretically have a tracking error of zero, where an enhanced index fund should exhibit TE below 2%, and an actively managed fund should bear between 5% and 10% tracking error compared to its designated benchmark. Although it is hard to place a label, we essentially form multiple 'enhanced index funds' through our screened portfolios. We should therefore keep in mind that below or around 2% ex-post tracking error is reasonable. We see in Table 6 that this threshold is exceeded after the 65% exclusion level.

It was argued that tracking error is a symmetrical phenomenon which means that it can be both a positive and a negative occurrence. We do however not see the translation towards negative returns, moreover we observe a decrease in maximum drawdown (MDD), almost perfect negatively correlated with the increase in tracking error (Figure 8). At first this leads us to think that the tracking error seems to be on the positive side as annualized returns increase with it and maximum drawdown decreases. However, when we take a look at the information ratio, which is the measure of portfolio returns beyond the returns of the benchmark, we see only a slightly positive number with little to no change. This means that the returns made by the exclusion portfolios are indeed in excess of the benchmark, but are largely the result of tracking error.

Turnover

Finally, we observe a steady increase in turnover, the more exclusions take place. The lower turnover for the benchmark is because it only accounts for index rebalancing semi-annually and constituents that start reporting ESG scores. However, when all firms are reweighted to market capitalization, the sum of the change in weights (which is the turnover) should still be the same as the benchmark as no optimization (=extra trading) is involved. The increase in turnover may be explained by the limited number of securities that is left when exclusion increases. Because exclusion happens on a monthly basis, a security that has a much larger weight in the screened portfolio but gets kicked out after a few months because other constituents simply have a better ESG score causes a higher turnover.

3.6 Conclusion

At the beginning of this chapter, we set the intention to answer the first two research questions. These were:

1. *What performance metrics are most relevant and practical to quantify risk/return and diversification of a portfolio in the context of institutional investing?*
2. *How do the performance metrics of the screened portfolios differ from the benchmark portfolio?*

In section 3.1 we discerned between performance measures as used in previous research, such as the return, risk, Sharpe ratio and tracking error. But we also added measures originating from MN, such as the information ratio and the maximum drawdown (MDD) measure as indicators that look deeper into relative performance against the benchmark versus the absolute measures such as risk and return. Also, related research was found (Alessandrini & Jondeau, 2021) that coincided with the wishes of MN to also measure turnover, and regional and sectoral exposures. Consequently, performing an exploratory data analysis on the benchmark portfolio with some of these measures would later assist us in explaining some of the effects of exclusion (see 3.5).

After the introduction on the ESG data and the screening procedure itself, we were able to answer the second research question through section 3.5, where we discussed the effects of screening the MSCI ACWI benchmark portfolio. When we plot the average performance of each metric against the level of exclusion, we clearly see the development of these metrics the further we take the screening. The results are explained in detail in the previous section, but what we can conclude overall is that over the time period measured (2013-2021), our findings match the positive academic findings from section 2.3. That is, exclusion has *positively* impacted performance on an *absolute* basis thus in terms of risk, return and therefore the Sharpe ratio. It seems to hold that eliminating lower-end ESG stocks contributes to positive performance, although we think that performance is mostly sector-attributed. In terms of risk, we see an almost unchanged evolution the more exclusion is taking place. We do see higher risk when 95% of securities is excluded, but we also observed that while exclusion leads to higher risk compared to the benchmark in relatively calm market environments, risk is lower than the benchmark when market volatility is high. This observation also fits in previous mentioned academic literature.

As earlier discussed, the effects of screening are subjective to the timeframe of the data used, and the performance measures against which performance is assessed. The majority of academic research has focused their work on portfolio performance against *absolute measures*, where we indeed find a positive relationship between screening and performance.

Looking beyond the *absolute* measures and focusing on *relative* measures and exposures, two categories that we earlier marked as important for MN and institutional investors, we see a different perspective. The information ratio is slightly positive but stays relatively unchanged through exclusion levels, which makes us believe that the increase in returns is for a large part attributed to the increase in tracking error, i.e. neutralizing the tracking error would also neutralize the returns. Practically speaking, the deviation from the benchmark has not been detrimental in the measured period, even quite the opposite, but the mere fact that there *is a deviation* still serves as a warning sign. Not only the deviation in residual risk and returns, but also in sectoral and regional exposure and in turnover could jeopardize an investor when the market turns to the wrong side of returns or when a regional or sectoral specific shock event occurs.

4 Optimizing the Screened Portfolios

In the previous chapter we discovered the effects of screening a portfolio and concluded that when looking deeper into a wider variety of performance measures, some unwanted exposures arise as a result. For institutional investors that follow a passive mandate, it is more important to follow a benchmark rather than to beat it (Alessandrini & Jondeau, 2021). As much as exclusions seem beneficial in terms of overall risk and return over the measured periods, they expose the portfolio towards an increased tracking error, which means that the returns beyond the benchmark that are adjusted for risk (information ratio) remain relatively unchanged. At the same time, the screening exposes the portfolios to unwanted sectoral and regional exposures, and an increase in turnover.

The goal of this chapter is to find out how we can use optimization on the screened portfolios, to maintain performance relative to the unscreened benchmark, while only having a subset of the original index to invest in after exclusion. Accordingly, we try to answer the third and fourth research question:

3. *What methods can be used to optimize our screening target while accounting for the performance measures?*
4. *Can we reweigh the screened portfolios in such a way that we maintain performance while improving the aggregated ESG score through exclusion?*

Where the fourth research question is also partly answered in Chapter 5 where the results will be discussed.

In order to answer the third research question, this chapter starts with a section that looks back at the related optimization methods that we discussed in section 2.2, and how they can help us to find our optimization objectives, where the second section discusses ways to solve these objectives. According to these objectives, we present the optimization design in section 4.3, along with the practical implementation of the optimization (section 4.4) where we verify research question 4. The definitive answer to research question 4 will consequently be formed after Chapter 5, where we discuss the results of the optimization.

4.1 Rationale

Reproducing the performance of a benchmark index while only having a subset of the original index to invest in is an Index Tracking Problem, as we discussed in paragraph 1.3.3. We see an analogy with our own problem context, where we screen a benchmark portfolio and try to maintain its performance with a portfolio that contains less securities as a result of this exclusion. Although we did not extensively research the various solution strategies that come with the ITP, we notice that the main goal of an ITP is to minimize the tracking error of the subset portfolio.

Minimizing the tracking error fits our optimization goal as well, while tracking error is also an important measure used by institutional investors when comparing performance against a benchmark. This brings us to the earlier discussed work of Branch et al. (2019) (see section 2.2), where exclusion is performed on the MSCI ACWI, and consequently optimized with minimization of the tracking error. However, the exclusion is done based on industry rather than ESG-ratings, which would result in unwanted sector exposure (section 3.5) even more. Moreover, they argue that there is a clear trade-off between excess tracking error when performing cap-weighted exclusion (screening a portfolio and reweighting the remaining constituents on market capitalization) on the one hand, and risk of unwanted exposures when optimizing tracking error on the other hand.

This is where the work of Alessandrini and Jondeau (2021) comes in, where the ESG score of a portfolio is maximized and constraints are imposed on the tracking error, turnover, regional and sectoral tilts and on factor exposures. These exposures are known to develop the more screening is taking place (see 3.5).

It turns out that both works described above are not directly useful within our problem context, but a combination of them is. First, Branch et al. only perform optimized exclusion and do not impose constraints on other variables. Moreover, there is no mentioning of the exact optimization design, and neither is there any form of systematic exclusion via ESG ratings. The work of Alessandrini and Jondeau, on the contrary, does control for the exposures, as we intend to as well. However, this research performs optimization on only the 200 firms with the largest market capitalization for MSCI ACWI and some of its subregions. This is different from systematic optimization for each exclusion level as we intend to do. Next to that, their program allows stock weights to be set on 0 when it deems beneficial to do so. This is inherently different from our work, where we have a *fixed* subset of securities that are subject to optimization. The only variable change in our optimization should be the weights of the securities, not the amount of securities.

We furthermore propose an optimization that rebalances the weights on a *monthly basis*. Similarly, we performed our exclusions on a monthly frequency. Not only is this the most compatible to our monthly return data and ESG-ratings, but it also allows institutional investors to process new market information into the portfolio construction quickly.

The combination that we propose is an optimization program that employs a target function of minimizing the tracking error of the screened portfolio, while accounting for several other constraints on linear performance measures that are shown to fluctuate when a portfolio is screened (section 3.5), based on the optimization design of Alessandrini and Jondeau. These are the turnover, sector weight and regional weight.

To the best of our knowledge, treating an ESG-screened portfolio as an index tracking problem by minimizing the tracking error relative to its original investment universe or benchmark, while accounting for other variables of which previous research has shown that they are inclined to fluctuate while screening or rebalancing, has not been researched before. Most optimizations rather focus on a more integral optimization, where an algorithm is executed on a complete portfolio, and thus the algorithm chooses which securities are optimal to exclude at that point in time.

The reason we resort to optimized exclusion rather than integral exclusion is simple. The exclusion policy, i.e. the ranking and screening of securities based on ESG score, is an inflexible task as it is imposed through strategical decision making. We can not decide for ourselves which securities to exclude first. We do not necessarily want to create the optimal ESG portfolio, but rather whether an optimization program can help 'soothe' the deviations resulting from screening.

4.2 Tracking Error Minimization

Following Lezmi, Roncalli, & Xu (2022); Roncalli (2013) and Perrin & Roncalli (2020), we consider an extension of the classic Markowitz mean-variance framework (which can be cast into a Quadratic Programming function) that performs a multi-period portfolio optimization, while minimizing the ex-ante tracking error. The interest in reformulating this as a Quadratic Programming problem is twofold. First, the ex-ante tracking error is a quadratic measure. Second, so that we can benefit from the convex optimization framework, that is, can be sure that the solution will be the global minimum. Here, it is also easy to add linear constraints such as the region and sector exposures, and to adapt some constraints that are neither linear nor quadratic, but absolute (turnover).

In section 3.1., we have shown that the *ex-ante* tracking error can be written as:

$$TE = \sqrt{Var(r_p - r_b)} = \sqrt{E\left[(r_p - r_b)^2\right] - (E[r_p - r_b])^2} = \sqrt{(w_p - w_b)^\top \Sigma (w_p - w_b)}$$

Now with a reformulation of the quadratic mean-variance optimization (see Appendix A.1.), Roncalli (2013) shows that a multi-period mean-variance objective function:

$$g(w) = \sum_{t=0}^{T} \left\{ \frac{1}{2} w_t^\top \Sigma_t w_t - \gamma w_t^\top \mu_t \right\}$$

Can be rewritten to:

$$g(x) = \sum_{t=0}^{T} \left\{ \frac{1}{2} w_t^\top \Sigma_t w_t - w_t^\top (\Sigma_t b_t + \gamma \mu_t) \right\}$$

Which can be cast into a quadratic programming problem:

$$g(x) = \frac{1}{2} w_t^\top Q_t w_t - w_t^\top R_t$$

Where w_t are the decision variables (portfolio weights), $Q_t = \Sigma_t$ is the covariance matrix of returns, and R_t is respectively equal to $\gamma\mu_t$, $\Sigma_t b_t$ and $\Sigma_t b_t + \gamma\mu_t$ as respectively the mean-variance, tracking error, and benchmark optimization problems. We use the second, tracking error minimization, for our objective function.

Perrin & Roncalli (2013) furthermore show us that we can impose a turnover constraint on a quadratic optimization:

If we note $\bar{x}$ as the current portfolio and x as the new portfolio, the turnover of Portfolio x with respect to portfolio $\bar{x}$ is the sum of purchases and sales:

$$\tau\,(w \mid \bar{w}) = \sum_{i=1}^{n} (w_i - \bar{w}_i)^+ + \sum_{i=1}^{n} (\bar{w}_i - w_i)^+ = \sum_{i=1}^{n} |w_i - \bar{w}_i|$$

Adding a turnover constraint in a long only MVO quadratic program then leads to the following problem:

$$w^* = \arg\min_{w} \frac{1}{2} w^\top \Sigma w - \gamma w^\top \mu$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^{n} w_i = 1 \\ \sum_{i=1}^{n} |w_i - \bar{w}_i| \le \tau^+ \\ 0 \le w_i \le 1 \end{cases}$$

Where τ^+ is the maximum turnover with respect to the current portfolio $\bar{w}$. Scherer (2007) introduces additional variables w_i^+ and w_i^- such that:

$$w_i = \bar{w}_i + w_i^+ - w_i^-$$

With $w_i^- \ge 0$ that indicates a negative weight change with respect to the initial weight $\bar{w}_i$ and $w_i^+ \ge 0$ indicates a positive weight change. The expression of the turnover becomes:

$$\sum_{i=1}^{n} |w_i - \bar{w}_i| = \sum_{i=1}^{n} |\, w_i^+ - w_i^- \,| = \sum_{i=1}^{n} w_i^+ + \sum_{i=1}^{n} w_i^-$$

Because one of the variables w_i^+ and w_i^- is necessarily equal to zero due to the minimization problem. The problem now becomes:

$$w^* = \arg\min_w \frac{1}{2} w^\top \Sigma w - \gamma w^\top \mu$$

$$\text{s.t.} \quad \begin{cases} \sum_{i=1}^n w_i = 1 \\ w_i = \bar{w}_i + w_i^+ - w_i^- \\ \sum_{i=1}^n w_i^+ + \sum_{i=1}^n w_i^- \le \tau^+ \\ 0 \le w_i, w_i^-, w_i^+ \le 1 \end{cases}$$

4.3 Optimization Design

Combining the derivations in the previous section on tracking error minimization and imposing a turnover constraint, together with our own constraints on sector and region exposure, using our own notation, we get:

$$w_{opt} = arg \min_{w_t} \frac{1}{2} w_t{}^\top \Sigma\, w_t - w_t{}^\top \Sigma b_t$$

s.t.

$$\sum_i w_{t,\mathrm{i}} = 1 \tag{1}$$

$$\boldsymbol{w_{opt}}\, \boldsymbol{w_{screened}} = \boldsymbol{0} \tag{2}$$

$$w_{min} \le w_{t,i} \le w_{max} \tag{3}$$

$$w_{t,i} = w_{t-1,i} + \Delta w_{t,i}{}^+ - \Delta w_{t,i}{}^- \tag{4}$$

$$\sum_i \left(\Delta w_{t,i}{}^+ + \Delta w_{t,i}{}^- \right) \le T \tag{5}$$

$$0 \le w_{t,i}{}^+, w_{t,i}{}^- \le w_{max} \tag{6}$$

$$Score\ (w_i) \ge ESG_{min} \tag{7}$$

$$\widetilde{w}_{t,s}^S - \theta^S \le \sum_i w_{t,i} 1_{i \in S} \le \widetilde{w}_t^S + \theta^S \qquad \textit{for sector } s = 1, \ldots, S \tag{8}$$

$$\widetilde{w}_t^R - \theta^R \le \sum_i w_{t,i} 1_{i \in R} \le \widetilde{w}_t^R + \theta^R \qquad \textit{for region } r = 1, \ldots, R \tag{9}$$

Where w_{opt} is the vector of optimized weights. These weights are the initial weights for month t+1, calculated using the weights and restrictions from month t (end of month), which is the previous month. These weights are depicted as w_t, which is the weight vector consisting of securities weights $w_{t,\ i}$ where i = 1,2,.... represent the individual securities. The reason for this is to prevent forward-looking bias. In practice, one wants to execute the optimization at the start of a new month (initial weights) to form an optimized portfolio out of the screened portfolio of the previous month. The weights and region and sector exposures can only be known from the previous month. Therefore, all data from the end of month t is used to calculate the 'new' and optimized portfolio initial weights.

$\sum$ is the covariance matrix of historic returns. The covariance matrix is calculated with a 10-year moving return window. The reason for this is that we wanted the returns sample size to be as large as possible. Since we start optimization on January 2013 (t=0) and our return data traces back to January 2003, 10 years is the maximum size we can obtain. Furthermore; $\boldsymbol{w_{screened}}$ *(2)* is a vector consisting of 1's on the positions in the matrix corresponding to the weights that are excluded i.e. are *not* in the portfolio and zeros on the positions of the weights that *are* excluded in the portfolio. The constraint that multiplying this vector elementwise with the optimized weights should result in a vector consisting of only 0's ensures that the optimization only chooses the constituents that are really in the screened portfolio. Furthermore, w_{min} *(3)* is the minimum weight assigned to the constituents, which is set at an amount of 5% divided by the number of constituents that the screened portfolio has. In other words, 5% of the 'room' is reserved for the screened portfolio, and 95% is reserved for the optimized portfolio. This is done so that constituents that should be in the index according to the screening, but that are of little use for the objective function, are given tiny amounts of weight. The maximum weight w_{max} at *(3)* and *(6)* is set at 5% for each constituent, to prevent the overweighting of high-scoring ESG securities.

Then, T is the turnover constraint for each month *(5)*. The optimization is programmed in such a way that by default, the turnover constraint for the first month (transition from the MSCI benchmark to the optimized screened portfolio) is set as high as is necessary, and the subsequent months the turnover restriction is set at 0.5% per month. When the solution is infeasible due to this strict restriction, the turnover constraint is upped with 0.5% until the solution becomes feasible.

The ESG_{min} *(7)* is the minimum aggregate ESG score that the portfolio has to obtain and is a simple linear constraint. The value of the constraint is set equal to the ESG score to that of the basic screened portfolio on which the optimization is based. Finally, $\widetilde{w}_{s,\,t}^{S}$ *(8)* and $\widetilde{w}_{t}^{R}$ *(9)* are respectively the sector weight and the region weight of the MSCI ACWI benchmark, which are bounded by respectively 1% and 5% in the optimization.

4.4 Practical Implementation

4.4.1 Estimating the Covariance Matrix

Calculating the ex-ante tracking error requires the calculation of the forecast covariance matrix of asset returns. The covariance is a measure of directional relationship, in this case, between two asset returns. Covariance is therefore a valuable tool for selecting constituents of a portfolio that complement each other in price movement, which in turn can help reducing risk and increase return.

A dataset of returns of 100 constituents and n amount of observations (months, weeks, days) of those returns, results in a covariance matrix of 100x100, whereby each entry represents the covariance of that combination of two constituents. The diagonal axis represents the variance of each constituent.

It is at the amount of constituents that the first problem arises in calculating the covariance matrix. In order to calculate a robust covariance matrix, one needs [N(N-1)/2] correlation parameters, which means a sample size of 5000 observations given the return set of a 100-stock portfolio. This translates to roughly 20 years of daily frequency data. Besides that we do not possess daily frequency data but only monthly-frequency data, the average amount of constituents for the MSCI ACWI is around 2642, making it virtually impossible to construct a robust estimation of the covariance matrix.

An effective way to reduce the amount of correlation parameters needed to obtain the covariance matrix is the Constant Correlation Model (Elton & Gruber, 1973). The assumption with the CCM is that all correlation parameters are the same. This means that rather than needing the correlation between two constituents in order to calculate the covariance, one uses the same correlation for the covariance calculations of all stock combinations. Although reducing the sample risk, the model leads to a higher model risk.

This means that each element in the covariance matrix i.e. the estimated covariance between stock *i* and j ($\hat{\sigma}_{i,j}^{CC}$) under the constant correlation model is calculated as follows:

$$\hat{\sigma}_{i,j}^{CC} = \hat{\sigma}_i \hat{\sigma}_j \hat{\rho}$$

where $\hat{\rho}$ is the best estimate for the constant correlation parameter:

$$\hat{\rho} = \frac{1}{N(N-1)} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \hat{\rho}_{i,j}$$

The tradeoff made here is that we are either able to estimate a single parameter accurately (a single correlation parameter) or estimate *all* correlation parameters poorly (as we lack the frequency of the data necessary to do so). However, Elton and Gruber (1973) show that using the constant correlation-based estimate for the covariance matrix out-of-sample, resulted in more meaningful minimum-variance portfolios in favor of using the regular sample estimate.

Another methodology for estimating the covariance between stock returns includes the use of factor models. This method reduces the amount of parameters needed to the amount of securities times the amount of factors: per security one needs the beta with respect to each factor. These are still a lot of parameters but much less than the regular sample estimate previously described. The variance of a given stock can consequently be obtained as a function of the stock's beta with respect to the factors, and also as a function of the variance of the factors. The covariance between two stocks can be obtained in the same way, while assuming that the covariance between the specific returns of the stocks (the part not explained by the factors) is equal to zero.

Factor models that can be used to explain the asset returns are *explicit* factor models on the one hand, divided into macro factors or micro factors (country, industry, size), and *implicit* factor models on the other hand. The latter makes use of statistical factors through for instance principal component analysis. The drawback from this model is that whatever part of an assets return is not explained by the factor model, is assumed to be fully stock-specific and therefore uncorrelated with other stocks. Using only one or two factors could therefore induce a lot of model risk.

So while estimating the covariance matrix using a factor model is less ad-hoc than the constant correlation method, the former requires determining coefficients for multiple factors for all constituents. Under the pretext of practicality, time, and the insufficient amount of data to determine enough explicit factors (and the implicit factor analysis being complex and time-consuming) to make the model viable, we use the easier implementable constant correlation model. We earlier withdrew from using factor exposures in our optimization design for this same reason.

4.4.2 Model Robustness

Because we achieve high model risk by shrinking the covariance matrix using the CCM, we should first run and compare the optimization with just the minimization of the tracking error, and the basic constraints such as no short-selling, the minimum *(3)* and maximum *(6)* weight constraint, as well of course the constraint that the optimization for month t is not allowed to use any other securities that are not in the screened index at month t *(2)*. We performed this optimization on the 25%, 50% and 75% portfolios. The results are depicted Table 7 on the next page.

The columns start again with the performance measures for the MSCI ACWI and the Pure index, similar to section 3.5. The other six columns are the portfolio's where respectively 25%, 50%, and 75% of constituents is screened out, and their subsequent optimized counterparts. The first item that draws our attention is that for the optimized portfolios, the ex-post tracking error differs significantly from the ex-ante tracking error. The ex-ante tracking error is the minimized target function, using the estimated covariance matrix. Although the ex-ante tracking error of the optimized portfolios is lower than their screened originals, indicating that the optimization works, we can clearly see the negative implications of the model risk taken through estimating the covariance matrix using the CCM. A second reason for the distortion between ex-ante and ex-post tracking error is that the ex-ante tracking error is calculated for 1 month only, while the ex-post tracking error is calculated using a 3-year rolling window. However, the ex-post and ex-ante distortion for the screened portfolios (where no optimization has taken place) should then also be large, but we observe the opposite. The small difference that is still there is either attributable to the 1-month ex-ante vs. 3-year ex-post calculation method, or because of the following: the ex-ante tracking error, just as in the optimization design, is calculated for the *initial weights* of month t, and the closing weights of the benchmark for month t-1. The ex-post tracking errors are calculated using the *closing weights* of month t, which are drifted because of the return that the securities made in month t.

Further examining the results show us that not restricting the turnover leads to intolerable turnover. Where the turnover of the MSCI ACWI is only 10,87% annually, the 50% screened portfolio bears a turnover of 59%, which is manageable, but its optimized counterpart bears a turnover of 444,38%, roughly meaning that annually, the entire holding of stocks within the portfolio is traded (bought or sold) four times. Finally, we see that the aggregate ESG scores of the screened and optimized portfolios remain roughly the same, while the regional and sectoral exposures, which are still unconstrained, even tilt more than the screened counterparts.

	MSCI ACWI	Pure	Excl 25%	Opt 25%	Excl 50%	Opt 50%	Excl 75%	Opt 75%
Average Constituents #	2642	2477	1857	1857	1238	1238	619	619
Average Marketcap %	100	95.58	80.12	80.12	58.5	58.5	31.15	31.15
Annual. Return %	11.11	11.42	11.68	11.82	11.93	10.2	12.54	10.16
Annual. Volatility %	13.13	13.16	13.17	13.04	13.04	12.42	13.13	12.61
Annual. Sharpe Ratio	0.29	0.31	0.33	0.34	0.35	0.24	0.39	0.23
Annual. Tracking Error (Ex-Post) %	-	0.36	0.63	2.35	1.22	3.71	2.2	3.6
Annual. Tracking Error (Ex-Ante) %	-	-	0.67	0.58	1.2	0.95	2.07	1.37
Annual. Information Ratio	-	0.23	0.24	0.08	0.17	-0.07	0.17	-0.08
Annual. Turnover %	10.87	17.5	39.06	302.26	59.04	444.38	97.43	356.66
Max Drawdown %	-21.25	-21.25	-20.75	-18.85	-19.82	-18.63	-19	-20.22
Average Portfolio IAA Score	5.28	5.52	6.2	6.27	7.01	7.05	8.09	8.12
Regional Exposures %								
US	56.61	57.81	55.44	50.66	51.3	43.13	47.85	34.84
EMEA	20.77	20.62	23.79	25.15	28.59	29.44	34.14	36.06
Asia-Pacific	11.29	11.63	12.47	12.47	13	14.56	12.64	17.77
Emerging Markets	11.34	9.94	8.31	11.73	7.11	12.87	5.38	11.33
Sector Exposures %								
Energy	6.57	6.38	6.16	5.48	4.75	4.04	3.17	3.7
Industrials	10.38	10.61	11.37	12.85	12.68	14.86	14.06	16.6
Consumer Discretionary	11.98	12.02	11.21	11.35	10.38	11.37	9.17	10.54
Health Care	11.45	11.89	11.11	12	10.93	11.56	7.37	8.65
Financials	18.41	17.84	17.38	16.63	16.05	15.12	17.22	14.99
Information Technology	16.52	16.84	18.14	16.94	20.31	16.26	23.12	13.23
Materials	5.25	5.34	5.32	5.44	5.53	5.79	5.21	6.36
Consumer Staples	8.99	9.11	9.42	9.35	10.02	10.24	11.06	12.29
Communication Services	5.5	5.3	5.14	4.76	4.34	4.28	4.16	4.81
Utilities	3.18	3.27	3.46	3.83	3.8	5.05	4.44	7.02
Real Estate	1.77	1.4	1.28	1.35	1.2	1.42	1.02	1.8

Table 7: Performance of screened portfolios against their counterparts, optimized only for tracking error, without constraints.

The conclusion from this model validation is simple. Although we find that our estimation of the covariance matrix is not sufficient, we see that the model does its job in minimizing the ex-ante tracking error as compared to the plain screened portfolios. Furthermore, our view in that the turnover and regional and sectoral exposures are important to constrain while optimizing for tracking error is substantiated. The next chapter discusses the results of the complete model.

4.5 Conclusion

The goal of this chapter was to find an optimization method that we could deploy on the screened portfolios, in order to maintain their performance relative to the unscreened benchmark while not changing the securities amount in the screened portfolio. All this while keeping the aggregated ESG score that was increased through systematically screening out the securities with the lowest ESG rating each time.

Achieving this goal would answer our third and part of the fourth research question. The third research question, *what methods can be used to optimize our screening target while accounting for the performance measures*, was answered by first summarizing our needs, wishes and observations drawn from the screening results in section 3.5. together with practical essentials from MN's consideration. Second, we compared these wishes with already familiar solutions to similar problems that we distilled from previous academic research in section 2.2. Although we did not consider multiple methods of optimization - as the research question suggests -, we did find a suitable method by combining two earlier academic works and tailoring and improving them to our own problem context. The result is a quadratic optimization method that minimizes the ex-ante tracking error (which is a quadratic function) of a screened portfolio on a monthly basis, while controlling for unwanted exposures in regions, sectors, and turnover. In the final sections, we tested the model for robustness and partly verified that we can reweigh the screened portfolios to at least minimize the ex-ante tracking error as compared to the screened portfolios. In the next chapter, the additional constraints will be added to the optimization and its results will be presented, so that we can answer the rest of research question 4: *Can we reweigh the screened portfolios in such a way that we maintain performance while improving the aggregated ESG score through exclusion?*

5 Results

In this chapter, we will discuss the results from the optimization that was presented in section 4.3. In the conclusion of the previous chapter, we already found that our optimization is capable of minimizing the ex-ante tracking error as compared to the screened portfolios, but that there is a large divergence between ex-ante and ex-post tracking error. Here, we will provide the results of optimization while controlling for the full set of constraints, thereby fully answering our fourth research question, which is whether we can reweigh the screened portfolios so that we maintain performance. This will be done by presenting the same table structure as we did in paragraph 4.4.2 on the model robustness and compare the results of unconstrained optimization versus constrained optimization.

Consequently, we will focus on our fifth and final research question:

5. *At what stage of exclusion can diversification or proper performance no longer be maintained, for the screened portfolios and the resulting optimized portfolios?*

We will answer this research question by overlaying the performance results of the optimized portfolios per screening level on top of the performance results of the screened portfolios. The latter we already discussed in section 3.5, but overlaying them with the optimization performance will provide us a clear view of whether optimization helps, for what measures it helps and at what stage of exclusion it helps. We will assess this on a total performance level, i.e. average performance over the measured sample, as well as performance through time. Section 1 of this chapter illustrates the former, while section 2 describes the latter. Subsequently, we will form a conclusion where we try to answer research question 5.

When we mention the 'optimized' portfolios, note that these portfolios are also screened, and *subsequently* optimized. Thus when we compare screened and optimized portfolios of exclusion level x%, both portfolios hold exactly the same securities, only the one has been optimized according to our optimization program, and the other is 'only' screened.

5.1 Overall Performance

We start by examining the performance measures for the 25%, 50% and 75% screened and optimized portfolios, so we can assess the difference between optimizing for only tracking error (paragraph 4.4.2), and optimizing for tracking error while adding additional constraints. The table on the next page is the same as Table 7 in the previous chapter, except that the 'Opt' columns are now the full optimizations that minimize tracking error, while constraining turnover, regional and sectoral exposures. The first remarkable item here, is that now the ex-ante and ex-post tracking error bear a much smaller difference than the results in the previous chapter on model validation. At the same time, only the 75% optimization really sees an improvement in both ex-ante and ex-post tracking error. While the ex-ante tracking error on all three levels improved under the optimization without constraints, it worsens with constraints.

	MSCI ACWI	Pure	Excl 25%	Opt 25%	Excl 50%	Opt 50%	Excl 75%	Opt 75%
Average Constituents #	2642	2477	1857	1857	1238	1238	619	619
Average Marketcap %	100	95.58	80.12	80.12	58.5	58.5	31.15	31.15
Annual. Return %	11.11	11.42	11.68	11.55	11.93	11.47	12.54	11.39
Annual. Volatility %	13.13	13.16	13.17	13.06	13.04	12.89	13.13	12.9
Annual. Sharpe Ratio	0.29	0.31	0.33	0.32	0.35	0.32	0.39	0.31
Annual. Tracking Error (Ex-Post) %	-	0.36	0.63	1.07	1.22	1.41	2.2	1.99
Annual. Tracking Error (Ex-Ante) %	-	-	0.67	0.85	1.2	1.25	2.07	1.79
Annual. Information Ratio	-	0.23	0.24	0.11	0.17	0.06	0.17	0.03
Annual. Turnover %	10.87	17.5	39.06	30.44	59.04	49.73	97.43	75.42
Max Drawdown %	-21.25	-21.25	-20.75	-20.86	-19.82	-20.7	-19	-21.03
Average Portfolio IAA Score	5.28	5.52	6.2	6.4	7.01	7.17	8.09	8.2
Regional Exposures %								
US	56.61	57.81	55.44	53.31	51.3	52.55	47.85	52.16
EMEA	20.77	20.62	23.79	24.02	28.59	24.62	34.14	24.95
Asia-Pacific	11.29	11.63	12.47	14.02	13	13.8	12.64	14.76
Emerging Markets	11.34	9.94	8.31	8.64	7.11	9.03	5.38	8.13
Sector Exposures %								
Energy	6.57	6.38	6.16	6.31	4.75	5.94	3.17	5.79
Industrials	10.38	10.61	11.37	11.03	12.68	10.88	14.06	11.12
Consumer Discretionary	11.98	12.02	11.21	11.46	10.38	11.52	9.17	11.42
Health Care	11.45	11.89	11.11	10.82	10.93	11.64	7.37	11.06
Financials	18.41	17.84	17.38	18.38	16.05	17.93	17.22	18.14
Information Technology	16.52	16.84	18.14	16.6	20.31	16.51	23.12	16.47
Materials	5.25	5.34	5.32	5.48	5.53	5.53	5.21	5.12
Consumer Staples	8.99	9.11	9.42	9.41	10.02	9.36	11.06	9.79
Communication Services	5.5	5.3	5.14	5.37	4.34	4.91	4.16	4.91
Utilities	3.18	3.27	3.46	3.66	3.8	4	4.44	4.15
Real Estate	1.77	1.4	1.28	1.48	1.2	1.78	1.02	2.03

Table 8: Performance of the 25%, 50%, and 75% screened portfolios and their optimized counterparts, full constrained optimization.

Although turnover increases when optimizing due to the rebalancing that the algorithm imposes, we still manage to keep the turnover of the optimized portfolios lower than the screened portfolios. At the end of this section, Table 9 with the full optimization results is shown. We are primarily interested in the comparison between the optimized portfolios and the screened portfolios, which are illustrated through the graphs below.

Figure 14 shows the graphed evolution of the most important measures: ex-post tracking error and turnover. It can be seen clearly that the ESG scores remain intact, as they should, and that turnover is minimized with the exception of the 95% exclusion portfolio. It should be noted that turnover difference is minimal, but when minimizing tracking error on the 50% exclusion portfolio *without* a turnover constraint, the annualized turnover is as high as 444.38% (see paragraph 4.4.2). Also note that the total annualized turnover includes the turnover for the first month, which is the 'transitional' month from a full MSCI ACWI portfolio towards the screened or optimized portfolios. For this first month, no turnover constrained is used in the optimization to allow the optimized portfolio to structure itself properly. This 1-month turnover is exceedingly high: would this month be excluded from measurements, both the screened and optimized turnover numbers would be around 15% lower. However, because in reality this transitional month also exists, we chose to include it in the performance measurement.

Although tracking error is minimized, we see that prior to the 65% exclusion level, the ex-post tracking error is higher than its corresponding screened portfolio. The ex-ante tracking error follows about the same curve as the ex-post tracking error, as seen in Figure 15, while inverting slightly earlier at the 55% exclusion level.

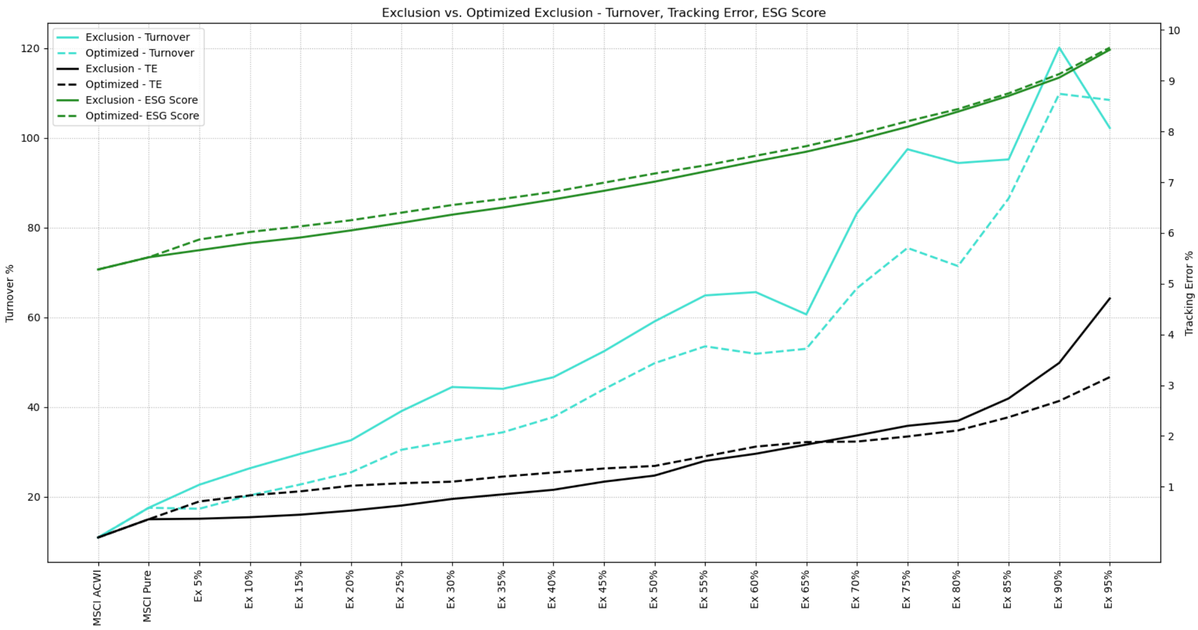


Figure 14: Comparison of the turnover, ex-post tracking error and aggregate ESG score for screened portfolios and their subsequently optimized counterparts, per level of exclusion on the x-axis.

This leads us to conclude that in the early stages of exclusion, the model does not perform its job of minimizing the tracking error *while* other constraints are imposed. As seen in section 4.5, it does its job well (minimizing the ex-ante tracking error compared to the screened portfolios) when no further restrictions are imposed (in terms of ex-ante minimization, not in terms of ex-post). Because the ex-ante and ex-post tracking error do not differ a large amount in the full optimization, we can also argue that the lousy performance of the optimized portfolios relative to the screened portfolios is at least not fully due to the sub optimally estimated covariance matrix. When evaluating the difference between the ex-ante and ex-post tracking error for the optimized portfolios over time, we see the same.

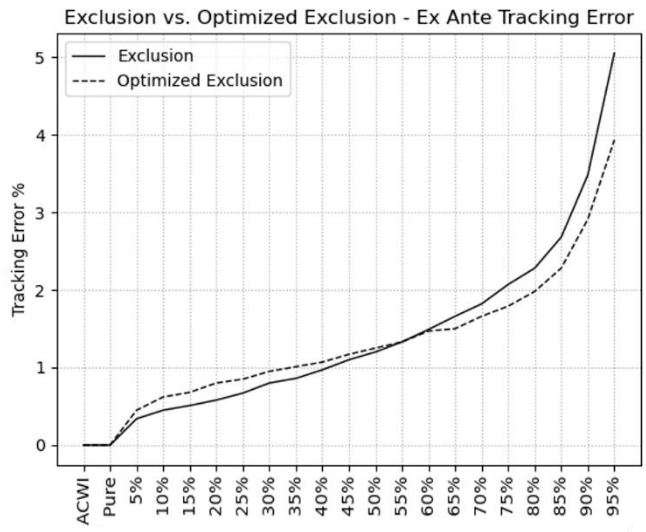


Figure 15: Ex-ante tracking error for each level of exclusion and optimization.

The reason for the poor minimization of the tracking error can be two-fold. First, the turnover constraint might be too strict. Consulting the results of Alessandrini & Jondeau (2021) when optimizing ESG score while restricting turnover and tracking error, we see that their turnover reaches a value of 99% annualized with a tracking error of 2.6. But again, we do not know the number of securities that their optimized portfolio has.

A second view could be that screening on ESG scores has delivered a significant outperformance over the sample period due to the regional and especially sectoral tilts, as can be seen from the exploratory benchmark analysis in section 3.2. Restricting these tilts has turned out to be disadvantageous to return performance, so this could also influence the minimization of the tracking error. However, tracking error is the deviation in return from the portfolio as compared to the benchmark, so restricting returns should intuitively lead to a lower tracking error.

We still see a decrease in outperformance, demonstrated by the lower information ratio for the optimized portfolios (Figure 18). Although the information ratio is still positive (except for the 95% optimized portfolio), it is lower than the screened portfolios, while we only see a notable change in tracking error from the 65% exclusion onwards. Also, maximum drawdown (Figure 16) for the optimized portfolios, shows no decrease, opposite from the screened portfolios, and even increases at the later exclusion stages, suggesting a deterioration in performance.

In the figures below, the performance regarding the other performance measures is depicted, where we indeed see that the return deteriorates relative to the screened portfolio, although it stays constant and close to the benchmark return. The same goes for volatility and thus the Sharpe ratio.

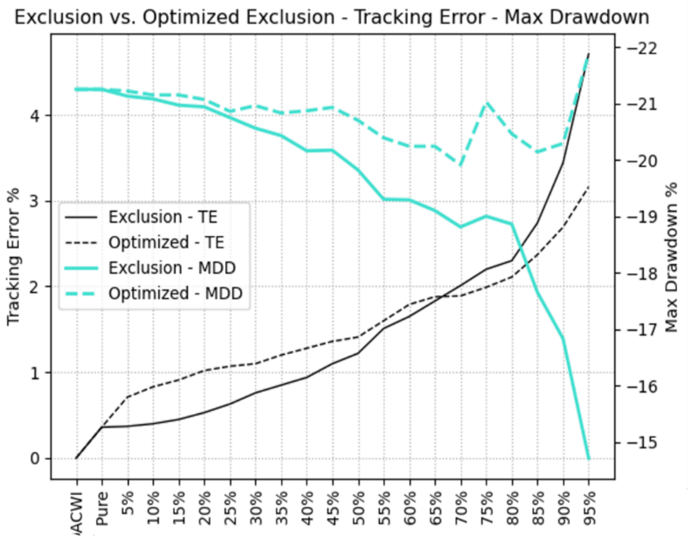


Figure 16: Evolution of the annualized ex-post tracking error and the maximum drawdown for each level of exclusion and subsequent optimization.

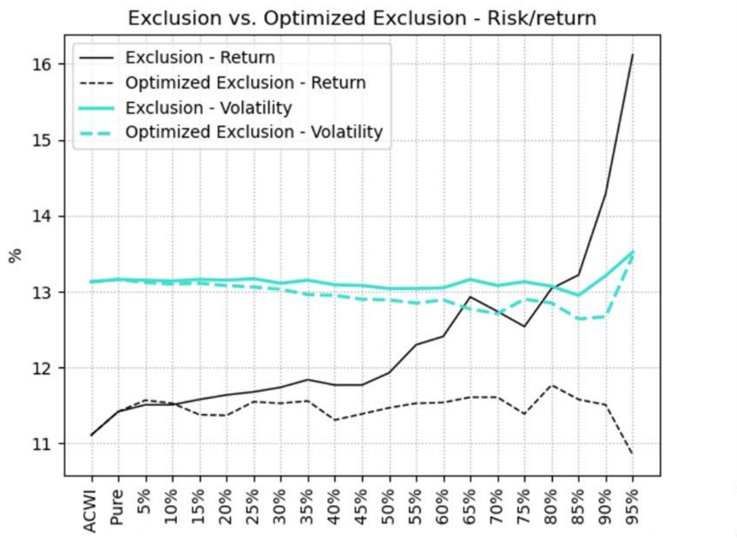


Figure 17: Evolution of the annualized return and volatility for each level of exclusion and subsequent optimization.

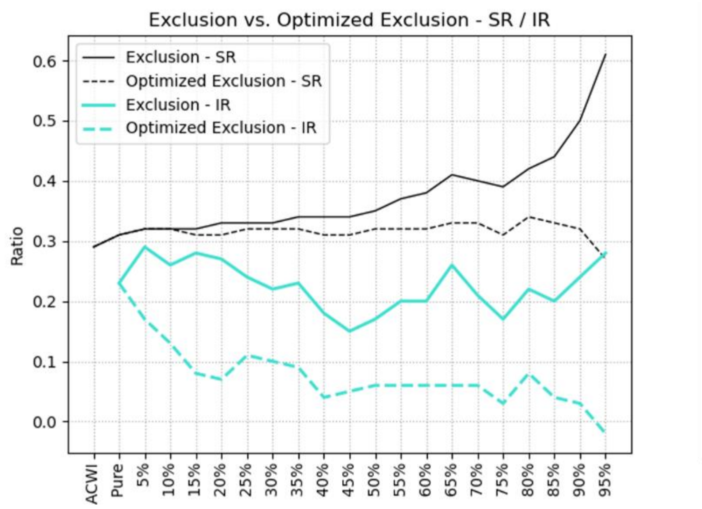


Figure 18: Evolution of the annualized Sharpe and information ratio for each level of exclusion and subsequent optimization.

	MSCI ACWI	Pure	Opt 5%	Opt 10%	Opt 15%	Opt 20%	Opt 25%	Opt 30%	Opt 35%	Opt 40%	Opt 45%	Opt 50%	Opt 55%	Opt 60%	Opt 65%	Opt 70%	Opt 75%	Opt 80%	Opt 85%	Opt 90%	Opt 95%
Average Constituents #	2642	2477	2352	2229	2105	1981	1857	1733	1610	1486	1362	1238	1114	990	867	743	619	495	371	248	124
Average Marketcap %	100	95.58	92.97	90.03	87.34	84.11	80.12	76.02	72.3	67.99	63.26	58.5	53.21	47.85	42.94	37.22	31.15	25.02	19.75	14.04	7.85
Annual. Return %	11.11	11.42	11.57	11.53	11.38	11.37	11.55	11.53	11.56	11.31	11.39	11.47	11.53	11.54	11.61	11.61	11.39	11.77	11.58	11.51	10.85
Annual. Volatility %	13.13	13.16	13.12	13.1	13.11	13.08	13.06	13.03	12.96	12.95	12.9	12.89	12.85	12.89	12.77	12.71	12.9	12.85	12.64	12.67	13.47
Annual. Sharpe Ratio	0.29	0.31	0.32	0.32	0.31	0.31	0.32	0.32	0.32	0.31	0.31	0.32	0.32	0.32	0.33	0.33	0.31	0.34	0.33	0.32	0.27
Annual. Tracking Error (Ex-Post) %	-	0.36	0.71	0.83	0.91	1.02	1.07	1.1	1.2	1.28	1.36	1.41	1.6	1.79	1.88	1.89	1.99	2.11	2.37	2.69	3.16
Annual. Tracking Error (Ex-Ante) %	-	-	0.45	0.61	0.68	0.8	0.85	0.94	1,00	1.06	1.16	1.24	1.32	1.46	1.48	1.65	1.77	1.96	2.26	2.89	3.89
Annual. Information Ratio	-	0.23	0.17	0.13	0.08	0.07	0.11	0.1	0.09	0.04	0.05	0.06	0.06	0.06	0.06	0.06	0.03	0.08	0.04	0.03	-0.02
Annual. Turnover %	10.87	17.5	17.28	20.27	22.7	25.41	30.44	32.42	34.32	37.73	43.93	49.73	53.49	51.83	52.94	66.42	75.42	71.38	86.46	109.78	108.4
Max Drawdown %	-21.25	-21.25	-21.22	-21.15	-21.15	-21.07	-20.86	-20.96	-20.83	-20.87	-20.93	-20.7	-20.39	-20.24	-20.24	-19.91	-21.03	-20.46	-20.14	-20.29	-21.88
Average Portfolio IAA Score	5.28	5.52	5.87	6.02	6.13	6.25	6.4	6.55	6.67	6.81	6.99	7.17	7.33	7.52	7.71	7.94	8.2	8.44	8.75	9.13	9.65
Regional Exposures %																					
US	56.61	57.81	56.6	55.82	55.17	54.3	53.31	53.36	53.15	52.64	52.72	52.55	52.4	52.24	52.37	52.19	52.16	52.03	52.07	52.11	52.19
EMEA	20.77	20.62	22.43	22.95	23.34	23.51	24.02	24.11	24.02	24.67	24.6	24.62	24.94	24.9	25.28	24.97	24.95	25.33	25.58	25.78	25.97
Asia-Pacific	11.29	11.63	11.85	12.14	12.8	13.29	14.02	13.55	13.86	14.09	13.85	13.8	14.04	13.98	14.19	14.23	14.76	14.97	14.74	14.31	14.89
Emerging Markets	11.34	9.94	9.12	9.09	8.68	8.9	8.64	8.98	8.96	8.61	8.84	9.03	8.61	8.89	8.16	8.61	8.13	7.66	7.62	7.8	6.95
Sector Exposures %																					
Energy	6.57	6.38	6.44	6.45	6.54	6.51	6.31	6.22	6.17	6.46	6.18	5.94	5.9	5.94	5.78	5.89	5.79	5.88	5.84	5.94	5.87
Industrials	10.38	10.61	10.9	10.94	11.02	11,00	11.03	10.99	10.76	10.81	10.88	10.88	10.98	11,00	11.09	11.12	11.12	11.1	11.15	11.31	11.29
Consumer Discretionary	11.98	12.02	11.76	11.58	11.45	11.58	11.46	11.53	11.58	11.51	11.51	11.52	11.45	11.43	11.46	11.37	11.42	11.3	11.26	11.27	11.56
Health Care	11.45	11.89	11.43	11.47	10.95	10.81	10.82	10.93	10.99	11.13	11.23	11.64	11.34	11.38	11.01	11.34	11.06	11.21	11.21	11.2	10.99
Financials	18.41	17.84	18.52	18.59	18.75	18.74	18.38	18.3	18.24	18.07	18.01	17.93	17.84	17.85	18.05	18.02	18.14	18.08	18.03	17.79	17.76
Information Technology	16.52	16.84	16.55	16.45	16.5	16.57	16.6	16.63	16.54	16.44	16.38	16.51	16.49	16.63	16.53	16.6	16.47	16.29	16.01	16.16	16.22
Materials	5.25	5.34	5.3	5.28	5.22	5.33	5.48	5.61	5.73	5.67	5.75	5.53	5.51	5.33	5.41	5.16	5.12	5.34	5.56	5.96	6.05
Consumer Staples	8.99	9.11	9.12	9.11	9.38	9.19	9.41	9.13	9.28	9.27	9.5	9.36	9.74	9.68	9.74	9.63	9.79	9.69	9.63	9.17	8.93
Communication Services	5.5	5.3	5.41	5.42	5.36	5.31	5.37	5.3	5.19	5.15	5.02	4.91	4.84	4.69	4.82	4.76	4.91	4.88	5.27	5.31	4.99
Utilities	3.18	3.27	3.22	3.35	3.47	3.48	3.66	3.87	3.9	3.93	3.99	4,00	4.09	4.1	4.13	4.13	4.15	4.17	4.18	4.02	4,00
Real Estate	1.77	1.4	1.34	1.34	1.36	1.48	1.48	1.5	1.61	1.56	1.56	1.78	1.83	1.96	1.99	1.98	2.03	2.06	1.86	1.86	2.33

Table 9: Portfolio performance for each level of optimization.

5.2 Results over Time

This section takes a closer look at some performance measures of interest by analyzing them over time. Although we can already say a lot from the overall results in the previous section, those numbers are still an average, while performance can fluctuate significantly over time depending on market conditions.

To keep things simple and clean, we will only assess the performance of the 50% and 95% exclusion and optimized portfolios. In the previous section, we can clearly see that up until the 50% exclusion mark, no significant differences in performance exist, and that the interesting observations start from there on, especially for the 95% exclusion mark.

We start with the differences in volatility. As we discovered in section 3.5, the volatility of screened portfolios over time tends to be higher than the benchmark during periods of relatively low market volatility, but this inverses when market volatility is high. Taking a look at Figure 19 below, we see that the 50% optimized portfolio bears risk that is closer to the benchmark than its screened counterpart, and also slightly lower than its screened counterpart during times of high volatility. Looking at the 95% optimized portfolio, we see that the volatility over the 2016-2020 period is closer to the benchmark than the screened 95% portfolio, with virtually no difference in times of higher volatility when looking at the 2020-2021 period.

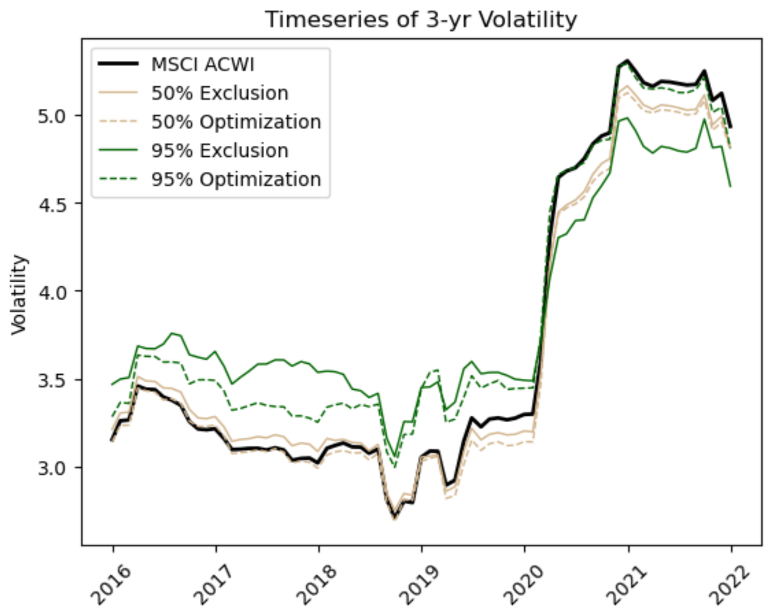


Figure 19: Timeseries of the 3-year volatility comparing the 50% and 95% screened versus optimized portfolios.

The figures below represent both the 3-year tracking error and the 3-year information ratio for the 50% and 95% screened and optimized portfolios. As seen in the overall results, there is almost no difference in tracking error on the 50% exclusion level when optimizing for it, but the effect becomes more significant, the higher the screened tracking error is, as can be seen in the difference between the 95% screened and 95% optimized portfolio. Furthermore, at the 95% level, we see that the optimized portfolio exhibits a more stable tracking error compared to its screened counterpart over time.

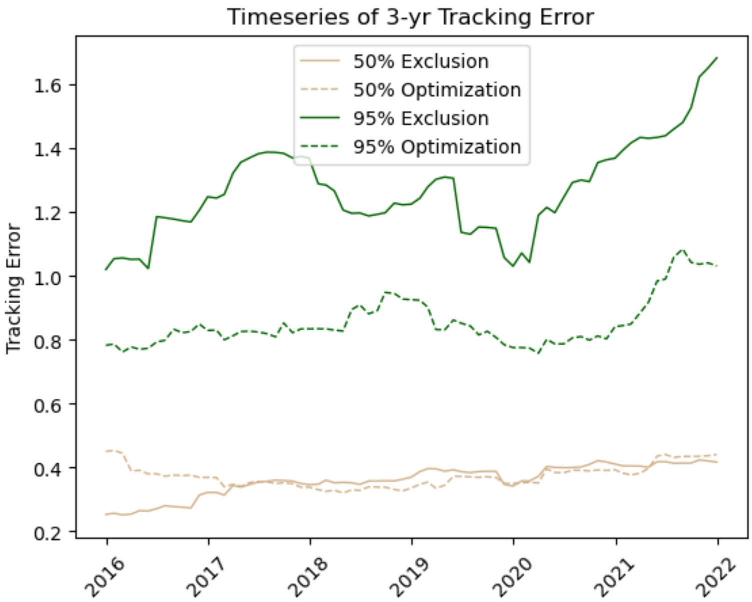


Figure 21: Timeseries of the 3-year ex-post tracking error comparing the 50% and 95% screened versus optimized portfolios.

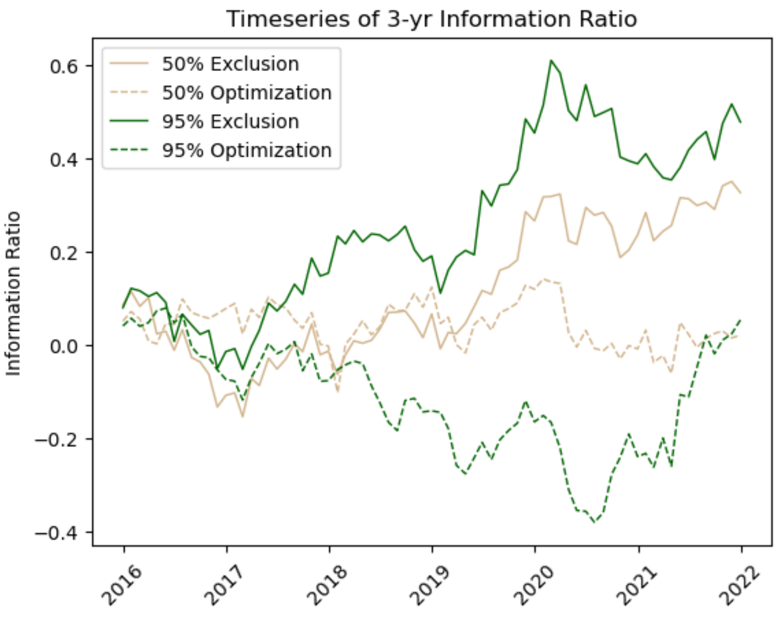


Figure 20: Timeseries of the 3-year information ratio comparing the 50% and 95% screened versus optimized portfolios.

As we saw in the previous section, lower returns and lower tracking error exhibited by the optimized portfolios when compared to the screened portfolios cause a lower information ratio. As Figure 20 shows, the difference in information ratio becomes higher, the more securities are excluded.

Regional and Sector Exposure

The optimization is designed in such a way that besides the minimization of the tracking error, it should control for exposures in region and sector to not exceed 5% and 1% respectively compared to the benchmark exposures in region and sector. Figures 23 below and 24 on the next page showcase the development in exposure over time for the two largest regions, namely US and EMEA. As we saw in section 3.5, one of the large effects of screening is an unwanted exposure towards sectors and regions that exhibit higher ESG scores. The optimization design prevents this so that the optimized portfolios still exhibit high ESG scores but are less exposed to sector and regional shocks than their screened counterparts. Figures 23 and 24 clearly illustrate that the optimization works, and results in much less volatility in region exposures than for instance the 95% screened portfolio.

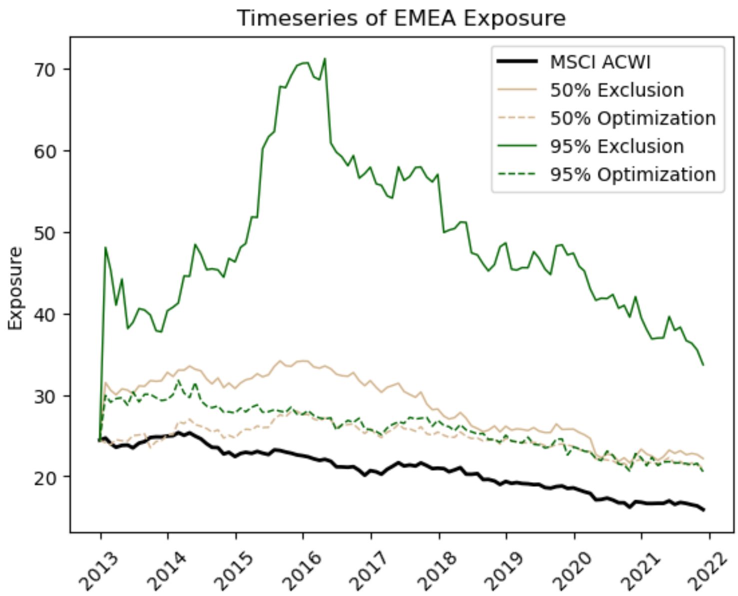


Figure 22: Timeseries of the portfolio exposure to EMEA stocks for the 50% and 95% screened and optimized portfolios.

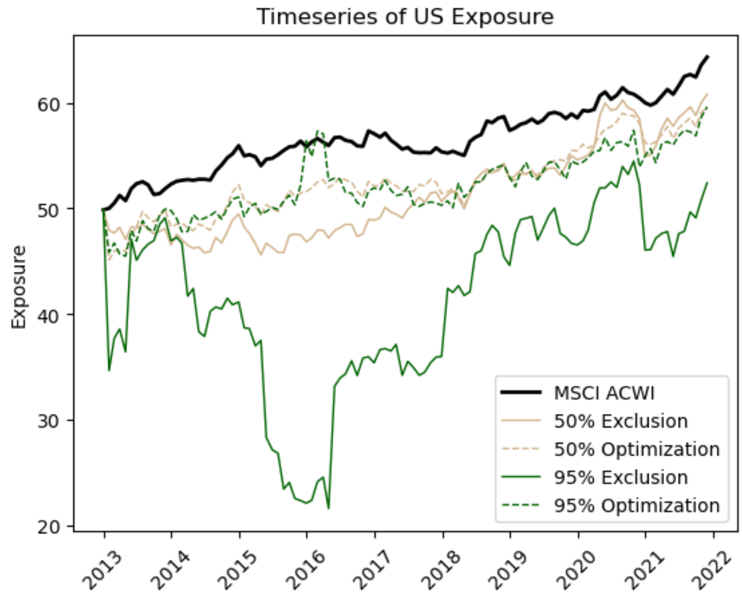


Figure 23: Timeseries of the portfolio exposure to US stocks for the 50% and 95% screened and optimized portfolios.

As far as the sector exposures concerned, we selected the 4 sectors that in the screening results showed the largest divergence from the benchmark exposures, namely Information Technology, Energy, Healthcare and Real Estate. Figures 25-28 show these exposures. As we can see, while the screened portfolios, specifically the 95% screened portfolios, exhibit substantial changes in exposure, the optimized portfolios manage to keep exposure close to the benchmark. The containing of these exposures is what we think primarily causes the degradation in returns as compared to the screened portfolios.

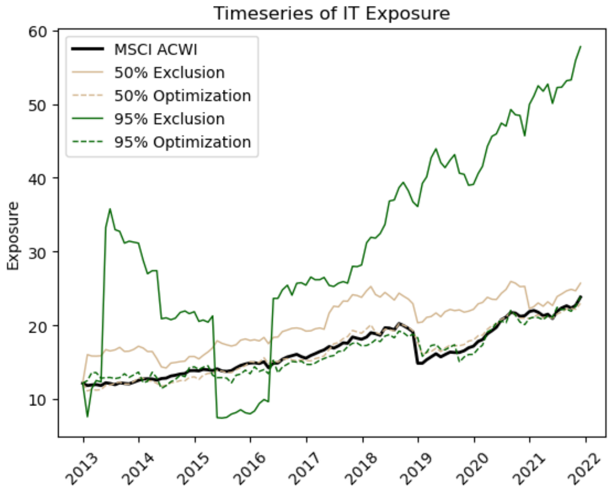


Figure 24

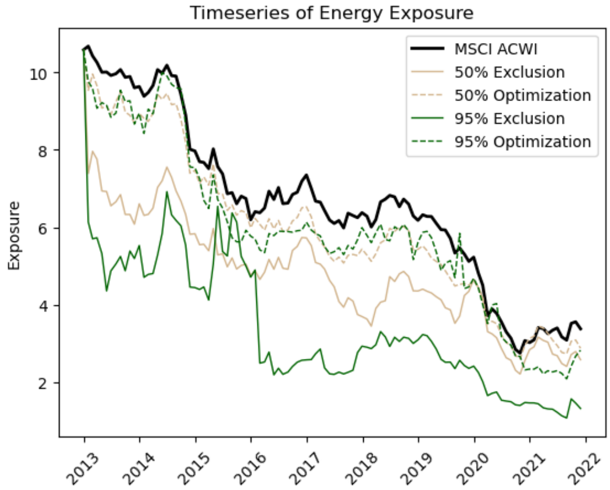


Figure 25

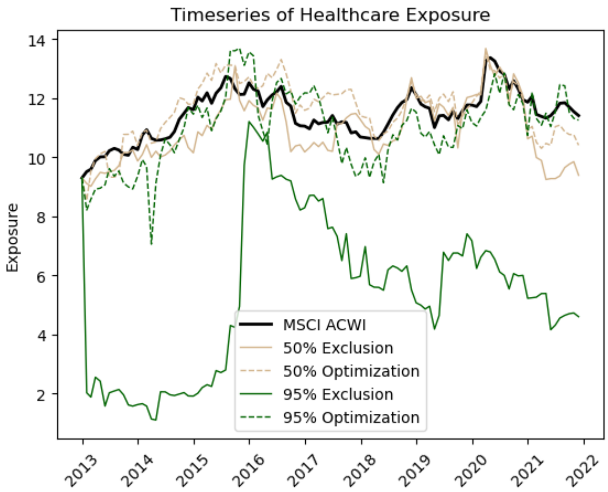


Figure 26

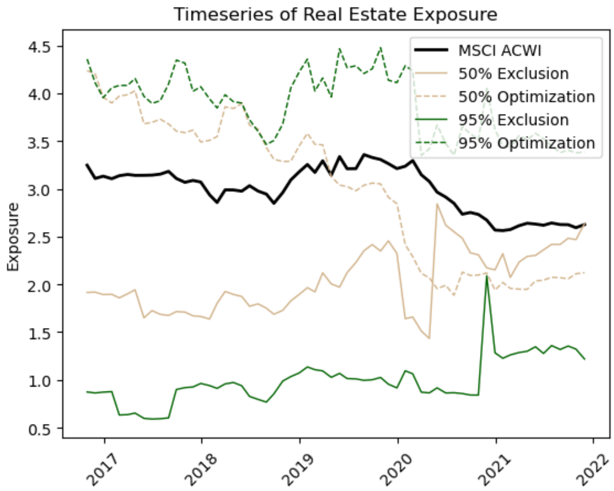


Figure 27

Figures 24-27: Timeseries of the portfolio exposure to 4 sectors exhibiting the largest drift when screened for. For the 50% and 95% screened and optimized portfolios.

5.3 Conclusion

The goal of this chapter was to present the results from the optimization design in comparison with the results of the plain screened portfolios. First, this would allow us to answer our fourth research question, that is *can we reweigh the screened portfolios in such a way that we maintain performance while improving the aggregated ESG score through exclusion?*

Through the model robustness check in paragraph 4.4.2 -where we ran the model without its constraints- and a subsequent comparison with the model running including the linear constraints, we were able to conclude that the model works and also keeps constraints in line over time (see 5.2).

Having compared the other results and comparisons throughout sections 5.1 and 5.2, both in overall performance as well as performance over time, we should now be able to answer our final research question, which is *at what stage of exclusion can diversification or proper performance no longer be maintained, for the screened portfolios and the resulting optimized portfolios?*

The answer to this question i.e. the 'point of no return', is subjective to one's preference towards risk and performance, as well as the exact definition of diversification. However, we have tried to provide a framework for each individual to answer it, by plotting each performance measure against 20 levels of exclusion. This way, an investor can decide for himself what the 'point of no return' is, i.e. at which a certain performance measure spirals out of control, and to what extent this divergence is worth the higher ESG profile.

When we give it our own thoughts, we refer to the performance measures that we deem important: the relative measures of divergence against the benchmark, such as the tracking error and the information ratio, as well as sector and regional exposures. Looking back to section 3.5, we see steady increase or decrease along these measures, up until the 80% exclusion mark, where the differences become significantly larger. This for us is the point where we think that the changes in performance and exposure do not anymore stand up to the benefits of a higher ESG profile.

When we run the optimization across the different screened portfolios, the sector and regional exposures obviously stay within the predetermined ranges of relatively 1 and 5% to the benchmark. We see that tracking error is not minimized by our model up until the 65% exclusion mark, and only shows significant improvement from the 80% mark onwards. However, the 80% exclusion mark is also where the information ratio and the maximum drawdown show underperformance as related to the excluded portfolios, which makes us believe that the restrictions on tracking error and region and sector exposure hinder the performance, and thus that the excess performance in the later stages of exclusion can be attributed largely to tilted regional and even more to sector exposures, as we concluded earlier based on the exploratory benchmark analysis in section 3.2.

We do note that, like the effects of screening, also optimization effectiveness depends on the timeframe of the data. The data period used (2013-2021) happens to carry positive performance for 7 out of 9 years, with only 2015 and 2018 showing negative returns (2.22% and 9.12% respectively). Would this be the other way around and total annualized performance be negative, thus having the tracking error on the 'wrong side' of returns, our tracking error minimization would have probably resulted in higher risk adjusted performance as compared to the benchmark.

We conclude that screening and optimization go relatively hand-in hand up until the 80% exclusion mark. Before this mark, just screening provides a better performance in terms of absolute returns, whereas optimization lacks the increase in return but still provides returns slightly better than the benchmark, while obviously also controlling regional and sectoral exposures. Also, turnover percentages are somewhat 10%-25% lower than the screened portfolios. Beyond the 80% exclusion mark we see the differences increasing faster. If an investor prefers absolute performance regardless of tracking error or regional and sector exposures that are different from the benchmark, then optimization beyond the 80% mark is more devastating to your performance than before the 80% mark. When an investor wants to keep performance close to the benchmark performance, which was our intention in the first place, optimization does help from 80% onwards, and maybe even from 65% onwards where the screened and optimized tracking error results invert.

Beyond the comparison with screening versus optimization, we can still conclude that the overall performance of the optimization is solid. At a 90% exclusion level, the optimization manages to construct a portfolio with only 14% of the original MSCI ACWI market capitalization (equal to an average of 248 constituents, from 2642 constituents at 100%) while exhibiting an increase in ESG score from 5,28 to 9,13. All while maintaining sector and regional exposure within close range to the MSCI ACWI index, and a tracking error of only 2.9%, which would be 3.44% when not optimized for. Furthermore, maximum drawdown, annualized return, and annualized volatility and thus the Sharpe ratio are either equal or slightly higher than the benchmark.

6 Conclusion

To conclude, this research has investigated whether there is a trade-off between portfolio performance and ESG screening, and to what extent the rebalancing portfolio weights through optimization can help to suppress deviations in performance from the benchmark.

The first section in this chapter summarizes the key findings from each chapter, thereby repeating the answers to the research questions. Subsequently, we will be able to draft a conclusion in regard to the main research question. Section 6.2 will then discuss the credibility of these results and discuss the meaning and relevance along with the limitations of this research. Finally, we provide recommendations for further research on this topic.

6.1 Findings

According to paragraph 1.4.1, following from the problem context, our research goals were first to understand the important drivers behind the broadly diversified equity benchmark and to identify the performance measures that related to risk, return and diversification that are relevant for institutional investors to maintain. Secondly, we aimed to apply a single pre-set screening policy accounting for ESG-scores and analyze the performance measures of the benchmark portfolio and the screened portfolios. Then, we developed a model that reweighs the remaining securities in the screened portfolios, leading to an optimized portfolio for each screened portfolio that accounts for the performance measures being within bounds of the original benchmark portfolio. Although it might seem counter-intuitive to track the performance of a portfolio that is 'less green' (i.e. ACWI) than the screened portfolios, the nature of our problem lies in investigating what can be done exposure- and performance-wise, *given* that we already performed exclusion on this index (see 1.4). The results of the optimization would finally enable us to summarize practical insights in the dynamics of screening and optimization on portfolio performance.

Starting from Chapter 3, we have tried to gain a good perspective on the effects of screening, by plotting various performance measures as found in literature and through practical experience by MN against 20 levels of exclusion. This would allow us to promote an objective viewpoint on where the 'point of no return' for exclusion is, so that the investor can use the results in line with his or her own ESG preference and risk appetite. The reason that we researched this in the first place is because there was no straight answer to be found on this question from theoretical and empirical financial literature. On the one hand, the theory suggests that shrinking an investment set is detrimental to diversification, on the other hand this effect can be compensated through a negative correlation between ESG rating and stock-specific volatility (see 2.1.2). Also, the effect on returns is not singularly described. There is ample evidence on high corporate financial performance for high-scoring ESG assets, but on the other side; companies that are excluded from portfolios have shown to bear significant outperformance in the past due to the increase in expected return by investors demanding compensation for lower ESG profiles. This means that excluding these stocks results in missed opportunities. Moreover, results tend do vary along the performance measures used for assessing portfolio performance, as well as the timeframe of the data used.

Our own results show that, given the timeframe of 2013 – 2021, where the coverage of ESG ratings is high enough to base solid conclusions on (>95% across all regions), exclusion on the most diversified tradeable equity benchmark MSCI ACWI has been beneficial in terms of absolute performance. This means in terms of annualized return, annualized volatility, maximum drawdown, and annualized Sharpe ratio. Looking beyond *absolute* performance however, towards the performance *relative* to the benchmark from which exclusion is performed, reveals a significant increase in tracking error. Although the tracking error seems to be on the positive side of returns, as we also experience a positive and balanced information ratio; the mere fact that there *is* a deviation is detrimental for institutional investors that generally prefer following index performance rather than beating it. Moreover, we see significant tilts and exposures towards sectors and regions that exhibit higher aggregated ESG scores. We believe that the increase in annualized return is mostly attributable to sector shifts, as sectors that bear high ESG scores, such as the IT sector, also showed significantly higher annualized risk-adjusted returns than the benchmark in the past. Finally, portfolio turnover increases as a result of screenings.

The model that we subsequently developed is a combination of two optimization designs found in earlier research. The model minimizes the tracking error of the screened portfolios by reweighing the weights of the portfolio's securities on a monthly basis, while controlling for regional and sectoral exposure, as well as turnover. The minimization of the tracking error originates from the Index Tracking Problem, of which the purpose is analogous to our problem context, namely to replicate the performance of an index (MSCI ACWI) with only a subset of securities (screened portfolios). Tracking error is furthermore an important performance measure for institutional investors, which is the perspective that we have used throughout this research.

The final results show that optimizing the screened portfolios counters the positive absolute performance in one way but helps the relative performance in another way. It all depends on the perspective and risk appetite of the investor. The main tradeoff is therefore between absolute and relative performance. Absolute performance in terms of risk, return, Sharpe ratio and maximum drawdown, and relative performance in terms of tracking error, the information ratio, and regional and sector exposures. There is also a tradeoff in turnover. While the turnover is high (but not unreasonable) when a portfolio is only screened, optimizing without a turnover constraint is merely impossible due to the frequent rebalancing that is allowed and necessary to maintain performance. The turnover constraint consequently implied when optimizing a portfolio leads to restriction that could detriment the optimizations performance.

We ultimately conclude that, given the used time period, screening and optimization go relatively hand-in hand up until the 80% mark of exclusion. Up until this point, there is a tradeoff between absolute performance on the one hand, which is better when evaluating only the screened portfolios, and relative performance on the other hand for optimization. That is because optimization until the 80%-point lacks in terms of absolute returns (still above benchmark but underperforming screened portfolios) but maintains performance and exposure relative to the benchmark i.e. holds well in terms of relative performance. Optimization furthermore reduces the turnover compared to screened portfolios.

Beyond the 80% exclusion mark we see the gap becoming larger. If an investor prefers absolute performance regardless of tracking error or regional and sector exposures that are different from the benchmark, then optimization beyond the 80% mark is more devastating to your performance than before the 80% mark. When an investor wants to keep performance close to the benchmark performance, which was our intention in the first place, optimization does help from 80% onwards, and maybe even from 65% onwards where the screened and optimized tracking error results invert.

Overall, this research shows that even with an optimization design that is not very robust, it is still possible to construct a quality ESG portfolio with a significantly smaller subset of constituents than the original benchmark, while maintaining exposure close to the benchmark and not underperforming in terms of risk and return. At a 90% exclusion level, the optimization manages to construct a portfolio with only 14% of the original market capitalization (248 from 2642 constituents) while increasing the aggregate ESG score from 5.28 to 9.13 out of 10.

Our research therefore contributes in terms of both academic and practical relevance. On the academic side, we extend earlier research on the effects of screening on a portfolio by broadening performance measurement and using more relevant data. On the practical side, we are the first to provide institutional investors with an objective viewpoint on the effects of screening, by using the most widely used equity benchmark index as a starting point, and evaluating both absolute and relative performance measures that were plotted against 20 sequential exclusion levels. Furthermore, we combined best practices and explored to what extent optimized exclusion can assist in tilting the screened portfolio towards one's goal of risk preference, which we think, including all its limitations and feedback, is a solid starting point for further research on this topic.

6.2 Discussion

Ironically, one of the biggest limitations of academic research on the effects of screening that we highlight, also poses a big limitation for our own research. Although the time-period was as extensive as it could be, our research findings are still limited to the relatively small timeframe due to the availability of properly covered ESG ratings. The ESG ratings themselves also provide a major limitation for the reliability of this thesis. Next to MSCI, several ESG data providers exist such as Refinitiv, Sustainalitics/Morningstar, S&P Global, RobecoSAM, and Asset4 to name a few. Each provider employs its own rating methodology. Given the many (sometimes subjective) aspects along which ESG ratings are determined, there exists a lot of uncertainty. For instance, Berg et al. (2022) researched ESG ratings among six major rating providers (including MSCI) and document an average pairwise correlation of only 0.54. Similarly, Brandon et al. (2021) reports an average pairwise correlation of 0.46, using data for S&P500 firms from seven rating providers between 2010 and 2017.

The conclusions on the final research question, that is to what extent diversification is no longer possible, is subjective to the risk appetite and definition of diversification for each investor. As stated in the beginning of this research, we adopted the perspective of an institutional investor, and with that the preference towards maintaining performance relative to a benchmark in favor of improving performance. We also pivoted our optimization design around this perspective, but nevertheless tried to also highlight performance from an improvement-seeking perspective.

Furthermore, comparing the results of optimization versus screening could have been more nuanced. Oftentimes we referred to some performance measures in the optimized portfolios as 'improved' towards the plain screened portfolios. However, the screened portfolios did not have the same linear constraints as we applied to the optimized portfolios, such as maximum weights, and sector and regional exposure. This makes the two not as easily comparable as two similar optimization approaches, but as the goal of this research was not necessarily to compare the two approaches but rather to see how optimization can complement screening, we find the comparison justified.

Regarding the optimization: at first, we have not extensively studied complex portfolio optimization methods. Although related to our research, the topic of portfolio optimization represents a whole other field of financial literature which is too extensive to study within this thesis. Rather, this thesis served more as a foundation or guideline for researchers and institutional investors to choose relevant next-step optimizations that are more complex and integrated towards one's specific risk appetite or investment believe. The optimization is also not fully robust due to the suboptimal covariance matrix. The size of the dataset and properties of a covariance matrix required data with a much higher frequency than we had to our disposal. The solution that we used for estimating the covariance matrix, the Elton Gruber Constant Correlation Model, is one of a few methods that can be used. We do not claim that the model we used suits our purpose best. Although some research has been done to explore other methods, we still selected this straightforward and easy to implement model. Future research should prioritize to replace or improve this limitation by using a more sophisticated method for estimation or by focusing on the frequency of the data as the root problem.

Finally, we did not try a sensitivity analysis on the strictness of constraints. The (adaptive) constraint parameters used for the turnover and sector and regional constraints are somewhat ad-hoc. As mentioned in section 5.1, we believe that one of the reasons of the poor minimization of both the ex-ante and the ex-post tracking error is the strict turnover constraint. A too strict sector constraint (1%) could also be a possible cause. Performing a sensitivity analysis on the difference in results by varying the constraint parameters could have led to a more optimal constraint setting. Furthermore, all sectors and regions have the same constrained range. It would have been more realistic to set a constraint proportional to the benchmark range. For example: the US exposure of MSCI ACWI is 56.61%, whereas APAC exposure is 11.29%. Setting a constraint of 5% above or below the benchmark allows the US to only move 8.8% of its original exposure, whereas APAC exposure can move 44.3% of its original exposure. 5% for US exposure is decent, but 44.3% for APAC exposure is too much. Although we do not see both APAC and EM approach the boundary in the optimization, it is still something to consider.

6.3 Scope & Limitations

In addition to the discussion on the results and methods used in this research (previous section), we had to account for some *upfront* limitations that are inherent to the research context and company problem. Nevertheless, these limitations should serve as a disclaimer on the results, or also as a starting point for future research. Items for future research based on the discussion and results are presented in the next section.

- We have only researched the effects of screening on a single index. Although ACWI has the largest market coverage of all available indices, some conclusions produced by this research may not be applicable to indices with a different market focus or other characteristics.

- The purpose of this research is to specifically quantify the effects of *negative* screening, or systematic securities exclusion. While similar, we will not discuss *positive* screening or other approaches of sustainable investing.

- The research is focused on exclusion based on ESG-scores solely. Other scores, such as CO2 emission or separate E, S, and G scores were not considered.

- As per our methodology, we resorted to *optimized exclusion*. That is, we first shrunk the investment set through screening, and then optimized on the remaining assets. We did not consider shrinking the benchmark universe (screening) through a multi-period optimization model, what we previously referred to as integral optimization (see 2.2).

6.4 Future Research

Based on the discussion and limitations of our work and items that we discovered by studying similar works (that were too extensive to include in this research), we have formed some improvements for future research on this topic.

First and foremost, it would be insightful to rerun this research with different and improved estimations of the covariance matrix (see 4.4.1). The best solution to this dimensionality problem would be to use higher-frequency data, for instance weekly or even daily data. If this is difficult to obtain, one could opt to estimate the covariance matrix using factor models as described in section 4.4.1.

Second, it would be interesting to reproduce the research on the effects of screening and subsequent optimization on various other indices that represent world-and sector-wide equities. Along with varying the index, varying different ESG rating providers could also benefit the robustness of this research field.

Third, some extensions or changes could be made to the optimization program. First of all, it would be interesting to see the optimization design including linear constraints on risk factors. Alessandrini and Jondeau accounted for the Fama-French factors of market, size, and value. Specifically, an extension towards the Fama-French 5 factor model would be interesting, as the factors Profitability and Investment have been shown to explain outperformance of sin-stocks (Blitz & Fabozzi, 2017). Dimson et al. (2020) similarly advocate for the replication of factor exposures of excluded stocks by increasing the weights of the 'remaining' constituents that have exposure to the same factors driving the returns of the screened out 'sin-stocks'. Next, the optimization design and the software that we used led to a computational time of around 30 hours for 20 screened portfolios. Exploring faster optimization methods could allow one to use penalty parameters and shadow values to research the 'degree of constraining' of the various constraints. Also, exploring quarterly, semi-annually, or annual rebalancing and/or screening and optimization could fasten the calculation time. Lowering the frequency of optimization from a monthly timeframe to a longer time period could also naturally result in a lower turnover since there would be less rebalancing and trading. However, this can lead to unwanted deviations in-between constraining/optimization periods. It is therefore up to further research to investigate what the proper balance is between the timely use of new market information and turnover costs and calculation speed.

Finally, it would be interesting to further research the possibilities of derivatives or structured products that hedge the tracking error, for instance through synthetic exposure towards excluded securities or underweighted regions and sectors, without the deterioration of the ESG quality of a portfolio.

7 References

Alessandrini, F., & Jondeau, E. (2020). ESG Investing: From Sin Stocks to Smart Beta. *The Journal of Portfolio Management, 46*(2), 75-94.

Alessandrini, F., & Jondeau, E. (2021). Optimal Strategies for ESG Portfolios. *Journal of Portfolio Management, 47*(6), 114-138.

Baracchini, C., & Addessi, M. E. (2012). Ethical Portfolio Theory: A New Course. *Journal of Management & Sustainability, 2*, 35-42.

Barnett, M., & Salomon, R. (2006). Beyond dichotomy: the curvilinear relationship between social responsibility and financial performance. *Strategic Management Journal, 27*(11), 1101-1122.

Bauer, R., Koedijk, K., & Otten, R. (2005). International evidence on ethical mutual fund performance and investments tyle. *Journal of Banking & Finance, 29*(7), 1751-1767.

Beasley, J., Meade, N., & Chang, T.-J. (2003). An evolutionary heuristic for the index tracking problem. *European Journal of Operational Research, 148*(3), 621-643.

Berg, F., Kölbel, J., & Rigobon, R. (2022). Aggregate Confusion: The Divergence of ESG Ratings. *Review of Finance*.

Blankenberg, A.-K., & Gottschalk, J. (2018). Is Socially Responsible Investing (SRI) in Stocks a Competitive Capital Investment? A Comparitive Analysis Based on the Performance of Sustainable Stocks. *SSRN Working Paper 3186094*.

Blitz, D., & Fabozzi, F. (2017). Sin Stocks Revisited: Resolving the Sin Stock Anomaly. *The Journal of Portfolio Management, 44*(1), 105-111.

Blitz, D., & Swinkels, L. (2020). Is Exclusion Effective? *The Journal of Portfolio Management, 46*(3), 42-48.

Blitz, D., & Swinkels, L. (2021). Does excluding sin stocks cost performance? *Journal of Sustainable Finance & Investment*, 1-18.

Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *The Journal of financial Economics, 142*(2), 517-549.

Branch, B., & Cai, L. (2012). Do Socially Responsible Index Investors Incur an Opportunity Cost? *The Financial Review, 47*, 617-630.

Branch, M., Goldberg, L. R., & Hand, P. (2019). A Guide to ESG Portfolio Construction. *Journal of Portfolio Management, 45*(4), 61-66.

Brandon, R. G., Krueger, P., & Schmidt, P. S. (2021). ESG Rating Disagreement and Stock Returns. *Financial Analysts Journal, 4*, 104-127.

Broccardo, E., Hart, O. D., & Zingales, L. (2021). Exit vs. Voice. *European Corporate Governance Institute - Finance Working Paper No. 694/2020.*

Cai, L., Cooper, R., & He, D. (2022). Socially Responsible Investing and Factor Investing, Is There an Opportunity Cost? *The Journal of Portfolio Management Quantitative Special Issue, 48*(2), 181-197.

Campbell, J., Lettau, M., Malkiel, B., & Xu, Y. (2001). Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *The Journal of Finance, 56*(1), 1-43.

Chava, S. (2014). Environmental Externalities and Cost of Capital. *Management Science, 60*(9), 2223-2247.

Chen, M., & Mussalli, G. (2020). An Integrated Approach to Quantitative ESG Investing. *The Journal of Portfolio Management, 46*(3).

Coqueret, G. (2022). *Perspectives in Sustainable Equity Investing.* CRC Press.

De Spiegeleer, J., Höcht, S., Jakubowski, D., Reyners, S., & Schoutens, W. (2020). ESG: a new dimension in portfolio allocation. *Journal of Sustainable Finance & Investment,* 1-41.

De, I., & Clayman, M. R. (2015). The Benefits of Socially Responsible Investing: An Active Manager's Perspective. *The Journal of Investing, 24*(4), 49-72.

Dimson, E., Marsh, P., & Staunton, M. (2020). Exclusionary Screening. *Journal of Impact and ESG Investing, 1*(1), 66-75.

Drut, B. (2010). Social Responsibility and Mean-Variance Portfolio Selection. *Working Papers CEB, 10.*

Eccles, R. G., Ioannou, I., & Serafeim, G. (2014). The Impact of Corporate Sustainability on organizational Processes and Performance. *Management Science, 60*(11), 2835-57.

El Ghoul, S., Guedhami, O., Kwok, C., & Mishra, D. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking and Finance, 35*(9), 2388-2406.

Elton, E., & Gruber, M. (1973). Estimating the Dependence Structure of Share Prices - Implication for Portfolio Selection. *Journal of Finance, 28,* 1203-1232.

Eurosif. (2018). *European SRI Study 2018.*

Fabozzi, F., Ma, K., & Oliphant, B. (2008). Sin Stock Returns. *The Journal of Portfolio Management, 35*(1), 82-94.

Fish, A., Kim, D., & Venkatraman, S. (2019). The ESG Sacrifice. *SSRN Working Paper, 3488475.*

Fisher, L., & Lorie, J. (1970). Some Studies of Variability of Returns on Investments in Common Stocks. *The Journal of Business, 2,* 99-134.

Friede, G., Busch, T., & Bassen, A. (2015). ESG and Financial Performance: Aggregated Evidence from More Than 2000 Empirical Studies. *Journal of Sustainable Finance and Investment, 5*(4), 210-233.

Gasser, S. M., Rammerstorfer, M., & Weinmayer, K. (2017). Markowitz Revisited: Social Portfolio Engineering. *European Journal of Operations Research, 258*(3), 1181-90.

Geczy, C., & Guerard, J. (2021). ESG and Expected Returns on Equities: The Case of Environmental Ratings. *SSRN Working Paper 3903480.*

Global Sustainable Investment Alliance. (2021). *Global Sustainable Investment Review 2020.*

Haensly, P. (2020). Risk decomposition, estimation error, and naïve diversification. *The North American Journal of Economics and Finance, 52.*

Hoepner, A. G. (2010). Portfolio Diversification and Environmental, Social or Governance Criteria: Must Responsible Investments Really Be Poorly Diversified? *SSRN Working Paper.*

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

Hornuf, L., & Yüksel, G. (2022). The Performance of Socially Responsible Investments: A Meta-Analysis. *CESifo Working Papers, no. 9724.*

Kumar, N., Smith, C., Badis, L., Wang, N., Ambrosy, P., & Tavares, R. (2016). ESG factors and risk-adjusted performance: a new quantitative model. *Journal of Sustainable Finance & Investment, 6*(4), 292-300.

Kurtz, L. (1997). No effect, or no net effect? Studies on socially responsible investing. *Journal of Investing, 6*(4), 37-49.

Le Sourd, V. (2007). *Performance Measurement for Traditional Investment.* EDHEC Risk and Management Research Centre.

Lezmi, E., Roncalli, T., & Xu, J. (2022). Multi-Period Portfolio Optimization. *SSRN Working Paper, 4078043.*

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance, 7*(1), 77-91.

Menchero, J., & Morozov, A. (2011). *The Relative Strenghts of Industry and Country Factors in Global Equity Markets.* MSCI.

Norges Bank Investment Management. (2019). *Country and Industry Effects in Global Equity Returns.*

Pedersen, L., Fitzgibbons, S., & Pomorski, L. (2021). Responsible Investing: The ESG-efficient Frontier. *Journal of Financial Economics, 142*(2), 572-597.

Peiris, D., & Evans, J. (2010). The relationship between environmental social governance factors and U.S. stock performance. *The Journal of Investing, 19*(3), 104-112.

Perrin, S., & Roncalli, T. (2020). Machine Learning Optimization Algorithms & Portfolio Allocation. In E. Jurczenko, *Machine Learning for Asset Management: New Developments and Financial Applications* (pp. 261-328). ISTE & John Wiley & Sons, Inc.

Rachev, S. T., Stoyanov, S., & Fabozzi, F. J. (2007). *Advanced Stochastic Models, Risk Assessment, and Portfolio Optimization: The Ideal Risk, Uncertainty, and Performance Measures.* John Wiley, Finance.

Renneboog, L., Ter Horst, J., & Zhang, C. (2008). Socially Responsible Investments: Institutional Aspects, Performance, and Investor Behavior. *Journal of Banking & Finance, 32*(9), 1723-1742.

Roncalli, T. (2013). *Introduction to Risk Parity and Budgeting.* Chapman & Hall/CRC Financial Mathematics Seriex.

Rudd, A. (1981). Social Responsibility and Portfolio Performance. *California Management Review, 23*(4), 55-61.

Scherer, B. (2007). *Portfolio Construction & Risk Budgeting* (Third edition ed.). Risk Books.

Schmidt, A. B. (2020). Optimal ESG Portfolios: An Example for the Dow Jones Index. *Journal of Sustainable Finance & Investment*, 1-7.

Statman, M. (1987). How Many Stocks Make a Diversified Portfolio? *The Journal of Financial and Quantitative Analysis, 22*(3), 353-363.

Statman, M. (2004). The Diversification Puzzle. *Financial Analysts Journal, 60*(4), 44-53.

Surz, R., & Price, M. (2000). The Truth About Diversification by the Numbers. *The Journal of Investing, 9*(4), 93-95.

Trinks, P., & Scholtens, B. (2017). The Opportunity Cost of Negative Screening in Socially Responsible Investing. *Journal of Business Ethics*, 193-208.

Vardharaj, R., Fabozzi, F., & Jones, F. (2004). Determinants of Tracking Error for Equity Portfolios. *The Journal of Investing, 13*(2), pp. 37-47.

Verheyden, T., Eccles, R., & Feiner, A. (2016). ESG for All? The Impact of ESG Screening on Return, Risk and Diversification. *Journal of Applied Corporate Finance, 28*(2), 47-55.

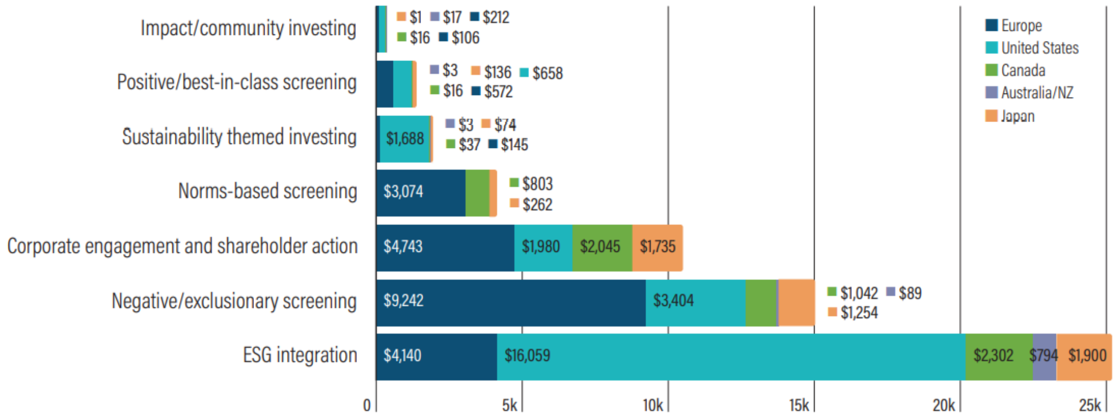
Verheyden, T., Eccles, R., & Feiner, A. (2016). ESG for All? The Impact of ESG Screening on Return, Risk, and Diversification. *Journal of Applied Corporate Finance, 28*(2), 47-55.

Whelan, T., & Atz, U. (2021). *ESG and Financial Performance: Uncovering the Relationship by Aggregating Evidence from 1,000 Plus Studies Published between 2015 - 2020.* NYU Stern Center for Sustainable Business.

Zaimovic, A., Omanovic, A., & Arnaut-Berilo, A. (2021). How Many Stocks Are Sufficient for Equity Portfolio Diversification? A Review of the Literature. *Journal of Risk and Financial Management, 14*(11), 1-30.

Appendix A: Sustainable Investing Assets by Strategy & Region 2020

Source: Global Sustainable Investment Alliance, 2021



NOTE: Asset values are expressed in billions of US dollars. European sustainable investing strategy data is based on extrapolation from historic data from the 2018 GSIR report and applying the same proportion to 2020 sustainable investing data across the different sustainable investing strategies. US SIF data extrapolates from numbers provided by a subset of overall respondents in its 2020 Trends report. US and Australasia did not report on the category of norms-based screening and Australasia on the category positive/best-in-class screening.

Appendix B: Software

For the solving of the quadratic optimization we make use of the MOSEK Optimization Tools version 10 optimizer for Python 3.8.3. MOSEK is a commercial solver that we used through an academic license. It allowed us to solve quadratic problems in matrix-form multiple times faster than standard convex optimization packages.

The solver makes use of a quadratic to conic reformulation in order to solve the problem using an interior point method.

Appendix C: Derivation of the TE minimization

The following sections are quoted from Perrin & Roncalli (2019).

Quadratic Programming: Primal Formulation

A quadratic program (QP) problem is an optimization problem with a quadratic objective function and linear inequality constraints:

$$x^* = arg \min_x \frac{1}{2} x^\top Q x - x^\top R \qquad \textit{Eq. A.1}$$

$$s.t. \quad Sx \leq T$$

where x is a $n \times 1$ vector, Q is a $n \times n$ matrix and R is a $n \times 1$ vector. We note that the system of constraints $Sx \leq T$ allows us to specify linear equality constraints $Ax = B$ or box constraints $x^- \leq x \leq x^+$. Most numerical packages then consider the following formulation:

$$x^* = arg \min_x \frac{1}{2} x^\top Q x - x^\top R \qquad \textit{Eq. A.2}$$

$$s.t. \begin{cases} Ax = B \\ Gx \leq H \quad 22 \\ x^- \leq x \leq x^+ \end{cases}$$

because the problem (*Eq. A.2*) is equivalent to the canonical problem (*Eq. A.1*) with the following system of linear inequalities:

$$\begin{bmatrix} -A \\ A \\ G \\ -I_n \\ I_n \end{bmatrix} x \leq \begin{bmatrix} -B \\ B \\ H \\ -x^- \\ x^+ \end{bmatrix}$$

If the space Ω defined by $Sx \leq T$ is non-empty and if Q is a symmetric positive definite matrix, the solution exists because the function $f(x) = \frac{1}{2} x^\top Q x - x^\top R$ is convex. In the general case where Q is a square matrix, the solution may not exist.

The Markowitz Framework

Consider an investment universe of n assets. Let $x = (x_1, \dots, x_n)$ be the vector of weights in the portfolio. We consider the wealth to be fully invested so that:

$$\sum_{i=1}^{n} x_i = 1_n^\top x = 1$$

We denote $R = (R_1, \dots, R_n)$ as the vector of asset price returns where R_i is the return of asset i. The total weighted portfolio return is then equal to $R(x) = \sum_{i=1}^{n} x_i R_i = x^\top R$.

Let $\mu = \mathrm{E}[R]$ and $\sum = \mathrm{E}[(R-\mu)(R-\mu)^\top$ be the vector of expected returns and the covariance matrix of asset returns. The expected return of the portfolio is equal to:

$$\mu(x) = \mathrm{E}[R(x)] = x^\top \mu$$

Whereas the portfolio variance is equal to:

$$\sigma^2(x) = \mathrm{E}[(R(x)-\mu(x))(R(x)-\mu(x))^\top] = x^\top \sum \mathrm{x}$$

Markowitz (1952) formulated the investor's financial problem as follows:

1. Maximizing the expected return of the portfolio under a volatility constraint (σ-problem):

$$\max \mu(x) \quad s.t. \quad \sigma(x) \leq \sigma^* \qquad \textit{Eq. A.3}$$

2. Or minimizing the volatility of the portfolio under a return constraint (μ-problem):

$$\min \sigma(x) \quad s.t. \quad \mu(x) \geq \mu^* \qquad \textit{Eq. A.4}$$

Markowitz's idea was to consider a quadratic utility function:

$$U(x) = x^\top \mu - \frac{\Phi}{2} x^\top \sum \mathrm{x}$$

where Φ ≥ 0 is the risk aversion. Since maximizing *U(x)* is equivalent to minimizing *-U(x)*, the Markowitz problems (*Eq. A.3*) and (*Eq. A.4*) can be cast into a QP problem:

$$x^*(\gamma) = \arg\min_x \frac{1}{2} x^\top \sum \mathrm{x} - \gamma x^\top \mu \qquad \textit{Eq. A5}$$

$$s.t.\ \mathbf{1}_n^\top x = 1$$

where $\gamma = \Phi^{-1}$. Therefore, solving the μ-problem or the σ-problem is equivalent to finding the optimal value of γ such that $\mu(x^*(\gamma)) = \mu^*$ or $\sigma(x^*(\gamma)) = \sigma^*$. We know that the functions $\mu(x^*(\gamma))$ and $\sigma(x^*(\gamma))$ are increasing with respect to γ and are bounded. The optimal value γ can then be easily computed using the bisection algorithm.

Note that (*Eq. A5*) above corresponds to the QP problem (*Eq. A.2*) where $Q = \sum$, $R = \gamma\mu$, $A = \mathbf{1}_n^\top$ and $B = 1$. Moreover, it is easy to include bounds on the weights, inequalities between asset classes, etc.

Portfolio Optimization with a Benchmark

We now consider a benchmark b. We note $\mu(x \mid b) = (x - b)^\top \mu$ as the expected excess return and $\sigma(x \mid b) = \sqrt{(x-b)^\top \Sigma (x-b)}$ as the tracking error volatility of portfolio x with respect to benchmark b. The objective function corresponds to a trade-off between minimizing the tracking error volatility and maximizing the expected excess return (or the alpha):

$$f(x \mid b) = \frac{1}{2}\sigma^2(x \mid b) - \gamma\mu(x \mid b)$$

In the following box, this problem is rewritten to a working QP problem following Roncalli (2013):

The excess return $\tilde{R}(x \mid b)$ of Portfolio x with respect to benchmark b is the difference between the return of the portfolio and the return of the benchmark:

$$\tilde{R}(x \mid b) = \tilde{R}(x) - \tilde{R}(b) = (x-b)^\top \tilde{R}$$

It is easy to show that the expected excess return is equal to:

$$\mu(x \mid b) = \mathrm{E}[\tilde{R}(x \mid b)] = (x-b)^\top \mu$$

whereas the volatility of the tracking error is given by:

$$\sigma(x \mid b) = \sigma(\tilde{R}(x \mid b)) = \sqrt{(x-b)^\top \Sigma (x-b)}$$

the objective function is then:

$$\begin{aligned} f(x \mid b) &= \frac{1}{2}(x-b)^\top \Sigma (x-b) - \gamma(x-b)^\top \mu \\ &= \tfrac{1}{2}\, x^\top \Sigma \mathrm{x} - x^\top(\gamma\mu + \Sigma b) + \left(\tfrac{1}{2} b^\top \Sigma b + \gamma b^\top \mu\right) \\ &= \tfrac{1}{2}\, x^\top Q \mathrm{x} - x^\top R + C \end{aligned}$$

where C is a constant which does not depend on Portfolio x. We recognize a QP problem where $Q = \Sigma$ and $R = \gamma\mu + \Sigma b$

Appendix D: Benchmark, Screened & Optimized portfolio results

	MSCI ACWI	Energy	Materials	Industrials	Consumer Discretiona	Consumer Staples	Health Care	Financials	Information Technology	Comm. Services	Utilities	Real Estate
Average Constituents #	2642	143	240	395	341	216	192	503	258	119	135	96
Average Marketcap %	100	6.57	5.25	10.38	11.98	8.99	11.45	18.41	16.52	5.5	3.18	1.77
Annual. Return %	11.11	0.41	6.05	10.22	13.05	8.4	13.69	7.75	22.35	8.55	8.36	4.15
Annual. Volatility %	13.13	22.81	17.5	15.03	15.31	11.09	12.55	16.91	15.39	13.11	12.47	11.07
Annual. Sharpe Ratio	0.29	-0.19	0	0.22	0.38	0.11	0.48	0.09	0.89	0.12	0.11	-0.24
Max Drawdown %	-21.25	-53.01	-32.56	-26.21	-21.22	-13.69	-15.02	-33.97	-17.48	-16.11	-18.74	-23.63
Average Portfolio IAA Score	5.28	4.54	5.32	6.04	4.78	5.59	4.95	4.96	6.17	4.98	6.06	2.35

	MSCI ACWI	US	EMEA	APAC	EM	AAA	AA	A	BBB	BB	B	CCC
Average Constituents #	2642	716	450	459	1016	122	316	458	538	490	397	153
Average Marketcap %	100	56.61	20.77	11.29	11.34	7.86	14.84	21.52	22.09	16.33	9.63	3.31
Annual. Return %	11.11	15.49	7.89	4.91	1.92	16.38	11.11	12.29	10.03	10.5	10.52	4.79
Annual. Volatility %	13.13	13.49	15.08	13.16	15.92	13.59	12.69	13.74	13.2	14.06	13.47	16.08
Annual. Sharpe Ratio	0.29	0.57	0.08	-0.13	-0.26	0.62	0.3	0.36	0.22	0.25	0.25	-0.09
Max Drawdown %	-21.25	-20.13	-23.76	-23.41	-31.47	-15.05	-19.1	-20.84	-23.09	-24.44	-21.34	-39.66
Average Portfolio IAA Score	5.28	5.06	6.52	5.71	3.68	9.61	7.95	6.53	5.07	3.6	2.16	0.6

Table D.1: Annualized performance measures per sector, region, and ESG-rating

	MSCI ACWI	Pure	Ex 5%	Ex 10%	Ex 15%	Ex 20%	Ex 25%	Ex 30%	Ex 35%	Ex 40%	Ex 45%	Ex 50%	Ex 55%	Ex 60%	Ex 65%	Ex 70%	Ex 75%	Ex 80%	Ex 85%	Ex 90%	Ex 95%
Average Constituents #	2642	2477	2352	2229	2105	1981	1857	1733	1610	1486	1362	1238	1114	990	867	743	619	495	371	248	124
Average Marketcap %	100	95.58	92.97	90.03	87.34	84.11	80.12	76.02	72.3	67.99	63.26	58.5	53.21	47.85	42.94	37.22	31.15	25.02	19.75	14.04	7.85
Annual. Return %	11.11	11.42	11.51	11.51	11.58	11.64	11.68	11.74	11.84	11.77	11.77	11.93	12.3	12.41	12.93	12.74	12.54	13.04	13.22	14.29	16.12
Annual. Volatility %	13.13	13.16	13.15	13.14	13.16	13.15	13.17	13.11	13.15	13.09	13.08	13.04	13.04	13.05	13.16	13.08	13.13	13.07	12.95	13.21	13.52
Annual. Sharpe Ratio	0.29	0.31	0.32	0.32	0.32	0.33	0.33	0.33	0.34	0.34	0.34	0.35	0.37	0.38	0.41	0.4	0.39	0.42	0.44	0.5	0.61
Annual. Tracking Error (Ex-Post) %	-	0.36	0.37	0.4	0.45	0.53	0.63	0.76	0.85	0.94	1.1	1.22	1.51	1.65	1.83	2.01	2.2	2.3	2.74	3.44	4.71
Annual. Tracking Error (Ex-Ante) %	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Annual. Information Ratio	-	0.23	0.29	0.26	0.28	0.27	0.24	0.22	0.23	0.18	0.15	0.17	0.2	0.2	0.26	0.21	0.17	0.22	0.2	0.24	0.28
Annual. Turnover %	10.87	17.5	22.63	26.31	29.53	32.55	39.06	44.42	44.02	46.6	52.39	59.04	64.84	65.57	60.61	83.16	97.43	94.35	95.16	120.07	102.16
Max Drawdown %	-21.25	-21.25	-21.13	-21.08	-20.97	-20.94	-20.75	-20.56	-20.43	-20.16	-20.17	-19.82	-19.3	-19.29	-19.1	-18.81	-19	-18.86	-17.65	-16.83	-14.71
Average Portfolio IAA Score	5.28	5.52	5.66	5.8	5.91	6.05	6.2	6.36	6.5	6.66	6.83	7.01	7.21	7.41	7.6	7.83	8.09	8.39	8.7	9.06	9.61
Regional Exposures %																					
US	56.61	57.81	57.5	57.21	56.78	56.18	55.44	54.72	54.09	53.23	52.26	51.3	50.39	50.13	50.18	48.91	47.85	45.95	45.43	44.94	41.38
EMEA	20.77	20.62	21.15	21.68	22.23	22.9	23.79	24.56	25.41	26.27	27.32	28.59	29.79	30.73	31.38	32.95	34.14	36.59	38.36	41.22	48.9
Asia-Pacific	11.29	11.63	11.8	11.93	12.14	12.34	12.47	12.71	12.81	13.04	13.13	13	13.01	12.8	12.76	12.68	12.64	12.51	12	10.23	7.99
Emerging Markets	11.34	9.94	9.55	9.17	8.86	8.57	8.31	8.01	7.69	7.46	7.28	7.11	6.81	6.33	5.68	5.46	5.38	4.95	4.21	3.61	1.73
Sector Exposures %																					
Energy	6.57	6.38	6.46	6.52	6.39	6.32	6.16	6.27	6.26	6.01	5.37	4.75	4.32	3.98	3.36	3.2	3.17	3.09	2.52	2.85	3.33
Industrials	10.38	10.61	10.71	10.89	11.06	11.21	11.37	11.65	11.89	12.14	12.44	12.68	13.06	13.49	13.79	14.07	14.06	14.65	15.18	16.29	14.87
Consumer Discretionary	11.98	12.02	11.81	11.59	11.48	11.47	11.21	10.66	10.49	10.43	10.5	10.38	9.95	9.44	9.15	9.31	9.17	9.12	8.5	8.41	8.57
Health Care	11.45	11.89	11.62	11.65	11.59	11.19	11.11	11.24	11.11	11.03	10.91	10.93	10.62	10.13	9.19	8.49	7.37	7.1	7.03	5.5	5.32
Financials	18.41	17.84	17.87	17.78	17.72	17.63	17.38	16.95	16.61	16.09	16.09	16.05	16.12	16.1	16.31	16.49	17.22	17.49	17.83	16.11	13.39
Information Technology	16.52	16.84	17.25	17.37	17.53	17.85	18.14	18.49	18.94	19.42	19.73	20.31	21.33	22.13	23.53	23.38	23.12	22.75	23.8	27.12	31.21
Materials	5.25	5.34	5.25	5.24	5.23	5.26	5.32	5.45	5.5	5.54	5.62	5.53	5.44	5.22	5.05	5.01	5.21	5.09	5.12	5.09	5.16
Consumer Staples	8.99	9.11	9.03	9.05	9.09	9.17	9.42	9.49	9.56	9.75	9.86	10.02	10.2	10.43	10.7	11.02	11.06	10.46	10.03	8.75	6.85
Communication Services	5.5	5.3	5.33	5.23	5.25	5.22	5.14	4.98	4.81	4.75	4.58	4.34	3.8	3.69	3.71	3.8	4.16	4.41	4.21	4.07	5.19
Utilities	3.18	3.27	3.3	3.33	3.35	3.38	3.46	3.55	3.57	3.61	3.7	3.8	3.96	4.1	4.05	4.19	4.44	4.74	4.88	4.96	5.54
Real Estate	1.77	1.4	1.39	1.34	1.31	1.29	1.28	1.28	1.25	1.23	1.2	1.2	1.21	1.28	1.15	1.05	1.02	1.1	0.9	0.84	0.58

Table D.2: Portfolio performance for each level of exclusion.

.

	MSCI ACWI	Pure	Opt 5%	Opt 10%	Opt 15%	Opt 20%	Opt 25%	Opt 30%	Opt 35%	Opt 40%	Opt 45%	Opt 50%	Opt 55%	Opt 60%	Opt 65%	Opt 70%	Opt 75%	Opt 80%	Opt 85%	Opt 90%	Opt 95%
Average Constituents #	2642	2477	2352	2229	2105	1981	1857	1733	1610	1486	1362	1238	1114	990	867	743	619	495	371	248	124
Average Marketcap %	100	95.58	92.97	90.03	87.34	84.11	80.12	76.02	72.3	67.99	63.26	58.5	53.21	47.85	42.94	37.22	31.15	25.02	19.75	14.04	7.85
Annual. Return %	11.11	11.42	11.57	11.53	11.38	11.37	11.55	11.53	11.56	11.31	11.39	11.47	11.53	11.54	11.61	11.61	11.39	11.77	11.58	11.51	10.85
Annual. Volatility %	13.13	13.16	13.12	13.1	13.11	13.08	13.06	13.03	12.96	12.95	12.9	12.89	12.85	12.89	12.77	12.71	12.9	12.85	12.64	12.67	13.47
Annual. Sharpe Ratio	0.29	0.31	0.32	0.32	0.31	0.31	0.32	0.32	0.32	0.31	0.31	0.32	0.32	0.32	0.33	0.33	0.31	0.34	0.33	0.32	0.27
Annual. Tracking Error (Ex-Post) %	-	0.36	0.71	0.83	0.91	1.02	1.07	1.1	1.2	1.28	1.36	1.41	1.6	1.79	1.88	1.89	1.99	2.11	2.37	2.69	3.16
Annual. Tracking Error (Ex-Ante) %	-	-	0.45	0.61	0.68	0.8	0.85	0.94	1,00	1.06	1.16	1.24	1.32	1.46	1.48	1.65	1.77	1.96	2.26	2.89	3.89
Annual. Information Ratio	-	0.23	0.17	0.13	0.08	0.07	0.11	0.1	0.09	0.04	0.05	0.06	0.06	0.06	0.06	0.06	0.03	0.08	0.04	0.03	-0.02
Annual. Turnover %	10.87	17.5	17.28	20.27	22.7	25.41	30.44	32.42	34.32	37.73	43.93	49.73	53.49	51.83	52.94	66.42	75.42	71.38	86.46	109.78	108.4
Max Drawdown %	-21.25	-21.25	-21.22	-21.15	-21.15	-21.07	-20.86	-20.96	-20.83	-20.87	-20.93	-20.7	-20.39	-20.24	-20.24	-19.91	-21.03	-20.46	-20.14	-20.29	-21.88
Average Portfolio IAA Score	5.28	5.52	5.87	6.02	6.13	6.25	6.4	6.55	6.67	6.81	6.99	7.17	7.33	7.52	7.71	7.94	8.2	8.44	8.75	9.13	9.65
Regional Exposures %																					
US	56.61	57.81	56.6	55.82	55.17	54.3	53.31	53.36	53.15	52.64	52.72	52.55	52.4	52.24	52.37	52.19	52.16	52.03	52.07	52.11	52.19
EMEA	20.77	20.62	22.43	22.95	23.34	23.51	24.02	24.11	24.02	24.67	24.6	24.62	24.94	24.9	25.28	24.97	24.95	25.33	25.58	25.78	25.97
Asia-Pacific	11.29	11.63	11.85	12.14	12.8	13.29	14.02	13.55	13.86	14.09	13.85	13.8	14.04	13.98	14.19	14.23	14.76	14.97	14.74	14.31	14.89
Emerging Markets	11.34	9.94	9.12	9.09	8.68	8.9	8.64	8.98	8.96	8.61	8.84	9.03	8.61	8.89	8.16	8.61	8.13	7.66	7.62	7.8	6.95
Sector Exposures %																					
Energy	6.57	6.38	6.44	6.45	6.54	6.51	6.31	6.22	6.17	6.46	6.18	5.94	5.9	5.94	5.78	5.89	5.79	5.88	5.84	5.94	5.87
Industrials	10.38	10.61	10.9	10.94	11.02	11,00	11.03	10.99	10.76	10.81	10.88	10.88	10.98	11,00	11.09	11.12	11.12	11.1	11.15	11.31	11.29
Consumer Discretionary	11.98	12.02	11.76	11.58	11.45	11.58	11.46	11.53	11.58	11.51	11.51	11.52	11.45	11.43	11.46	11.37	11.42	11.3	11.26	11.27	11.56
Health Care	11.45	11.89	11.43	11.47	10.95	10.81	10.82	10.93	10.99	11.13	11.23	11.64	11.34	11.38	11.01	11.34	11.06	11.21	11.21	11.2	10.99
Financials	18.41	17.84	18.52	18.59	18.75	18.74	18.38	18.3	18.24	18.07	18.01	17.93	17.84	17.85	18.05	18.02	18.14	18.08	18.03	17.79	17.76
Information Technology	16.52	16.84	16.55	16.45	16.5	16.57	16.6	16.63	16.54	16.44	16.38	16.51	16.49	16.63	16.53	16.6	16.47	16.29	16.01	16.16	16.22
Materials	5.25	5.34	5.3	5.28	5.22	5.33	5.48	5.61	5.73	5.67	5.75	5.53	5.51	5.33	5.41	5.16	5.12	5.34	5.56	5.96	6.05
Consumer Staples	8.99	9.11	9.12	9.11	9.38	9.19	9.41	9.13	9.28	9.27	9.5	9.36	9.74	9.68	9.74	9.63	9.79	9.69	9.63	9.17	8.93
Communication Services	5.5	5.3	5.41	5.42	5.36	5.31	5.37	5.3	5.19	5.15	5.02	4.91	4.84	4.69	4.82	4.76	4.91	4.88	5.27	5.31	4.99
Utilities	3.18	3.27	3.22	3.35	3.47	3.48	3.66	3.87	3.9	3.93	3.99	4,00	4.09	4.1	4.13	4.13	4.15	4.17	4.18	4.02	4,00
Real Estate	1.77	1.4	1.34	1.34	1.36	1.48	1.48	1.5	1.61	1.56	1.56	1.78	1.83	1.96	1.99	1.98	2.03	2.06	1.86	1.86	2.33

Table D.3: Portfolio performance for each level of optimization

Appendix E: Country-region Mapping

Abbreviation	Country	Region	*Comments*
AR	Argentina	EM	May 2009: from Emerging to Frontier Markets | May 2019: from Frontier to Emerging Markets | November 2021: From Emerging Markets to Standalone
AU	Australia	PACIFIC	
AT	Austria	EMEA	
BE	Belgium	EMEA	
BR	Brazil	EM	
CA	Canada	US	
CL	Chile	EM	
CN	China	EM	May 2018: China-A Shares included in the Emerging Markets
CO	Colombia	EM	
CZ	Czechia	EM	
DK	Denmark	EMEA	
EG	Egypt	EM	
FI	Finland	EMEA	
FR	France	EMEA	
DE	Germany	EMEA	
GR	Greece	EM	May 2001: from Emerging to Developed Markets | November 2013: from Developed to Emerging Markets
HK	Hong Kong	PACIFIC	
HU	Hungary	EM	
IN	India	EM	
ID	Indonesia	EM	
IE	Ireland	EMEA	
IL	Israel	EMEA	May 2010: from Emerging to Developed Markets
IT	Italy	EMEA	
JP	Japan	PACIFIC	
JO	Jordan	EM	November 2008: from Emerging to Frontier Markets
KR	Korea (the Republic of)	EM	
KW	Kuwait	EM	November 2020: from Frontier Markets to Emerging Markets
MY	Malaysia	EM	
MX	Mexico	EM	
MA	Morocco	EM	November 2013: from Emerging to Frontier Markets
NL	Netherlands (the)	EMEA	
NZ	New Zealand	PACIFIC	
NO	Norway	EMEA	
PK	Pakistan	EM	December 2008: from Emerging Markets to Standalonee | May 2009: from Standalone to Frontier Markets | May 2017: from Frontier Markets to Emerging Markets | November 2021: from Emerging to Frontier Markets
PE	Peru	EM	
PH	Philippines (the)	EM	
PL	Poland	EM	
PT	Portugal	EMEA	
QA	Qatar	EM	May 2014: from Frontier to Emerging Markets
RU	Russian Federation (the)	EM	March 2022: MSCI announced that it would reclassify the MSCI Russia Indexes from Emerging Markets to Standalone Markets status in one step as of the close of March 9, 2022
SA	Saudi Arabia	EM	May 2019: from Standalone to Emerging Markets
SG	Singapore	PACIFIC	
ZA	South Africa	EM	
ES	Spain	EMEA	
SE	Sweden	EMEA	
CH	Switzerland	EMEA	
TW	Taiwan (Province of China)	EM	
TH	Thailand	EM	
TR	Turkey	EM	
AE	United Arab Emirates (the)	EM	May 2014: from Frontier to Emerging Markets
GB	United Kingdom of Great Britain and Northern Ireland (the)	EMEA	
US	United States of America (the)	US	
VE	Venezuela (Bolivarian Republic of)	EM	May 2006: from Emerging to Standalone | January 2008: Index discontinued