Non-contour efficient fronts for identifying most preferred portfolios in sustainability investing

Ralph E. Steuer[a,*], Sebastian Utz[b]

[a] University of Georgia, Department of Finance, Athens, Georgia 30602, USA

[b] University of Augsburg, Faculty of Business Administration and Economics and Centre for Climate Resilience, Universitaetsstrasse 12, 86159 Augsburg, Germany

1. Introduction

In a world of climate change, environmental problems, concerns about working conditions, and so forth, this has caused many investors to evaluate what their investment dollars support and what they do not support. This has given rise to the broad investment area, variously referred to as socially responsible investing, sustainable investing, ethical investing, and so forth, that we call *ESG investing*. ESG stands for environmental, social, and governance and is a common measure for assessing such characteristics.

With surveys showing high percentages[1] of investors favorably disposed toward ESG, ESG is now a major item of interest in investing today (Hartzmark & Sussman, 2019). But among the favorably disposed, there is also the matter of *degree* – the degree to which an investor is motivated by the goals of ESG. Consider all investors with connections to ESG as being placed into a range ordered by degree. At the lower end of the range would be investors that have only casual interests in ESG. Sure ESG is a nice idea, but only if it won't hurt returns.

However, at the top of the range are investors for whom ESG is serious business. From these investors, this paper focuses on those we call *serious ESG investors*. These are investors whose commitment to ESG is such that they are willing to compromise on returns for ESG. By being willing to do this, we use the following as our definition of a serious ESG investor. A serious ESG investor is an investor whose strength of interest in ESG makes that consideration a criterion competitive on the playing field of portfolio selection with risk and return, thus causing the investor's efficient frontier to become an *efficient surface*. This places the paper within the area of multiple criteria portfolio selection (Xidonas, Mavrotas, Krintas, Psarras, & Zopounidis, 2012), and with our interests in sustainability-related issues narrowing the field, we join the portion of this area populated by Ballestero, Bravo, Pérez-Gladish, Arenas-Parra, & Plà-Santamaría (2012); Bilbao-Terol, Arenas-Parra, & Cañal-Fernández (2012); Cabello, Ruiz, Pérez-Gladish, & Méndez-Rodríguez (2014); Garcia-Bernabeu, Salcedo, Hi-

* Corresponding author.

E-mail addresses: rsteuer@uga.edu (R.E. Steuer), sebastian.utz@uni-a.de (S. Utz).

[1] According to the 2020 report of the Forum for Sustainable and Responsible Investment (www.ussif.org), of the $51 trillion in US investments under management that they track, nearly one-third is managed under the auspices of ESG strategies with the proportion having grown each year since 2010.

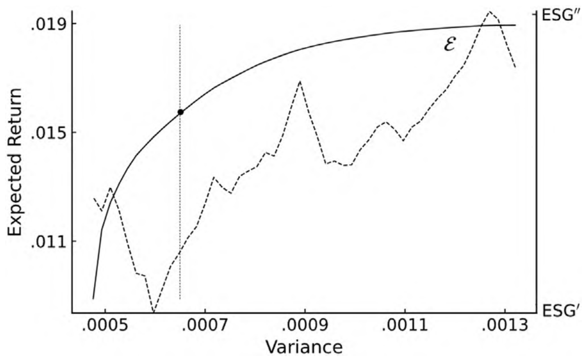

Fig. 1. Efficient frontier along with ESG scores over the efficient frontier. The line labeled $\mathcal{E}$ is the efficient frontier of a 100-security problem in which no more than 3% can be invested in any security. The dot on it is an example of an investor's most preferred selection along it. With ESG′ the lowest ESG value over the efficient surface, ESG′ to ESG″ is the range of ESG over the efficient frontier. Numbers for ESG′ and ESG″ in this example are 69.07 and 72.31, respectively. Since the maximum value of ESG over the efficient surface is 85.21, this means that 79.93% of the range of ESG over the efficient surface cannot be seen from the efficient frontier, a typical figure. Observe the non-constant nature of the dashed line over the efficient frontier.

lario, Plà-Santamaría, & Herrero (2019); Liagkouras, Metaxiotis, & Tsihrintzis (2020), and Chen & Mussalli (2020) in trying to work out improved methods for computing efficient surfaces and locating most preferred portfolios within them in problems in which ESG-related issues as a third criterion cannot be ignored.

Because of the resources and expertise involved in keeping up to date with ESG-qualified investments, investors interested in ESG find that they have few options but to place their money in vehicles representing themselves as ESG mutual funds. The hope of a serious ESG investor is that the intermediaries offering such funds possess the knowledge to run the funds in a fashion worthy of the faith they wish to place in them. This leads to questions like:

(1) Are today's ESG mutual funds doing all that can be done to offer investment solutions with the greatest ESG content for their investors?
(2) Does the mutual fund industry possess the knowledge to run an ESG mutual fund optimally for serious ESG investors?

Quick answers to the questions are "No" and "No". The difficulty is that the portfolio problem of a serious ESG investor is no longer a bi-criterion program to minimize risk and to maximize return as its two objectives, but a *tri-criterion* program with to minimize risk, to maximize return, and to maximize ESG score as its three objectives. With the efficient frontier now an efficient surface, this means that the investor's task is no longer to identify the point of best risk/return tradeoff on the efficient frontier, but to identify the point of best risk/return/ESG tradeoff on the efficient surface.

A reason behind the position that the portfolio problem of a serious ESG investor is to be treated as a tri-criterion program is shown in Fig. 1. In the standard approach for computing portfolios in ESG investing, securities are screened for ESG characteristics in a first stage and portfolios formed from the survivors that maximize expected return for each level of risk are computed in a second stage (see Utz, Wimmer, Hirschberger, & Steuer, 2014). This results in an efficient frontier shown as $\mathcal{E}$ in Fig. 1. With the reasoning of the standard approach being that ESG is now out of the way in the first stage, it is then only necessary in the second stage to compute the efficient frontier and pick the best point off it to be done. While the investor is free to pick any point along $\mathcal{E}$, let us say that in the problem of Fig. 1 the investor picks the point whose variance is 0.00065 as indicated by the vertical line. While things might seem fine, after a dashed line is drawn showing the ESG scores of portfolios along the efficient frontier (which would not be done in the standard approach as it is assumed that ESG is already behind us at this point), it is seen that things are not fine.

With the dashed line in this example showing most portfolios along the efficient frontier as having higher ESG scores, and with there being no reason to expect the point of maximum ESG score to have the same variance as one's point of most preferred risk/return tradeoff, one can see how easily there can be a need for further tradeoffs. This means that ESG cannot be considered as being taken care of by just screening, but that further ESG involvement in the portfolio construction process must be provided for post screening.

This leads to the definition of *ESG integration* used in this paper. In this paper, any ESG involvement in the portfolio construction process beyond the screening process in the first stage is ESG integration such that the greater the involvement of ESG after the screening stage, the greater the ESG integration. But how to carry out ESG integration is not an obvious matter. Suppose the investor, wishing to share in the higher ESG values along $\mathcal{E}$, considers nudging the point indicated by the dot a little further up the efficient frontier. This shows that we may not be done when we are supposed to be done.

But there is something even bigger than nudging one's most preferred point. It involves the interval [ESG′, ESG″] which shows how much ESG varies over the efficient frontier. In an empirical analysis of a large sample of conventional and sustainable mutual funds, Utz, Wimmer, & Steuer (2015) show that applying a technique that maximizes the ESG score generates portfolios with substantially higher ESG scores than the ESG scores on the efficient frontier, i.e., than those ESG scores in the interval [ESG′, ESG″]. The experiments conducted in this paper reveal that there is significant room for improving ESG compared to existing optimization approaches. This means that a large part of the range of ESG over the efficient surface is not visible from the standard procedure's efficient frontier $\mathcal{E}$. Moreover, it is not part of the standard procedure to tell you this, nor provide you with a means for accessing it. This is what we address in this paper, that is, how to access regions in the ignored ESG range and how to search them so as not to leave large amounts of ESG on the table that serious ESG investors are in the market to have.

The way in which regions of ignored ESG are searched in this paper is by means of the use of non-contour (NC)-efficient fronts. These are special lines, constructed to stretch across the efficient surface from one side to the other, which are then dragged sideways in a longline fishing fashion from one end of the efficient surface to the other. The idea is for nothing to escape, and with NC-efficient fronts, we use this approach to make sure nothing gets missed when looking for the investor's portfolio of optimal tri-criterion utility.

With regard to papers having any closeness to this paper, there is only the paper by Pedersen, Fitzgibbons, & Pomorski (2022). While both this and that paper start with the same basic two assumptions, namely, that in sustainability investing there are top-end investors ("serious ESG investors" in our case and "motivated" investors in the Pedersen et al., 2022 case), and that there are the three objectives risk, return and ESG, the papers diverge significantly from there.

The divergence takes place because Pedersen et al. (2022) transform the tri-criterion problem of a serious (or motivated) ESG investor into a bi-criterion one. That is, they take the problem with the three objectives of to minimize risk, to maximize return, and to maximize ESG and turn it into a bi-criterion one whose two objectives are to maximize the Sharpe ratio[2] and to maximize ESG.

[2] A point of maximum Sharpe ratio along a line traversing the surface of the feasible region is the point on it at which the value of expected return minus the riskfree rate divided by the square root of variance is greatest.

The point of doing this, from the Pedersen et al. (2022) point of view, is so that only a line is to be searched as opposed to a surface to find one's optimal portfolio. But the price to be paid for this is that, by cutting down on the number of objectives in the way of Pedersen et al. (2022), the assumption is made that in the criterion vector of an investor's optimal portfolio there will be a maximum Sharpe ratio relationship between the criterion vector's first two components of risk and return. What is problematic about this assumption is that in being a priori, if it is not spot-on, it is guaranteed that the investor will miss his or her optimal solution. This is worrisome because, on one hand, there are many conservative investors who would not want as much risk as comes with a maximum Sharpe ratio point, and on the hand, there are aggressive investors who would prefer to be further up the efficient frontier than a maximum Sharpe ratio point.

Regardless, it is not the purpose of this paper to talk about the research of others beyond clarifying the contribution of this paper and showing where this paper fits into the literature. In fact, we merely wish to present our approach as an alternative as it is good in multiple criteria decision making (see, for instance, Belton & Stewart, 2002; Ehrgott, 2005; Miettinen, 1999) for there to be several competing approaches for any problem so that users will always have the opportunity to choose from them the approach that best matches their decision-making style.

With the NC-efficient fronts approach of this paper designed to solve, by interacting with it, the three-objective problem of a serious ESG investor, what is the relationship of this method to the procedures of interactive multiple objective programming such as covered in Gardiner & Steuer (1994) and other places? These would include the procedures of Korhonen & Wallenius (1988); Korhonen & Laakso (1986); Wierzbicki (1980) and more recently by Figueira, Liefooghe, Talbi, & Wierzbicki (2010) and Miettinen, Eskelinen, Ruiz, & Luque (2010). The advantage is that points on NC-efficient fronts have a one-to-one correspondence with points on the efficient surface. Thus when using the NC-efficient fronts method, one is able to visualize the whole efficient surface at all times. This is better than a sequence of points (each normally better than the previous), which is typically the result of existing methods as many users in finance are not able to recognize their optimal solution in the absolute. They are only able to recognize their optimal solution by seeing that everything is worse.

One more thing to be noted is that in Pedersen et al. (2022) the focus is on investors that are more of an institutional variety in that they are willing to take on leverage, whereas in this paper the focus is more on retail-level mutual funds that serious ESG investors would find suitable for, say, their retirement accounts, a much larger market. This means that, in contrast to Pedersen et al. (2022), in this paper no short selling is allowed and on each security there is an upper bound.

Due to ESG integration still being in its infant stages, it is assumed that few retail-level ESG mutual funds have had the chance to install anything advanced in the way of ESG integration in their mutual funds, thus causing them to rely mostly upon the standard two-stage procedure that we have been referring to. With most mutual funds having similar second stages, the thing then that primarily distinguishes one ESG mutual fund from another in today's environment is the screening that takes place in the first stage. Thus, if a fund wants to be an ESG mutual fund, it just has to apply appropriate ESG screens in the first stage as the second stage is essentially the same one mutual fund to the next. While this is the way that most ESG mutual funds operate, it is, as we will see, a wasteful way of infusing a portfolio with ESG. While the waste might not make much difference to investors looking for only enough ESG in their investing that it can be handled within the screening operations of Stage I, in which case they don't need ESG integration, it can make a lot of difference to serious ESG investors who are in the market for as much ESG as they can get. This is because with ESG as an objective, serious ESG investors are *optimizers*, not *satisficers*. With satisficers, after a certain amount of ESG has been achieved, their interests level off, but with optimizers, their interests keep going.

The question of this paper is this: If the two-stage approach of a currently operated ESG mutual fund is not good enough for serious ESG investors, is there a better way? The thesis of the paper is that there is, and it is to replace the standard two-stage approach with another two-stage approach whose only difference with the first is that it is equipped with a more-capable Stage II, one built around efficient surfaces and the use of NC-efficient fronts for searching them. Stage I can stay as is.

To compare the new approach, we set up two modules, Mod 1 and Mod 2. The purpose of Mod 1 is to represent the Stage II of a non-integrated ESG mutual fund, that is, one that follows the standard two-stage procedure, to have something baseline against which to compare Mod 2 whose purpose is to represent a more powerful Stage II whose ESG integration is accomplished by taking efficient surfaces and the use of NC-efficient fronts into account. Furthermore, within Mod 2 we develop an ESG integration index which ascribes 0 to all portfolios on the Mod 1 efficient frontier, 100 to the portfolio of global maximum ESG score, and values in between to all other portfolios depending upon where they are located on the efficient surface. In the computational experiments part of the paper (parameterized with data from stocks in the S&P500) we show, for instance, how giving up 1 basis point per month in expected return generally leads to an increase in ESG integration of up to about 10%, how giving up 5 basis points per month in expected return generally leads to about a 25% increase in ESG integration, how giving up 10 basis points per month in expected return generally leads to about a 35% increase in ESG integration, and so forth.

The rest of the paper is as follows. In Section 2, after placing Mods 1 and 2 in perspective, preliminaries are reviewed with regard to the tri-criterion nature of the paper. Section 3 discusses efficient surfaces and the unique tool of this paper, NC-efficient fronts, that enable the ESG integration accomplishments of the paper. In Section 4, computational tests are conducted to show the additional amounts of ESG that are achievable via Mod 2 versus Mod 1. Section 5 illustrates the straightforwardness of Mod 2's application, and Section 6 concludes the paper with final remarks.

2. Preliminaries

In this section we do three things. With regard to the different items and concepts discussed, which include non-integrated versus integrated ESG mutual funds, additional screening, and efficient frontiers and surfaces, we place them all in perspective so it can be seen how they interrelate. Also, since it is Markowitz's mean-variance approach that gets extended to enable us to work with efficient surfaces, we review Markowitz's mean-variance approach to be clear about what in it gets extended and how. Finally, we explain why the dashed line in Fig. 1, along with other lines seen in this paper, are jagged as opposed to being smooth in appearance.

2.1. Perspective

With regard to the things we have been talking about, placing them in perspective we have Fig. 2. In the first row is a conventional mutual fund which is only included as a point of reference. After conventional screening, the securities surviving Stage I are passed to Stage II where from them, in terms of Markowitz, the M-V efficient frontier of the mutual fund is constructed. After picking the best point on this frontier, the mutual fund's most preferred conventional mutual fund portfolio is found.

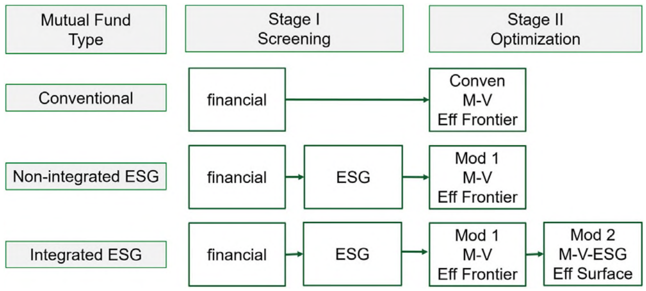

Fig. 2. Three two-stage procedures. In the first row is the two-stage procedure of a conventional mutual fund with its resulting "Conventional" M-V efficient frontier indicated in Stage II. In the second row is the two-stage procedure of a non-integrated ESG mutual fund with its extra Stage I screening box, and its resulting "Mod1" M-V efficient frontier indicated in Stage II. In the third row is the two-stage procedure of an integrated ESG mutual fund. Following the same three boxes of a non-integrated fund, the fund becomes integrated due to the fourth box with its resulting "M-V-ESG" efficient surface indicated also in Stage II.

In the second row is a non-integrated ESG mutual fund. It is also a diagram of the standard procedure discussed previously, but it now is called a non-integrated ESG mutual fund and its efficient frontier is called the Mod 1 M-V efficient frontier (in accordance with the two models set up). This mutual fund type is indeed non-integrated as in it there is no ESG involvement in the portfolio construction process after the ESG screening box. In this way, the only difference between a conventional mutual fund and a non-integrated ESG mutual fund is that in a non-integrated ESG mutual fund there is a second ESG screening box. Since this causes the securities passed from Stage I to Stage II to be different, the Mod 1 M-V efficient frontier that results will then, of course, be different from the M-V efficient frontier of a conventional mutual fund.

In the third row is the two-stage procedure of the integrated ESG mutual fund type that is the object of attention in this paper for serious ESG investors. Through its first three blocks, it is the same as a non-integrated ESG mutual fund but becomes integrated with the addition of the fourth block. Thus, instead of stopping with the best point on the Mod 1 M-V efficient frontier, the procedure continues to only stop for good at the best point on the M-V-ESG efficient surface. This causes three questions to arise.

(1) How is the M-V-ESG efficient surface to be searched as efficient surfaces are not easy to search?
(2) What is the connection, if any, between the Mod 1 M-V efficient frontier and the Mod 2 M-V-ESG efficient surface?
(3) How has it been possible for the ESG mutual fund industry to maintain customers given that most ESG funds are run in a non-integrated fashion?

With regard to the first, the M-V-ESG efficient surface is searched by means of the special lines called NC-efficient fronts which are new to the literature with this paper, and about which more is discussed in Section 3. With regard to the second question, consider the periphery of the Mod 2 M-V-ESG efficient surface. The Mod 1 M-V efficient frontier is a portion of this periphery, namely, the portion of the periphery furthest away from the point of maximum ESG score (as a result of ESG not being taken into account when computing an efficient frontier). Thus, when an investor picks a point on the Mod 1 efficient frontier, that investor is picking a point in one of the worst places that a serious ESG investor could look for the portfolio that optimizes his or her utility function in the efficient set. Consequently, the chance of a non-integrated ESG mutual fund being able to identify a portfolio that optimizes the risk/return/ESG tradeoff of a serious ESG investor is highly remote.

We say this because in a non-integrated ESG mutual fund, ESG is only addressed in Stage I. With the two-stage procedure of a non-integrated ESG fund assuming that Mod 1 is enough, the idea is that by constructing the Mod 1 M-V efficient frontier and picking the best risk/return point on it, the point will have, due to the second screening box in Stage I, enough ESG in it to satisfy mutual fund customers. While this might be the case for investors in the lower part of the range, it definitely shortchanges serious ESG investors who are not in the market to be shortweighted. As optimizers for ESG, they want all the ESG that can be obtained for any returns given up, not some lesser amount that other investors might feel is not worth complaining about.

As for the third question, since virtually all ESG mutual funds follow the non-integrated approach of the second row in Fig. 2 in one way or another, ESG investors have no choice but to invest in one of them no matter how badly they want to save the world. What this means is that Mod 1 leaves considerable amounts of ESG on the table which Mod 2, as shown in the computational experiments of Section 4, is able to prevent from happening. Now for a review of M-V portfolio selection to set the stage for the rest of the paper.

2.2. Review of M-V portfolio selection

Expressed in *multiple criteria* format, the M-V model of portfolio selection, due to Markowitz (1952, 1959), is as follows

$$\min\{z_1 = \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}\} \tag{1}$$
$$\max\{z_2 = \boldsymbol{\mu}^T \mathbf{x}\}$$
$$\text{s.t.} \quad \mathbf{x} \in S = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{1}^T\mathbf{x} = 1,\ 0 \le x_i \le \omega_i \text{ for all } i\}$$

The convenience of this format is that it makes clear the objectives that are to be simultaneously optimized and, in this instance, shows M-V to be a bi-criterion problem. In (1), n is the number of securities, S is the feasible region in decision space, and any $\mathbf{x} \in S$ is a portfolio. Furthermore, with $\boldsymbol{\Sigma}$ the covariance matrix, $z_1 = \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x}$ is portfolio return variance, and $z_2 = \boldsymbol{\mu}^T \mathbf{x}$ is expected portfolio return. As for the constraints defining S, short selling is not allowed and all securities have upper bounds on the proportions of capital that can be invested in them.

In any problem with more than one objective, there is a second version of the feasible region designated $Z \subset \mathbb{R}^k$ where k is the number of objectives. Under a problem's k objective functions, Z is the set of images of all points in S. It is called the feasible region in *criterion space*. In the case of (1), with its two objectives, we have

$$Z = \{\mathbf{z} \in \mathbb{R}^2 \mid z_1 = \mathbf{x}^T \boldsymbol{\Sigma} \mathbf{x},\ z_2 = \boldsymbol{\mu}^T \mathbf{x},\ \mathbf{x} \in S\}. \tag{2}$$

With members of Z called *criterion vectors*, a criterion vector $\bar{\mathbf{z}} \in Z$ is *nondominated* in (1) if and only if there exists no $\mathbf{z} \in Z$ such that $z_1 \le \bar{z}_1$, $z_2 \ge \bar{z}_2$ with $\mathbf{z} \ne \bar{\mathbf{z}}$. As for portfolios, $\bar{\mathbf{x}} \in S$ is *efficient* if and only if its criterion vector $\bar{\mathbf{z}} \in Z$ is nondominated.

Although the set of all nondominated criterion vectors is known in some quarters as the Pareto front or nondominated frontier, we follow tradition in finance and call it the *efficient frontier*. The significance of an efficient frontier is that it is the set of M-V combinations of all potentially optimal portfolios. This means that if a portfolio can be optimal in (1), its M-V combination will be on the efficient frontier, but if a portfolio cannot be optimal, its M-V combination will not be on the efficient frontier. Following Markowitz (1952, 1959), the procedure for solving (1) is to compute the efficient frontier and then select the most preferred M-V combination on it. By taking the inverse image of the selected M-V combination, we then have the decision maker's optimal portfolio. This describes the Markowitz-based M-V procedure of Mod 1 that will be expanded upon to form the Markowitz-like M-V-ESG procedure of Mod 2.

For computing the efficient frontier, the method used is to solve the following formulation repetitively

$$\max\{\boldsymbol{\mu}^T\mathbf{x}\} \tag{3}$$
$$\text{s.t.} \quad \mathbf{x}^T\boldsymbol{\Sigma}\,\mathbf{x} \le \upsilon \qquad \upsilon \in [\sigma^2_{\min}, \sigma^2_{\max}]$$
$$\mathbf{x} \in S$$

for 100 equally-spaced values of υ from $[\sigma^2_{\min}, \sigma^2_{\max}]$ where $\sigma^2_{\min}$ and $\sigma^2_{\max}$ are the minimum and maximum variances over the efficient frontier obtained by solving $\min\{\mathbf{x}^T\boldsymbol{\Sigma}\,\mathbf{x} \mid \mathbf{x} \in S\}$ and $\max\{\boldsymbol{\mu}^T\mathbf{x} \mid \mathbf{x} \in S\}$ respectively. The purpose of the repetitive optimizations is to produce enough M-V combinations so that when plotted as dots and connected with straight lines the result appears as the continuous curve that the efficient frontier is. In the remainder of this paper, we refer to the fine representation of the curve/surface as the efficient frontier/surface.

2.3. Jaggedness

Let us now dispose of a question that might have been asked in the Introduction. It is about why the dashed line in Fig. 1 isn't more steady in value. For this, we go to bi-criterion formulation (1). Should the feasible region S of (1) be specified by only an equality constraint as in Merton (1972), the efficient frontier is only from a single parabola. But should inequality constraints be involved in the specification of S, the efficient frontier will be *piecewise parabolic*, meaning that it consists of a connected string of parabolic line segments, each coming from a different parabola. Since, in the form of $0 \le x_i \le \omega_i$ we have many inequality constraints, our efficient frontiers will consist of many parabolic line segments, typically in the neighborhood of a hundred or two (Hirschberger, Qi, & Steuer, 2010).

Also, it is to be noted that each parabolic line segment is defined by its own subset of securities, just in different proportions as one moves along the parabolic line segment. We call this a parabolic line segment's *defining subset* of securities. Thus, when moving from one parabolic line segment to the next along an efficient frontier, we move from one defining subset to another, and with different securities going in and out of the subsets, this results in no guaranteed steady value in ESG scores when moving along an efficient frontier. In this way, a steady dashed line would be an exception rather than the rule.

3. NC-efficient fronts

Not enthused by some of the aspects of Mod 1, let us now deal more directly with the third-criterion status of ESG and begin discussions about the details involved in the operation of the Mod 2 M-V-ESG optimization box in Fig. 2. The portfolio selection problem of this box, which is a tri-criterion problem, is

$$\min\{z_1 = \mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x}\} \tag{4}$$
$$\max\{z_2 = \boldsymbol{\mu}^T\mathbf{x}\}$$
$$\max\{z_3 = \boldsymbol{\nu}^T\mathbf{x}\}$$
$$\text{s.t.} \quad \mathbf{x} \in S$$

where $\boldsymbol{\nu} \in \mathbb{R}^n$ is the vector of ESG scores associated with the different securities and $z_3 = \boldsymbol{\nu}^T\mathbf{x}$ is portfolio ESG. While the only difference between this formulation and that of (1) is that this formulation has a third objective, the third objective has ramifications in that (4)'s feasible region Z in criterion space is now in three dimensions. Here, with $Z \subset \mathbb{R}^3$, $\bar{\mathbf{z}} \in Z$ is nondominated in (4) if and only if there exists no $\mathbf{z}$ a member of $Z \subset \mathbb{R}^3$ such that $z_1 \le \bar{z}_1$, $z_2 \ge \bar{z}_2$, $z_3 \ge \bar{z}_3$ with $\mathbf{z} \ne \bar{\mathbf{z}}$. As for a portfolio $\bar{\mathbf{x}} \in S$, it is efficient in (4) if and only if its 3-dimensional criterion vector $\bar{\mathbf{z}} \in Z$ is nondominated. Here, the set of all nondominated criterion vectors, being a portion of the surface of $Z \subset \mathbb{R}^3$, is the efficient surface.

3.1. Dealing with efficient surfaces

To deal with the efficient surface caused by the third criterion of (4), we introduce the idea of *NC-efficient fronts*. NC-efficient fronts are not to be confused with contours or iso-quants as on contours and iso-quants a criterion value is held fixed. Along NC-efficient fronts all criterion values change. It is just that NC-efficient fronts are constructed differently for the purposes they are to serve. And it is by means of NC-efficient fronts that the uniqueness of the ESG integration of this paper is carried out.

It is noted that while the Mod 1 M-V efficient frontier is a line across the efficient surface, it is only minimally the case. This is because it is only in the sense that it coincides with the boundary of the efficient surface between the minimum variance (minVar) and maximum return (maxRet) points. Since an efficient surface also has a point[3] of maximum ESG score (maxESG), the Mod 1 M-V efficient frontier then constitutes the portion of the boundary of the efficient surface furthest away from the point of maxESG. More about this is discussed in Section 5. In this way, NC-efficient fronts can be viewed as variants of the Mod 1 M-V efficient frontier that have been dragged sideways across the efficient surface toward the maxESG point in a kind of a search-party fashion. Whereas M-V efficient frontiers tell us about tradeoffs between mean and variance, they tell us nothing about tradeoffs between mean and ESG. However, NC-efficient fronts tell us about such tradeoffs.

As for constructing NC-efficient fronts, let $\mathbf{x}^i$ (*current portfolio*) be a Mod 1 M-V efficient frontier portfolio. In this capacity, $\mathbf{x}^i$ is a portfolio whose mean-variance combination (μ_i, σ_i^2) is on the periphery of the efficient surface furthest away from the point of maxESG. Then, using the μ_i and σ_i^2 of the current portfolio's M-V combination, we solve

$$\max\{z_3 = \boldsymbol{\nu}^T\mathbf{x}\} \tag{5}$$
$$\text{s.t.} \quad \mathbf{x}^T\boldsymbol{\Sigma}\mathbf{x} = \sigma_i^2$$
$$\boldsymbol{\mu}^T\mathbf{x} \ge \mu_i - bp$$
$$\mathbf{x} \in S$$

where bp is some number of basis points by which the μ_i of the current portfolio is relaxed in the above. The solution $\mathbf{x}^{bp,i}$ then obtained by solving (5) is an efficient portfolio whose criterion vector, being nondominated, is on the M-V-ESG efficient surface (verified in Appendix A). Then, for a given bp, the collection of M-V combinations of all portfolios $\mathbf{x}^{bp,i}$ obtained by solving (5) for all (μ_i, σ_i^2) combinations on the Mod 1 M-V efficient frontier forms the *bp-NC-efficient front* of (4). For instance, a 5 basis point NC-efficient front is a line across the efficient surface that is 5 basis points lower in expected return than that of the efficient surface's minVar/maxRet boundary. In practice, we compute the NC-efficient front for a given bp value by solving (5) for, say, 100 different (μ_i, σ_i^2) combinations on the Mod 1 M-V efficient frontier. In this way, enough information is produced so that when plotted a nice graph of the bp-NC-efficient front can be obtained. In general, we have found 100 points to be plenty, but if not, (5) can always be solved for more points.

With respect to a current portfolio's ESG score $\boldsymbol{\nu}^T\mathbf{x}^i$, the solution of (5) tells us how much extra ESG can be obtained by relaxing expected return bp basis points, holding variance σ_i^2 constant. This amounts to $\boldsymbol{\nu}^T\mathbf{x}^{bp,i} - \boldsymbol{\nu}^T\mathbf{x}^i$. However, such amounts are not part of Mod 1 as formulation (5) is outside the scope of Mod 1. Nevertheless, unknown to Mod 1 because they cannot be seen from a bi-criterion point of view, the $\boldsymbol{\nu}^T\mathbf{x}^{bp,i} - \boldsymbol{\nu}^T\mathbf{x}^i$ are the amounts of

[3] In this paper, "points" and "portfolios" are used in the following sense. Points are M-V combinations on an efficient frontier or M-V-ESG combinations on an efficient surface, whereas portfolios are in n-space whose images are in M-V or M-V-ESG space.

ESG, as a function of the *bp* value involved, alluded to earlier as getting left on the table with Mod 1.

3.2. ESG integration index

With regard to an $\mathbf{x}^i$ portfolio on the Mod 1 M-V efficient frontier and any of its *bp*-NC-efficient portfolios $\mathbf{x}^{bp,i}$, we now introduce the idea of a $\Delta\boldsymbol{\nu}$ *ESG integration index* given by

$$\Delta\boldsymbol{\nu} = \frac{\boldsymbol{\nu}^T\mathbf{x}^{bp,i} - \boldsymbol{\nu}^T\mathbf{x}^i}{\nu_{\max} - \boldsymbol{\nu}^T\mathbf{x}^i}(100) \qquad (6)$$

where, in terms of ESG,

(1) the numerator tells us how far $\mathbf{x}^{bp,i}$ is across the efficient surface from $\mathbf{x}^i$
(2) the denominator tells us how far the portfolio of maxESG is across the efficient surface from $\mathbf{x}^i$, where $\nu_{\max}$ is the ESG score of the portfolio of maxESG.

The calculation $\Delta\boldsymbol{\nu}$ functions as an ESG integration index as follows. Without ESG playing a role in Stage II as in the case of the second mutual fund type of Fig. 2, there is no ESG integration, in which case the index should be zero. But when ESG plays a role in Stage II as it does in the case of the third mutual fund type of Fig. 2, $\Delta\boldsymbol{\nu}$ should be of some positive value (as for instance given in percent) depending upon the degree to which ESG is incorporated into a Stage II portfolio. Let us see if $\Delta\boldsymbol{\nu}$ operates in this way. In the case of the second mutual fund type where ESG is only taken into account in Stage I, we only wind up with some portfolio $\mathbf{x}^i$ on the Mod 1 M-V efficient frontier. With $\mathbf{x}^i$ then being the end of the line for this mutual fund type, no relaxations are carried out. But this is equivalent to solving (5) with $bp = 0$ in which case the resulting $\mathbf{x}^{bp,i}$ is $\mathbf{x}^i$ all over again. With this making the numerator of (6) zero, the $\Delta\nu$ ESG integration index then takes on a value 0% as it is supposed to in the case of ESG non-integration. Now when ESG plays a role in Stage II, *bp*-NC-efficient front portfolios $\mathbf{x}^{bp,i}$ are computed for positive numbers of basis points for *bp*. Since the current point $\mathbf{x}^i$ is efficient, this causes the numerator of (6) to be positive in which case the $\Delta\nu$ ESG integration index takes on a positive percentage, with the positive percentage growing with *bp* as it should in this case.

In this paper, we use (6) in relative terms, since ESG metrics of different providers differ in their range (i.e., Bloomberg, Sustainalytics, and Thomson Reuters use percentile ranks from 0 to 100; Amundi, FTSE, ISS, and MSCI have other ranges) and a relative quantity is able to capture different scales. The relative quantity reveals how much of the unused ESG potential can be achieved by relaxing expected return by a certain number of basis points. Nevertheless, there might well be times when (6) is useful in absolute terms (achieved when using only the numerator). Reporting absolute values would document the increase in ESG points that correspond to a basis point relaxation in expected return. For a given scale of an ESG score, an absolute quantity would document the ESG improvement of a basis point expected return relaxation independent of the current ESG level of the mutual fund. However, it does not reveal the size of the maximum possible absolute ESG improvement.

Without there being any other ESG integration measures known to us, it appears that the $\Delta\boldsymbol{\nu}$ ESG integration index described here is the first proposed for assigning a scalar value to the degree of ESG integration manifesting itself in the portfolio of an integrated ESG mutual fund.

To visualize the $\Delta\boldsymbol{\nu}$ ESG integration index as a function of *bp*, consider Fig. 3. In the figure, 0% on the vertical axis means that no ESG integration has occurred whereas a figure like, say, 40% means that (5) is able to compute portfolios on the efficient surface that close the gap in ESG between points on the Mod 1 M-V efficient frontier and the point of maxESG by 40%.

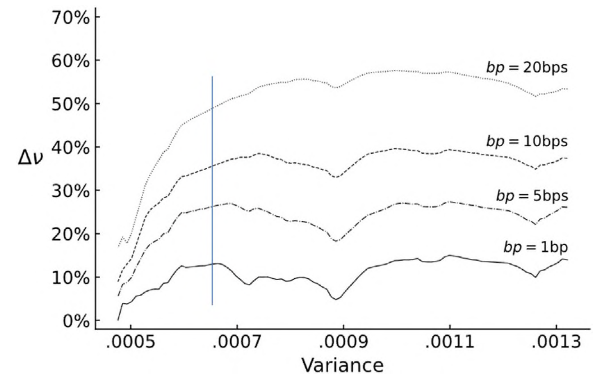

Fig. 3. NC-efficient fronts graph. For the 100-security problem of Fig. 1, the lines show the degrees of ESG integration made possible by Mod 2 for the *bp* relaxations in expected return of 1, 5, 10 and 20 bps as a function of variance. The lines show initially large gains in $\Delta\nu$ integration for small relaxations in *bp*, but then gains in $\Delta\nu$ slow down as the *bp* relaxations become larger.

In the figure, for the 100×0.03 tri-criterion problem of Fig. 1, the bottommost plot pertains to a 1 bp/month reduction in expected return. In this problem, for this reduction in expected return, the graph shows us that for almost all Mod 1 M-V efficient portfolios, there are corresponding NC-efficient front portfolios able to capture up to 10% or even a little more of the unutilized differences in ESG between portfolios on the Mod 1 M-V efficient frontier and the portfolio of maxESG regardless of variance. As for the other three lines with 5, 10 and 20 bp relaxations in expected return, they show gains in the neighborhoods of 25, 35 and 50%, respectively, for nearly all values of variance except near the minimum variance point.[4]

A question similar to the one asked in Section 2.3 is why aren't the lines in Fig. 3 more steady? While the formulation here is (4), the answer is very much the same as with (1), just in 3D rather than 2D. That is, instead of a connected string of parabolic line segments making up the efficient frontier in 2D, a collection of connected paraboloidic platelets makes up the efficient surface in 3D (Hirschberger, Steuer, Utz, Wimmer, & Qi, 2013).

Given ESG's inclusion as a third criterion, the four lines in Fig. 3 reflect curvature of the efficient surface. Whereas large improvements in ESG for small relaxations in expected return take place as we depart sideways from the Mod 1 M-V efficient frontier, improvements in ESG drop off at an ever increasing rate for further relaxations in expected return as we move closer to the point of maxESG. This is due to the essentially zero slope of the efficient surface near the efficient frontier and the essentially negative infinity slope of the efficient surface in locales near the point of maxESG. Such behavior is noticed in Fig. 3 with the gains in ESG for *bp* going from 0 to 5 bps being on the order of the gains in ESG when going from 5 to 20 bps for virtually all values of variance.

4. Computational experience

What is seen in Fig. 3 are the gains in ESG that can be anticipated from the types of concessions that serious ESG investors would almost surely grant fund managers for pursuing ESG. For example, over the 5-year data period of this paper, mean monthly returns for stocks in the S&P500 were 89 basis points. Would a se-

[4] A difficulty at the minimum variance point is that there is often not enough room in the feasible region for expected return to be relaxed by the full amount of the *bp* specified, but this situation disappears quickly by moving only a small amount away from the point of minimum variance.

Table 1
Parameter settings for configuring the first part of the random problem generator that generates all but vector $\boldsymbol{\nu}$.

	Parameter	Statistic	Value
1	number of problems		100
2	approved list size (number of securities) n		50, 100, 150, 250, 500
3	security upper bound ω_i for all $i = 1 \dots n$		0.02, 0.03, 0.04, 0.05, 0.08, 0.12
4	riskfree rate	mean	0.0014
5	covariance matrix $\boldsymbol{\Sigma}$ main diagonal elements	mean	0.0055
6		stdev	0.0066
7	covariance matrix $\boldsymbol{\Sigma}$ off-diagonal elements	mean	0.0012
8		stdev	0.0011
9	expected return vector $\boldsymbol{\mu}$ elements	mean	0.0089
10		stdev	0.0093

rious ESG investor be willing to give up 1, 5 or 10 of these 89 basis points for the better ESG figures seen in Fig. 3? Probably, but what about 20? With no data on this, one can only speculate, but based upon the market's performance over the last 5 years, this would still leave 69 basis points, so there would probably be a few.

But with Fig. 3 representing only one instance of a 100-security problem, the question is how general are the results of this one instance for problems with different numbers of securities and under different constraint conditions? To answer this question, we ran four batteries of computational experiments over a 5×6 matrix of scenarios with the first dimension of the matrix involving problems of different sizes and the second dimension involving problems with different upper bounds on how much can be invested in each security. The four batteries involve running the matrix of scenarios over the four levels of expected return relaxation of $bp = 1, 5, 10$ and 20. Because of the number of portfolio selection problems needed to carry out the experiments with a sample size of 100 for each scenario in each battery, it was necessary to generate the problems needed by a random problem generator. The random portfolio selection problem generator employed is as follows.

4.1. Random portfolio selection problem generator

For realism, our random portfolio selection problem generator is parameterized with real-world descriptive data derived from stocks in the S&P500 as of the end of December 2019. The problems generated are of form (4). The advantage of working with randomly generated samples is the flexibility afforded to study sensitivities with regard to different approved list sizes and differently restricted feasible regions while keeping within the correlation structure of actual S&P500 constituents.

The random problem generator used exists in two parts, with the first part generating all but $\boldsymbol{\nu} \in \mathbb{R}^n$, and second part just generating $\boldsymbol{\nu}$. The ten parameters used for operating the first part of the random problem generator, and the values used for them over the course of the experiments conducted, are given in Table 1.

As for the settings of the first three parameters, they imply that 100 problems are to be generated for each useable combination of number of stocks n and upper bounds ω_i. But while there are thirty $n \times \omega_i$ combinations, only 29 are usable as the 50×0.02 combination is not useable. This is because the 50×0.02 combination generates only one portfolio, that being with all 50 stocks each at its upper bound of 0.02. This does not give us a usable efficient frontier to work with in this paper. But with 29 usable combinations, this causes 2900 problems to be generated.

The value of parameter four (riskfree rate) is the one-month US Treasury Bill rate as taken from the website of Kenneth French for December 2019. The values specified for the next six parameters are means taken from monthly data for the S&P500 over the 5-year period January 1, 2015 to December 31, 2019. For instance,

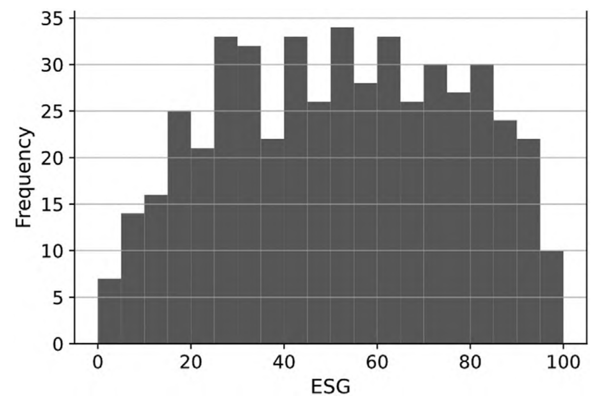

Fig. 4. Histogram of how the ESG scores of the stocks in the S&P500 at the end of 2019 were distributed.

0.0066 is a mean standard deviation value. Parameters five through eight are quantities required by the procedure of (Hirschberger, Qi, & Steuer, 2007) that is built into the random problem generator for the generation of realistic portfolio-selection covariance matrices $\boldsymbol{\Sigma}$. Parameters nine and ten pertain to the generation of $\boldsymbol{\mu}$ vectors.

Concerning elements to comprise $\boldsymbol{\nu}$, there is competition in the ratings business resulting in ESG-type scores being available from a number of vendors including RepRisk,[5] MSCI,[6] Sustainalytics,[7] and others. In addition, some mutual funds, not satisfied with what the vendors have to offer, generate their own in-house ratings. Consequently, ratings may be letter-based, binary in nature, integer, or continuous over a range. In this paper, we have chosen ESG ranks from Sustainalytics as it is a known vendor whose product is representative. From this source, ESG scores range from 0 (worst) to 100 (best) and are available on virtually all tradeable securities. Thus, for the elements of $\boldsymbol{\nu}$ which constitute the second part of the random problem generator, they are generated by sampling from the histogram of Fig. 4 where the histogram shows how the ESG scores by Sustainalytics of the stocks in the S&P500 were distributed at the end of 2019. Actually, in this paper we only sample from the upper 50% of the histogram as it is assumed that stocks in the lower 50% would be of the type that ESG mutual funds would screen out in their Stage I's.

4.2. Experimental results

The four batteries of experiments give us Tables 2–5. Each table presents three quantities for the different experimental setups

[5] https://www.reprisk.com/.
[6] https://www.msci.com/.
[7] https://www.sustainalytics.com/.

Table 2
Relaxation 1 bp: $\Delta\boldsymbol{\nu}$ ESG integration potential. All entries in the table are in percents. With regard to a 1 basis point relaxation in expected return, reported for each of the 29 useable $n \times \omega_i$ combinations are (a) av$\Delta\nu$, the average $\Delta\nu$ ESG integration potential that portfolios along the Mod 1 M-V efficient frontier have when using Mod 2, (b) av25, the average ESG integration potential that points 25% along the Mod 1 M-V efficient frontier have when using Mod 2, and (c) av60, the potential that points 60% along the frontier have.

Problem size n	Quantity	Upper bound ω_i					
		0.02	0.03	0.04	0.05	0.08	0.12
50	av$\Delta\nu$	–	15.84	13.30	10.20	8.58	9.12
	av25	–	17.94	14.77	10.69	9.13	9.73
	av60	–	16.04	13.35	10.86	9.00	9.59
100	av$\Delta\nu$	13.20	11.36	10.34	8.90	8.03	8.76
	av25	14.73	12.46	11.02	10.08	8.92	8.97
	av60	13.51	11.54	10.65	8.88	8.16	9.33
150	av$\Delta\nu$	11.61	10.72	9.48	8.28	7.77	8.98
	av25	12.97	12.17	9.91	9.00	8.40	9.64
	av60	11.31	10.86	10.17	8.98	7.93	9.02
250	av$\Delta\nu$	10.04	10.59	8.87	7.86	8.01	8.99
	av25	10.91	11.50	9.59	8.72	8.26	9.58
	av60	10.20	11.34	9.60	8.32	8.56	9.25
500	av$\Delta\nu$	8.98	10.55	9.13	8.28	7.56	8.58
	av25	9.72	11.40	9.75	9.20	7.82	8.67
	av60	9.48	10.89	9.34	8.23	8.08	9.24

Table 3
Relaxation 5 bps: $\Delta\boldsymbol{\nu}$ ESG integration potential. Same as with Table 2, but with regard to 5 basis points in relaxation in expected return.

Problem size n	Quantity	Upper bound ω_i					
		0.02	0.03	0.04	0.05	0.08	0.12
50	av$\Delta\nu$	–	35.02	29.95	23.68	19.78	21.15
	av25	–	39.58	33.55	25.66	21.51	22.99
	av60	–	35.51	30.21	24.62	20.19	21.94
100	av$\Delta\nu$	29.95	25.69	23.46	20.26	18.90	20.27
	av25	33.15	28.29	25.37	22.53	20.76	21.29
	av60	30.86	26.11	24.17	20.59	19.15	21.38
150	av$\Delta\nu$	25.92	24.41	21.69	18.96	18.33	20.76
	av25	28.90	27.02	23.13	20.70	19.77	22.55
	av60	25.96	24.93	22.81	19.67	19.02	21.24
250	av$\Delta\nu$	22.71	23.93	20.62	18.37	18.67	20.82
	av25	24.69	25.67	22.34	19.96	19.51	22.17
	av60	23.29	25.23	21.81	19.09	19.76	21.52
500	av$\Delta\nu$	20.56	23.76	21.11	18.15	18.11	19.98
	av25	22.06	25.82	22.77	20.81	18.70	20.44
	av60	21.46	24.29	21.64	19.46	18.30	20.87

Table 4
Relaxation 10 bps: $\Delta\boldsymbol{\nu}$ ESG integration potentials. Same as with Table 2, but with regard to 10 basis points of relaxation in expected return.

Problem size n	Quantity	Upper bound ω_i					
		0.02	0.03	0.04	0.05	0.08	0.12
50	av$\Delta\nu$	–	48.30	41.56	33.66	28.07	30.02
	av25	–	55.07	46.70	36.83	30.69	32.64
	av60	–	48.94	42.18	35.10	28.53	31.17
100	av$\Delta\nu$	41.88	36.15	33.14	28.65	26.99	28.93
	av25	46.66	40.07	36.14	31.49	29.80	30.69
	av60	43.18	36.65	34.15	29.31	27.26	30.29
150	av$\Delta\nu$	36.14	34.33	30.85	26.83	26.29	29.52
	av25	40.10	37.87	33.24	29.20	28.11	31.94
	av60	36.64	35.35	32.06	27.70	27.23	30.33
250	av$\Delta\nu$	32.05	33.62	29.49	26.13	26.65	29.86
	av25	34.86	36.11	32.06	28.47	27.93	31.79
	av60	33.08	35.13	30.72	27.00	28.01	31.09
500	av$\Delta\nu$	29.23	33.50	30.07	27.29	26.17	28.88
	av25	31.47	36.34	32.64	29.52	27.26	29.83
	av60	30.46	34.35	31.17	27.91	27.57	29.88

Table 5
Relaxation 20 bps: $\Delta\boldsymbol{\nu}$ ESG integration potential. Same as with Table 2, but with regard to 20 basis points in relaxation in expected return.

Problem size n	Quantity	Upper bound ω_i					
		0.02	0.03	0.04	0.05	0.08	0.12
	av$\Delta\nu$	–	64.18	56.35	46.78	39.67	42.09
50	av25	–	73.79	63.75	51.85	43.71	46.04
	av60	–	64.68	57.43	48.25	40.49	43.71
	av$\Delta\nu$	56.98	49.98	46.12	40.27	37.97	40.77
100	av25	64.24	55.55	50.67	44.13	41.99	43.43
	av60	58.86	50.96	47.70	41.34	38.41	42.55
	av$\Delta\nu$	49.74	47.50	43.31	37.83	37.40	41.47
150	av25	55.49	52.47	47.01	41.04	40.01	44.87
	av60	50.99	49.04	44.86	26.28	39.10	42.79
	av$\Delta\nu$	48.88	46.58	41.70	36.91	37.70	42.23
250	av25	49.01	50.25	45.50	40.14	39.76	45.09
	av60	46.53	48.20	43.24	38.02	39.29	43.92
	av$\Delta\nu$	41.21	46.60	42.35	38.52	37.39	41.44
500	av25	44.55	50.53	45.54	41.51	39.40	43.35
	av60	42.95	47.94	43.45	39.57	38.98	42.93

(i.e., combinations of size n and upper bound ω_i). We refer to the location in the table with the three quantities of an experimental setup as a "cell" (where the 5 rows and 6 columns in a table intersect). Of the 30 cells, only 29 have entries in them (for the 29 useable scenarios). The quantities reported in each are av$\Delta\nu$, av25, and av60. The way the av$\Delta\nu$ values in the cells are computed is as follows. Consider the *bp* value of the cell. Then for each of the 100 test problems of the cell, we compute the average $\Delta\boldsymbol{\nu}$ amount by which points along that problem's Mod 1 M-V efficient frontier can be advanced across the efficient surface toward the point of maxESG when relaxing expected return by the *bp* amount of the cell. Then by averaging the 100 average amounts, we have av$\Delta\nu$. For the av25 entry in each cell, it is obtained by averaging over the 100 test problems of the cell the $\Delta\boldsymbol{\nu}$ amount by which the point 25% variance-wise along a problem's Mod 1 M-V efficient frontier can be advanced across the efficient surface when relaxing expected return by the *bp* amount of the cell. Similarly, but for the point 60% along the Mod 1 M-V efficient frontier, we have av60.

Concerning computer times, consider for instance the 500×0.04 cell in Table 2. Along with 9.13% for av$\Delta\nu$, 9.75% for av25, and 9.34% for av60, only 35.10 seconds per problem on average were required for doing all of the work to compute the numbers of this cell demonstrating that CPU times are not an obstacle to this research.[8]

The av25 and av60 points are included to show how well the results indicated by av$\Delta\nu$ hold for points typical of being in the realm of a conservative investor, and hold for points typical of being in the realm of an aggressive investor, respectively. In other words, the av$\Delta\nu$, av25 and av60 index the amounts of ESG that our experiments indicate are being left on the table when non-integrating. That is, they are being left on the table as a consequence of the fact that they lie in regions of a serious ESG investor's portfolio problem that are unfindable by the bi-criterion techniques used by today's ESG mutual funds as the tri-criterion efficient surface and NC-efficient front techniques of this paper are unknown to them at the moment. In line with our definition of the ESG integration index, av$\Delta\nu$, av25 and av60 in the tables are reported in percentage values.

Comparing av$\Delta\nu$ values in cells across the tables, we get a sense of what different relaxations are able to buy in terms of increased ESG. Consider, for instance, the 250×0.04 cells in the tables. In Table 2, 1 bp in relaxation buys for us on average 8.87% more ESG, and in Table 3, 5 bps in relaxation buy for us on average 20.62% more in ESG. Continuing, in Table 4, 10 bps in relaxation buy for us on average 29.49% more in ESG, and in Table 5, 20 bps in relaxation buy for us on average 41.70% more in ESG. For ESG investors, it is useful to interpret $\Delta\boldsymbol{\nu}$ values in a Laffer Curve sense (Buchanan & Lee, 1982). Whereas 0% ESG integration is not good, neither would be 100% ESG integration (because of the likely condition of risk and return at the point of maxESG), but somewhere in between is where one would find a happier home.

Looking again into the tables we see two secondary effects. One is the tendency for ESG integration to decrease when security upper bounds are increased. The other is the tendency for ESG integration to decrease when problem size increases. While both might seem counter-intuitive as one might think the efficient surface, which stretches from its Mod 1 M-V boundary to the point of maximum ESG score, would get bigger, it actually tends to get smaller. This is because the efficient surface moves with these two changes in condition, but the maximum ESG score end of the efficient surface generally moves less than the M-V end, creating a compression effect on the efficient surface.

Given the $\Delta\boldsymbol{\nu}$ ESG integration gains possible for the amounts of expected return to be given up, it is hard to see how any serious ESG investor wouldn't find ample tradeoffs in the tables worth pursuing in the interest of gaining further portfolio satisfaction and increased utility. While we used Sustainalytics ESG scores in this paper, the results reported in Tables 2–5 have shown themselves to be robust in the sense that we have experimented with other schemes for scoring securities for use in $\boldsymbol{\nu} \in \mathbb{R}^n$ (such as discretely on the basis of 1 to 5 to simulate Morningstar globes, and continuously on the range of 0 to 1 to simulate total randomness), but we have noted little difference in the resulting percentages. Thus, for instance, a relaxation of 1 basis point always seems to get us about 10% more $\Delta\nu$, a relaxation of 5 basis points always seems to get us about 25% more $\Delta\nu$, a relaxation of 10 basis points always seems to get us about 35% more $\Delta\nu$, and so forth.

5. Mod 2 M-V-ESG implementation

In the light of the tradeoffs seen, this section shows how to work with the Mod 2 M-V-ESG procedure to take advantage of them. With a serious ESG investor's optimal portfolio on the efficient surface, and most likely in the center somewhere, the challenge is how to search an efficient surface for wherever we may

[8] All of the computational experiments of this paper were run using Python calling Gurobi 9.5 on an HP Z2 Tower G5 Desktop with 32 GB of RAM.

Table 6
Steps of Mod 2 M-V-ESG Procedure. Mod 2 seen as a tri-criterion add-on to bi-criterion Mod 1.

Markowitz-like M-V-ESG procedure of Mod 2	
2.0.	Perform the steps of Markowitz (1952) as outlined in Section 2.2 to obtain a most preferred portfolio on the Mod 1 M-V efficient frontier noting its variance as υ^o.
2.1.	For the n securities on the approved list, load (4) and (5) with the problem's ESG coefficient vector $\boldsymbol{\nu} \in \mathbb{R}^n$.
2.2.	For bp values of the user's choice, construct an NC-efficient fronts graph (as for instance in Figs. 3 and 6).
2.3.	Using υ^o as a reference variance, look upward into the graph to evaluate its contents.
2.4.	Select the most preferred point from among those on the NC-efficient fronts.
2.5.	After taking the inverse image of the point chosen, the decision maker's M-V-ESG optimal portfolio is obtained.

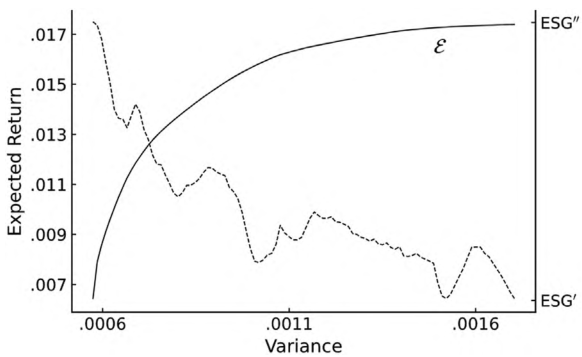

Fig. 5. Efficient frontier $\mathcal{E}$ and dashed line of the second 100×0.03 problem. Note how the dashed line of this problem is not only considerably different from the one of Fig. 1, but its minimum and maximum points occur at completely different places relative to the problem's Mod 1 M-V efficient frontier as there appears to be no predicable pattern for dashed lines in any problem.

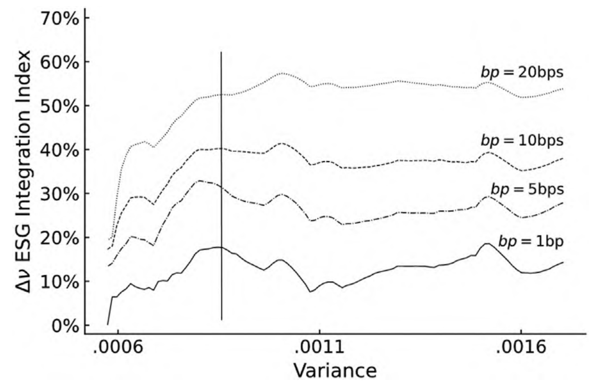

Fig. 6. NC-efficient fronts graph for the second 100×0.03 problem. While different, note the similarity of the $bp = 1, 5, 10$ and 20 NC-efficient fronts of this problem to the NC-efficient fronts of the problem of Fig. 3.

need to go on it. With little in the literature on this, we use the NC-efficient front approach developed in this paper.

With NC-efficient fronts, steps for implementing the Mod 2 M-V-ESG procedure of this paper are given in Table 6. For getting started, let us do so in Step **2.0** from the Mod 1 of a different 100×0.03 problem to show how problems can vary in detail but be the same in principle. For the Mod 1 of the new problem, consider Fig. 5. While its efficient frontier is quite similar, note how its dashed line is considerably different from the one in Fig. 1, but as discussed earlier, this is no surprise because it is a different problem. Let the selection from $\mathcal{E}$ be the point of 25% variance-wise along the efficient frontier whose variance is $\upsilon^o = 0.00082$. With this variance as a reference point, we enter Step **2.1**. With Fig. 5, but without the dashed line, fairly well describing how today's ESG mutual funds operate, mutual funds don't have to do anything different from what they are currently doing before formally entering Mod 2 in Step **2.1**. At this point it is just a matter of whether they are willing to do more, that is, let Mod 2 guide them into regions of the problem that cannot be sighted from Mod 1.

After loading (4) and (5) with the problem's ESG vector $\boldsymbol{\nu} \in \mathbb{R}^n$, we proceed to Step **2.2** for the selection of bp values for NC-efficient fronts we would like to consider. One way to help in the selection would be to thumb through the 100×0.03 cells of Tables 2–5 as these cells match our problem. Let's say the $\Delta\nu$ values in the 100×0.03 cell of Table 4 look good. This is for 10 bps of relaxation. Wishing to see a few surrounding relaxations for context, let's say the investor in Step **2.2** asks for an NC-efficient fronts graph with NC-efficient fronts of 1, 5, 10 and 20 basis points displayed on it. In this instance, Fig. 6 is returned. Note how the plots in it are a bit different from those in Fig. 3, but this is normal and to be expected.

Now, in Step **2.3** with $\upsilon^o = 0.00082$, looking up into Fig. 6 from this variance we see $\Delta\boldsymbol{\nu}$ values of 17%, 30%, 39% and 52% on the four NC-efficient fronts we chose to display. With the 39% on the $bp = 10$ NC-efficient front looking ideal, let's say the investor in Step **2.4** decides to go with this point on the graph as most preferred. Note that it is not mandatory to look up into the graph from υ^o. Actually, we could look up into the graph from any variance. It is just that υ^o seems like a good suggestion to keep things lined up. Furthermore, we need not stay on the straight line for any υ^o. A user is free to veer off the line to the right or left as appropriate when making his or her final selection.

For determining the inverse image of our final point, that is, the **x**-vector generating it, we solve two optimization problems. The first, parameterized with υ^o for σ_i^2, is (3). The second, parameterized with υ^o, bp and the expected return obtained from the solution of (3) for μ_i, is (5). Then, the resulting **x**-vector solution gives us in Step **2.5** our optimal portfolio. Now, in knowing this portfolio, we know exactly how we got to it. Thus, in the event we might wish to go back and see what might result if we changed anything, we would know how to do that.

Mod 2 enables us to display a graph showing the projection of the efficient surface of our second 100×0.03 problem onto the mean-variance plane in Fig. 7. The northwest boundary of the projection is the Mod 1 M-V efficient frontier $\mathcal{E}$ shown in Fig. 5. While standard procedures would only reveal the efficient frontier $\mathcal{E}$, the rest of the projection of the efficient surface consists of portfolios made efficient by virtue of the third (i.e., ESG) objective and are thus only knowable by means of Mod 2. Thus, the efficient surface is no trivial enlargement of the Mod 1 M-V efficient frontier as the areas alluded to earlier as cannot be seen from Mod 1 are indeed of considerable size.

In interpreting Fig. 7, the efficient surface, with its curvature, is like an octant-portion of the surface of a tennis ball. Thus, when looking at it sideways as in the figure, most of what is seen is the portion of the surface that is most perpendicular to us. In a problem in which all objectives are of equal or near equal importance (not unreasonable for serious ESG investors), the optimal solution would be expected to be somewhere near the geometric center of the efficient surface. This means that if we were to leave from a middle point on the Mod 1 M-V efficient frontier and then ad-

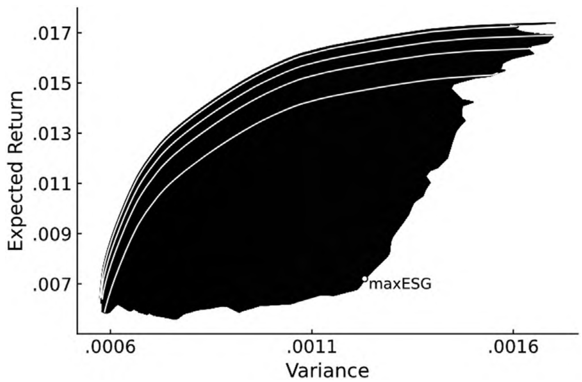

Fig. 7. Efficient surface and four NC-efficient fronts as projected onto the M-V plane. For the second of our two 100×0.03 illustrative problems, this is its efficient surface and its $bp = 1, 5, 10$ and 20 NC-efficient fronts projected onto the M-V plane. The northwest boundary of the projection is the Mod 1 M-V efficient frontier $\mathcal{E}$ shown in Fig. 5. With the Mod 1 M-V efficient frontier furthest away from the reader, the $bp = 1$ NC-efficient front is the white line closest to it, and the $bp = 20$ NC-efficient front is the white line closest us. Note that the white lines, being NC-efficient fronts and not contours or iso-quants, also bulge outward toward the viewer in reality.

vance across the efficient surface in the direction of the maxESG point, we wouldn't need to go much beyond one-third of the way to reach the geometric center. In looking again at the tables of the previous section, particularly Tables 4 and 5, one can deduce that for most problem sizes with most upper bounds, advances up to over about 35% can generally be achieved by relaxations in expected return in the neighborhood of about 10–12 bps/month, or 1.2 to 1.44%/year.

In Fig. 7, the white lines show the $bp = 1, 5, 10$ and 20 NC-efficient fronts of our 100×0.03 problem as projected onto the M-V plane. While they appear flat to the M-V plane, not being contours or iso-quants, they actually bulge outward toward the viewer. Consequently, projections as in Fig. 7 are not highly usable for analytical purposes. But since every efficient point of a projection corresponds to a point on an NC-efficient fronts graph, graphs such as in Fig. 6 turn out to be much more productive when attempting to pinpoint a final solution since an NC-efficient fronts graph is like a map spread out flat on a table.

6. Concluding remarks

For serious ESG investors, optimal portfolios are not on the M-V efficient frontier, but are to be found in the interior of the M-V-ESG efficient surface. Unfortunately, today's standard procedures for constructing ESG portfolios, screening in the first stage and optimization over the securities permitted by the first stage for risk and return in the second stage are unable to compute solutions in the interior of an efficient surface. At most, they are able to compute a portion of an edge of the efficient surface. It is like trying to fight a war with a map upon which one can only see a line across it. The difficulty here is that the portfolio problem of a serious ESG investor is in 3D-space, but today's customary procedures are only able to probe in 2D.

In this paper, the problem of a serious ESG investor is conceptualized from the outset as a three-criterion problem. Using special techniques called NC-efficient fronts (not available anyplace else), we are not only able to characterize an efficient surface, but we are able to dragnet it for points of best risk/return/ESG tradeoff on it.

Extensive computational tests are conducted to show the effectiveness of searching efficient surfaces in this fashion across problems with up to 500 securities and differently restricted feasible regions. Also, while ESG integration has been only a concept in ESG investing, in this paper we provide an operationally applicable definition of ESG integration and have developed an index for quantitatively assessing the degree of ESG integration present in a mutual fund based upon how deeply the mutual fund's portfolio is situated within the interior of its efficient surface. Clearly, it would be ideal to be able to accommodate cardinality constraints in our models, but portfolio problems with cardinality constraints are NP-hard and thus only heuristics are available, so this is something for the future. Another area for future research is the investigation of the feasibility of NC-efficient portfolios based on real-world data and the comparison of the performance of NC-efficient portfolios in an out-of-sample analysis.

With regard to this research leading to other things, there are life-cycle and target-date mutual funds with various horizon dates. Now, with NC-efficient fronts, there can be ESG mutual funds possessing various NC-exposure levels (such as at 5, 10, 15 and 20 basis points), so that investors can select the particular mutual fund that best matches their ESG intentions as we now have, with efficient surfaces and NC-efficient fronts, a theoretical and methodological basis upon which to construct such mutual funds. Also, while we have only focused on ESG, it appears that the efficient surface/NC-efficient front content of this paper could well be applied to other more specific third criterion portfolio situations such as in decarbonization (Trinks, Scholtens, Mulder, & Dam, 2018), renewable energy (Rezec & Scholtens, 2017), and even Shariah compliance in Islamic investing (Masri, 2017).

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A

In mathematical programming, Eq. (5) is recognized as an e-constraint program. Then from e-constraint results covered for instance in Miettinen (1999) and other places, it is known that the criterion vector of any $\mathbf{x} \in S$ optimizing (5) is weakly nondominated, where a criterion vector $\bar{\mathbf{z}} \in Z \subset \mathbb{R}^3$ in (5) is *weakly-nondominated* if and only if there exists no other $\mathbf{z}$ in Z such that $z_1 < \bar{z}_1$, $z_2 > \bar{z}_2$ and $z_3 > \bar{z}_3$. Observing the strictness of the inequalities, it is seen that the concept of a weakly-nondominated criterion vector is only slightly more general than that of a nondominated criterion vector. In this way, the set of all weakly-nondominated criterion vectors of (5) subsumes the set of all nondominated criterion vectors of (5) where a nondominated criterion vector of (5) is as defined in Section 3.

Furthermore, it is known from Miettinen (1999) that from among the solutions that optimize an e-constraint program there is at least one whose criterion vector is nondominated. Since it is known from Hirschberger et al. (2013) that the efficient surface of a tri-criterion program of form (4) is platelet-wise paraboloidic, (5) with its linear objective function will only ever have one optimal solution. Thus, the criterion vector of the optimal solution of (5), being nondominated, is on the M-V-ESG efficient surface. Should one be concerned that there could possibly be a flat platelet, one could add extra variables to the formulation as described in Ehrgott & Ruzika (2008).

References

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

Belton, V., & Stewart, T. J. (2002). *Multiple criteria decision analysis: An integrated approach*. Kluwer, Boston.

Bilbao-Terol, A., Arenas-Parra, M., & Cañal-Fernández, V. (2012). A fuzzy technology multiobjective approach for sustainable investments. *Expert Systems with Applications, 39*(12), 10904–10915.

Buchanan, J. T., & Lee, D. R. (1982). Politics, time, and the Laffer curve. *Journal of Political Economy, 90*(4), 816–819.

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

Chen, M., & Mussalli, G. (2020). An integrated approach to quantitative ESG investing. *Journal of Portfolio Management, 46*(3), 65–74.

Ehrgott, M. (2005). *Multicriteria optimization* (2nd ed.). Springer.

Ehrgott, M., & Ruzika, S. (2008). Improved e-constraint method for multiobjective programming. *Journal of Optimization Theory and Applications, 138*, 375–396.

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

Garcia-Bernabeu, A., Salcedo, J. V., Hilario, A., Plà-Santamaría, D., & Herrero, J. M. (2019). Computing the mean-variance-sustainability nondominated surface by ev-MOGA. *Complexity*. https://doi.org/10.1155/2019/6095712.

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

Hartzmark, S. M., & Sussman, A. B. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *The Journal of Finance, 74*(6), 2789–2837.

Hirschberger, M., Qi, Y., & Steuer, R. E. (2007). Randomly generating portfolio-selection covariance matrices with specified distributional characteristics. *European Journal of Operational Research, 177*(3), 1610–1625.

Hirschberger, M., Qi, Y., & Steuer, R. E. (2010). Large-scale MV efficient frontier computation via a procedure of parametric quadratic programming. *European Journal of Operational Research, 204*(3), 581–588.

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

Korhonen, P., & Wallenius, J. (1988). A Pareto race. *Naval Research Logistics, 35*(6), 615–623.

Korhonen, P. J., & Laakso, J. (1986). A visual interactive method for solving the multiple criteria problem. *European Journal of Operational Research, 24*, 277–287.

Liagkouras, K., Metaxiotis, K., & Tsihrintzis, G. (2020). Incorporating environmental and social considerations into the portfolio optimization process. *Annals of Operations Research*. https://doi.org/10.1007/s10479-020-03554-3.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*(1), 77–91.

Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification in investments*. John Wiley, New York.

Masri, H. (2017). A Shariah-compliant portfolio selection model. *Journal of the Operational Research Society, 14*(4), 368–378.

Merton, R. C. (1972). An analytical derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis, 7*(4), 1851–1872.

Miettinen, K., Eskelinen, P., Ruiz, F., & Luque, M. (2010). NAUTILUS method: An interactive technique in multiobjective optimization based on the nadir point. *European Journal of Operational Research, 206*, 426–434.

Miettinen, K. M. (1999). *Nonlinear multiobjective optimization*. Kluwer, Boston.

Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2022). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics, 142*(2), 572–597.

Rezec, M., & Scholtens, B. (2017). Financing energy transformation: The role of renewable energy equity indices. *International Journal of Green Energy, 14*(4), 368–378.

Trinks, A., Scholtens, B., Mulder, M., & Dam, L. (2018). Fossil fuel divestment and portfolio performance. *Ecological Economics, 146*(April), 740–748.

Utz, S., Wimmer, M., Hirschberger, M., & Steuer, R. E. (2014). Tri-criterion inverse portfolio optimization with application to socially responsible mutual funds. *European Journal of Operational Research, 234*(2), 491–498.

Utz, S., Wimmer, M., & Steuer, R. E. (2015). Tri-criterion modeling for constructing more-sustainable mutual funds. *European Journal of Operational Research, 246*(1), 331–338.

Wierzbicki, A. P. (1980). The use of reference objectives in multiobjective optimization. In G. Fandel, & T. Gal (Eds.), *Lecture notes in economics and mathematical systems: vol. 177* (pp. 468–486). Springer-Verlag, Berlin.

Xidonas, P., Mavrotas, G., Krintas, T., Psarras, J., & Zopounidis, C. (2012). *Multicriteria portfolio management*. Springer-Verlag, Berlin.