On Imposing ESG Constraints of Portfolio Selection for Sustainable Investment and Comparing the Efficient Frontiers in the Weight Space

Yue Qi1 and Xiaolin Li2

Abstract
Sustainable investment is typically fulfilled by screening of environmental, social, and governance (ESG); the screening strategies are practical and expedite sustainable-investment development. However, the strategies typically build portfolios by a list of good stocks and ignore portfolio completeness. Moreover, there has been limited literature to study the portfolio weights of sustainable investment in the weight space. In such an area, this article contributes to the literature as follows: We extend a conventional portfolio-selection model and impose ESG constraints. We analytically solve our model by computing the efficient frontier and prove that the frontier’s portfolio weights all lie on a ray (half line). By the ray structure, we prove that portfolio selection for sustainable investment and conventional portfolio selection fundamentally possess highly different portfolio weights. Overall, our aim is comparing the portfolio weights of sustainable portfolio selection and of conventional portfolio selection; the comparison result has been unknown until now. The result is important for sustainable investment because portfolio weights are the foundation of portfolio selection and investments. We sample the component stocks of Dow Jones Industrial Average Index from 2004 to 2013 and find that our efficient frontier and the conventional efficient frontier are quite similar. Therefore, in plain financial language, investors can still obtain risk-return performance similar to conventional portfolio selection after imposing strong ESG requirements, although the portfolio weights can be totally different. The result is both an endorsement and a reminder for sustainable investment.

Keywords
sustainability investment, environmental and social and governance (ESG), portfolio selection, portfolio optimization, constraint, efficient frontier, parabola, ray

Introduction
Sustainable Investment
The Forum of Sustainable and Responsible Investment defines sustainable investment as “an investment discipline that considers environmental, social, and governance (ESG) criteria to generate long-term competitive financial returns and positive societal impact.” Sustainable investment has been rapidly growing since 2000. The Global Sustainable Investment Review reports that global sustainable-investment assets reached $30.7 trillion in 2018 with a 34% increase from 2016.2

Sustainable investment is typically fulfilled by ESG screening as J. P. Morgan Asset Management reports that “Investing sustainably can be achieved by avoiding companies with poor practices in this area, actively selecting leaders in each field.” The Forum of Sustainable and Responsible Investment and the Principles for Responsible Investment Association report similar strategies.

Literature for Sustainable Investment
A large body of sustainable-investment research focuses on the following:

1. Effect of the screening strategies,
2. Comparison between the returns of sustainable-investment mutual funds and conventional mutual funds, and
3. Relationship between ESG performance and financial performance.

1Nankai University, Tianjin, China
2Tianjin University of Finance and Economics, China

Corresponding Author:
Yue Qi, China Academy of Corporate Governance, Department of Financial Management, Business School, Nankai University, 94 Weijin Road, Tianjin 300071, China.
Email: yorkche@nankai.edu.cn
For Research Line 1, Auer (2016) systematically utilizes the strategies, compares sustainable investments against passive investments, and finds mixed results. Trinks and Scholtens (2017) and Bertrand and Lapointe (2015) utilize the strategies and analyze the strategies and risk-based allocation strategies for opportunity and report some positive risk-adjusted performance. In contrast, Amel-Zadeh and Serafeim (2018) survey mainstream investment organizations, find the strategies relatively ineffective, and recommend ESG-integration strategies.

For Research Line 2, the scholars typically compare the returns of sustainable-investment mutual funds and the returns of conventional mutual funds. For example, Capelle-Blancard and Monjon (2014) report that “As of 2011, more than fifty academic papers have examined this issue, all using similar methodology. They almost unanimously show that the financial performance of SRI funds does not differ significantly from their conventional peers” (p. 495).

For Research Line 3, for example, Verheyden et al. (2016) document that “one recent review by Arabesque and Oxford University of over 200 studies reports that 90% of those studies found a positive link between ESG and the cost of capital” (p. 47).

**Strength and Weakness of the Literature**

Because this article is related to Research Lines 1 and 2, we focus on their strength and weakness. On one side, the screening strategies are highly practical and thus substantially expedite sustainable-investment development. However, on the other side, the strategies typically ignore the sustainability of the portfolios as Lydenberg (2016) emphasizes measuring portfolio ESG and integrating system risk in portfolio selection. Moreover, the screening strategies typically build portfolios by a list of good stocks and ignore portfolio completeness because Nobel laureate Markowitz (1959) argues that “we speak of ‘portfolio selection’ rather than ‘security selection.’ A good portfolio is more than a long list of good stocks and bonds. It is a balanced whole” (p. 3).

Research Line 2 basically focuses on only $z^{ESG}$ of sustainable-investment mutual funds and $z^c$ of conventional mutual funds in risk-return space—that is, $(z_1, z_2)$ space—in the upper left part of Figure 1. Therefore, the research methodology can be simplistic.

**Originality of This Article**

Up until now, there has been limited literature to explicitly impose ESG as constraints of portfolio selection for sustainable investment, analyze the (whole) efficient frontiers, and study the portfolio weights in the weight space. Therefore, it has not been answered whether portfolio selection for sustainable investment and conventional portfolio selection fundamentally possess different portfolio weights; that is, are the building blocks for sustainable investment different from those of conventional investment?

In such an area, this article contributes to the literature as follows:
1. We propose our portfolio-selection model for sustainable investment by explicitly imposing ESG as constraints.

2. We then compute the optimal solutions of our model as a thick black curve in the upper right part of Figure 1. We also depict the optimal solutions of the conventional model as a thick gray curve in the upper right part. Therefore, we compare the two curves instead of basically focusing on only \( x_{ESG}^{\text{opt}} \) and \( x_{12}^{\text{opt}} \) in the upper left part.

3. Fundamentally, we compute the thick black curve’s portfolio weights as a black ray in the weight space in the lower left part of Figure 1. We also compute the thick gray curve’s portfolio weights as a gray ray. Therefore, we compare the two rays. Specifically, we quantitatively compare the rays’ apexes in the lower right part (i.e., \( x_{ESG}^{\text{opt}} \) vs. \( x_{12}^{\text{opt}} \)).

We sample the component stocks of Dow Jones Industrial Average Index from 2004 to 2013. We find that the thick black curve and thick gray curve are similar but the black ray and gray ray are totally different. The practical implication is that sustainable investors can still obtain risk-return performance similar to that of conventional portfolio selection after imposing strong ESG requirements although the portfolio weights can be totally different from those of conventional portfolio selection. The result is both an endorsement and a reminder for sustainable investment. Endorsement means that sustainable investors can obtain risk-return performance similar to that of conventional investors even after imposing ESG constraints. Reminder means that sustainable investors should mind the highly different portfolio weights.

Overall, our aim is comparing the portfolio weights of sustainable portfolio selection and of conventional portfolio selection; the comparison result has been unknown until now. The result is important for sustainable investment because portfolio weights are the foundation of portfolio selection and investments. Our comparison methodology is portfolio-selection models of Markowitz (1959). No other scholars have conducted the comparison before. Therefore, our methodology is exploratory and based on modeling instead of testing hypotheses.

The rest of this article is organized as follows: We review portfolio selection and portfolio optimization in the next section. We briefly compare major methods of portfolio optimization in the next section. We impose ESG constraints and construct portfolio-selection models for sustainable investment in the next section. Computationally, we calculate the minimum-variance frontier in the next section. We then calculate the efficient frontiers of our model and conventional models in the weight space in the next section. We illustrate the analysis by an example in the next section. We conclude this article in the next section. We put proofs in the appendix.

Portfolio Selection and Portfolio Optimization

Portfolio Selection

Markowitz (1959) emphasizes that “Two objectives, however, are common to all investors . . . 1. They want ‘return’ to be high . . . 2. They want this return to be dependable” (p. 6). Then, for “a balanced whole,” he formulates portfolio selection by two-objective optimization as follows:

\[
\begin{align*}
\min & \quad z_1 = x^T \Sigma x \\
\max & \quad z_2 = \ldots^T x \\
\text{s.t.} & \quad x \in S,
\end{align*}
\]

where for \( n \) stocks, \( \boldsymbol{x} \) is a vector of stock expected-returns, \( \Sigma \) is a covariance matrix of stock returns, and superscript \( T \) denotes transposition. Brockwell and Davis (1987) define covariance matrices as symmetric and positive semidefinite. Moreover, \( \mathbf{z}_1 \) and \( \mathbf{z}_2 \), respectively, measure the variance and expectation of the return of portfolio \( \mathbf{x} \). Furthermore, \( \mathbf{S} \subset \mathbb{R}^n \) is a feasible region; \( \mathbf{Z} = \{ \mathbf{z}_1, \mathbf{z}_2 \} | x \in S \) is the feasible region in \((\mathbf{z}_1, \mathbf{z}_2)\) space.

Markowitz (1959) calls the optimal solutions of Equation 1 as an efficient frontier. For example, the efficient frontier is depicted as a thick gray curve in the upper right part of Figure 1.

Markowitz (1959) then suggests “first, separates efficient from inefficient portfolios; second, portrays . . . efficient portfolios; third, . . . select the combination . . . that best suits his needs; and fourth, determines the portfolio” (p. 7). He computes (whole) efficient frontiers and then asks investors to pinpoint their preferred portfolios from the frontiers, so investors never worry about missing some parts of the frontiers. Therefore, computing (whole) efficient frontiers is important for his suggestions.

Portfolio Optimization

Sharpe (1970) and Merton (1972) study the following model:

\[
\begin{align*}
\min & \quad z_1 = \mathbf{x}^T \Sigma \mathbf{x} \\
\max & \quad z_2 = \ldots^T \mathbf{x} \\
\text{s.t.} & \quad \mathbf{I}^T \mathbf{x} = 1,
\end{align*}
\]

where \( \mathbf{I} \) is a vector of ones. Merton (1972) analytically solves the model; “analytical” means in the form of formulae. Equation 2 is a fundamental model of portfolio selection and thus described in classic finance textbooks (e.g., those of Bodie et al., 2018; Campbell, 2018; and Huang and Litzenberger, 1988). Moreover, Equation 2 serves as a basis of capital asset pricing models.

Merton (1972) applies an \( \varepsilon \)—constraint method (as described by Steuer, 1986) to Equation 2 as follows:
\[
\min \{ z_i = x^T \Sigma x \}
\]
\[
s.t. \mu^T x = \epsilon
\]
\[
1^T x = 1,
\]
where \( \epsilon \) is the parameter. The \([z_1 z_2]^T\) set of all optimal solutions of Equation 3 with varying \( \epsilon \) is called minimum-variance frontier of Equation 2. The frontier is the boundary of the feasible region \( Z \) of Equation 2. Merton (1972) analytically computes the frontier and proves it as a parabola as follows:
\[
z_i = \frac{1}{af - c^2} \left( f z_2^2 - 2cz_1 + a \right),
\]
where symbols \( a, c, \) and \( f \) are introduced as follows:
\[
a = \mu^T \Sigma^{-1} \mu, c = 1^T \Sigma^{-1} \mu, f = 1^T \Sigma^{-1} 1.
\]

The minimum-variance frontier of Equation 4 is depicted as a gray curve in the upper right part of Figure 1. The efficient frontier is the minimum-variance frontier’s upper portion and depicted as a thick gray curve. Merton (1972) computes the minimum-variance portfolio of Equation 2 (i.e., the optimal solution of \( \min \{ x^T \Sigma x \} \), s.t. \( 1^T x = 1 \)) as follows:
\[
x_{mv} = \frac{1}{f} \Sigma^{-1} 1.
\]

Merton (1972) computes the \( x \) set of the efficient frontier of Equation 2 as:
\[
E(\frac{1}{2}) = \left\{ x \in \mathbb{R}^n \mid x = \lambda j(\frac{1}{2}) + x_{mv} \mid \lambda \in [0, \infty) \right\},
\]
where
\[
j(\frac{1}{2}) = \frac{1}{2} \left( \Sigma^{-1} \mu - \frac{c}{f} \Sigma^{-1} 1 \right).
\]

The set is a ray; that is, the ray is generated by \( j(\frac{1}{2}) \) and then translated to \( x_{mv} \). The ray is depicted in the lower left part of Figure 1.

**Comparing Major Portfolio-Optimization Methods**

We briefly compare major portfolio-optimization methods and depict the result, advantage, and disadvantage in Figure 2.

**Analytical Methods**

Analytical methods are based on calculus and linear algebra and thus understandable. Merton (1972) analytically solves Equation 2 and obtains complete optimal solution information. For example, the minimum-variance frontier \( z_1 = 0.350z_2^2 - 0.070z_2 + 0.011 \) is depicted as a parabola in the first part of Figure 2. The efficient frontier is the upper portion and depicted as a thick curve. Qi et al. (2017) and Qi and Steuer (2020) extend Equation 2 and analytically solve multiple-objective portfolio selection. Analytical methods’ unique advantage is to readily compute the minimum-variance frontier including the efficient frontier.

However, analytical methods are for models with equality constraints only. Unpractically unbounded portfolio weights are thus allowed. Nevertheless, with almost all results in formulae, analytical methods bypass the need of mathematical programming and bring convenience in research and instruction. Huang and Litzenberger (1988) emphasize Equation 2 as “popular because of its analytical tractability and its rich empirical implications” (p. 60).

**Parametric Quadratic Programming**

Markowitz and Todd (2000) deploy parametric quadratic programming and propose a critical-line algorithm to solve Equation 1. Markowitz and Todd (2000) prove that the efficient frontier is piece-wise made up by connected parabolic segments. For example, the frontier consists of two segments in the second part of Figure 2. The segment between \( z^1 \) and \( z^2 \) is a portion of parabola \( z_1 = 3z_2^2 - z_2 + 3 \); the segment between \( z^2 \) and \( z^3 \) is a portion of parabola \( z_1 = 9z_2^2 - z_2 + 9 \). Best (1996), Stein et al. (2008), and Hirschberger et al. (2010) also suggest their parametric quadratic programming algorithms. However, parametric quadratic programming is difficult to understand and thus rarely utilized.

**Repetitive Quadratic Programming**

Repetitive quadratic programming can be formulated as Equation 3. A group of \( \epsilon \) is preset. Equation 3 is then repetitively solved for each \( \epsilon \); the optimal solutions are obtained for approximating the efficient frontier.

Repetitive quadratic programming is easier to understand than parametric quadratic programming and thus described in classic textbooks (e.g., those of Bodie et al., 2018; Elton et al., 2014; and Best, 2010). However, repetitive quadratic programming obtains only incomplete information of efficient frontiers by discrete approximations and thus cannot reveal the frontier structure. In the third part of Figure 2, the gray curve is copied from the second part as a comparison; repetitive quadratic programming obtains three points: \( z^1, z^3, \) and \( z^6 \). The gray curve is insufficiently approximated by the points.

**Heuristic Methods**

Some models with additional nonlinear formulation are difficult to solve; heuristic methods thus become important provisional tools. Evolutionary algorithms, tabu search, and simulated annealing can be common approaches (as reported by Woodside-Oriakhi et al., 2011).
However, heuristic methods inherently provide suboptimal solutions and only incomplete information of efficient frontiers and cannot reveal the frontier structure. In the fourth part of Figure 2, the methods obtain three points $z^7$, $z^8$, and $z^9$. Three broken lines demonstrate the difference between suboptimal solutions and optimal solutions.

**Imposing ESG Constraints and Constructing Portfolio-Selection Models for Sustainable Investment**

We extend Equation 2, explicitly impose ESG as constraints, and obtain the following general model:
\[ \min \{ z_1 = x^T \Sigma x \} \]
\[ \max \{ z_2 = \mu^T x \} \]
\[ \text{s.t.} \ 1^T x = 1, \]
\[ e^T x = e \]
\[ s^T x = s \]
\[ g^T x = g \]
\[ \vdots \]

where \( e, s, \) and \( g \) are vectors of expected ESG of the stocks, respectively. \( e, s, \) and \( g \) are expected environmental, social, and governance of the portfolio, respectively. Symbol \( \vdots \) denotes other constraints (e.g., thematic exposure) in Equation 10.

We introduce the following symbols:

\[
A = \begin{bmatrix} 1^T \\ e^T \\ s^T \\ \vdots \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ e \\ s \\ g \end{bmatrix}. \]

We then rewrite Equation 9 in matrix format as follows:

\[
\min \{ z_1 = x^T \Sigma x \} \]
\[
\max \{ z_2 = \mu^T x \} \]
\[ \text{s.t.} \ Ax = b. \]  

We inherit the assumptions of Merton (1972) and extend them as follows:

**Assumption 1.**
*As a covariance matrix, \( \Sigma \) is invertible and thus positive definite.*

**Assumption 2.**
*Vector \( \mu \) and all row vectors of \( A \) (altogether \( m+1 \) vectors) are linearly independent.*

The assumptions are not strict prerequisite because of the following:

- By index models in finance (e.g., Bodie et al., 2018, Chapter 8), we obtain \( \Sigma = \beta \Omega \beta^T + D \) where for \( n \) stocks and \( i \) indexes, \( \beta \) is an \( n \times i \) coefficient matrix, \( \Omega \) is an \( i \times i \) covariance matrix of the \( i \) indexes, and \( D \) is an \( n \times n \) diagonal matrix with positive diagonal elements. With \( \beta \Omega \beta^T \) as positive semidefinite and \( D \) as positive definite, \( \Sigma \) is positive definite.
- If the \( m+1 \) vectors are dependent, some vectors are linearly combined and the corresponding objective function or constraints can be consequently deleted.

With the vectors for Equation 9 as \( n \)-vectors, we know \( (m+1) \leq n \) by Assumption 2.

**Analytically Deriving the Minimum-Variance Frontier of Equation 9**

We follow Merton (1972) and also apply an \( \epsilon \)-constraint method to Equation 11 and obtain

\[
\min \{ z_1 = x^T \Sigma x \} \]
\[ \text{s.t.} \mu^T x = \epsilon \]
\[ Ax = b, \]  

where \( \epsilon \) is the parameter. The application’s financial interpretation is taking \( \epsilon \) as a target expected portfolio return and minimizing portfolio variance with \( \mu^T x = \epsilon \) and \( Ax = b \). The set of \( \{ z_1, z_2 \} \) vectors of all optimal solutions of Equation 12 with varying \( \epsilon \) is the minimum-variance frontier of Equation 9 and is an extension of the minimum-variance frontier of Equation 2. We introduce the following symbols:

\[
H = \begin{bmatrix} \mu^T \\ A \end{bmatrix}^{(m+1) \times n}, \quad f = \begin{bmatrix} \epsilon \\ b \end{bmatrix}^{(m+1) \times 1}. \]

We then rewrite all constraints of Equation 12 as \( Hx = f \). The system of linear equations \( Hx = f \) is always solvable by Assumption 2. To solve Equation 12, we apply a Lagrange-multiplier method and construct:

\[
L(x, \ell) = x^T \Sigma x + \ell^T (Hx - f),
\]

where \( \ell \) is an \( (m+1) \)—Lagrange-multiplier vector. Because \( L(x, \ell) \) is a convex function with respect to \( x \) by Assumption 1, \( x \) is the optimal solution of Equation 12 if and only if:

\[
\frac{\partial L}{\partial x} = 2\Sigma x + H^T \ell = 0,
\]

\[
\frac{\partial L}{\partial \ell} = Hx - f = 0,
\]

We premultiply the first partial-derivative equation by \( \Sigma^{-1} \) and obtain

\[
x = -\frac{1}{2} \Sigma^{-1} H^T \ell.
\]

We then substitute \( x \) into the second partial-derivative equation and obtain

\[
(H\Sigma^{-1}H^T)(m+1)(m+1) \ell = -2f.
\]

We propose the following lemma about the invertibility of \( (H\Sigma^{-1}H^T)_{m+1}(m+1) \) for later computations:
Lemma 1.
Matrix \((\Sigma^{-1}H^T)\) is a covariance matrix and positive definite and thus invertible.

Then, we premultiply \((\Sigma^{-1}H^T)f = -2f\) by \((\Sigma^{-1}H^T)^{-1}\) and obtain \(\ell = -2(\Sigma^{-1}H^T)^{-1}f\). We substitute \(\ell\) into \(x = \frac{1}{2}\Sigma^{-1}H^T\ell\) and obtain the optimal solution of Equation 12 as:
\[
x = \Sigma^{-1}H^T(\Sigma^{-1}H^T)^{-1}f.
\]

We substitute \(x\) into \(z = x^T\Sigma x\) of Equation 9 and compute the minimum-variance frontier as:
\[
z = (\Sigma^{-1}H^T(\Sigma^{-1}H^T)^{-1}f)\Sigma(\Sigma^{-1}H^T(\Sigma^{-1}H^T)^{-1}f)\]
\[
= (f^T(\Sigma^{-1}H^T)^{-1}\Sigma(\Sigma^{-1}H^T)^{-1}f)
= f^T(\Sigma^{-1}H^T)^{-1}(\Sigma^{-1}H^T)(\Sigma^{-1}H^T)^{-1}f
= f^T(\Sigma^{-1}H^T)^{-1}f.
\]

We use symbols \(z_2\) instead of \(\epsilon\) of Equation 13 to present the frontier in \((z_1, z_2)\) space of Equation 9 as:
\[
z_1 = \left[ z_2 - b^T(\Sigma^{-1}H^T)^{-1}\right]
\]
\[
\begin{bmatrix}
z_2 \\
b
\end{bmatrix}
\].

The minimum-variance frontier of Equation 14 is more complicated than Equation 4. Moreover, we prove the frontier also as a parabola in the following theorem:

Theorem 1.
The minimum-variance frontier of Equation 14 is a parabola in \((z_1, z_2)\) space of Equation 9.

By Theorem 1 and Equation 12, the parabola of Equation 14 is the boundary of the feasible region \(Z\) of Equation 9. Moreover, the efficient frontier of Equation 9 is the parabola’s upper portion. We have already depicted the parabola as a black curve in the upper right part of Figure 1.

Analytically Deriving the Efficient Frontier of Equation 9

We follow Sharpe (1970) and also apply a weighted-sums method (as described by Steuer, 1986) to Equation 11 and obtain:
\[
\min \left\{ x^T\Sigma x - \lambda \mu^T x \right\}
\]
\[
s.t. Ax = b,
\]
where \(\lambda\) is the parameter. The application’s financial interpretation is taking \(\lambda\) as a risk-aversion parameter, computing the expected utility (i.e., \(-x^T\Sigma x + \lambda \mu^T x\)), and maximizing the expected utility. We then apply a Lagrange-multiplier method to Equation 15 and construct:
\[
L(x, \ell) = x^T\Sigma x - \lambda \mu^T x + \ell^T (Ax - b),
\]
where \(\ell\) is an \(m\)—Lagrange-multiplier vector. For notation-consistency reasons, we still utilize symbols \(L(x, \ell)\) and \(\ell\), but they are different from those in the previous section. Because \(L(x, \ell)\) is a convex function with respect to \(x\) by Assumption 1, \(x\) is the optimal solution of Equation 15 if and only if:
\[
\frac{\partial L}{\partial x} = 2\Sigma x - \lambda \mu + A^T\ell = 0,
\]
\[
\frac{\partial L}{\partial \ell} = Ax - b = 0.
\]

We premultiply the first partial-derivative equation by \(\Sigma^{-1}\) and obtain:
\[
A\Sigma^{-1}\lambda = (A\Sigma^{-1}A^T)\ell - 2b = 0
\]
\[
(A\Sigma^{-1}A^T)\ell = A\Sigma^{-1}\lambda - 2b.
\]

We know \((A\Sigma^{-1}A^T)\) as invertible by the proof of Lemma 1 and premultiply the equation above by \((A\Sigma^{-1}A^T)^{-1}\). We then obtain:
\[
\ell = (A\Sigma^{-1}A^T)^{-1}(A\Sigma^{-1}\lambda - 2b),
\]
and substitute \(\ell\) into \(x = (1/2)\Sigma^{-1}(\lambda - A^T\ell)\) as:
\[
x = \frac{1}{2}\Sigma^{-1}\lambda - \frac{1}{2}\Sigma^{-1}A^T(A\Sigma^{-1}A^T)^{-1}(A\Sigma^{-1}\lambda - 2b)
\]
\[
= \frac{1}{2}\Sigma^{-1}\lambda - \frac{1}{2}\Sigma^{-1}A^T(A\Sigma^{-1}A^T)^{-1}(A\Sigma^{-1}\lambda - 2b)
\]
\[
\Sigma^{-1}A^T(A\Sigma^{-1}A^T)^{-1}b.
\]

where \(I_n\) is an \(n \times n\) identity matrix. The optimal solution of Equation 15 is:
\[
x = \frac{\lambda}{2}\Sigma^{-1}(I_n - A^T(A\Sigma^{-1}A^T)^{-1}A\Sigma^{-1})\mu + \Sigma^{-1}A^T(A\Sigma^{-1}A^T)^{-1}b.
\]
To demonstrate the \( x \) set’s properties, we introduce an \( n \times 1 \) symbol as follows:

\[
\hat{x}_0 = \frac{1}{2} \sum (I_n - A^T \left( A \Sigma^{-1} A^T \right)^{-1} A \Sigma^{-1}) \mu.
\]  

(16)

where subscript (9) is for Equation 9. We then rewrite the optimal solution as:

\[
x = \langle /b \rangle \lambda \hat{x}_0 + \hat{x}_0\text{mv},
\]

(17)

\[
x_0\text{mv} = \Sigma A^T \left( A \Sigma^{-1} A^T \right)^{-1} b,
\]

(18)

where \( x_0\text{mv} \) is the minimum-variance portfolio of Equation 9 with \( \lambda = 0 \) for Equation 15; that is, \( x_0\text{mv} \) is the optimal solution of \( \min \{ x^T \Sigma x \} \), s.t. \( Ax = b \). On the basis of Equation 17, we summarize the result in the following theorem:

Theorem 2.

The \( \langle /b \rangle \lambda \hat{x}_0 + \hat{x}_0\text{mv} \) set of the efficient frontiers of Equation 9 is:

\[
E_9 = \{ x \in \mathbb{R}^n | x = \lambda \hat{x}_0 + \hat{x}_0\text{mv}, \lambda \in [0, \infty) \}.
\]

(19)

The set is a ray; that is, the ray is generated by \( \hat{x}_0 \) and then translated to \( \hat{x}_0\text{mv} \).

We have already depicted the ray in the lower left part of Figure 1.

Comparing the Efficient Frontiers of Equations 2 and 9 in the Weight Space

We have already computed the \( x \) sets of Equations 7 and 19 of the efficient frontiers of Equations 2 and 9. We also have demonstrated Equations 7 and 19 as rays. By the ray structure, we compare the \( x \) sets:

Comparing \( x_{20} \text{mv} \) and \( x_{90} \text{mv} \).

We compare \( x_{20} \text{mv} \) and \( x_{90} \text{mv} \) by proposing a changing ratio as follows:

\[
\delta = \frac{||x_{20} \text{mv} - x_{90} \text{mv}||}{||x_{20} \text{mv}|| + ||x_{90} \text{mv}||},
\]

(20)

where \( || \cdot || \) denotes lengths by \( l^2 \)-norm. We have already depicted the ratio in the lower right part of Figure 1. Because \( ||x_{20} \text{mv}|| + ||x_{90} \text{mv}|| \geq ||x_{20} \text{mv} - x_{90} \text{mv}|| \), we obtain \( 0 \leq \delta \leq 1 \).

- If \( x_{20} \text{mv} \) and \( x_{90} \text{mv} \) completely overlap (i.e., \( x_{20} \text{mv} = x_{90} \text{mv} \)), we obtain \( \delta = 0 \).
- If \( x_{20} \text{mv} \) and \( x_{90} \text{mv} \) are completely opposite (i.e., \( x_{20} \text{mv} = -x_{90} \text{mv} \)), we obtain \( \delta = 1 \).

Comparing the \( x \) Sets

Although \( x_{20} \text{mv} \) of Equation 6 is simpler than \( x_{90} \text{mv} \) of Equation 18, and although \( \hat{x}_0 \) of Equation 8 is simpler than \( \hat{x}_0 \) of Equations 16, 2, and 9, they have similar efficient frontier structure (i.e., rays of Equations 7 and 19). By the ray structure, we prove that portfolio selection for sustainable investment and conventional portfolio selection fundamentally possess highly different portfolio-weight vectors in the following theorem:

Theorem 3.

For the relationship between the \( x \) sets of the efficient frontiers of Equations 2 and 9, there are three exhaustive and exclusive cases: complete overlap, only one intersection, and no intersection.

Complete-overlap case requires \( x_{20} = x_{90} \) and \( \hat{x}_2 = \hat{x}_9 \). It is relatively easy to detect the case by comparing Equation 6 against Equation 18 and comparing Equation 8 against Equation 16.

We then compare the two \( x \) sets by basically checking whether they intersect. If they intersect, they have one identical portfolio-weight vector. Otherwise, they have completely different portfolio-weight vectors (i.e., no similarity). By the ray structure, we formulate the intersection as follows:

\[
\lambda v_{20} \hat{x}_2 + x_{20} \text{mv} = \lambda v_{90} \hat{x}_9 + x_{90} \text{mv}, \quad \lambda_{20}, \lambda_{90} \geq 0.
\]

(21)

We rearrange Equation 21 into the following linear-equation-system format:

\[
\begin{pmatrix} \lambda_{20} \hat{x}_2 - \lambda_{90} \hat{x}_9 \\ \lambda_{20} \end{pmatrix} = (x_{20} \text{mv} - x_{90} \text{mv}) \lambda_{20}
\]

(22)

where \( \lambda_{20} \) and \( \lambda_{90} \) are the unknowns. We construct the augmented matrix for Equation 22 as follows:

\[
\begin{pmatrix} \hat{x}_2 - \hat{x}_9 \lambda_{20} \hat{x}_9 \end{pmatrix} = (x_{20} \text{mv} - x_{90} \text{mv}) \lambda_{20}
\]

We propose the following theorem for the intersection possibility under three exhaustive and exclusive situations:

Theorem 4.

If \( \text{rank} \left( \begin{pmatrix} \hat{x}_2 - \hat{x}_9 \lambda_{20} \hat{x}_9 \end{pmatrix} \right) = \text{rank} \left( \begin{pmatrix} J \end{pmatrix} \right) \), Equation 22 is solvable. Moreover, if there exist \( \lambda_{20} \geq 0 \) and \( \lambda_{90} \geq 0 \), the \( x \) sets of the efficient frontiers of Equations 2 and 9 intersect.

If \( \text{rank} \left( \begin{pmatrix} \hat{x}_2 - \hat{x}_9 \lambda_{20} \hat{x}_9 \end{pmatrix} \right) = \text{rank} \left( \begin{pmatrix} J \end{pmatrix} \right) \) but there do not exist \( \lambda_{20} \geq 0 \) and \( \lambda_{90} \geq 0 \), the \( x \) sets of the efficient frontiers of Equations 2 and 9 do not intersect.

If \( \text{rank} \left( \begin{pmatrix} \hat{x}_2 - \hat{x}_9 \lambda_{20} \hat{x}_9 \end{pmatrix} \right) \neq \text{rank} \left( \begin{pmatrix} J \end{pmatrix} \right) \), Equation 22 is insolvable. Therefore, the \( x \) sets of the efficient frontiers of Equations 2 and 9 do not intersect.
An Illustration

We utilize the following model on the basis of Equation 9:

\[
\begin{align*}
\min \{ z_1 = x^T \Sigma x \} \\
\max \{ z_2 = \mu^T x \} \\
st. \begin{cases} 
1^T x = 1 \\
e^T x = e \\
s^T x = s \\
g^T x = g.
\end{cases}
\end{align*}
\]

Measuring ESG

Carroll (2000) argues for the reason and method to theoretically measure corporate social responsibility (CSR) which typically subsumes ESG. In practice, there are several professional organizations to measure ESG or more broadly measure CSR, for example, RobecoSAM AG, Sustainalytics, the KLD Research & Analytics (KLD), ARESE, and Thomson Reuters ESG data.

On one side, Hopkins (2005) criticizes the measurement systems by demonstrating several indicators as subjectively chosen. Entine (2003) draws similar conclusions. On the other side, Waddock and Graves (1997) review CSR-measure problems and conclude KLD as relatively reliable. Kang (2015), Hart and Sharfman (2015), and Waddock (2003) draw similar conclusions. Overall, Mattingly and Berman (2006), Chatterji et al. (2009), and Hart and Sharfman (2015) relatively comprehensively evaluate KLD ratings. Turkker (2009) realizes the measurement methods’ limitation and proposes a new method on the basis of stakeholders’ responsibility. Graafland et al. (2004) utilize similar methods.

KLD adopts the following categories: alcohol, community, corporate governance, diversity, employee relations, environment, firearms, gambling, human rights, military, nuclear power, product, and tobacco. KLD typically divides a category into two subcategories: strengths and concerns, but some categories (e.g., alcohol) have only concerns. KLD then constructs binary variables under all subcategories. Therefore, a category measurement can be computed by adding all the strengths variables under the category and then subtracting all the concerns variables under the category. Specifically,

- For environmental of ESG, the category environment can be utilized.
- For social of ESG, the summation of categories community, diversity, employee relations, human rights, and product can be utilized.
- For governance of ESG, the category corporate governance can be utilized.

Dyck et al. (2019), Halbritter and Dorfbein (2015), and Galema et al. (2009) utilize similar methods.

Sampling

For Equation 9, we select the following 27 component stocks of Dow Jones Industrial Average Index because they have complete observations whereas the other component stocks do not: Apple (AAPL), American Express (AXP), Boeing (BA), Caterpillar (CAT), Cisco (CSCO), Chevron (CVX), Disney (DIS), Dow Chemical (DOW), Goldman Sachs (GS), Home Depot (HD), IBM (IBM), Intel (INTC), Johnson & Johnson (JNJ), JPMorgan Chase (JPM), Coca-Cola (KO), McDonald’s (MCD), 3M (MMM), Merck (MRK), Microsoft (MSFT), Nike (NKE), Pfizer (PFE), Procter & Gamble (PG), United Health (UNH), United Technologies (UTX), Verizon (VZ), Wal-Mart (WMT), and Exxon Mobil (XOM). (data source: https://us.spindices.com/indices/equity/dow-jones-industrial-average, S&P Dow Jones Indices LLC., October 22, 2019). To save space, we list just brief company names with ticker symbols above.

The sample is made up of the 27 stocks’ monthly return and annual ESG data from January 1, 2004, to December 31, 2013 (data source: database “CRSP” and database “MSCI ESG KLD STATS” via Wharton Research Data Services, https://wrds-web.wharton.upenn.edu/wrds/, October 24, 2019). When we wrote this article, the latest data of “MSCI ESG KLD STATS” was for 2016. However, the ESG data from 2014 to 2016 have many missing values. Therefore, we sample from 2004 to 2013.

We compute the sample covariance matrix and sample mean vector of the stocks’ return and then assume the sample statistics as the population parameters (i.e., \(\Sigma\) and \(\mu\), respectively) because Elton et al. (2014) contend that “It is common to use historical risk, return, and correlation as a starting point in obtaining inputs for calculating the efficient frontier” (p. 87). We annualize \(\Sigma\) and \(\mu\) by the method of Bodie et al. (2018) and report them in Table 1.

For Equation 9, we take:

\[
\begin{align*}
\Sigma &= \frac{1}{T-1} \sum_{t=1}^{T} x_t x_t^T \\
\mu &= \frac{1}{T} \sum_{t=1}^{T} x_t
\end{align*}
\]

For example, the covariance between AAPL and AXP is .051 in the second row. The expected return of AAPL is .474 in the last row.

We adopt the ESG measurement for KLD, compute ESG from 2004 to 2013, and report them in Table 2. For example, the ESG of AAPL in 2004 is 0, 1, and −1, respectively. Moreover, we compute the sample mean vectors of the ESG and then assume the sample statistics as the population parameters (i.e., \(e\), \(s\), and \(g\), respectively). We also report \(e\), \(s\), and \(g\) in Table 2. For example, the expected ESG of AAPL are 1.2, −0.2, and −0.8, respectively.

For Equation 9, we take:

\[
\begin{align*}
e &= \text{the 75th percentile of } e = 1.275 \\
s &= \text{the 75th percentile of } s = 5.450 \\
g &= \text{the 75th percentile of } g = -0.100.
\end{align*}
\]
Table 1. The $\Sigma$ and $\mu$ of the 27 Component Stocks of Equation 9.

|   | AAPL | AXP | BA | CAT | CSCO | CVX | DIS | DOW | GS | HD | IBM | INTC | JNJ | JPM | KO | MCD | MMM | MRK | MSFT | NKE | PFE | PG | UNH | UTX | VZ | WMT | XOM |
|---|------|-----|----|-----|------|-----|-----|-----|----|----|-----|------|-----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| $\Sigma$ | 0.134 | 0.051 | 0.028 | 0.044 | 0.032 | 0.030 | 0.029 | 0.058 | 0.049 | 0.016 | 0.025 | 0.042 | 0.008 | 0.020 | 0.009 | 0.020 | 0.022 | 0.009 | 0.037 | 0.023 | 0.001 | 0.011 | 0.023 | 0.022 | 0.014 | 0.001 | 0.022 |
| $\mu$ | 0.474 | 0.149 | 0.173 | 0.174 | 0.037 | 0.158 | 0.158 | 0.128 | 0.119 | 0.133 | 0.104 | 0.038 | 0.096 | 0.117 | 0.089 | 0.177 | 0.095 | 0.081 | 0.095 | 0.199 | 0.046 | 0.086 | 0.141 | 0.127 | 0.130 | 0.071 | 0.129 |

Note. AAPL = Apple; AXP = American Express; BA = Boeing; CAT = Caterpillar; CSCO = Cisco; CVX = Chevron; DIS = Disney; DOW = Dow Chemical; GS = Goldman Sachs; HD = Home Depot; IBM = IBM Company; INTC = Intel; JNJ = Johnson & Johnson; JPM = JPMorgan Chase; KO = Coca-Cola; MCD = McDonald's; MMM = 3M; MRK = Merck; MSFT = Microsoft; NKE = Nike; PFE = Pfizer; PG = Procter & Gamble; UNH = United Health; UTX = United Technologies; VZ = Verizon; WMT = Wal-Mart; XOM = Exxon Mobil.
Table 2. The e, s, and g of the 27 Component Stocks of Equation 9.

| Year | AAPL | AXP | BA | CAT | CSCO | CVX | DIS | DOW | GS | HD | IBM | INTC | JNJ | JPM | KO | MCD | MMM | MRK | MSFT | NKE | PFE | PG | UNH | UTX | VZ | WMT | XOM |
|------|------|-----|----|-----|------|-----|-----|-----|----|----|-----|------|-----|-----|----|-----|-----|-----|------|-----|-----|----|-----|-----|----|-----|-----|
| 2004 | 0    | 0   | -2 | 0   | 0    | -4  | 1   | -1  | 0  | 0  | -2  | 1    | 0   | 0   | -1 | 1   | -1  | -2  | 0   | 1   | -1  | 0  | 0  | -1  | -1  | -1  | -5  |
| 2005 | 1    | 0   | -2 | -1 | 1    | -4  | 1   | -1  | 0  | 1  | 2   | 2    | 0   | 0   | 0  | 0   | -3  | 0   | 3   | 0   | -1  | 0  | 0  | -1  | -1  | -1  | -5  |
| 2006 | 1    | 0   | -2 | -1 | 1    | -4  | 1   | -1  | 0  | 1  | 1   | 2    | 0   | 1   | 0   | -1  | 0   | 3   | 0   | 3   | 0   | 1  | 0  | 0   | -1  | -1  | -5  |
| 2007 | 1    | 0   | -2 | -1 | 1    | -3  | 1   | -1  | 0  | 0  | 2   | 2    | 2   | 0   | 1   | 0   | 1   | -1  | 0   | 3   | 1   | 0  | 0  | 0   | -1  | -1  | -4  |
| 2008 | 1    | 0   | -4 | 0   | 1    | -3  | 1   | -1  | 0  | 1  | 2   | 2    | 2   | 0   | 2   | 2   | 0   | 1   | 0   | -1  | 0   | 3   | 1   | 0   | -1  | -4  |
| 2009 | 1    | 0   | -4 | 0   | 1    | -3  | 1   | -1  | 0  | 1  | 2   | 2    | 2   | 0   | 2   | 2   | 0   | 1   | 0   | -1  | 0   | 3   | 1   | 0   | -1  | -4  |
| 2010 | 3    | 2   | 1  | 1   | 4    | -3  | 3   | 0   | 2  | 0  | 3   | 4    | 4   | 2   | 0   | 1   | 2   | 1   | 3   | 3   | 3   | 2   | 1   | 1   | 2   | 1   | -1  |
| 2011 | 3    | 2   | 0  | 1   | 4    | -3  | 3   | 0   | 2  | 0  | 3   | 4    | 4   | 4   | 0   | 1   | 3   | 2   | 3   | 3   | 3   | 3   | 2   | 1   | 2   | 2   | 0   |
| 2012 | 0    | 1   | 0  | 1   | 3    | 0   | 1   | 1   | 1  | 1  | 2   | 5    | 3   | 1   | 0   | 0   | 4   | 1   | 2   | 2   | 3   | 4   | 0   | 2   | 1   | 0   | -2  |
| 2013 | 1    | 1   | 4  | 1   | 2    | -3 | 2   | 1   | 1  | 1  | 1   | 5    | 5   | 4   | 0   | 1   | 2   | 4   | 4   | 2   | 1   | 5   | 0   | 4   | 2   | 2   | -1  |

\[(e)^T \begin{pmatrix} 1.2 & 0.6 & -1.1 & 0.1 & 1.8 & -3 & 1.5 & -0.3 & 0.8 & 0.7 & 2.1 & 2.9 & 2.4 & 0.7 & 0.5 & 0.6 & 1.2 & -0.1 & 1.2 & 2.6 & 1.2 & 1.2 & 0.2 & 1.3 & 0.5 & -0.3 & -3.2 \end{pmatrix}

\[(s)^T \begin{pmatrix} -0.2 & 6.8 & 2.9 & -0.8 & 6.9 & 1.7 & 1.1 & 1.2 & 4.2 & 0.7 & 9.2 & 10.9 & 6.2 & 2.1 & 1.5 & 0.9 & 3.7 & 5.3 & 5.5 & 5.2 & 7.2 & 1.3 & -0.2 & 3.1 & -3.6 & -0.8 \end{pmatrix}

\[(g)^T \begin{pmatrix} -0.8 & -0.5 & -1.9 & -0.4 & -0.5 & -0.8 & 0.1 & -1.6 & -2 & -1.1 & 0.5 & -0.2 & -1.9 & -1.2 & 0.3 & 0.4 & -0.9 & 1.1 & 1.1 & -0.1 & -0.1 & -1.5 & -0.2 & -0.2 & -1.7 & -1.2 \end{pmatrix}

Note. AAPL = Apple; AXP = American Express; BA = Boeing; CAT = Caterpillar; CSCO = Cisco; CVX = Chevron; DIS = Disney; DOW = Dow Chemical; GS = Goldman Sachs; HD = Home Depot; IBM = IBM Company; INTC = Intel; JNJ = Johnson & Johnson; JPM = JPMorgan Chase; KO = Coca-Cola; MCD = McDonald's; MMM = 3M; MRK = Merck; MSFT = Microsoft; NKE = Nike; PFE = Pfizer; PG = Procter & Gamble; UNH = United Health; UTX = United Technologies; VZ = Verizon; WMT = Wal-Mart; XOM = Exxon Mobil.
**Solving Equation 9**

By Equation 14, we compute the minimum-variance frontier of Equation 9 as follows:

\[
 z_1 = \begin{bmatrix} z_2 \\ 1 \\ 1.275 \\ 5.450 \\ -0.100 \end{bmatrix}^T
\]

\[
\begin{bmatrix}
0.387 & -0.039 & -0.003 & 0.002 & 0.004 \\
-0.039 & 0.013 & 0.000 & 0.000 & 0.001 \\
-0.003 & 0.000 & 0.001 & 0.000 & -0.001 \\
0.002 & 0.000 & 0.000 & 0.000 & 5.450 \\
0.004 & 0.001 & -0.001 & 0.000 & 0.002 \\
\end{bmatrix}
\begin{bmatrix}
z_2 \\
1 \\
1.275 \\
5.450 \\
-0.100 \\
\end{bmatrix}
\]

\[= 0.387z_2^2 - 0.069z_2 + 0.011.\]

We have already plotted the frontier as a black parabola in Figure 1. The frontier’s upper half is the efficient frontier of (29) and is plotted as a thick curve.

We compute \(\text{I}_{(9)}\) by Equation 16, compute \(x_{(9)}^{mv}\) by Equation 18, and report them in Table 3. We substitute \(x_{(9)}^{mv}\) into Equation 9 to obtain the minimum-variance portfolio in \((z_1, z_2)\) space as follows:

\[z_{(9)}^{mv} = [0.008, 0.089]^T.\]

We then compute \(E_{(9)}\) by Equation 19.

**Solving Equation 2**

As a comparison, we also solve Equation 2 for the 27 component stocks. By Equation 4, we compute the minimum-variance frontier of Equation 2 as follows:

\[
z_1 = 0.350z_2^2 - 0.070z_2 + 0.011.\]

We compute \(\text{I}_{(2)}\) by Equation 8, compute \(x_{(2)}^{mv}\) by Equation 6, and report them in Table 4. We substitute \(x_{(2)}^{mv}\) into Equation 2 to obtain the minimum-variance portfolio in \((z_1, z_2)\) space as follows:

\[z_{(2)}^{mv} = [0.007, 0.100]^T.\]

We then compute \(E_{(2)}\) by Equation 7.

**Comparing the \(x\) Sets of the Efficient Frontiers of Equations 9 and 2**

By Equation 20, we compute \(\delta = 0.287\). For Equation 22, we compute \(\text{rank}(\text{I}_{(2)} - \text{I}_{(9)}) = 2\) and \(\text{rank}(\text{J}) = 3\). Therefore, by Theorem 4, the \(x\) sets of the efficient frontiers of Equations 9 and 2 do not intersect.

**Changing \(e, s, \) and \(g\) to the 50th Percentiles and 90th Percentiles**

In addition to the 75th percentile setting for \(e, s, \) and \(g\), we also reset them as the 50th percentiles and then reset them as the 90th percentiles. We depict the minimum-variance frontiers (including the efficient frontiers) for the 50th percentiles, 75th percentiles, and 90th percentiles in the upper left, upper right, and lower parts of Figure 3, respectively. Overall, the results of the 50th and 90th percentiles are similar to those of the 75th percentiles.

**Conclusion and Future Direction**

Researching sustainable investment is challenging; difficulties in extending conventional portfolio selection are encountered in this article. However, the researching is worthwhile because sustainable investment will no longer be practical experience and will become a theoretically and quantitatively robust tool for investment on the basis of portfolio selection.
Table 3. $j_{(9)}$ of Equation 16 and $(x^{(9)}_{n})$ of Equation 18.

|       | AAPL | AXP | BA | CAT | CSCO | CVX | DIS | DOW | GS | HD | IBM | INTC | JNJ | JPM | KO | MCD | MMM | MRK | MSFT | NKE | PFE | PG | UNH | UTX | VZ | WMT | XOM |
|-------|------|-----|----|-----|------|-----|-----|-----|----|----|-----|-----|-----|-----|----|-----|-----|-----|------|-----|----|----|-----|-----|----|-----|-----|
| $j_{(9)}$  | 1.70 | 0.23 | 0.79 | 0.70 | -0.78 | 0.79 | 0.33 | -0.75 | -0.71 | 0.60 | 0.05 | -1.45 | 1.08 | 0.75 | -1.51 | 0.74 | -0.79 | 1.01 | -0.69 | 0.81 | -1.54 | -0.09 | -0.24 | -0.64 | 0.74 | -1.40 | -0.73 |
| $(x^{(9)}_{n})$  | -0.03 | 0.02 | -0.02 | -0.13 | -0.05 | 0.01 | -0.05 | -0.06 | 0.02 | 0.05 | 0.23 | 0.00 | 0.12 | -0.08 | -0.02 | 0.02 | 0.14 | -0.02 | 0.08 | 0.14 | 0.06 | 0.24 | 0.00 | 0.06 | 0.04 | 0.12 | 0.11 |

Note. AAPL = Apple; AXP = American Express; BA = Boeing; CAT = Caterpillar; CSCO = Cisco; CVX = Chevron; DIS = Disney; DOW = Dow Chemical; GS = Goldman Sachs; HD = Home Depot; IBM = IBM Company; INTC = Intel; JNJ = Johnson & Johnson; JPM = JPMorgan Chase; KO = Coca-Cola; MCD = McDonald’s; MMM = 3M; MRK = Merck; MSFT = Microsoft; NKE = Nike; PFE = Pfizer; PG = Procter & Gamble; UNH = United Health; UTX = United Technologies; VZ = Verizon; WMT = Wal-Mart; XOM = Exxon Mobil.
Table 4. \( j_{(2)} \) of Equation 8 and \( (x_{(3)}^{m}) \) of Equation 6.

|        | AAPL | AXP | BA  | CAT | CSCO | CVX | DIS | DOW | GS  | HD  | IBM | INTC | JNJ | JPM | KO  | MCD | MMM | MRK | MSFT | NKE | PFE | PG  | UNH | UTX | VZ  | WMT | XOM |
|--------|------|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| \( j_{(2)} \) | 2.14 | 0.01| 0.53| 0.92| -0.51| -0.12| 0.60| -0.73| -0.82| 0.83| -0.54| -1.96| 1.19| 0.75| 0.57| 0.81| -0.99| 0.76| -1.02| 1.54| -1.49| -1.11| -0.07| 0.83| -0.54| -1.96| 1.19| 0.75| 0.57|
| \( (x_{(3)}^{m}) \) | -0.01| 0.01| -0.01| -0.12| -0.03| 0.00| -0.06| -0.04| 0.01| 0.08| 0.13| -0.05| 0.02| -0.08| 0.10| 0.03| 0.10| -0.02| 0.05| 0.09| 0.01| 0.17| 0.04| 0.15| 0.00| 0.24| 0.18|

Note. AAPL = Apple; AXP = American Express; BA = Boeing; CAT = Caterpillar; CSCO = Cisco; CVX = Chevron; DIS = Disney; DOW = Dow Chemical; GS = Goldman Sachs; HD = Home Depot; IBM = IBM Company; INTC = Intel; JNJ = Johnson & Johnson; JPM = JPMorgan Chase; KO = Coca-Cola; MCD = McDonald's; MMM = 3M; MRK = Merck; MSFT = Microsoft; NKE = Nike; PFE = Pfizer; PG = Procter & Gamble; UNH = United Health; UTX = United Technologies; VZ = Verizon; WMT = Wal-Mart; XOM = Exxon Mobil.
Appendix

Proving Theorem 1

Proof. We rewrite $(H\Sigma^{-1}H^T)^{-1}$ as:

$$(H\Sigma^{-1}H^T)^{-1} = \begin{bmatrix}
    h_{11} & \ldots & h_{1(m+1)} \\
    \vdots & & \vdots \\
    h_{(m+1)1} & \ldots & h_{(m+1)(m+1)}
\end{bmatrix}.$$  

By matrix multiplication, we further compute Equation 14 as $z_1 = h_1 z_2^2 + q_1 z_2 + q_0$ where $q_1$ and $q_0$ are computed as functions of the elements of $(H\Sigma^{-1}H^T)^{-1}$. By Lemma 1, $(H\Sigma^{-1}H^T)^{-1}$ is positive definite, so its main-diagonal elements are all positive with $h_{11} \geq 0$. Then with $h_{11} \geq 0$, $z_1 = h_1 z_2^2 + q_1 z_2 + q_0$ is a parabola in $(z_1, z_2)$ space.

Proving Lemma 1

Proof. With $\Sigma$ as positive definite by Assumption 1, so is $\Sigma^{-1}$ as documented by Lax (2007). Thus, as documented by Brockwell and Davis (1987), there exists a random vector $s \in \mathbb{R}^n$ such that the covariance matrix of $s$ is $\Sigma^{-1}$. Then, $(H\Sigma^{-1}H^T)$ is the covariance matrix of the random vector $Hs$.

$\forall y \neq 0 \in \mathbb{R}^{m+1}, \quad y^T (H\Sigma^{-1}H^T) y = (H^T y)^T \Sigma^{-1} (H^T y)$. Let $u = H^T y$. We know $u \neq 0$ because $H$ has full row rank by Assumption 2. Then, because $\Sigma^{-1}$ is positive definite:

$$y^T (H\Sigma^{-1}H^T) y = (H^T y)^T \Sigma^{-1} (H^T y) = u^T \Sigma^{-1} u > 0.$$  

Therefore, $(H\Sigma^{-1}H^T)$ is positive definite by definition and thus invertible.

Proving Theorem 3

Proof. The proof is based on Theorem 2 and the relationship between two lines in $\mathbb{R}^n$.

Proving Theorem 4

Proof. The proof is based on basic theorems for solvability status of linear-equation systems (e.g., Lax, 2007).

Acknowledgments

The authors greatly appreciate four anonymous reviewers’ highly constructive comments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This
research was supported by the National Social Science Fund of China 2018 (grant number 18BGL063).

Notes
1. Data source: https://www.ussif.org/sribasics, October 10, 2019.
2. Data source: http://www.gsi-alliance.org/trends-report-2018/, October 16, 2019.
3. Data source: https://am.jpmorgan.com/se/en/asset-management/adv/investment-themes/esg/#, September 2, 2019.
4. Data source: https://www.ussif.org/esg and https://www.unpri.org/ fixed-income/esg-screening-in-fixed-income-investing/36/article, September 4, 2019.
5. Steuer (1986) and Evans (2017) describe multiple-objective optimization. For notations, normal symbols (e.g., $n$) denote scalars; bold-face symbols (e.g., $x$ or $\Sigma$) denote vectors or matrices.

ORCID iD
Yue Qi [i] https://orcid.org/0000-0002-6486-0232

References
Amel-Zadeh, A., & Serafeim, G. (2018). Why and how investors use ESG information: Evidence from a global survey. *Financial Analysts Journal, 74*(3), 87–103.
Auer, B. R. (2016). Do socially responsible investment policies add or destroy European stock portfolio value? *Journal of Business Ethics, 135*(2), 381–397.
Bertrand, P., & Lapointe, V. (2015). How performance of risk-based strategies is modified by socially responsible investment universe? *International Review of Financial Analysis, 38*(3), 175–190.
Best, M. J. (1996). An algorithm for the solution of the parametric quadratic programming problem. In H. Fischer, B. Riedmüller, & S. Schäffler (Eds.), *Applied mathematics and parallel computing: Festschrift for Klaus Ritter* (pp. 57–76). Physica-Verlag.
Best, M. J. (2010). *Portfolio optimization*. Taylor & Francis.
Bodie, Z., Kane, A., & Marcus, A. J. (2018). *Investments* (11th ed.). McGraw-Hill Education.
Brockwell, P. J., & Davis, R. A. (1987). *Time series: Theory and methods* (1st ed.). Springer-Verlag.
Campbell, J. Y. (2018). *Financial decisions and markets: A course in asset pricing*. Princeton University Press.
Capelle-Blancard, G., & Monjon, S. (2014). The performance of socially responsible funds: Does the screening process matter? *European Financial Management, 20*(3), 494–520.
Carroll, A. B. (2000). A commentary and an overview of key questions on corporate social performance measurement. *Business & Society, 39*(4), 466–478.
Chatterji, A. K., Levine, D. I., & Toffel, M. W. (2009). How well do social ratings actually measure corporate social responsibility? *Journal of Economics & Management Strategy, 18*(1), 125–169.
Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? *International evidence*. *Journal of Financial Economics, 131*(3), 693–714.
Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2014). *Modern portfolio theory and investment analysis* (9th ed.). John Wiley.
Entine, J. (2003). The myth of social investing: A critique of its practice and consequences for corporate social performance research. *Organization & Environment, 16*(3), 352–368.
Evans, G. W. (2017). *Multiple criteria decision analysis for industrial engineering: Methodology and applications*. CRC Press.
Galema, R., Plantinga, A., & Scholtens, B. (2009). The cost of socially responsible portfolios: Testing for mean-variance spanning [Working paper]. Department of Finance, University of Groningen.
Graafland, J. J., Eijffinger, S., & Smid, H. (2004). Benchmarking of corporate social responsibility: Methodological problems and robustness. *Journal of Business Ethics, 53*(1), 137–152.
Halbritter, G., & Dorflreiter, G. (2015). The wages of social responsibility—Where are they? A critical review of ESG investing. *Review of Financial Economics, 26*, 25–35.
Hart, T. A., & Sharfman, M. (2015). Assessing the concurrent validity of the revised Kinder, Lydenberg, and Domini corporate social performance indicators. *Business & Society, 54*(5), 575–598.
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.
Hopkins, M. (2005). Measurement of corporate social responsibility. *International Journal of Management and Decision Making, 6*(3/4), 213–231.
Huang, C., & Litzenberger, R. H. (1988). *Foundations for financial economics*. Prentice Hall.
Kang, J. (2015). Effectiveness of the KLD social ratings as a measure of workforce diversity and corporate governance. *Business & Society, 54*(5), 599–631.
Lax, P. D. (2007). *Linear algebra and its applications* (2nd ed.). John Wiley.
Lydenberg, S. (2016). Integrating systemic risk into modern portfolio theory and practice. *Journal of Applied Corporate Finance, 28*(2), 56–61.
Markowitz, H. M. (1959). *Portfolio selection: Efficient diversification in investments* (1st ed.). John Wiley.
Markowitz, H. M., & Todd, G. P. (2000). *Mean-variance analysis in portfolio choice and capital markets*. Frank J. Fabozzi Associates.
Mattingly, J. E., & Berman, S. L. (2006). Measurement of corporate social action: Discovering taxonomy in the Kinder Lydenberg Domini ratings data. *Business & Society, 45*(1), 20–46.
Merton, R. C. (1972). An analytical derivation of the efficient portfolio frontier. *Journal of Financial and Quantitative Analysis, 7*(4), 1851–1872.
Qi, Y., & Steuer, R. E. (2020). On the analytical derivation of efficient sets in quad-and-higher criterion portfolio selection. *Annals of Operations Research, 293*, 521–538. https://doi.org/10.1007/s10479-018-3101-y
Qi, Y., Steuer, R. E., & Wimmer, M. (2017). An analytical derivation of the efficient surface in portfolio selection with three criteria. *Annals of Operations Research, 251*(1–2), 161–177.
Sharpe, W. F. (1970). *Portfolio theory and capital markets* (1st ed.). McGraw-Hill.
Stein, M., Branke, J., & Schmeck, H. (2008). Efficient implementation of an active set algorithm for large-scale portfolio selection. *Computers & Operations Research, 35*(12), 3945–3961.

Steuer, R. E. (1986). *Multiple criteria optimization: Theory, computation, and application*. John Wiley.

Trinks, P. J., & Scholtens, B. (2017). The opportunity cost of negative screening in socially responsible investing. *Journal of Business Ethics, 140*(2), 193–208.

Turker, D. (2009). Measuring corporate social responsibility: A scale development study. *Journal of Business Ethics, 85*(4), 411–427.

Verheyden, T., Eccles, R. G., & 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.

Waddock, S. A. (2003). Myths and realities of social investing. *Organization & Environment, 16*(3), 369–380.

Waddock, S. A., & Graves, S. B. (1997). The corporate social performance-financial performance link. *Strategic Management Journal, 18*(4), 303–319.

Woodside-Oriakhi, M., Lucas, C., & Beasley, J. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research, 213*(3), 538–550.