Assessing the performance of exchange traded funds in the energy sector: a hybrid DEA multiobjective linear programming approach

Carla Oliveira Henriques1,2,3 · Maria Elisabete Neves1,4 · Licínio Castelão1 · Duc Khuong Nguyen5,6

Accepted: 5 October 2021 / Published online: 23 January 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
This paper proposes a two-step approach to build portfolio models. The first step employs the Data Envelopment Analysis (DEA) to select assets attaining efficient financial performance according to a set of indicators used as inputs and outputs. The second step builds interval multiobjective portfolio models to obtain the optimal composition of efficient portfolios previously identified with respect to investor preferences. The usefulness of this proposed methodology is illustrated through a selected sample of diversified Exchange Traded Funds (ETFs) operating in the US energy sector. Our results with respect to all models and time horizons mainly show that: (i) ETFs related to nuclear energy are more often viewed as efficient according to all DEA models considered; (ii) the efficient portfolios do not contain any ETFs related to the renewable energy sector; and (iii) natural gas and oil are the sectors that have the most ETFs represented in efficient portfolios.

Keywords ETF · DEA · Multi-objective portfolio models · Energy sector

Carla Oliveira Henriques
chenriques@iscac.pt

1 Polytechnic of Coimbra, Coimbra Business Research CentreISCAC, Quinta Agrícola – Bencanta, 3045-601 Coimbra, Portugal
2 INESC Coimbra - DEEC, University of Coimbra, Polo 2, 3030-290 Coimbra, Portugal
3 Faculty of Economics, CeBER, University of Coimbra, Av Dias da Silva 165, 3004-512 Coimbra, Portugal
4 University of Trás-Os-Montes and Alto DouroCETRAD, Vila Real, Portugal
5 IPAG Business School, Paris, France
6 International School, Vietnam National University, Hanoi, Vietnam
1 Introduction

The portfolio model suggested by Markowitz (1952) is based on a reward and risk perspective, where the reward is measured by expected return and the risk is measured by the variance of return. The portfolio diversification is assessed through the return correlation, while portfolio optimization relies on either maximizing the return for a given risk level or minimizing the risk for a given return target.

Accordingly, this classical Markowitz approach disregards the fact that, more often than not, the investor’s position in relation to risk is not coincident with symmetry or normal distribution. A small loss may thus be enough to make a non-risk prone investor unsatisfied (Henriques & Neves, 2019). In this sense, the assumptions in the Markowitz approach only hold if the expected returns follow multivariate normal distributions and the investor is risk averse (Papahristodoulou & Dotzauer, 2004; Xidonas et al., 2018). In this context, Markowitz (1968) suggests the semi-variance as an alternative measure to assess risk to consider negative deviations from a certain level of return.

It is therefore clear that since the formulation of the problems originally proposed by Markowitz is quadratic, the computational complexity required to obtain solutions may easily become NP-hard (i.e., non-deterministic polynomial-time hard), particularly if the number of assets under consideration is high and there are cardinality constraints (Escudero et al., 2006; Liagkouras & Metaxiotis, 2018). For this reason, several authors have proposed alternative mathematical formulations to improve the computational effort to obtain efficient portfolios.1 Our current study contributes to this literature and proposes a two-step approach to address the above-mentioned technical problem by combining two distinct mathematical programming methodologies. The first methodology uses the Data Envelopment Analysis (DEA) to select the set of assets that have the best financial performance, according to the indicators used as inputs and as outputs. The second one applies the interval multiobjective portfolio models to choose the possible efficient portfolio compositions, which explicitly consider the investor preferences (i.e., assuming more aggressive, conservative, and/or combined strategies).

Our proposed approach is advantageous in that it allows, via the DEA, to identify the sources of inefficiency through contrasting the performance of each inefficient decision-making unit (DMU) with those that are viewed as benchmarks, i.e., efficient DMUs that serve as a reference in terms of best practices (Cooper et al., 2007). It also offers a convenient way to treat the uncertainty related to portfolio selection via the interval multiobjective optimization model (or interval programming). Indeed, most conventional models consider average return and risk as the factors that mostly influence investment decisions, but the use of average values of the coefficients restrains the possibility to capture the model uncertainty. Henriques and Neves (2019) further argue that, in the context of portfolio selection, the exact and complete information related to several parameters that the decision-maker needs to obtain is not always available. The subjectivity underlying the human being and the imprecision existing in real-world problems are, for example, some factors that can give rise to the uncertainty present in portfolio models (Gupta et al., 2014). Comparatively, the interval programming may be more useful than the stochastic and fuzzy programming in modeling the model uncertainty because it only requires information about the range of variation of the coefficients needed to instantiate the model (Henriques et al., 2020).2 This particular characteristic is of interest for decision-makers in practice (Oliveira & Antunes, 2007).

---

1 See Biglova et al. (2004) and Ortobelli et al. (2005) for an extensive review of the literature regarding risk measures and models that can be applied in portfolio selection.

2 In stochastic programming, the coefficients of the optimization problem are regarded as uncertain, and their probability distributions are known. In fuzzy programming, the coefficients of the problem are given as...
At the empirical stage, the usefulness of our proposed approach will be validated through its application to a set of Exchange Traded Funds (ETFs) of the energy sector traded in the United States, since this sector deserves special attention due to its size and impact on the global economy and environment. Our interest in the study of ETFs lies in the fact that they offer portfolio diversification opportunities with the convenience of trading individual stocks and have been gaining greater attention from investors over recent years. The choice of the US ETF market is related to the fact that this market is the most representative worldwide, corresponding in 2018 to 72% of the 4.7 trillion ETF assets traded worldwide (Pagano et al., 2019). More importantly, the approach proposed appears as an interesting contribution in the construction of portfolios, since the composition of the portfolios scrutinized will be composed of funds that, by nature, are already diversified.

The remainder of this paper is structured as follows. Section 2 provides a brief literature review on portfolio selection involving energy-related ETFs. Section 3 describes the methodological approach. Section 4 presents the dataset used in the study. Section 5 discusses the obtained results. Section 6 concludes the paper.

2 Literature review

In the traditional portfolio selection problem, the arithmetic average of past returns is generally viewed as a proxy for the expected return of an asset and is thus a certain value. However, in real-world problems, asset prices and returns are driven by a set of variables whose behavior cannot simply be predicted from past events (Gupta et al., 2014; Henriques & Neves, 2019; Xidonas et al., 2017). Moreover, the use of the arithmetic average of historical returns as the expected return has two major shortcomings. On the one hand, if historical data for a long period is considered, the influence of the earlier historical data is the same as that of recent past data. On the other hand, if the historical data of an asset are not available due to lack of information, the estimation of statistical parameters would be inadequate. With the foregoing in mind, the interval programming approach, where the coefficients are assumed to lie within a range of variation can be particularly resourceful (Gupta et al., 2014; Henriques et al., 2020; Oliveira & Antunes, 2007). An extensive review of portfolio selection models with interval coefficients can be obtained in Henriques and Neves (2019). Generally, the portfolio selection models identified therein considered a small number of assets for illustrative purposes disregarding the reasons for their choice and just employed interval coefficients in the objective functions (either risk or return). Besides, the computational complexity associated with these models’ solutions highlights the need of developing filtering procedures to select the assets that should be considered as the starting portfolios under assessment (Henriques & Neves, 2021).

In this context, we suggest a two-step approach to tackle the aforementioned problem by coupling two mathematical programming approaches. First, we use DEA to select the set of assets that will be used in the next step of the analysis. Then, we build interval multiobjective portfolio models to compute the possible efficient portfolio compositions, exploring distinct fuzzy numbers with known possibilistic functions. However, the amount of information needed to obtain these functions may not exist and may be very expensive.

Footnote 2 continued

The total net assets of the US ETFs in 2019 reached US$4.4 trillions (https://www.statista.com/statistics/295632/etf-us-net-assets/), compared to only US$2.1 trillions recorded in 2015. According to recent estimates of Bank of America, the US ETF market size is expected to hit US$50 trillions in 2050 (https://markets.businessinsider.com/news/stocks/etf-market-grow-50-trillion-assets-2030-bank-america-passive-2019-12-1028763048).
investor preferences towards risk. The usefulness of the proposed approach is subsequently validated through its application to a sample of energy-related ETFs traded in the United States.

2.1 Performance evaluation of ETFs

ETFs are open-end investment funds that attempt to replicate the return and risk of their reference indexes (Poterba & Shoven, 2002; Zopounidis et al., 2010). These investment vehicles couple stock flexibility with the diversification provided by mutual funds. After their establishment, ETFs have become progressively popular as they provide investors diversification benefits, greater tax efficiency, and lower expenses. The performance evaluation of these funds usually relies on risk-adjusted measures (Blitz et al., 2012), from which two of the most used ones are the Sharpe ratio and the Jensen’s alpha (Jensen, 1972). However, they have some drawbacks including, among others, the necessity of identifying benchmarks which can be used for comparison purpose and the role of market timing. For example, the Jensen’s alpha’s main limitation is linked to the selection of a benchmark asset pricing model (e.g., CAPM and APT) which, in turn, depends on the correct identification of the market portfolio or the systematic risk factors (Choi, 1995; Elton et al., 1993; Roll, 1978). Market timing may lead to biased estimate of the Jensen’s alpha (Admati & Ross, 1985; Dybvig & Ross, 1985). For its part, the Sharpe ratio is sensitive to the market index used as benchmark (Roll, 1978).

The performance of ETFs has been drawing scholars’ attention, particularly in the energy finance field in recent years. Murphy and Wright (2010) examine the capacity that 12 leveraged commodity-based ETFs (including oil and gas) had to meet their investment objectives. The results of this study indicate that the ETFs under consideration have been underperforming since their inception, considering their exposure (long vs short term), desired leverage and expense ratios, mainly due to the volatility of the commodity prices and indices. However, despite their poor performance, the authors find that one third of ETFs exceeded expectations since their launch in 2007/2008. Guedj et al. (2011) investigate three types of oil ETFs to assess possible deviations between the returns of commodity ETFs, which essentially use future contracts as an investment strategy and the change in commodity prices. They document that the returns of ETFs show deviations from the change in commodity prices due to the change in the spot price of oil. Moreover, this divergence is essentially due to the practice of rollover of future contracts when the latter approach the maturity date. Accordingly, the authors do not advise investment to uninformed investors and investors who wish to hold futures-based commodity ETFs for the long term. Sabbaghi (2011a) looks at the behavior of returns and volatility of a sample of green ETFs (i.e., ETFs based on firms engaged directly or indirectly in research, development, production, and supply of alternative energy) using the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) methodology. The accumulated returns for the green ETFs showed positive values for the period from 2005 to 2008. Additionally, Sabbaghi (2011a) shows strong evidence of the persistence of ETF volatility. Going one step further, Sabbaghi (2011b) aims to build a portfolio based on green ETFs and compares its performance in terms of returns with the S&P 500 stock index. The obtained results indicate that, until the collapse of the 2008 financial market, the portfolio composed of green ETFs outperformed the S&P 500. Afterwards, the green ETFs were highly volatile, causing a lower performance than the S&P 500. Alexopoulos (2018) studies the performance of portfolios based on renewable and conventional energy ETFs during bear and bull market.
periods and finds that a portfolio aggregating all ETFs exceeds two disaggregated (renewable and conventional energy ETFs) portfolios, through providing higher returns and lower risk. In the same vein, La Monaca et al. (2018) use the Markowitz’s classical approach to test whether including renewable energy ETFs in a portfolio provides diversification benefits over a period of 2, 7 and 9 years, respectively. The authors also evaluate the benefits of investing in yields (i.e., companies characterized as having operational renewable energy assets). Their result show that renewable energy ETFs provide only minimal diversification benefits, while yields do not provide any diversification benefits. However, after evaluating the performance of yields outside the US, these authors obtained superior results with respect to the Sharpe Index. Miralles-Quirós and Miralles-Quirós (2019) compare the performance of renewable energy ETFs with that of conventional energy ETFs, based on their respective returns and volatility, using an Asymmetric Dynamic Conditional Correlations—GARCH model. Their results show the superiority of the renewable energy ETF portfolio over the conventional one, which is essentially due to the constant increase in oil and natural gas prices. Recently, Yang et al. (2021) model jump events in the S&P500 index, crude oil futures price and 24 energy ETFs using multivariate Hawkes Processes. Analyzing the estimated model parameters, they can interpret the complex relationship of various jump events concerning their rollover, excitation, time-decaying effect, and accumulated contagion effects. They find that the self-contagion in negative jumps is stronger than the self-contagion in the positive jumps across all ETF groups.

2.2 Financial literature focusing on the construction of portfolios based on DEA

Despite the distinct characteristics of the existing methodologies employed in the evaluation of ETFs performance, there is one common feature: they all encompass different risk-adjusted performance measures. Therefore, one of the challenges inherent to the appraisal of ETFs performance is the combination of the multiple metrics available towards this end into a single indicator in such a way that it provides an effective support to investors. Hence, a suitable methodology to select the ETFs with the best risk-adjusted performance is needed to instantiate the starting interval multiobjective portfolio models and to find optimal portfolio compositions.

In this framework, the DEA methodology originally proposed by Charnes et al. (1978) allows to assess the efficiency of a set of units called DMUs, whose performance depends on the inputs and outputs selected to perform the evaluation. With this methodology, DMUs are classified into efficient and inefficient according to a single efficiency score. Additionally, the DEA approach enables managers to obtain the benchmarks of inefficient DMUs, providing them with information regarding the best practices to be followed. In finance, this approach has been used to measure the financial performance of ETFs because it allows to overcome the intrinsic limitations of traditional performance measures (Choi & Min, 2017; Murthi et al., 1997). On the one hand, this tool does not involve any theoretical reference model (e.g., CAPM or APT). The efficiency of each ETF (DMU) is measured against a set of efficient ETFs within the same category. On the other hand, the model is flexible and can simultaneously contemplate different inputs and outputs.

Hence, in this study, we use DEA to overcome the above-mentioned limitations of traditional performance measures. Table 1 shows a list of relevant studies documenting the usefulness and suitability of the DEA approach in assessing the performance of stock portfolios, mutual funds, and ETFs, among others.
| References                  | Objective                                      | Model                        | Inputs                                                                 | Outputs                                      |
|-----------------------------|------------------------------------------------|------------------------------|------------------------------------------------------------------------|----------------------------------------------|
| Murthi et al. (1997)        | Assess the performance of mutual funds         | Charnes, Cooper and Rhodes (CCR) | Standard deviation, expense ratio, turnover and loads                  | Return                                       |
| Basso and Funari (2001)     | Evaluate the performance of mutual funds       | CCR                          | Standard deviation, square root of the half-variance and Beta          | The expected return and stochastic dominance indicator |
| Choi and Murthi (2001)      | Assess the performance of 6 categories of mutual funds | CCR and Bunker, Charnes and Cooper (BCC) | Standard deviation, expense ratio, loads and turnover                 | Return                                       |
| Haslem and Scheraga (2003)  | Assess the performance of 80 mutual funds      | CCR                          | Cash, stocks, P/E, Price Book (P/B) and total assets                  | Sharpe Index                                 |
| Daraio and Simar (2006)     | Assess the performance of mutual funds         | CCR and BCC                  | Standard deviation, expense ratio, loads and turnover                 | Return                                       |
| Chen (2008)                 | Build portfolios of stocks                     | CCR and BCC                  | Average assets, average assets and sales costs                        | Revenues, operating profits, net result      |
| Dia (2009)                  | Build portfolios of stocks                     | CCR and BCC                  | Beta                                                                   | Return and Exchange flow ratio              |
| Hsu (2014)                  | Build portfolios of stocks                     | CCR                          | Total equities, total assets, sales costs, and operational expenses   | Net sales and net return                    |
| Gregoriou and Henry (2015)  | Assess the performance of commodities trading advisers (CTAs) | BCC                          | Margin-to-equity, round turn, incentive fee e management fee          | bad: standard deviation, asymmetry, maximum drawdown, semi-variance good: return and Sharpe Index |
| References                      | Objective                                                                 | Model            | Inputs                                | Outputs                                    |
|---------------------------------|---------------------------------------------------------------------------|------------------|---------------------------------------|--------------------------------------------|
| Huang et al. (2015)             | Assess the performance of stocks and optimize portfolios                  | BCC              | Downside risk and beta                | Return and Sharpe index                    |
| Tsolas and Charles (2015)       | Assess the performance of Green ETFs                                     | Range Adjusted Measure (RAM-BCC) | P/CF and P/B                           | Sharpe Index and Jensen’s Alpha            |
| Basso and Funari (2016)          | Analyze the efficiency of European mutual funds from different countries  | BCC              | Standard deviation and Beta           | Final investment value (average return)    |
| Choi and Min (2017)              | Assess the performance of an index, 8 ETFs and 200 stocks                | RDM              | Standard deviation and beta           | Return                                     |
| Zhang and Chen (2018)            | Assess the performance of portfolios composed of oil, natural gas and coal futures | RDM              | Standard deviation and VaR            | Return                                     |
| Isakov (2019)                    | Assess the performance of ETFs of the platform Xetra (Germany)           | CCR              | Expense ratio and downside risk       | Return and upper deviation                 |
| Tsolas (2019)                    | Measure the efficiency of 10 “Utility ETFs”                               | Grey Relational Analysis, BCC, Additive model, GPDF | P/E ratio and expense ratio            | Sharpe ratio                              |
The financial literature focusing on the construction of portfolios based on the DEA methodology is rather scarce. The studies of Hsu (2014) and Huang et al. (2015) deserve special mention. The first one applies DEA in a first stage to select stocks with the best potential to generate profit and then find, in the second phase, the investment proportions allocated to each stock. Similarly, the second study first makes a stock selection based on certain criteria obtained from the companies’ financial statements and then calculates the efficiency of stocks using the DEA model. Finally, it determines the investment weight for each stock.

Using this type of approach, we evaluate the efficiency of a set of ETFs in the energy sector considering inputs and outputs according to the related literature (Table 1). By doing so, we are able to establish an initial filtering procedure of the ETFs that could compose the possibly efficient portfolios, thus resulting in a smaller computational effort.

Overall, our current study proposes an innovative approach to select the ETFs with the best risk-adjusted performance as well as to evaluate the impact of investor preferences on the construction of a portfolio of ETFs.

### 3 Methodology

In what follows, a brief description of the methodological approaches used herein is provided, specifically regarding the DEA model and the multiobjective portfolio problems with interval coefficients.

#### 3.1 The DEA model

As it can be seen from Table 1, several authors employ the CCR and BCC models in a financial context. However, these two models do not allow to tackle negative inputs and outputs because the factors typically used to evaluate investments, namely returns, can take negative values. To overcome this limitation, Chu et al. (2010), Tsolas and Charles (2015) and Choi and Min (2017) use DEA models to evaluate the performance of ETFs while considering the presence of negative data. Our study employs the Weighted Russell Directional Distance Model (WRDDM) approach, which is inspired by the Range Directional Model (RDM +) as in Portela et al. (2004), to handle negative data.

#### 3.1.1 The WRDDM

We consider the WRDDM which, unlike the Directional Distance Function (DDF) model, allows to consider different rates of contraction and expansion of inputs and outputs. The model formulation can therefore be generalized as follows (Chen et al., 2015):

\[
\max \beta_o^R = \max \left( w_y \left( \sum_r \sigma^r y r \sigma^r_o \right) + w_x \left( \sum_i \sigma^i x i \sigma^i_o \right) \right).
\]

s.t. \( \sum_{j=1}^{n} \lambda_j y_{rj} \geq y_{r\alpha}^c + \alpha^r g_{yr}, \ r = 1, ..., s, \)

\( \sum_{j=1}^{n} \lambda_j x_{ij} \leq x_{i\alpha} - \varepsilon_{io} \sigma_{xi}, \ i = 1, ..., m, \)
where the vectors of inputs and outputs of DMU_o are given by x_o and y_o, respectively; parameters \( \alpha^*_o \) and \( \zeta^*_o \) are the individual measures of inefficiency for each output and input, respectively; and all the other variables are non-negative except for \( \beta_o \). The value \( \zeta^*_o g_{xi} \) indicates the level by which DMU_o must reduce its \( i \)-th input to become efficient. Similarly, \( \alpha^*_o g_{yr} \) provides information on the level by which DMU_o must increase its \( r \)-th output in order to become efficient. The values \( w_y \) and \( w_x \) are the priorities given to the outputs and inputs, and their sum should be one. On the other hand, different priorities may be assigned to the evaluation factors, i.e.,

\[
\sum_r \varpi^r y_r = 1, \quad \sum_i \varpi^i x_i = 1.
\]

If the inefficiency measure is zero (\( \beta^R_o = 0 \)), the DMU will be fully efficient. Further developments regarding the WRDDM approach can be found in Chen et al. (2015).

The set of reference DMUs of the inefficient DMU_o is obtained by solving the following LP problem, assuming that \( \alpha^*_o \) and \( \zeta^*_o \) are obtained in the optimal solution of the system (2):

\[
\begin{align*}
\max & \quad \sum_r s^+_r + \sum_i s^-_i, \\
\text{s.t} & \quad \sum_{j=1}^n \lambda_j y_{rj} - s^+_r = y_{ro} + \alpha^*_o g_{yr}, \quad r = 1, \ldots, s, \\
& \quad \sum_{j=1}^n \lambda_j x_{ij} + s^-_i = x_{io} - \zeta^*_o g_{xi}, \quad i = 1, \ldots, m, \\
& \quad \sum_{j=1}^n \lambda_j = 1, \quad \lambda_j \geq 0 \quad (\forall j), \\
& \quad s^+_r \geq 0 \quad (\forall r), \\
& \quad s^-_i \geq 0 \quad (\forall i)
\end{align*}
\]

Let \( (\alpha^*_o, \zeta^*_o, s^+_r, s^-_i, \lambda^*_j) \) the optimal solution to (2). The set of efficient DMUs that serves as a reference (in terms of best practices) for the inefficient DMU_o is obtained as follows:

\[ E_o = \{ j : \lambda^*_j > 0, \quad j = 1, \ldots, n \} \]

The point of the efficient frontier that can be seen as a linear combination of the reference DMUs’ inputs and outputs is given by \( (\hat{x}_o, \hat{y}_o) = (\sum_{j \in E_o} \lambda^+_j x_j, \sum_{j \in E_o} \lambda^+_j y_j) \). In this case, directional vectors must be measured according to the same units of measurement as the original input and output vectors.

The maximum possible improvement region for DMU_o (Portela et al., 2004) is given by:

\[
R_{ro} = \max \{ y_{rj} \} - y_{ro}, \quad r = 1, \ldots, s, \quad \text{and} \quad R_{io} = x_{io} - \min \{ x_{ij} \}, \quad i = 1, \ldots, m
\]

The values computed in this way are always nonnegative. The treatment of our data will be based on the RDM+ model proposed by Portela et al. (2004), which has two important features: it is translation invariant under the assumption of VRS and units invariant. However,
rather than using the DDF model as in Portela et al. (2004), we consider the following WDDRM:

$$\begin{align*}
\max_{\beta} & \quad R_{ro}' = \max \left( w_y \left( \sum_r \varphi_{y_0}^r \alpha_{y_0}^r \right) + w_x \left( \sum_i \varphi_{x_0}^i \xi_{x_0}^i \right) \right) \\
\text{s.t} & \quad \sum_{j=1}^n \lambda_j y_{rj} \geq y_{ro}^c + \alpha_{y_0}^c R_{ro}, \quad r = 1, \ldots, s, \\
& \quad \sum_{j=1}^n \lambda_j x_{ij} \leq x_{io} - \zeta_{i0}^i R_{io}, \quad i = 1, \ldots, m, \\
& \quad \sum_{j=1}^n \lambda_j = 1 \\
& \quad \lambda_j \geq 0 \quad (\forall j)
\end{align*}$$

The efficient units given by the RDM + model (4) have a value of $\beta_{y_0}^{R'} = 0 = 0$. Nevertheless, a $\beta_{y_0}^{R'} = 0$ is not a sufficient condition to guarantee the efficiency of a DMU. For a DMU to be efficient, two conditions must be met simultaneously: (i) $\beta_{y_0}^{R'} = 0$ and (ii) all the constraints of model (4) are satisfied as equalities, i.e., all slack variables are zero.

### 3.2 The multiobjective portfolio problems with interval coefficients

Let us assume that investors allocate their wealth among $n$ assets offering random rate of returns and that the portfolio selection problem is based on a single period model of investment.

#### 3.2.1 Objective functions

A portfolio is composed of two or more assets represented by an ordered $n$-tuple $\Theta = (x_1, x_2, \ldots, x_n)$, where $x_i$ is the proportion of the total funds invested in the $i$-th asset.

#### 3.2.2 Return

The rate of return on each asset is computed as follows:

$$r_{it} = \ln(p_{it}) - \ln(p_{it-1}) = \ln \frac{p_{it}}{p_{it-1}}$$

where $p_{it}$ is the closing price of the $i$-th asset during the period $t$, and $p_{it-1}$ is the closing price during the period $t - 1$.

All in all, due to the uncertainty, the expected return of an asset should be rather given as an interval number. Therefore, the return of the portfolio is expressed as:

$$\sum_{i=1}^n \left[ r^L_i, r^U_i \right] x_i = \left[ \sum_{i=1}^n r^L_i x_i, \sum_{i=1}^n r^U_i x_i \right]$$

where $\left[ r^L_i, r^U_i \right]$ is the interval valued return.
3.2.3 Risk

Usually, an investor would rather prefer to have the portfolio return as large as possible and at the same time with minimum possible risk. In this context, Speranza (1993) proposes the semi-absolute deviation as an alternative measure to quantify risk. The expected mean semi-absolute deviation of return of the portfolio below the expected return can be then given by:

$$\sum_{t=1}^{T} \left[ \sum_{i=1}^{n} (r_{it} - r_i) x_i \right] + \sum_{i=1}^{n} \left( r_i - r_{it} \right) x_i$$

(7)

If the expected returns of assets are given as interval numbers, the expected semi-absolute deviation of return of the portfolio below the expected return is an interval number as well:

$$\left[ \sum_{t=1}^{T} \left[ \sum_{i=1}^{n} (r_{it} - r_i^L) x_i \right] + \sum_{i=1}^{n} \left( r_i^L - r_{it} \right) x_i \right], \sum_{t=1}^{T} \left[ \sum_{i=1}^{n} (r_{it} - r_i^U) x_i \right] + \sum_{i=1}^{n} \left( r_i^U - r_{it} \right) x_i$$

(8)

3.2.4 Constraints

Capital budget constraint. Since $x_i$ is the proportion of the total funds invested, the capital budget constraint on the assets is expressed as:

$$\sum_{i=1}^{n} x_i = 1$$

(9)

Maximum proportion of capital that can be invested. The maximum proportion of capital allocated to the assets in the portfolio depends upon several factors. Since investors differ in their interpretation of the available information, we impose the upper bound to be within interval, $[u_{iL}, u_{iU}]$ to achieve a sufficient diversification of investments and thus obtain the following interval constraint:

$$\sum_{i=1}^{n} x_i \leq [u_{iL}, u_{iU}], i = 1, ..., n,$$

(10)

No short selling is allowed. In portfolio mathematical modeling, short selling is not allowed, i.e., the values of $x_i$ are not negative and hence,

$$x_i \geq 0 \text{ for all } i = 1, 2, \ldots, n$$

(11)

3.2.5 Method used for obtaining efficient portfolios

This subsection presents three mathematical models based on the approach proposed by Henriques and Neves (2019), which allow us to explore three different strategies: a conservative, aggressive and/or a combination of the two strategies.

---

4 If short selling is allowed, investors establish a market position by selling an asset they do not actually own in anticipation that its price will fall. Mathematically, this situation implies that the number of assets owned by the investor is negative, which requires the use of negative variables in our model. However, we opted for considering that no short selling is allowed because, as pointed by Miralles-Quirós and Miralles-Quirós (2019), the performance of ETFs is improved when no short-selling constraints are included.
The problem previously described is given as:

\[
\begin{align*}
& \text{Min} \left[ \frac{1}{T} \sum_{t=1}^{T} \left| \sum_{i=1}^{n} (r_{it} - r^L_i) x_i \right| + \sum_{i=1}^{n} (r^L_i - r_{it}) x_i \right] + \frac{1}{T} \left| \sum_{t=1}^{T} \sum_{i=1}^{n} (r_{it} - r^U_i) x_i \right| + \sum_{i=1}^{n} (r^U_i - r_{it}) x_i \right] \\
& \text{Max} \left[ \sum_{i=1}^{n} r^L_i x_i, \sum_{i=1}^{n} r^U_i x_i \right].
\end{align*}
\]

s.t. \( \sum_{i=1}^{n} x_i = 1 \)

\[
x_i \leq [u^L_i, u^U_i], \quad i = 1, \ldots, n,
\]

\[
x_i \geq 0, \quad i = 1, \ldots, n,
\]

\((12)\)

and has the following surrogate:

\[
\begin{align*}
& \text{Min} \left[ \frac{1}{T} \sum_{t=1}^{T} p^1_t, \frac{1}{T} \sum_{t=1}^{T} p^2_t \right] \\
& \text{Max} \left[ \sum_{i=1}^{n} r^L_i x_i, \sum_{i=1}^{n} r^U_i x_i \right],
\end{align*}
\]

s.t. \( \sum_{i=1}^{n} x_i = 1 \)

\[
p^1_t + \sum_{i=1}^{n} (r_{it} - r^L_i) x_i \geq 0, \quad t = 1, 2, \ldots, T,
\]

\[
p^2_t + \sum_{i=1}^{n} (r_{it} - r^U_i) x_i \geq 0, \quad t = 1, 2, \ldots, T,
\]

\[
x_i \geq 0, \quad i = 1, \ldots, n,
\]

\[
p^1_t \geq 0, \quad t = 1, 2, \ldots, T,
\]

\[
p^2_t \geq 0, \quad t = 1, 2, \ldots, T,
\]

\((13)\)

where \( p^1_t = \left| \sum_{i=1}^{n} (r_{it} - r^L_i) x_i \right| = \max \{ \sum_{i=1}^{n} (r_{it} - r^L_i) x_i, -\sum_{i=1}^{n} (r_{it} - r^L_i) x_i \} \) and \( p^2_t = \left| \sum_{i=1}^{n} (r_{it} - r^U_i) x_i \right| = \max \{ \sum_{i=1}^{n} (r_{it} - r^U_i) x_i, -\sum_{i=1}^{n} (r_{it} - r^U_i) x_i \}. \)

Problem \((13)\) can then be solved through the weighted-sum method (Gupta et al., 2014; Henriques & Neves, 2019). Next, we suggest distinct optimization models for portfolio selection according to three types of investment strategies, i.e., conservative, aggressive, and combined.

\(i\) **Conservative strategy**

In this type of strategy, the investor is more risk averse. In this way, when facing the worst possible scenario of coefficients (as indicated by problem 3A in Appendix 1 of the Online Supplementary Materials, knowing that maximizing an objective function is the
same as minimizing its symmetrical value\(^5\), the investor wants to solve the following model:

\[
\text{Max} \left( \phi \sum_{i=1}^{n} r_i x_i - \tau \frac{1}{T} \sum_{t=1}^{T} P_2^t \right) \\
\text{s.t.} \sum_{i=1}^{n} x_i = 1, \\
x_i \leq u_i^L, \quad i = 1, \ldots, n, \\
\sum_{i=1}^{n} x_i \leq 1, \\
x_i \leq u_i^U, \quad i = 1, \ldots, n, \\
\sum_{t=1}^{T} P_2^t \geq 0, \quad t = 1, \ldots, T, \\
P^2_t + \sum_{i=1}^{n} \left( r_{it} - r_i^U \right) x_i \geq 0, \quad t = 1, \ldots, T, \\
x_i \geq 0, \quad i = 1, \ldots, n, \\
P^1_t \geq 0, \quad t = 1, \ldots, T, \\
\sum_{i=1}^{n} x_i \geq 1, \\
P^1_t \geq 0, \quad t = 1, \ldots, T, \\
\sum_{i=1}^{n} x_i \geq 1.
\] (14)

where \( \phi \) and \( \tau \) are the weights that the investor assigns to each objective function. Their values range from zero to one.

(ii) **Aggressive strategy**

This type of strategy is associated with a risk-loving investor. When facing the best possible scenario of coefficients (as indicated by problem 2A in Appendix 1 of the Online Supplementary Materials, knowing that maximizing an objective function is the same as minimizing its symmetrical value), the investor wants to solve the following model:

\[
\text{Max} \left( \phi \sum_{i=1}^{n} r_i x_i - \tau \frac{1}{T} \sum_{t=1}^{T} P_1^t \right) \\
\text{s.t.} \sum_{i=1}^{n} x_i = 1, \\
x_i \leq u_i^L, \quad i = 1, \ldots, n, \\
\sum_{t=1}^{T} P_1^t \geq 0, \quad t = 1, \ldots, T, \\
P^1_t + \sum_{i=1}^{n} \left( r_{it} - r_i^L \right) x_i \geq 0, \quad t = 1, \ldots, T, \\
x_i \geq 0, \quad i = 1, \ldots, n, \\
P^1_t \geq 0, \quad t = 1, \ldots, T. \\
\] (15)

(iii) **Combined strategy**

All the previous strategies, including a combined strategy with a more balanced approach regarding risk and return, can be obtained through the following mathematical formulation:

\[
\text{Max} \left( \phi \sum_{i=1}^{n} r_i x_i - \tau \frac{1}{T} \sum_{t=1}^{T} P_2^t \right) + (1 - \rho) \left( \phi \sum_{i=1}^{n} r_i x_i - \tau \frac{1}{T} \sum_{t=1}^{T} P_1^t \right)
\] (16)

The concepts of best optimum value and worst optimum value are further developed in Oliveira and Antunes (2007) and are given in Appendix 1 of the Online Supplementary Materials.
\[ s.t. \sum_{i=1}^{n} x_i = 1, \]
\[ x_i + \left( -u U_i + \delta_i \left( -u L_i + u U_i \right) \right) \leq 0, i = 1, \ldots, n, \]
\[ P^1_t + \sum_{i=1}^{n} (r_{it} - r L_i) x_i \geq 0, t = 1, \ldots, T, \]
\[ P^2_t + \sum_{i=1}^{n} (r_{it} - r U_i) x_i \geq 0, t = 1, \ldots, T, \]
\[ x_i \geq 0, i = 1, \ldots, n, \]
\[ P^1_t \geq 0, t = 1, \ldots, T, \]
\[ P^2_t \geq 0, t = 1, \ldots, T. \]

where \( \rho \) and \( \delta_i, i = 1, \ldots, n \), are pessimistic indices that range between zero (aggressive strategy) to one (conservative strategy).

### 4 Data

#### 4.1 Section of energy-related ETFs

In a first phase, a sample of ETFs linked to the energy sector was gathered. The identification of the ETFs was possible through the ETF database, and 60 funds were identified.\(^6\) The objective was to evaluate the efficiency of each fund in periods of 1 year (2018), 3 years (from 2016 to 2018) and 5 years (from 2014 to 2018) as of 31/12/2018. Table A1 of the Online Supplementary Materials shows 60 ETFs of the energy sector traded in the United States, the respective reference index and the composition of each fund. The sample refers to petroleum, natural gas and alternative energy ETFs that represent several categories such as alternative energy equities, China equities, energy equities, global equities, inverse commodities, leveraged commodities, leveraged equities, and oil & gas equities.

When analyzing the composition of each fund, we find that they are essentially composed of shares of U.S. and non-U.S. companies, bonds, cash and cash equivalents, and derivative products such as futures and swap contracts. As far as the benchmark is concerned and given the great multiplicity of indices that could serve as a reference, those described in this study are the same ones used by the Financial Times.

#### 4.2 Selection of the inputs and outputs

The selection of inputs and outputs is not consensual in the literature. In our analysis, we based on the scientific literature (Table 1) to choose the beta parameters and the standard deviation as inputs and the Jensen’s Alpha, Sharpe index, mean annual return, and trailing total return inputs.

\(^6\) https://etfdb.com. Due to space constraint, we show the list of these 60 funds in Table A1 of the Online Supplementary Materials.
as outputs. The beta parameter evaluates the sensitivity of a given fund to market movements and beta values were calculated by regressing excess returns of a fund over treasury bills to excess returns of the market index over treasury bills. A typical beta of one indicates that the ETF moves perfectly with the market, while its systematic risk is higher if the beta is above one. Regarding the standard deviation, it shows the volatility with which a fund’s returns vary over a certain time horizon. Investors use the standard deviation to predict the volatility of a fund’s returns. The standard deviations of ETFs were calculated using the monthly total returns for the corresponding periods. All monthly standard deviations are annualized.

As for the Sharpe index, also called the reward-to-variability measure, it indicates the reward per unit of risk. The higher the Sharpe index, the better the historical risk-adjusted performance of the fund. To compute the values for the Sharpe index, we use data of the last 36 months and divide the annualized excess returns of a fund over the risk-free rate by the annualized standard deviation of the fund’s returns.

Concerning Jensen’s Alpha, this parameter evaluates the excess return that the portfolio generates over the expected return. A higher Jensen’s Alpha indicates better risk-adjusted returns. The values for Jensen’s Alpha were calculated through the difference between the fund return and the reference return. In this case, the reference return is given by the Capital Asset Pricing Model (CAPM), i.e., the risk-free asset return plus the beta multiplied by the market risk premium (differential between market return and risk-free asset return).

The total trailing return evaluates the return on assets for specific past periods, which allows to track, on a specific date, the performance of the fund over 1 year, 3 years, 5 years, etc. We calculate the trailing total return by determining the changes on the net asset value (NAV) of a fund during the considered period, assuming the reinvestment of all income distributions and capital gains (on the actual reinvestment rate of the fund), and divide it by the initial NAV.

The mean annual return is a percentage that is used to report historical returns, such as average returns over 3, 5 and 10 years for a fund. It is calculated net of the fund’s expense ratio and does not include sales charges or brokerage commissions from investment transactions.

To analyze the efficiency of each DMU (in our case ETFs), we propose three DEA models. The first model (DEA 1) uses beta and standard deviation as inputs (in the spirit of DEA, the less the better) and trailing total return and mean annual return as outputs (in the spirit of DEA the more the better). The second model (DEA 2) uses the same inputs, and the trailing total return and the Sharpe index as outputs (in the spirit of DEA, the more the better). The third model (DEA 3) considers the same inputs again, and the trailing total return and Jensen’s Alpha as outputs (in the spirit of DEA the more the better).

After a first phase of the analysis where the objective is to select the ETFs that can be part of the efficient portfolios through the DEA models, the next phase aims at determining the proportion to be invested in each fund through using multiobjective portfolio models with interval coefficients. After obtaining the average returns, it was necessary to eliminate the outliers, and then, confidence intervals were built, assuming a 95% confidence level. The results of these confidence intervals were considered as the upper and lower limits of the logarithmic returns. To guarantee a certain level of diversification of the portfolio, we have

---

7 To conserve spaces, the values for the inputs and outputs used in our DEA models are presented in Tables A2 and A3 of the Online Supplementary Materials.
Table 2 Efficient DMUs – DEA 1 model

| DMU | Benchmark | Trailing total return | Beta | Standard deviation | Mean annual return |
|-----|-----------|-----------------------|------|--------------------|-------------------|
| KOL | 29        | 31.05                 | 1.14 | 26.01              | 2.54              |
| NLR | 54        | 7.36                  | 0.32 | 9.75               | 0.63              |
| DGAZ| 1         | −32.74                | 3.16 | 140.23             | 3.68              |
| DTO | 4         | 23.62                 | 3.34 | 61.55              | 3.31              |
| KOLD| 2         | −1.51                 | 1.90 | 85.76              | 2.68              |
| NLR | 52        | 4.26                  | 0.41 | 10.14              | 0.39              |
| SZO | 4         | 16.37                 | 1.67 | 30.77              | 1.65              |

This table presents the performance characteristics of the efficient ETFs selected by the DEA 1 model. Benchmark indicates the number of times a DMU is referred as benchmark for the remaining ETFs considered the interval $[u_i^L, u_i^U] = [20\%, 60\%]$. The efficient portfolios were obtained using Excel solver.

5 Results and discussions

The efficiency scores taking into account the different time horizons (i.e., 1, 3 and 5 years for DEA models 1, 2 and 3) are first presented. Then, the efficient portfolios are presented where we show the proportion to be invested in each asset according to aggressive, conservative and combined strategies that vary according to the risk profile of each investor.

The efficiency scores obtained from the DEA 1 model for 3- and 5-year horizons, show respectively 10 DMUs (ENY, FAN, FILL, GRID, KOL, NLR, PBD, PZD, QCLN and UNG) and 9 (DGAZ, DTO, FAN, KOLD, NLR, PZD, SZO, UNG and UNL) with an efficiency score of 1. Since the score of 1 does not guarantee the efficiency of DMUs, we have to assess whether there is a need for each DMU to make improvements to become efficient. This is carried out by examining whether all the constraints of model (4) are satisfied as equalities (i.e., all slack variables are zero).

Table 2 allows us to conclude that, for the DEA 1 model with a horizon of 3 years, there are only two efficient ETFs (KOL and NLR). The KOL ETF belongs to the energy equities category and aims to follow as closely as possible, before taxes and expenses, the price and performance of its underlying index. This fund typically invests 80% of its assets in securities that make up the index. In general, the underlying index is composed of companies in the global coal sector. Concerning the ETF NLR, this fund refers to alternative energy equities category and invests about 80% of its total assets in securities that make up the underlying index (primarily companies involved in uranium and nuclear energy). Note that the ETF KOL differs from the others as it has the highest total trailing return and the mean annual return in

---

8 These intervals guarantee a maximum proportion of capital allocated to assets in the portfolio within the defined range so as to achieve a sufficient diversification of investments, as in expression (10). In this particular case, this means that the investors cannot allocate their investment less than 20% and more than 60% to each ETF, respectively. This limit also allows reducing costs of transactions.

9 See more details in Table A4 of the Online Supplementary Materials.
Table 3 Efficient DMUs – DEA 2 model

| DMU | Benchmark | Trailing total return | Beta | Standard Deviation | Sharpe |
|-----|-----------|-----------------------|------|--------------------|--------|
|     |           |                       |      |                    |        |
| 1 year |         |                       |      |                    |        |
| NLR  | 52       | 5.15                  | 0.30 | 7.96               | 0.42   |
| SZO  | 5        | 14.87                 | 2.28 | 33.35              | 0.50   |
| UNL  | 5        | 10.79                 | 0.13 | 23.13              | 0.46   |
| 3 years |        |                       |      |                    |        |
| KOL  | 41       | 31.05                 | 1.14 | 26.01              | 1.13   |
| NLR  | 51       | 7.36                  | 0.32 | 9.75               | 0.66   |
| 5 years |        |                       |      |                    |        |
| NLR  | 54       | 4.26                  | 0.41 | 10.14              | 0.40   |
| SZO  | 6        | 16.37                 | 1.67 | 30.77              | 0.62   |

This table presents the performance characteristics of the efficient ETFs selected by the DEA 2 model. Benchmark indicates the number of times a DMU is referred as benchmark for the remaining ETFs.

According to the ETF.com report, the ETF KOL stood out in the energy sector with a strong increase from 2016 onwards. Regarding the ETF NLR, it has the lowest beta and standard deviation. Thus, the ETF KOL was indicated as a reference ETF 29 times, while the ETF NLR was considered 54 times as a benchmark for the remaining DMUs.

As for the 5-year horizon, five DMUs (DGAZ, DTO, KOLD, NLR and SZO) are efficient. Thus, these funds were considered as reference ETFs regarding the other inefficient ETFs, 1, 4, 2, 52, and 4 times, respectively. The ETF DGAZ corresponds to the leveraged commodities category and offers daily leverage of 3 times the natural gas prices, which is normally used to express a bear market perspective in the energy sector. The ETF DTO belongs to the same category as the previous one and seeks to offer a 2-times leverage over the underlying index, making it a good investment to express bear market periods of crude oil futures. The ETF KOLD also corresponds to the leveraged commodities category and offers 2 times daily leveraged reverse exposure to natural gas. It seeks to obtain extended returns from an index composed of natural gas future contracts. As for the ETF SZO which focuses on inverse commodities and seeks to obtain desired returns by betting on the devaluation of crude oil, it is highly volatile in view of its standard deviation in Table 2. We also see a large negative value of the total trailing return for the ETF DGAZ due to its stock splits. Even so, this ETF obtained the maximum value for the output mean annual return making it efficient. By analyzing the ETFs on a longer-term basis, we notice an increase in the number of efficient ETFs compared to the previous period. Indeed, a higher performance of the ETFs DGAZ, DTO, KOLD and SZO benefited from the fall in prices of future oil and natural gas contracts in 2014 and 2015. Again, the NLR is indicated as a reference most times (52).

According to the efficiency scores of the DEA 2 model for periods of 1, 3 and 5 years, there are 15 ETFs, 11 ETFs, and 7 ETFs with an efficiency score of 1 for the 1-year, 3-year, and 5-year horizons, respectively (see Table 3).10

The results of the DEA 2 model show the efficient ETFs for the horizons of 1 year (NLR, SZO, and UNL), 3 years (KOL and NLR), and 5 years (NLR and SZO). Again, the ETF NLR is the most important reference for the remaining ETFs regardless of the time horizons. The ETF NLR reaches efficiency for the 3 horizons analyzed, while the ETF SZO returns to

10 See more details in Table A5 of the Online Supplementary Materials.
Table 4 Efficient DMUs – DEA 3 model

| DMU | Benchmark | Trailing total return | Beta | Standard Deviation | Alpha |
|-----|-----------|-----------------------|------|--------------------|-------|
| 1 year |          |                       |      |                    |       |
| CHIE | 11        | −3.04                 | 1.23 | 18.67              | 16.86 |
| NLR  | 46        | 5.15                  | 0.30 | 7.96               | 6.58  |
| SZO  | 2         | 14.87                 | 2.28 | 33.35              | −5.09 |
| UNL  | 11        | 10.79                 | 0.13 | 23.13              | 12.35 |
| 3 years |        |                       |      |                    |       |
| KOL  | 34        | 31.05                 | 1.14 | 26.01              | 22.76 |
| NLR  | 53        | 7.36                  | 0.32 | 9.75               | 4.58  |
| 5 years |        |                       |      |                    |       |
| DTO  | 2         | 23.62                 | 3.34 | 61.55              | 10.47 |
| KOLD | 2         | −1.51                 | 1.90 | 85.76              | 15.23 |
| NLR  | 55        | 4.26                  | 0.41 | 10.14              | 2.33  |
| SZO  | 4         | 16.37                 | 1.67 | 30.77              | 4.85  |

This table presents the performance characteristics of the efficient ETFs selected by the DEA 3 model. Benchmark indicates the number of times a DMU is referred as benchmark for the remaining ETFs.

Efficiency after having already been efficient in the 1-year horizon. The UNL ETF, which was not considered efficient in the DEA 1 model, belongs to the oil & gas category, offers exposure to one of the most important US commodities (natural gas) and is used as a protection for inflation. This ETF diversifies into several maturities to mitigate the adverse impact of the contango effect (i.e., a situation where the future prices of a commodity is higher than the spot price at maturity). The fund is essentially composed of natural gas futures and has the least sensitivity to market movements for one year, given its lowest beta.

Regarding the DEA 3 model, the efficiency scores for the horizons of 1, 3 and 5 years indicate respectively 16 ETFs, 8 ETFs, and 9 ETFs with a score of 1. Table 4 shows the number of efficient ETFs: 4 (CHIE, NLR, SZO, and UNL) for the 1-year horizon, 2 (KOL and NLR) for the 3-year horizon, and 4 (DTO, KOLD, NLR, and SZO) for the 5-year horizon. The ETF NLR remains the most often indicated as a reference for the other ETFs regardless of the time horizons considered. The ETF CHIE appears for the first time as an efficient ETF within the DEA 3 model. It refers to China equities category and offers exposure to the energy sector by investing mainly in shares of companies in the Asian market. With respect to the 5-year period, the ETF KOLD obtained the highest Jensen’s Alpha, suggesting that its efficiency is essentially due to the expansion relative to this output, even though it has a negative total trailing return.

Table 5 provides information on the composition of the possibly efficient portfolios which are determined with respect to the different risk levels and different investment strategies for the ETFs considered to be efficient based on the DEA 1 model (see Table 2). Portfolios 1 and 2 are obtained by following a more aggressive approach (with a more favorable coefficients scenario to obtain a broader feasible admissible region, and a zero-pessimism index). They indeed seek to obtain the maximum value of the upper return and the minimum value of the lower risk. Portfolios 3 and 4 are obtained according to a more conservative strategy (with a more unfavorable coefficients scenario, and a pessimistic index equal to one). They seek to

11 See more details in Table A6 of the Online Supplementary Materials.
| Portfolio | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|---|---|---|---|---|---|---|
| Strategy  | Aggressive | Aggressive | Conservative | Conservative | Balanced | Combined | Combined |
| Risk propensity | Risk prone | Risk averse | Risk prone | Risk averse | Neutral | Risk prone | Risk averse |
| ETFs      | $\varphi = 1$ $\tau = 0$ | $\varphi = 0$ $\tau = 1$ | $\varphi = 1$ $\tau = 0$ | $\varphi = 0$ $\tau = 1$ | $\varphi = 0.5$ $\tau = 0.5$ | $\varphi = 0.7$ $\tau = 0.3$ | $\varphi = 0.3$ $\tau = 0.7$ |
| 3 years   | 0.6 | 0.6 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
| KOL       | 0.6 | 0.4 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
| NLR       | 0.4 | 0.4 | 0.6 | 0.4 | 0.2 | 0.2 | 0.2 |
| 5 years   | 0.4 | 0.4 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 |
| DGAZ      | 0.4 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
| DTO       | 0.4 | 0.2 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 |
| KOLD      | 0.4 | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
| NLR       | 0  | 0  | 0  | 0  | 0  | 0  | 0  |
| SZO       | 0  | 0  | 0  | 0  | 0  | 0  | 0  |

This table presents the composition of the possibly efficient portfolios according to different strategies and risk propensities departing from the efficient ETFs obtained with DEA 1 model.
obtain the maximum value of the lower return and the minimum value of the upper risk, with a feasible region defined with the less favorable coefficients (i.e., obtaining the tightest feasible region). The remaining portfolios (Portfolios 5, 6 and 7) were computed through a combined strategy, assigning a pessimistic index of 0.5 to all objective functions and constraints. For the 3-year horizon, the results suggest that, through a more aggressive strategy (Portfolios 1 and 2), the portfolio should be composed of 60% of the ETF KOL and 40% of the ETF NLR. For the remaining portfolios, the opposite is true with a larger proportion invested in the ETF NLR. Aggressive strategies (portfolios 1 and 2) lead to a higher return, reaching the value range \([-0.018898, 0.020110]\). In turn, a higher return is also implied by a higher risk.

The results for the 5-year horizon, Portfolio 1 (with a greater weight attributed to the return) is composed of 40% of ETF DGAZ, 20% of ETF DTO and 40% of ETF KOLD, while Portfolio 2 (with a greater weight attributed to risk) holds 40% of the ETF DGAZ, 40% of the ETF DTO and 20% of the ETF KOLD. The remaining portfolios have the same composition. Smaller weights are allocated to highly volatile ETFs (DGAZ, DTO, KOLD) and higher weights to less volatile ETFs (NLR and SZO). As expected, the return and risk of these portfolios are much lower than those presented by Portfolios 1 and 2.

It is worth noting that the efficient portfolios presented for the DEA 1 model (3-year horizon) are similar for the DEA 2 and 3 models based on the same horizon since the efficient ETFs were the same. In what follows, we only report the results of 1-year and 3-year horizons of the DEA models 2 and 3.

The efficient portfolios based on the DEA 2 model, reported in Table 6, show that the portfolio of an optimistic scenario (Portfolios 1 and 2) over the 1-year period is composed of 60% ETF UNL and 40% ETF NLR. For the remaining portfolios, 60% of the budget is allocated to the ETF NLR. Portfolios 1 and 2 with the same composition show a return that falls into a range of values \([-0.015653, 0.016250]\), which is slightly higher than the other portfolios \([-0.014792, 0.015218]\). The results for the 5-year period indicate that the ETF NLR has a weight of 60% in Portfolios 1 and 2, but only 40% in the remaining portfolios.

The composition of the efficient portfolios based on the DEA 3 model is reported in Table 7. For 1-year period, when considering the optimistic scenario and risk seeker profile (i.e., high weights assigned to funds with a higher return), we find that the ETFs CHIE and UNL are the ones with the highest return and have an optimal weight of 40% each in Portfolio 1. Concerning Portfolio 2 (still optimistic scenario, but higher weight assigned to funds with lower risk), its optimal composition changes relatively to Portfolio 1 as the ETF CHIE now has a lower weight and the ETF NLR a higher weight. The return of the two portfolios is different. While the first portfolio’s return reaches the range of values \([-0.016724, 0.016949]\), the second portfolio’s return varies in the range of \([-0.015758, 0.016083]\). Regarding a more pessimistic strategy but with a higher weight attributed to the return, Portfolio 3 is composed mostly of the ETF NLR (40%) and the ETF UNL (40%). Still, in a pessimistic scenario but with greater weight attributed to higher risk, Portfolio 4 is composed mostly of the ETF NLR (40%) and the CHIE (40%), thus suggesting that the ETF CHIE is less risky than the ETF UNL. Portfolio 5 adopts a balanced strategy with an equal proportion for return and risk, suggesting a higher allocation rate in the less risky ETF, i.e., the NLR and the one showing the highest return, the UNL. Portfolios 6 and 7 suggest the same proportions as Portfolios 3 and 4, respectively.

Regarding the results for the 5-year period, they indicate that the ETFs with higher returns in aggressive strategies should have a greater weight in the portfolio. For instance, the optimal allocation of Portfolio 1 is consisted of 40% ETF DTO, 40% ETF KOLD, and only 20% ETF NLR. As to the Portfolio 2 which gives more weight to lower risk ETFs, it no longer includes the ETF DTO (0%), but 40% ETF KOLD, 40% ETF NLR (40%) and 20% ETF SZO. For
| Portfolio | 1       | 2       | 3       | 4       | 5       | 6       | 7       |
|-----------|---------|---------|---------|---------|---------|---------|---------|
| Strategy  | Aggressive | Aggressive | Conservative | Conservative | Balanced | Combined | Combined |
| Risk propensity | Risk prone | Risk averse | Risk prone | Risk averse | Neutral | Risk prone | Risk averse |
| ETFs      | $\varphi = 1$ | $\varphi = 0$ | $\varphi = 1$ | $\varphi = 0$ | $\varphi = 0.5$ | $\varphi = 0.7$ | $\varphi = 0.3$ |
|           | $\tau = 0$ | $\tau = 1$ | $\tau = 0$ | $\tau = 1$ | $\tau = 0.5$ | $\tau = 0.3$ | $\tau = 0.7$ |
| 1 year    | NLR 0.4 | 0.4 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |
|           | UNL 0.6 | 0.6 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
| 5 years   | NLR 0.6 | 0.6 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |
|           | SZO 0.4 | 0.4 | 0.6 | 0.6 | 0.6 | 0.6 | 0.6 |

This table presents the composition of the possibly efficient portfolios according to different strategies and risk propensities departing from the efficient ETFs obtained with DEA 2 model.
Table 7 Efficient Portfolios – DEA 3 model

| Portfolio | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|---|---|---|---|---|---|---|
| Strategy  | Aggressive | Aggressive | Conservative | Conservative | Balanced | Combined | Combined |
| Risk propensity | Risk prone | Risk prone | Risk averse | Risk averse | Neutral | Risk prone | Risk averse |

| ETFs | $\varphi = 1$ | $\varphi = 0$ | $\varphi = 1$ | $\varphi = 0$ | $\varphi = 0.5$ | $\varphi = 0.7$ | $\varphi = 0.3$ |
|------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| $\tau = 0$ | 0.4 | 0.2 | 0.2 | 0.4 | 0.2 | 0.2 | 0.4 |
| $\tau = 1$ | 0.2 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 |

1 year

- CHIE: 0.4 0.2 0.2 0.4 0.2 0.2 0.4
- NLR: 0.2 0.4 0.4 0.4 0.4 0.4 0.4
- UNL: 0.4 0.4 0.4 0.2 0.4 0.4 0.4

5 years

- DTO: 0.4 0 0.2 0.2 0.2 0.2 0.2
- KOLD: 0.4 0.4 0 0 0 0 0
- NLR: 0.2 0.4 0.4 0.4 0.4 0.4 0.4
- SZO: 0 0.2 0.4 0.4 0.4 0.4 0.4

This table presents the composition of the possibly efficient portfolios according to different strategies and risk propensities departing from the efficient ETFs obtained with DEA 3 model.
Portfolios 3, 4, 5, 6, and 7, they have similar optimal composition and the largest allocation rate go to less risky ETFs, i.e., the ETF NLR (40%) and the ETF SZO (40%).

Overall, our findings show that renewable energy ETFs are never classified as efficient in any of the models and time horizons employed herein. This evidence corroborates the results of some previous studies. For example, Silva and Cortez (2016), Reboredo et al. (2017), and Rezec and Scholtens (2017) employ distinct linear regression models and found that renewable energies underperform the corresponding benchmarks and, thus, are not viewed as financially attractive portfolio investments. Similar conclusions were obtained by Inchauspe et al. (2015) who show evidence of the underperformance of the renewable energy sector since 2009. Moreover, La Monaca et al. (2018) established that renewable energy ETFs provide only minimal diversification benefits over distinct periods. Recent studies by Martí-Ballester (2019a, 2019b, c) on fund managers’ ability to create value for investors through renewable energy-focused investments also show that, on average, both European and U.S. fund managers were not able to outperform the conventional market benchmark, nor do they achieve higher performance than conventional mutual funds. All in all, these results are not surprising since the profitability of renewable energy ETFs substantially depends on government subsidies, and there has been a cut on the remuneration system for renewables after the recent global financial crisis.

Finally, in general, the application of our methodological framework to portfolio selection and optimization for the U.S. energy-linked ETFs allows us to not only identify the efficient ETFs but also to find optimal composition of the ETF portfolios with respect to risk profiles and investment strategies. The 2-step operational process has computational and time efficiency advantages.

6 Conclusion

We developed a two-step approach to build portfolio models. The first step uses the DEA methodology to obtain a set of assets that provided better financial performance, according to a set of indicators used as inputs and outputs. The second step relies on interval multiobjective portfolio models to find the composition of possibly efficient portfolios while considering three types of investment strategies (aggressive, conservative and combined).

We then tested the practical utility of the proposed methodology using a sample of 60 ETFs from the U.S. energy sector. The sustained growth of ETFs as well as their advantages in terms of investment facility and portfolio diversification is of great interest for both portfolio managers and individual investors. Empirically, three DEA models were employed to identify efficient ETFs, over periods of 1 year (2018), 3 years (2016, 2017 and 2018) and 5 years (2014, 2015, 2016, 2017 and 2018). Beta and standard deviation were used as inputs, and trailing total return, mean annual return, Sharpe index and Jensen’s Alpha as outputs. Once the efficient ETFs are selected, we run the portfolio optimization models under constraints to find their optimal weights.

Our results show that none of renewable energy ETFs was considered efficient. The natural gas and crude oil related ETFs are most often present in the efficient portfolios of ETFs. These findings are consistent with those of Inchauspe et al. (2015) who report evidence for underperformance of the renewable energy sector since 2009. The NLR ETF was always considered as efficient in all DEA models regardless of time horizons and most frequently chosen as a benchmark. The efficiency of this ETF was also documented in Tsolas and Charles (2015) over the period 2008–2010. Other interesting findings relate to a good performance of
the ETF KOL over the 3-year horizon, benefiting from good profitability, essentially obtained in 2016. Regarding the 5-year horizon, we uncover the good performance of the ETFs DGAZ, DTO and KOLD which are leveraged funds with high risk, bet on bear market periods, and ended up benefiting from the fall in future prices of crude oil and natural gas in 2014 and 2015.

In terms of portfolio optimization, we chose to perform the uncertainty treatment based on interval coefficients, both in the objective functions and constraints of the portfolio model. The objective functions considered in the models were the maximization of return and the minimization of risk (measured through the average semi-absolute deviation). In addition to the coherence constraints (budget constraints), constraints were also imposed on a certain level of diversification. The assets that allowed the instantiation of the interval portfolio models were obtained from the different DEA models with different time horizons. Our results suggest that more aggressive strategies give more weight to ETFs that exhibit higher returns but with higher risk, while conservative and balanced strategies are mostly composed of ETFs with lower associated risk but with lower associated returns.

Future research can improve our proposed framework by explicitly incorporating transaction and holding costs into the portfolio models and/or considering the multi-period portfolio selection and other business cycle reference dates. Given the uncertain context of the health crisis related to the COVID-19 pandemic, it would also be interesting to investigate whether these energy-related ETFs continue to perform well and how the optimal composition of the ETF portfolios has changed in response to the pandemic effects. Naturally, our approach can be applied to portfolio selection in sectors other than energy when the investors face the problem of identifying suitable assets to be included in portfolios and of finding their optimal allocation weights.

Supplementary Information  The online version contains supplementary material available at https://doi.org/10.1007/s10479-021-04323-6.

Acknowledgements Carla Oliveira Henriques acknowledges the Portuguese Foundation for Science and Technology through Projects UID/MULTI/00308/2020 and UIDB/05037/2020, and the European Regional Development Fund in the framework of COMPETE 2020 Programme within project T4ENERTEC (POCI-01-0145-FEDER-029820). Elisabete Neves also acknowledges the Portuguese Foundation for Science and Technology under the project UIDB/04011/2020.

References

Admati, A. R., & Ross, S. A. (1985). Measuring investment performance in a rational expectations equilibrium model. *Journal of Business*, 1-26.
Alexopoulos, T. A. (2018). To trust or not to trust? A comparative study of conventional and clean energy exchange-traded funds. *Energy Economics*, 72, 97–107.
Basso, A., & Funari, S. (2016). DEA performance assessment of mutual funds. In *Data Envelopment Analysis* (pp. 229–287). Springer
Basso, A., & Funari, S. (2001). A data envelopment analysis approach to measure mutual fund performance. *European Journal of Operational Research*, 135(3), 477–492.
Biglova, A., Ortobelli, S., Rachev, S. T., & Stoyanov, S. (2004). Different approaches to risk estimation in portfolio theory. *The Journal of Portfolio Management*, 31(1), 103–112.
Blitz, D., & Huij, J. (2012). Evaluating the performance of global emerging markets equity exchange-traded funds. *Emerging Markets Review*, 13(2), 149–158.
Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision-making units. *European Journal of Operational Research*, 2(6), 429–444.
Chen, H. H. (2008). Stock selection using data envelopment analysis. *Industrial Management & Data Systems*, 108(9), 1255–1268.
Chen, P. C., Yu, M. M., Chang, C. C., Hsu, S. H., & Managi, S. (2015). Nonradial directional performance measurement with undesirable outputs: An application to OECD and non-OECD countries. *International Journal of Information Technology & Decision Making, 14*(03), 481–520.

Choi, H. S., & Min, D. (2017). Efficiency of well-diversified portfolios: Evidence from data envelopment analysis. *Omega, 73*, 104–113.

Choi, Y. K. (1995). The sensitivity in tests of the efficiency of a portfolio and portfolio performance measurement. *The Quarterly Review of Economics and Finance, 35*(2), 187–206.

Choi, Y. K., & Murthi, B. P. S. (2001). Relative performance evaluation of mutual funds: A non-parametric approach. *Journal of Business Finance & Accounting, 28*(7–8), 853–876.

Cooper, W. W., Seiford, L. M., & Tone, K. (2007). Efficiency change over time. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*, 323-347.

Daraio, C., & Simar, L. (2006). A robust nonparametric approach to evaluate and explain the performance of mutual funds. *European Journal of Operational Research, 175*(1), 516–542.

Dybvig, P. H., & Ross, S. A. (1985). Differential information and performance measurement using a security market line. *Journal of Finance, 40*(2), 383–399.

Elton, E. J., Gruber, M. J., Das, S., & Hlavka, M. (1993). Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies, 6*(1), 1–22.

Gregoriou, G. N., & Henry, S. C. (2015). Undesirable outputs in commodities trading advisers: A data envelopment analysis approach. *Journal of Wealth Management, 17*(4), 85–92.

Guedj, I., Li, G., & McCann, C. (2011). Futures-based commodity ETFs. *Journal of Index Investing, 2*(1), 14–24.

Gupta, P., Mehlawat, M. K., Inuiguchi, M., & Chandra, S. (2014). Fuzzy portfolio optimization. *Studies in Fuzziness and SoftComputing, 316*.

Haslem, J. A., & Scheraga, C. A. (2003). Data envelopment analysis of Morningstar’s large-cap mutual funds. *The Journal of Investing, 12*(4), 41–48.

Henriques, C. O., Inuiguchi, M., Luque, M., & Figueira, J. (2020). New conditions for testing necessarily possibly efficiency of non-degenerate basic solutions based on the tolerance approach. *European Journal of Operational Research, 283*(1), 341–355.

Henriques, C. O., & Neves, M. E. D. (2019). A multiobjective interval portfolio framework for supporting investor’s preferences under different risk assumptions. *Journal of the Operational Research Society, 70*(10), 1639–1661.

Henriques, C. O., & Neves, M. E. D. (2021). Exploring the trade-off between liquidity, risk, and return under sectoral diversification across distinct economic settings. *The Journal of Risk Finance*. https://doi.org/10.1108/JRF-05-2020-0101

Hsu, C. M. (2014). An integrated portfolio optimisation procedure based on data envelopment analysis, artificial bee colony algorithm and genetic programming. *International Journal of Systems Science, 45*(12), 2645–2664.

Huang, C. Y., Chiou, C. C., Wu, T. H., & Yang, S. C. (2015). An integrated DEA-MODM methodology for portfolio optimization. *Operational Research, 15*(1), 115–136.

Inchauspe, J., Ripple, R. D., & Trück, S. (2015). The dynamics of returns on renewable energy companies: A state-space approach. *Energy Economics, 48*, 325–335.

Isakov, V. (2019). Performance appraisal of exchange-traded funds using clustering and data envelopment analysis. XETRA.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance. In G. Szego & K. Shell (Eds.), *Mathematical methods in investment and finance*. North Holland.

La Monaca, S., Assereto, M., & Byrne, J. (2018). Clean energy investing in public capital markets: Portfolio benefits of yields. *Energy Policy, 121*, 383–393.

Liagkouras, K., & Metaxiotis, K. (2018). A new efficiently encoded multiobjective algorithm for the solution of the cardinality constrained portfolio optimization problem. *Annals of Operations Research, 267*(1), 281–319.

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

Markowitz, H. M. (1968). *Portfolio selection: efficient diversification of investments* (Vol. 16). Yale University Press.

Martí-Ballester, C. P. (2019c). Analyzing alternative energy mutual fund performance in the Spanish market. In *Energy sustainability in built and Urban environments* (pp. 201–213). Springer

Martí-Ballester, C. P. (2019a). Do European renewable energy mutual funds foster the transition to a low-carbon economy? *Renewable Energy, 143*, 1299–1309.
Marti-Ballester, C. P. (2019b). The role of mutual funds in the sustainable energy sector. *Business Strategy and the Environment, 28*(6), 1107–1120.

Miralles-Quirós, J. L., & Miralles-Quirós, M. M. (2019). Are alternative energies a real alternative for investors? *Energy Economics, 78*, 535–545.

Murphy, R., & Wright, C. (2010). An empirical investigation of the performance of commodity-based leveraged ETFs. *Journal of Index Investing, 1*(3), 14–23.

Murthi, B. P. S., Choi, Y. K., & Desai, P. (1997). Efficiency of mutual funds and portfolio performance measurement: A non-parametric approach. *European Journal of Operational Research, 98*(2), 408–418.

Oliveira, C., & Antunes, C. H. (2007). Multiple objective linear programming models with interval coefficients—an illustrated overview. *European Journal of Operational Research, 181*(3), 1434–1463.

Ortobelli, S., Rachev, S. T., Stoyanov, S., Fabozzi, F. J., & Biglova, A. (2005). The proper use of risk measures in portfolio theory. *International Journal of Theoretical and Applied Finance, 8*(08), 1107–1133.

Pagano, M., Sánchez Serrano, A., & Zechner, J. (2019). Can ETFs contribute to systemic risk? (No. 9). Reports of the Advisory Scientific Committee, European Systemic Risk Board.

Papahristodoulou, C., & Dotzauer, E. (2004). Optimal portfolios using linear programming models. *Journal of the Operational Research Society, 55*(11), 1169–1177.

Portela, M. C. A. S., Thanassoulis, E., & Simpson, G. (2004). A directional linear distance approach to deal with negative data in DEA: An application to bank branches. *Journal of Operational Research Society, 55*(10), 1111–1121.

Poterba, J. M., & Shoven, J. B. (2002). Exchange-traded Funds: A New Investment Option for Taxable Investors. *American Economic Review, 92*(2), 422–427.

Reboredo, J. C., Rivera-Castro, M. A., & Ugolini, A. (2017). Wavelet-based test of co-movement and causality between oil and renewable energy stock prices. *Energy Economics, 61*, 241–252.

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

Roll, R. (1978). Ambiguity when performance is measured by the securities market line. *Journal of Finance, 33*(4), 1051–1069.

Sabbagh, O. (2011a). The behavior of green exchange-traded funds. *Managerial Finance, 37*(5), 426–441.

Sabbagh, O. (2011b). Do Green Exchange-Traded Funds Outperform the S&P500. *Journal of Accounting and Finance, 11*(1), 50–59.

Silva, F., & Cortez, M. C. (2016). The performance of US and European green funds in different market conditions. *Journal of Cleaner Production, 135*, 558–566.

Speranza, M. G. (1993). Linear programming models for portfolio optimization. *Finance, 14*(1), 107–123.

Tsolas, I. E. (2019). Utility exchange traded fund performance evaluation. A comparative approach using grey relational analysis and data envelopment analysis Modelling. *International Journal of Financial Studies, 7*(4): 67.

Tsolas, I. E., & Charles, V. (2015). Green exchange-traded fund performance appraisal using slacks-based DEA models. *Operational Research, 15*(1), 51–77.

Xidonas, P., Hassapis, C., Mavrots, G., Staikouras, C., & Zopounidis, C. (2018). Multiobjective portfolio optimization: Bridging mathematical theory with asset management practice. *Annals of Operations Research, 267*(1), 585–606.

Xidonas, P., Mavrots, G., Hassapis, C., & Zopounidis, C. (2017). Robust multiobjective portfolio optimization: A minimax regret approach. *European Journal of Operational Research, 262*(1), 299–305.

Yang, S. Y., Liu, Y., Yu, Y., & Mo, S. Y. K. (2021). Energy ETF return jump contagion: a multivariate Hawkes process approach. *The European Journal of Finance, 1*-23.

Zhang, Y. J., & Chen, M. Y. (2018). Evaluating the dynamic performance of energy portfolios: Empirical evidence from the DEA directional distance function. *European Journal of Operational Research, 269*(1), 64–78.

Zopounidis, C., Doumpos, M., & Pardalos, P. M. (Eds.). (2010). Handbook of financial engineering (Vol. 18). Springer Science & Business Media.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.