

Article

A Multicriteria Extension of the Efficient Market Hypothesis

Francisco Salas-Molina ^{1,*}, David Pla-Santamaria ², Fernando Mayor-Vitoria ²
and Maria Luisa Vercher-Ferrandiz ²

- ¹ Department of Management “Juan José Renau Piqueras”, Faculty of Economics, Universitat de València, Av. Tarongers s/n, 46022 Valencia, Spain
- ² Department of Economics and Social Sciences, Higher Polytechnic School of Alcoy, Universitat Politècnica de València, Ferrándiz y Carbonell, 03801 Alcoy, Spain; dplasan@upv.es (D.P.-S.); fermavi@upv.es (F.M.-V.); maverfer@epsa.upv.es (M.L.V.-F.)
- * Correspondence: francisco.salas-molina@uv.es

Abstract: Challenging the Efficient Market Hypothesis (EMH) has been a recurrent topic for researchers and practitioners since its formulation. Hundreds of empirical studies claim to either prove or disprove the EMH by means of a number of heterogeneous methods. Even though the EMH is usually adjusted to a measure of risk, there is a lack of a formal analysis within a multiple-criteria context. In this paper, we propose an extension of the EMH that accommodates the foundations of multiple-criteria decision analysis. To this end, we rely on a family of parametric signed dissimilarity measures to assess multidimensional performance differences. Since normalization is a critical step in our approach to avoid meaningless comparisons, we present two novel theoretical results connecting different normalization techniques. This multicriteria extension provides a common framework on which to add empirical evidence regarding the EMH testing.



Citation: Salas-Molina, F.; Pla-Santamaria, D.; Mayor-Vitoria, F.; Vercher-Ferrandiz, M.L. A Multicriteria Extension of the Efficient Market Hypothesis. *Mathematics* **2021**, *9*, 649. <https://doi.org/10.3390/math9060649>

Academic Editor: Antonio Jiménez-Martín

Received: 23 February 2021
Accepted: 15 March 2021
Published: 18 March 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: efficient market hypothesis; multiple-criteria test; signed measures; normalization

1. Introduction

A market is considered efficient when its prices show all the important information and, therefore, there is no possibility of using forecasts to beat it. In this context, the EMH appears as a postulate stating that investors can only increase their benefits if they take higher risks, because the prices of the shares show all the available information in the market.

Forecasting returns in financial markets has been a prominent issue since Bachelier [1], Samuelson [2] and Mandelbrot [3], to name a few. The theoretical and empirical research by Fama [4] generated some controversy. EMH supporters argue that investors benefit from a passive portfolio, while critics believe that it is possible to beat the market because shares can deviate from market values due to some anomalies [5]. Most criticisms highlight that the EMH is based on assumptions not related to rationality and other characteristics of human behavior. For instance, Sharma and Kumar [6] review the behavioral finance literature to understand emerging trends in behavioral finance and establish its future potential as a mainstream alternative theory of asset pricing. Peón et al. [7] show that several empirical studies provide evidence in favor of or against the EMH using methodologies based on data from specific stock markets [8,9].

However, it is hard to find a theoretical framework that reflects market-driven changes in a generic way. Furthermore, the existing literature does not clearly reflect all the criteria that can influence the measurement of the EMH. A multicriteria EMH would allow us to incorporate all the market contingencies and changes that can occur due to the economic conditions. Multiple criteria decision making (MCDM) has been widely used in all fields of study, including finance. Recent literature reviews, such as [10] or [11], show a remarkable lack of connection between the EMH and MCDM.

As a result, this paper aims to provide a formulation of the EMH in a multicriteria context. The rationale behind this approach is that, along the lines of the seminal work by Markowitz [12], financial performance is a multidimensional concept. Note that the concept of market efficiency introduced by Fama and the concept of financial efficiency by Markowitz are quite different notions. Efficiency in Markowitz refers to the combined return-risk performance obtained through diversification. In our multicriteria approach, we first describe the critical dimensions of the EMH: (1) a set of criteria; (2) an algorithm to select the best portfolio; (3) a benchmark; (4) a measure of multidimensional performance. We consider basic requirements such as exhaustiveness, cohesiveness, non-redundancy, consistency, compensability and robustness. Later, we describe the main steps of our multidimensional approach to test the EMH and propose a parametric family of dissimilarity measures of performance. Our EMH test proposal is illustrated by means of two numerical examples using two measures of return and two measures of risk.

In sum, the main contribution of this paper is a multidimensional formulation of the EMH based on a family of parametric signed dissimilarity measures. However, since the presence of multiple criteria in our EMH test requires the use of normalization techniques, we present two additional theoretical results:

- A theorem connecting normalization techniques and p -norm transformations;
- A theorem providing a sufficient condition for the percentile normalization to be equivalent to the linear max–min normalization.

The main benefits derived from this paper form a comprehensive tool to help future researchers, individuals and financial institutions to check if their investments have beaten the market from different points of view and not only from the returns perspective.

In addition to this introduction, the paper is organized as follows. Fama’s formulation of the efficient market hypothesis is presented in Section 2. A multidimensional approach to the EMH is described in Section 3. The main steps of our EMH test are considered in Section 4. A numerical example is presented in Section 5 in order to illustrate our theoretical proposal. Finally, we provide some concluding remarks in Section 6.

2. Fama’s Formulation of the Efficient Market Hypothesis and an Empirical Test

According to Fama [4], a market in which asset prices fully reflect all available information about them is called efficient. Since this definition is very general, there is a need to use a model to test efficiency within markets. The usual way to test the EMH is by means of a performance measure such as a time-indexed return for any asset j

$$r_{j,t+1} = \frac{p_{j,t+1} - p_{j,t}}{p_{j,t}}. \tag{1}$$

In practice, prices and returns are random variables denoted by \tilde{p}_t and \tilde{r}_t . The main implication of the EMH under the expected returns theory is that trading systems (algorithms) are unable to produce returns in excess of the equilibrium returns of the market. More formally, given a set of available information Φ_t (a dataset at time t), the realized return $z_{j,t+1}$ at time index $t + 1$ in excess of the equilibrium expected return (what the market expects) at time index t is defined as

$$z_{j,t+1} = r_{j,t+1} - E(\tilde{r}_{j,t+1}|\Phi_t). \tag{2}$$

Focusing now on the random variable representing the excess of returns, the EMH states that expected value of returns $\tilde{z}_{j,t+1}$ in excess of the market expectations given dataset Φ_t is zero

$$E(\tilde{z}_{j,t+1}|\Phi_t) = 0. \tag{3}$$

As a result, any portfolio selection model (or trading algorithm) that selects amount x_j to be allocated in each asset based on dataset Φ_t will obtain the following excess market value

$$v_{t+1} = \sum_{j=1}^q x_j [r_{j,t+1} - E(\tilde{r}_{j,t+1} | \Phi_t)]. \quad (4)$$

However, according to the returns form of the EMH encoded in Equation (3), the expected value in excess of the market obtained by the portfolio selection model is also null

$$E(\tilde{v}_{t+1} | \Phi_t) = \sum_{j=1}^q x_j \cdot E(\tilde{z}_{j,t+1} | \Phi_t) = 0. \quad (5)$$

In the expression of the expected excess value in Equation (5), the focus is placed on returns as a measure of performance. Thus, a simple way to empirically test the EMH is to compare two time-series of returns: one with the returns obtained by a portfolio selection model summarized in vector v , and another one with the returns obtained by a market benchmark in vector b . As long as returns in vector v are consistently above returns in vector b , we can label this EMH test as negative. Similarly, as long as returns in v are consistently below returns in b , we can label this EMH test as positive. The definition of consistency in this test is a real challenge and represents one of the goals of our paper. Our main argument is that consistency should be not only a question of the difference in performance within a given time horizon, but also a question of the number of performance criteria under consideration. Indeed, there are many other ways of measuring performance. Fama [4] states that the expected value is just one of many possible summary measures of a distribution of returns. Furthermore, expected returns are usually linked to a risk measure and we can find different approaches to incorporating risk in portfolio selection. A suitable way to consider multiple measures of efficiency, including returns and risk, but possibly many others, is multiple-criteria decision-making.

3. A Multidimensional Approach to the EMH

In this section, we first introduce the critical dimensions of the portfolio selection problem that we use to extend the EMH to a multiple-criteria context. We next set the basic requirements that our approach aims to fulfill.

3.1. Critical Dimensions to Test the EMH

In order to analyze the consistency of the EMH, we consider a general context, characterized by the following four critical dimensions:

- A n -dimensional set of criteria functions $\mathcal{G} = \{g_1(\cdot), \dots, g_j(\cdot), \dots, g_n(\cdot)\}$;
- An algorithm $\alpha(\Phi_t)$ that determines a candidate portfolio x_t at each time step t based on a given information sequence Φ_t indexed by $t \in \{1, 2, \dots, m\}$;
- A benchmark portfolio b_t that can also be evaluated in terms of set \mathcal{G} defining the equilibrium performance of the market;
- A dissimilarity measure $\mu(A, B)$, where A and B are $m \times n$ data matrices with elements set to $a_{ij} = g_j(x_t)$ and $b_{ij} = g_j(b_t)$, respectively.

Since the proposal by Markowitz of his portfolio selection model [12], both the average and variance of returns have become generally accepted performance measures in finance. Within our framework of analysis, testing the EMH using the Markowitz's model would be equivalent to set $g_1(x_t)$ to the average of returns and $g_2(x_t)$ to the variance of returns derived from a candidate portfolio x_t and compare to the performance of benchmark b_t given by $g_1(b_t)$ and $g_2(b_t)$ according to some dissimilarity measure $\mu(A, B)$.

The rationale behind our multidimensional approach is twofold. First, we aim to generalize the EMH testing to a context in which more than one criteria are considered. Our main argument here is that EMH testing consistency should not only be a question of the difference in performance within the time horizon considered, but also a question of the number of performance criteria under consideration. Second, we propose the use of a general dissimilarity measure that allows us to classify two financial time-series as

significantly different. Summarizing, the main elements of our multidimensional EMH test are the following:

1. Number of criteria under consideration.
2. The time horizon (sample size).
3. A dissimilarity measure.
4. A summary result to prove or disprove the EMH.

Given algorithm $\alpha(\Phi_t)$, determining portfolio x_t and benchmark b_t , the result of our multidimensional EMH test is a classification algorithm $\alpha(\Phi_t)$ as better or worse than benchmark b_t in terms of multiple criteria according to some dissimilarity measures $\mu(A, B)$. In this context, we believe that a multidimensional EMH test needs to fulfill some basic requirements.

3.2. Basic Requirements

With respect to the number of criteria under consideration in a multiple-criteria decision-making context, Roy [13] introduced the concept of coherent set of criteria based on the following logical requirements, which were recently used by [14] in the multicriteria economics context:

- **Exhaustiveness.** The set of criteria must minimize the loss of information that any model implies with respect to the reality that it is trying to represent;
- **Cohesiveness.** The set of criteria must ensure compatibility between the role that each criterion plays and the more comprehensive role that a set of criteria plays when integrating all preferences. Cohesiveness implies that if decision a results in degrading one criterion and decision b results in improving another criterion, then b must outrank a with respect to comprehensive preferences;
- **Non-redundancy.** A set of criteria is non-redundant when leaving out some criterion implies the infringement of either exhaustiveness or cohesiveness.

In order to both extend and accommodate a set of basic requirements to the context of multidimensional EMH testing, we consider the following additional features:

- **Consistency.** We argue that the reliability of an EMH test should be proportional to the number of criteria under consideration, but also to the size of the dataset used. Thus, we here introduce the concept of (time, criteria)-consistency of the test, or (m, n) -consistency, according to some dissimilarity measure $\mu(A, B)$, to describe the power of the test. An EMH test based on returns and risk is more consistent than another EMH based only on returns. This reasoning is behind the usual risk adjustment in regular EMH tests. Similarly, EMH tests considering additional criteria such as diversification, liquidity, dividends, social and environmental responsibility, and the amount of short selling, are more criteria-consistent, provided that exhaustiveness, cohesiveness and non-redundancy are respected. Furthermore, an EMH test based on larger datasets (covering longer periods of time and, ideally, including bull-market periods, bear-market conditions and sideways trends) are more time-consistent than other tests based on smaller datasets;
- **Compensability.** The selection of the aggregation of multiple criteria may have an impact in terms of the compensation among criteria. Here, compensability means the degree to which the achievement of any criterion is offset or balanced by the achievement of another one. This concept is a critical issue when categorizing aggregation methods such as the one described in this paper [15–17];
- **Robustness.** The presence of outliers or measurement errors in the data used to test the EMH may lead to distorted or wrong results. Even though (time-criteria)-consistency may significantly contribute to reducing the impact of outliers in a multidimensional EMH test due to the need for larger datasets, methods and dissimilarity measures that are more robust to the presence of outliers are preferred.

4. Main Steps of a Multidimensional EMH Test

Testing the EMH under our multicriteria approach requires the use of a dissimilarity measure between two matrices A and B to determine if the data contained provide enough empirical evidence to prove or disprove the EMH. Many different dissimilarity measures may be suitable for our end. However, in order to guarantee this suitability, we propose to follow the next main steps:

1. Normalize criteria;
2. Design a weight system;
3. Define a dissimilarity measure;
4. Determine the result for the test.

Next, we further elaborate on the previous steps to highlight the pros and cons of different alternative dissimilarity measures according to the basic requirements described above.

4.1. Normalize Criteria

Since we are possibly dealing with criteria measured in both different units and scales, there is a need for the normalization of criteria in order to avoid meaningless comparisons. Normalization is a data-transformation process that aims to provide a common scale to allow comparisons, and this may affect the obtained results. Vafei et al. [18] provide a description of the main normalization methods, including vector, linear and non-linear normalization, and logarithmic normalization. Both linear and vector normalization are the most widely methods [19]. Non-linear transformations may lead to more diminished values [20], and logarithmic normalization can be used when we want the sum of normalized values to add up to one, but it does not deal well with negative values and values close to zero.

Within the context of the multidimensional EMH, our goal is to measure the difference in performance for each criterion under consideration summarized in two $m \times n$ matrices, A and B , for a given algorithm and a benchmark, respectively. Without loss of generality, let us consider dissimilarity measure $\mu(A, B)$ based on the difference matrix D

$$D = A - B. \quad (6)$$

Linear normalization is achieved by transforming elements d_{tj} of matrix D , when more is better

$$\bar{d}_{tj} = \frac{d_{tj} - \min_t(d_{tj})}{\max_t(d_{tj}) - \min_t(d_{tj})}. \quad (7)$$

As a result, normalized values \bar{d}_{tj} in Equation (7) are restricted to the interval $[0, 1]$. Maximum normalization is achieved when we set $\min_t(d_{tj}) = 0$. Percentage or sum normalization is obtained when computing the ratio between a single value and the total sum of absolute values

$$\bar{d}_{tj} = \frac{d_{tj}}{\sum_{t=1}^m |d_{tj}|}. \quad (8)$$

On the other hand, vector normalization is achieved by means of the Euclidean norm

$$\bar{d}_{tj} = \frac{d_{tj}}{\sqrt{\sum_{t=1}^m d_{tj}^2}}. \quad (9)$$

Note that vector normalization preserves the sign of d_{tj} which may be useful in our context, since we aim to reflect a difference in performance. According to Equation (6), we know that the performance of algorithm α is below the benchmark because d_{tj} is negative. This feature adds interpretability to the normalization method.

By means of the concept of vector norm $\|v\|_p$ with metric p ranging in the closed interval $[0, \infty]$, we next generalize the linear and vector normalization through the following expression

$$\bar{d}_{tj} = \frac{d_{tj} - \min_t(d_{tj})}{\|\mathbf{d}_j\|_p - \min_t(d_{tj})} \tag{10}$$

where $\|\mathbf{d}_j\|_p$ is the p -norm of the j -th column vector of matrix D

$$\|\mathbf{d}_j\|_p = \left[\sum_{t=1}^m |d_{tj}|^p \right]^{1/p}. \tag{11}$$

Lemma 1. *The linear max-min normalization method described in Equation (7) is a special case of p -norm transformations in Equation (10) when $p = \infty$ and $d_{tj} \geq 0 \ \forall t$.*

Proof. Given an $m \times 1$ vector \mathbf{d}_j with its maximum absolute value component $d_{kj} \equiv \max_t(|d_{tj}|)$ placed in the k -th position, we know that p -norm $\|\mathbf{d}_j\|_p$ with $p = \infty$ leads to

$$\|\mathbf{d}_j\|_\infty = \lim_{p \rightarrow \infty} \left[\sum_{t=1}^m |d_{tj}|^p \right]^{1/p} = \lim_{p \rightarrow \infty} \left[\sum_{t=1}^m \left| \frac{d_{tj}}{d_{kj}} \right|^p \cdot d_{kj}^p \right]^{1/p} = \lim_{p \rightarrow \infty} \left[\sum_{\substack{t=1 \\ t \neq k}}^m \left| \frac{d_{tj}}{d_{kj}} \right|^p \cdot d_{kj}^p + d_{kj}^p \right]^{1/p}. \tag{12}$$

Since $|d_{tj}/d_{kj}| < 1$ because d_{kj} is a maximum absolute value component, the first term of the limit tends towards zero because p tends towards infinity and the limit equals d_{kj} , which, by definition, is the maximum absolute value component of vector \mathbf{d}_j

$$\lim_{p \rightarrow \infty} \left[d_{kj}^p \right]^{1/p} = d_{kj} \equiv \max_t(|d_{tj}|). \tag{13}$$

Finally, equality $\max_t(|d_{tj}|) = \max_t(d_{tj})$ holds when vector \mathbf{d}_j contains only non-negative values and Equations (7) and (10) are equivalent for $p = \infty$. \square

Lemma 2. *The linear sum normalization method described in Equation (8) is a special case of p -norm transformations as in Equation (10) when $p = 1$ and $\min_t(d_{tj}) = 0$.*

Proof. Given an $m \times 1$ vector \mathbf{d}_j , we know that p -norm $\|\mathbf{d}_j\|_p$ with $p = 1$ leads to

$$\|\mathbf{d}_j\|_1 = \sum_{t=1}^m |d_{tj}|. \tag{14}$$

Then, Equations (8) and (10) are equivalent when $p = 1$ and $\min_t(d_{tj}) = 0$. \square

Lemma 3. *The vector normalization method described in Equation (9) is a special case of p -norm transformations as in Equation (10) when $p = 2$ and $\min_t(d_{tj}) = 0$.*

Proof. Given an $m \times 1$ vector \mathbf{d}_j , we know that p -norm $\|\mathbf{d}_j\|_p$ with $p = 2$ leads to

$$\|\mathbf{d}_j\|_2 = \sqrt{\sum_{t=1}^m d_{tj}^2}. \tag{15}$$

Then, Equations (9) and (10) are equivalent when $p = 2$ and $\min_t(d_{tj}) = 0$. \square

The concept or p -norm normalization can be extended to the cost perspective when less is better. Furthermore, the implications derived from the generalization proposed in Equation (10) are not restricted to the special cases described in Lemmas 1–3 as the following theorem states.

Theorem 1. Normalization by p -norm transformation described in Equation (10) generalizes linear and vector normalization by attaching a weight which is an increasing exponential function of the absolute value of each component as its base and metric p as its exponent.

Proof. Given an $m \times 1$ vector \mathbf{d}_j , we can rewrite the p -norm $\|\mathbf{d}_j\|_p$ as follows

$$\|\mathbf{d}_j\|_p = \left[\sum_{t=1}^m |d_{tj}|^p \right]^{1/p} = \left[\sum_{t=1}^m |d_{tj}|^{p-1} \cdot |d_{tj}| \right]^{1/p}. \tag{16}$$

Then, terms of the form $|d_{tj}|^{p-1}$ can be viewed as weights controlling the impact of the absolute value of each component of vector \mathbf{d}_j . For $p = 1$, all weights $|d_{tj}|^0$ are equal to one, as in the case of Lemma 2. For $p = 2$, weights $|d_{tj}|^1$ are proportional to the absolute value of each component, as in the case of Lemma 3. For $p > 2$, weights $|d_{tj}|^{p-1}$ become increasing exponential functions of the absolute value of each component raised to metric p . The larger the value of p , the larger the weight attached to the larger components. In the limit, for $p = \infty$, the weight attached to the largest absolute value of components is one and the rest of weights are zero, as in the case of Lemma 1. \square

In search for robustness in the context of the EMH, the presence of outliers or measurement errors in the data used leads us to consider percentile (winsorized) normalization as a way to reduce their impact

$$\bar{d}_{tj}(\lambda) = \frac{f(d_{tj}, \lambda) - p(d_{tj}, \lambda)}{p(d_{tj}, 1 - \lambda) - p(d_{tj}, \lambda)} \tag{17}$$

where $f(d_{tj}, \lambda)$ is a winsorizing function of metric λ , defined as follows

$$f(d_{tj}, \lambda) = \begin{cases} p(\mathbf{d}_j, \lambda) & \text{if } d_{tj} < p(\mathbf{d}_j, \lambda) \\ d_{tj} & \text{if } p(\mathbf{d}_j, \lambda) \leq d_{tj} \leq p(\mathbf{d}_j, 1 - \lambda) \\ p(\mathbf{d}_j, 1 - \lambda) & \text{if } d_{tj} > p(\mathbf{d}_j, 1 - \lambda) \end{cases} \tag{18}$$

where $p(\mathbf{d}_j, \lambda)$ returns the λ -th percentile value of elements in vector \mathbf{d}_j and $p(\mathbf{d}_j, 1 - \lambda)$ returns the $(1 - \lambda)$ -th percentile value of elements in vector \mathbf{d}_j . Typical values for λ and $1 - \lambda$ are 0.05 and 0.95, respectively.

Theorem 2. Given an $m \times 1$ vector \mathbf{d}_j with elements sorted in increasing order, a sufficient condition for the percentile normalization described in Equations (17) and (18) to be equivalent to the linear max–min normalization in Equation (7) is that $\lambda < 1/m$.

Proof. Since the elements of vector \mathbf{d}_j are sorted, it holds that $d_{1j} \leq d_{2j} \leq \dots \leq d_{tj} \leq \dots \leq d_{mj}$ and we can compute the index t_λ of the λ -th percentile element as

$$t_\lambda = \min\{k \in \mathbb{Z} \mid \lambda \cdot m \leq k\}. \tag{19}$$

Then, a sufficient condition for $p(\mathbf{d}_j, \lambda) = \min(\mathbf{d}_j)$ is that $\lambda < 1/m$, since the smallest integer k satisfying $\lambda \cdot m \leq k$ when $\lambda < 1/m$ is 1 because $\lambda \cdot m$ is strictly less than 1 and the first element of our sorted vector is its minimum. Furthermore, if $\lambda < 1/m$ and we define $\beta = 1 - \lambda$, it follows that β satisfies condition $\beta > 1 - 1/m$. Then, the smallest integer k satisfying $\beta \cdot m \leq k$ is m because the following inequality holds:

$$m - 1 < \beta \cdot m < m. \tag{20}$$

Consequently, $p(\mathbf{d}_j, 1 - \lambda) = p(\mathbf{d}_j, \beta) = \max(\mathbf{d}_j)$ since the m -th element of our sorted vector is its maximum. As a result, the pair of Equations (17) and (18) are equivalent to Equation (7). \square

4.2. Design a Weight System

When considering different criteria, we need to design a weight system to assess the relative importance or priority of the criteria under consideration. To this end, several alternatives are available in practice, as described in [16] within the context of composite indicators. The simplest approach is likely that following the Laplace [21] system, which gives equal weights to each criterion. A second alternative is the absence of weights which does not necessarily imply equal weighting, since hierarchical structures may result in different weights for each criterion depending on the number of criteria within a given group. Another group of weight systems derives weights by means of judgments provided by experts. These judgments can be organized in hierarchical structures, as in the case of Analytic Hierarchical Process [22], and summarized by means of aggregations and voting systems [23,24]. Expert judgments or voting systems have the advantage of transparency and participation but may be difficult to apply when there is not a small number of criteria.

Some other systems in which weights are directly obtained from data are Principal Component Analysis, Data Envelopment Analysis [17,25]. One of the advantages of obtaining weights from data is the possibility of avoiding redundancies by studying the correlation among criteria. A different approach is suggested by Ballester [26], in which weights attached to each criterion are inversely proportional to the range of its observed values (maximum minus minimum). The rationale behind this proposal is that decision-makers may distrust criteria with large ranges or deviations due to the likely presence of outliers or measurement errors. Note that, in this case, the study of weights must be done in advance of normalization techniques. More general representations of preferences and weights can be found in [27].

In the context of the EMH testing in which time is a critical variable, weighting criteria may not be the only concern. Users interested in studying the most recent performance of algorithms may choose to overweigh recent observations and underweigh old data. Techniques such as exponential smoothing [28] or moving averages computed over the available data may result in different weights in terms of the time variable.

4.3. Define a Signed Dissimilarity Measure

Weighting and aggregation are usually related issues when considering multiple criteria in a decision-making context. However, it seems useful to separate these two steps to clearly focus on the particular issues of each step. An additional advantage of this approach is that allows for the full comprehension of the possible combinations of weighting and aggregation methods. Recall that our main goal is to study the difference in performance between some algorithm and a given market benchmark. As a result, a basic design requirement is that any aggregation procedure must preserve the sign of the difference in performance to allow for comparisons. This fact motivates the use of the term signed dissimilarity measure instead of aggregation to account for the more general concept of signed measures, which may take negative values [29].

Definition 1. *Tuple (X, Σ, μ) is called a measure space if μ is a measure on a σ -algebra Σ of subsets of set X .*

A σ -algebra Σ defined over X is a collection of subsets of X . An example of σ -algebra is the powerset of X , which is the set of all possible subsets of X , including the empty set \emptyset and X itself.

Definition 2. *Given the measure space (X, Σ, μ) , function $\mu : \Sigma \rightarrow \mathbb{R}$ is called a measure if it satisfies non-negativity, null-empty set and countable additivity:*

1. (Non-negativity) For all element $E \in \Sigma$, we have that $\mu(E) \geq 0$;
2. (Null-empty set) Measure μ over the empty set is null, $\mu(\emptyset) = 0$;

3. (Countable additivity) For all countable collections $\{E_k\}_{k=1}^\infty \in \Sigma$ of pairwise disjoint sets, it holds that

$$\mu\left(\bigcup_{k=1}^\infty E_k\right) = \sum_{k=1}^\infty \mu(E_k). \tag{21}$$

Definition 3. Given the measure space (X, Σ, μ) , function $\mu : \Sigma \rightarrow \mathbb{R} \cup \{-\infty, \infty\}$ is called a signed measure if it satisfies the null-empty set and countable additivity properties and μ takes, at most, one of the values $-\infty$ and ∞ .

A special case of signed measures are finite signed measures:

Definition 4. A signed measure $\mu : \Sigma \rightarrow \mathbb{R}$ defined on a measure space (X, Σ, μ) that takes only real values is called a finite signed measure.

A distinctive feature of our multidimensional approach to the EMH is that we are dealing with time-series data. Then, let us assume that we are given two $m \times n$ matrices A and B for a given algorithm and benchmark, respectively. Both matrices contain m time-indexed observations for n different criteria. For simplicity we assume that observations are aligned (all criteria are of the type the more the better), normalized and weighted in terms of both time and criteria.

Recall from Section 3.2 that consistency, compensability and robustness represent important features to be considered. Consistency mainly refers to the size of the available data including the time period and the number of criteria. Furthermore, both the concepts of compensability and robustness allow us to provide the appropriate semantics in the design of any dissimilarity measure. Although the use of the term aggregation is generalized in the MCDM literature, it may refer to both the sum (additive aggregation) and multiplication (geometric aggregation) of weighted criteria [16,25]. Additive aggregation allows full compensability by considering the weighted arithmetic mean of normalized criteria, while geometric aggregation implies limited compensability due to the multiplicative construction of geometric averaging methods [17].

In what follows, we build on the concept of generalized mean to propose a set of dissimilarity measures (finite signed measures) that best fit the task of testing the EMH with respect to consistency, compensability and robustness. Let us first consider the following generalized mean expression defined for any vector $v \in \mathbb{R}^n$ of non-negative numbers and parameter $p \in [0, \infty]$

$$\mu(v, p) = \left[\frac{1}{n} \sum_{i=1}^n v_i^p \right]^{(1/p)}. \tag{22}$$

The concept of generalized mean can be extended from vectors to matrices. Given a matrix $D \in \mathbb{R}^{m \times n}$ of non-negative numbers, an element-wise generalized mean can be defined relying on vectorization as follows:

$$\mu(D, p) = \mu(\text{vec}(D), p) = \left[\frac{1}{m \cdot n} \sum_{t=1}^m \sum_{j=1}^n d_{tj}^p \right]^{1/p} \tag{23}$$

where $\text{vec}(D)$ is a vectorization operator that transforms a matrix D of dimension $m \times n$ in a vector of dimension $m \cdot n$.

Since we want to consider negative numbers, we next transform generalized means into finite signed measures. First, we preserve the sign of d_{ij} . Second, we use the p -th power of the generalized mean since functions of negative numbers raised to a power $1/p$ are undefined. Note that for optimization and comparison purposes, the power of a measure and the measure itself yield the same result. Third, we remove constant $1/(m \cdot n)$

for simplicity, since the multiplication by a constant does not affect our comparison goal. Thus, we propose the following finite signed measure

$$\mu(D, p) = \sum_{t=1}^m \sum_{j=1}^n \text{sign}^{p+1}(d_{tj}) \cdot d_{tj}^p \quad 0 \leq p < \infty \tag{24}$$

where $\text{sign}(d_{tj})$ is a sign function that returns 1 if $d_{tj} \geq 0$ and -1 if $d_{tj} < 0$. By means of this sign function, Table 1 shows that the finite signed measure described in Equation (24) preserves the sign of term d_{tj} in the aggregation process.

Table 1. Finite signed measure in Equation (24) preserves the sign of d_{tj} .

d_{tj}	p	$\text{sign}(d_{tj}^p)$	$\text{sign}^{p+1}(d_{tj})$	$\text{sign}^{p+1}(d_{tj}) \cdot d_{tj}^p$
$d_{tj} \geq 0$	0, 2, 4, ...	1	1	1
$d_{tj} < 0$	0, 2, 4, ...	1	-1	-1
$d_{tj} \geq 0$	1, 3, 5, ...	1	1	1
$d_{tj} < 0$	1, 3, 5, ...	-1	1	-1

For $p = 1$, we obtain a linear sum of signed observations as follows

$$\mu(D, 1) = \sum_{t=1}^m \sum_{j=1}^n \text{sign}^2(d_{tj}) \cdot d_{tj} = \sum_{t=1}^m \sum_{j=1}^n d_{tj} \tag{25}$$

In addition to consistency, determined by the size of the available dataset, the expression of a finite signed measure in Equation (24) allows us to incorporate the concepts of compensability and robustness. Let us first consider compensability, defined as the degree to which some criteria performances can be compensated by other criteria performances. The maximum level of compensability is achieved by setting $p = 1$. Indeed, $\mu(D, 1)$ is equivalent to the average of signed Manhattan distances between a number of time-indexed performances. As a result, bad performances in one criterion can be fully offset by good performances in another one. The level of compensability can be reduced by increasing the value of p . For $p = 2$, $\mu(D, 2)$ is equivalent to a signed squared quadratic mean, and for $p > 2$, $\mu(D, p)$ is equivalent to a signed generalized mean raised to p , which reduces the level of compensability as long as p increases. The lowest degree of compensability is achieved by setting $p = \infty$ in Equation (23) to obtain the maximum of the elements of matrix D

$$\mu(D, \infty) = \lim_{p \rightarrow \infty} \left[\sum_{t=1}^m \sum_{j=1}^n d_{tj}^p \right]^{1/p} = \max\{d_{11}, d_{12}, \dots, d_{mn}\} \tag{26}$$

Then, a suitable counterpart of the generalized mean that leads to obtain the signed maximum absolute value is the following expression

$$\mu(D, \infty) = \text{sign}(d_{tj}) \cdot \max_{t,j}(|d_{tj}|) \tag{27}$$

If required, the finite signed measured described in Equation (27) yields no compensability at all, but presents two important limitations. First, most of the consistency derived from the $m \cdot n$ observations in matrix D is lost, since only one datapoint is considered. Second, as a consequence of the first limitation, this measure presents very low robustness because it is prone to becoming affected by the presence of outliers. By computing a column-wise maximum for each criterion, it is possible to increase the consistency of the measure, but the robustness is still very weak

$$\mu(D, \infty) = \sum_{j=1}^n \text{sign}(d_{tj}) \cdot \max_t(|d_{tj}|) \tag{28}$$

Finite signed dissimilarity measures derived from the concept of generalized means also allow to limit the degree of compensability along both the time index and the criteria index by considering multiplicative functions such as the geometric mean or Cobb–Douglas production functions. It can be shown that the generalized mean $\mu(v, p)$ with $p = 0$ of a vector v of positive numbers is the geometric mean of order n , which can also be expressed as the exponential of the arithmetic mean of the logarithms of the elements of vector v

$$\mu(v, 0) = \lim_{p \rightarrow 0} \left[\sum_{i=1}^n v_i^p \right]^{(1/p)} = \left[\prod_{i=1}^n v_i \right]^{(1/n)} = e^{(1/n) \sum_{i=1}^n \ln v_i}. \tag{29}$$

By taking advantage of this relationship to logarithms, we next propose a new signed dissimilarity measured of limited compensability derived from the geometric mean for an $m \times n$ matrix D

$$\mu(D, 0) = \sum_{t=1}^m \sum_{j=1}^n -\text{sign}(d_{tj}) \ln(|d_{tj}|) \quad |d_{tj}| < 1. \tag{30}$$

Note that the natural logarithm function is always negative for values of ($|d_{tj}|$) below one. Then, changing the sign allows us to preserve the sign d_{tj} when using normalized values below one, as in Equations (7) and (8).

It is also important to note that techniques based on ideal reference points such as TOPSIS and VIKOR techniques used to rank alternatives according to a set of n different criteria [17,19] are, in essence, signed measures. Consider, for example, the TOPSIS method, in which the i -th alternative is ranked according to an index C_i derived from the Euclidean distance d_i^* to an ideal point and distance d_i^- to an anti-ideal point for each criterion

$$C_i = \frac{d_i^-}{d_i^- - d_i^*}. \tag{31}$$

The higher the value of C_i , the better the rank of alternative i . In our context, with only two alternatives (the algorithm and the benchmark), the problem reduces to a comparison of two alternatives, 1 and 2, that can be expressed as the difference $C_1 - C_2$, which can be positive, negative or zero, as in the case of a finite signed dissimilarity measure. We must remark, however, that setting an ideal and an anti-ideal value as a reference point restricted to the dataset under consideration may be problematic due to the presence of outliers.

A final remark must be done in the sense that many other dissimilarity measures can be designed to test the EMH. However, because of the motives mentioned above, dissimilarity measures derived from the concept of generalized mean provide us with sufficient generality to cover different degrees of compensability and robustness.

4.4. Determine the Result of the EMH Test

Given $m \times n$ matrices A and B , summarizing time-indexed performances ($t = \{1, 2, \dots, m\}$) according to a set of n normalized, weighted and aligned criteria for some algorithm or trading rule and a benchmark, respectively, the next step is to conclude if the algorithm performed better than the benchmark as a result of a multidimensional EMH test by means of dissimilarity measure μ .

A direct way to establish a summary result of the difference in performance is computing matrix $D = A - B$ and applying a finite signed measure $\mu(D, p)$ that allows to control for the compensability of criteria and data robustness, as is usual in a multiple-criteria context. Consistency is determined by the size of matrix D establishing the (time-criteria)-consistency of the test. Other factors being equal, we argue that an (m_1, n_1) -test is more consistent than an (m_2, n_2) -test if, and only if, $m_1 > m_2$ and $n_1 > n_2$. Other factors may include aspects such as quality of data, correlation between criteria and time extension.

If $\mu(D, p) > 0$, we may conclude that the algorithm performed better than the benchmark, but we must quantify the difference in performance. A usual method to represent differences in finance is computing a percentage as in the following expression

$$\Delta(A, B, p) = \frac{\mu(D, p)}{\mu(B, p)} \cdot 100. \tag{32}$$

A clear advantage of using percentages is the possibility of allowing comparisons for different EMH tests. Finally, a summary result that is coherent and compatible with the previous assumptions is: “Algorithm $\alpha(\Phi_t)$ performed Δ % better than benchmark b_t according to dissimilarity measure μ with (time-criteria)-consistency”. An example of a summary result may be: “A buy-and-hold strategy on socially responsible companies in Spain performed 10 % better than market index IBEX35 according to signed Manhattan distances with (10 years of weekly data of risk and return)-consistency”. Note also that Equation (32) can be used to characterize the set of all algorithms that performed better than a given benchmark. As a result, this allows us to describe a family of efficient trading algorithms as a fuzzy subset in the space of all possible algorithms.

5. Illustrative Example

In this section, we illustrate our multicriteria extension of the EMH with some numerical examples. More precisely, we aim to compare the performances of two funds with respect to their respective benchmark indexes:

- The SPDR EURO STOXX 50 ETF seeks to provide investment results that correspond to the performance of the EURO STOXX 50 Index as a benchmark. The EURO STOXX 50 Index is designed to represent the performance of some of the largest companies across components of the 19 EURO STOXX supersectors in terms of capitalization of 11 European countries;
- The iShares STOXX Europe 600 ETF (DE) seeks to track the performance of the EURO STOXX 600 Index as a benchmark. The EURO STOXX 600 Index is designed to represent the performance of a fixed number of 600 components representing large, mid and small capitalization companies among 17 European countries.

By comparing the performances of these two funds with their respective benchmarks, we aim to illustrate how any financial analyst can apply our multicriteria EMH test to any other fund or investment strategy. We do not claim that the main goal of these particular funds is to beat their respective benchmark. On the contrary, we use these funds because we expect similar performances that allow us to illustrate our multicriteria EMH test. To this end, we use two measures of return performance and two measures of risk. Since our initial dataset contains 3130 daily market closing values V_t for six days in a week between 3 January 2011 and 31 December 2020 of two funds and two benchmark indexes, where t is a time index in days, we use the following measures of return:

- Weekly return (WR), computed as the difference in value during 6 days

$$WR = \frac{V_t - V_{t-6}}{V_{t-6}}. \tag{33}$$

- Compound annual growth rate (CAGR), as a proxy a constant rate of return over a given time period

$$CAGR = \left(\frac{V_t}{V_{t-6}} \right)^6 - 1. \tag{34}$$

Additionally, we use the following measures of risk:

- Sample standard deviation (SD) of six daily returns grouped by weeks

$$SD = \left[\frac{1}{5} \sum (r_t - r_w)^2 \right]^{1/2} \tag{35}$$

where r_w is the average return in the w -th week;

- Maximum drawdown (MD) is the peak-to-trough decline in value during a week

$$MD = \frac{V_{tL}}{V_{tP}} - 1 \quad (36)$$

where V_{tL} is the trough value and V_{tP} is the peak value during a week satisfying that $V_{tL} < V_{tP}$ and $tL > tP$. Otherwise, MD for a given week is set to zero. As a result, MD is always non-positive.

Once we have selected and computed the measures of return and risk that form a multidimensional EMH test, we follow the steps described in Section 4.

5.1. Alignment and Normalization

Our test is based on two measures of return (more is better) and two measures of risk (less is better). Then, when computing difference matrix $D = A - B$, we align SD by changing the sign of the subtraction and leave MD without change because it is always non-positive. As a result, matrix D is constructed by horizontally stacking four difference vectors in which each element is positive if the fund performed better than the benchmark, negative if the fund performed worse than the benchmark and zero if both the fund and the benchmark performed equally

$$D = [(WR_A - WR_B) \quad (CAGR_A - CAGR_B) \quad (SD_B - SD_A) \quad (MD_A - MD_B)]. \quad (37)$$

In order to avoid meaningless comparisons between measures of return and risk, we use percentage normalization, as described in Equation (8). The reason for choosing this type of normalization is that it preserves the sign of the initial value. This requirement is a need when using signed dissimilarity measures.

5.2. Neutral Weight System

For simplicity, in this example we use a neutral weight system. Then, all measures of return and risk are given the same importance.

5.3. Defining Eight Different Dissimilarity Measures

In our numerical example, we use eight dissimilarity measures. First, we consider four different combinations of return and risk measures:

- Case 1. Considering only a measure of return (WR);
- Case 2. Considering a measure of return (WR) and a measure of risk (SD);
- Case 3. Considering two measures of return (WR and $CAGR$) and a measure of risk (SD);
- Case 4. Considering two measures of return (WR and $CAGR$) and two measures of risk (SD and MD).

Second, from the general signed dissimilarity measure described in Equation (24), we consider metrics $p = 1$ and $p = 0$. By setting $p = 1$, we obtain Equation (25) that considers the signed sum of difference in multicriteria performance as a suitable alternative for full compensability among criteria. By setting $p = 0$, we obtain (30) that considers the signed sum of logarithm of performance as a suitable alternative for limited compensability. Equation (28) is discarded because the test would be based only on a single observation and it would be very prone to lack of robustness due to the presence of outliers. Combining the four different cases and the two metrics mentioned above, we perform eight EMH tests based on different dissimilarity measures.

5.4. Results of the EMH Test

The results of our EMH tests for the EMH test SPDR EURO STOXX 50 ETF using the EURO STOXX 50 index as a benchmark are summarized in Table 2. More precisely, we compute percentage differences in performance using Equation (32). The first interesting finding that we observe is that when considering only a measure of return and (Case 1), the fund performed better than the benchmark. However, when considering WR as a measure

of return and SD as a measure of risk (Case 2), the sign of the EMH test changed. Adding a new measure of return such as CAGR (Case 3) reduced the difference in performance. Finally, when considering two measures of return and two measures of risk, (Case 4) produced a new increase in the difference in global performance. In addition, there is a correlation between the EMH test with $p = 1$ (full compensability) and $p = 0$ (limited compensability). In sum, these results show that the SPDR EURO STOXX 50 ETF performed better than the benchmark EURO STOXX 50 in terms of WR as a measure of return, but worse in terms of risk and also globally when considering the rest of the measures. This change in the sign of performance when adding a measure of risk gives grounds for our multidimensional approach to the EMH in search for consistency.

Table 2. EMH test SPDR EURO STOXX 50 ETF versus EURO STOXX 50.

$\Delta(A, B, p)$	Case 1 (WR)	Case 2 (WR; SD)	Case 3 (WR; SD; CAGR)	Case 4 (WR; SD; CAGR; MD)
$p = 1$	8.6%	−109.4%	−49.6%	−75.5%
$p = 0$	24.8%	−80.5%	−42.0%	−68.6%

The results of our EMH tests for the iShares STOXX Europe 600 ETF (DE) using the EURO STOXX 600 index as a benchmark are summarized in Table 3. In this second numerical example, we observe that iShares STOXX Europe 600 ETF (DE) performed better during the period under investigation than the EURO STOXX 600 benchmark. This better performance occurred not only for the first measure of return as in the previous case, but for the whole range of measures of return and risk. However, there is a reduction for the difference when adding SD as a measure of risk (Case 2), and an increase in the difference in Cases 3 and 4. Again, we observe a correlation between differences in performance using $p = 1$ (full compensability) and $p = 0$ (limited compensability). Signs and direction of change are aligned in both metrics and less variability in performance differences in the case of $p = 0$. A possible explanation for this reduced variability could be caused by the effect of limited compensability introduced by the construction of this signed measure.

Table 3. iShares STOXX Europe 600 ETF (DE) versus EURO STOXX 600.

$\Delta(A, B, p)$	Case 1 (WR)	Case 2 (WR; SD)	Case 3 (WR; SD; CAGR)	Case 4 (WR; SD; CAGR; MD)
$p = 1$	39.7%	4.3%	7.3%	204.6%
$p = 0$	40.2%	26.7%	28.4%	41.6%

6. Concluding Remarks

The results of this paper show that our multicriteria approach to testing the EMH enriches financial analysis from a double perspective. From a theoretical perspective, a multicriteria EMH test extends existing approaches to consider not only a measure of return but several measures of return, risk and possibly many other measures of performance, such as liquidity, environmental, social and governance measures. The main motivation for following such an approach is an increase in consistency when testing the EMH. As a result, we link consistency in an EMH test to the temporal extension of the dataset and the number of criteria under consideration. The rationale behind this approach follows a simple logic: the longer the time window of available data the better, and the larger the number of criteria the better.

From an empirical perspective, our EMH test is based on signed dissimilarity measures that allow practitioners to deal with compensability and robustness. Both requirements can be adjusted by selecting a single metric. The numerical results presented in the illustrative example show how a multicriteria EMH test can be used to better characterize empirical claims regarding the performance of funds and trading algorithms.

The presence of multiple criteria in our EMH test requires the use of normalization techniques. In this context, we also present further theoretical results on normalization. By means of the concept of p -norm, we show in Theorem 1 that the most widely used methods of normalization are special cases of p -norm transformations. In addition, Theorem 2 provides a sufficient condition for percentile normalization to be equivalent to linear max–min normalization. In sum, this paper extends the EMH to a multiple criteria context from both a theoretical and empirical perspective. Analysts and practitioners have now the opportunity to provide further evidence on the multiple-criteria performance of investment algorithms and trading rules. An example of an interesting direction of future research is the empirical characterization of subsets of efficient trading algorithms in the space of all possible algorithms.

Author Contributions: Conceptualization, F.S.-M., D.P.-S., M.L.V.-F. and F.M.-V.; Data curation, F.S.-M., D.P.-S., M.L.V.-F. and F.M.-V.; Formal analysis, F.S.-M., M.L.V.-F. and F.M.-V.; Investigation, F.S.-M., D.P.-S., M.L.V.-F. and F.M.-V.; Methodology, F.S.-M., D.P.-S., M.L.V.-F. and F.M.-V.; Project administration, F.S.-M. and D.P.-S.; Resources, D.P.-S.; Supervision, F.S.-M., M.L.V.-F. and F.M.-V.; Validation, F.S.-M.; Visualization, F.M.-V.; Writing—original draft, F.S.-M., F.M.-V. and D.P.-S.; Writing—review and editing, F.S.-M., D.P.-S., M.L.V.-F. and F.M.-V. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Bachelier, L. Théorie de la spéculation. *Ann. Sci. L'Éc. Norm. Supér.* **1900**, *17*, 21–86. [[CrossRef](#)]
- Samuelson, P.A. Proof that properly anticipated prices fluctuate randomly. *Ind. Manag. Rev.* **1965**, *6*, 41–49.
- Mandelbrot, B.B. The variation of certain speculative prices. In *Fractals and Scaling in Finance*; Springer: Berlin/Heidelberg, Germany, 1997; pp. 371–418.
- Fama, E.F. Efficient capital markets: A review of theory and empirical work. *J. Financ.* **1970**, *25*, 383–417. [[CrossRef](#)]
- Naseer, M.; Bin Tariq, D. The efficient market hypothesis: A critical review of the literature. *IUP J. Financ. Risk Manag.* **2015**, *12*, 48–63.
- Sharma, A.; Kumar, A. A review paper on behavioral finance: study of emerging trends. *Qual. Res. Financ. Mark.* **2019**, *12*, 137–157. [[CrossRef](#)]
- Peón, D.; Antelo, M.; Calvo, A. A guide on empirical tests of the EMH. *Rev. Account. Financ.* **2019**, *18*, 268–295. [[CrossRef](#)]
- Borges, M.R. Efficient market hypothesis in European stock markets. *Eur. J. Financ.* **2010**, *16*, 711–726. [[CrossRef](#)]
- Jain, K.; Jain, P. Empirical study of the weak form of EMH on Indian stock market. *Int. J. Manag. Soc. Sci. Res.* **2013**, *2*, 52–59.
- Rossi, M. The efficient market hypothesis and calendar anomalies: a literature review. *Int. J. Manag. Financ. Account.* **2015**, *7*, 285–296. [[CrossRef](#)]
- Jethwani, D.; Ramchandani, K. Semi Strong Form of Efficiency of Stock Market: A Review of Literature. *Int. Multidiscip. Res. J.* **2017**, *4*, 2349–7637.
- Markowitz, H. Portfolio selection. *J. Financ.* **1952**, *7*, 77–91.
- Roy, B. *Multicriteria Methodology for Decision Aiding*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1996.
- Salas-Molina, F. A formal specification of multicriteria economics. *Oper. Res.* **2019**, *1*–24. [[CrossRef](#)]
- Munda, G. Multiple criteria decision analysis and sustainable development. In *Multiple Criteria Decision Analysis: State of the Art Surveys*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 953–986.
- Greco, S.; Ishizaka, A.; Tasiou, M.; Torrisi, G. On the methodological framework of composite indices: A review of the issues of weighting, aggregation, and robustness. *Soc. Indic. Res.* **2019**, *141*, 61–94. [[CrossRef](#)]
- Garcia-Bernabeu, A.; Hilario-Caballero, A.; Pla-Santamaria, D.; Salas-Molina, F. A Process Oriented MCDM Approach to Construct a Circular Economy Composite Index. *Sustainability* **2020**, *12*, 618. [[CrossRef](#)]
- Vafaei, N.; Ribeiro, R.A.; Camarinha-Matos, L.M. Normalization techniques for multi-criteria decision making: Analytical hierarchy process case study. In Proceedings of the Doctoral Conference on Computing, Electrical and Industrial Systems, Costa de Caparica, Portugal, 11–13 April 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 261–269.

19. Opricovic, S.; Tzeng, G.H. Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS. *Eur. J. Oper. Res.* **2004**, *156*, 445–455. [[CrossRef](#)]
20. Zavadskas, E.K.; Turskis, Z. A new logarithmic normalization method in games theory. *Informatica* **2008**, *19*, 303–314. [[CrossRef](#)]
21. Laplace, P.S. *Essai Philosophique sur les Probabilités*; H. Remy: Paris, France, 1825.
22. Saaty, T.L. *The Analytic Hierarchy Process*; Mc Graw-Hill: New York, NY, USA, 1980.
23. González-Pachón, J.; Romero, C. Bentham, Marx and Rawls ethical principles: In search for a compromise. *Omega* **2016**, *62*, 47–51. [[CrossRef](#)]
24. González-Pachón, J.; Diaz-Balteiro, L.; Romero, C. A multi-criteria approach for assigning weights in voting systems. *Soft Comput.* **2019**, *23*, 8181–8186. [[CrossRef](#)]
25. El Gibari, S.; Gómez, T.; Ruiz, F. Building composite indicators using multicriteria methods: A review. *J. Bus. Econ.* **2019**, *89*, 1–24. [[CrossRef](#)]
26. Ballesteros, E. Strict uncertainty: A criterion for moderately pessimistic decision makers. *Decis. Sci.* **2002**, *33*, 87–108. [[CrossRef](#)]
27. Salas-Molina, F.; Pla-Santamaria, D.; Garcia-Bernabeu, A.; Reig-Mullor, J. A Compact Representation of Preferences in Multiple Criteria Optimization Problems. *Mathematics* **2019**, *7*, 1092. [[CrossRef](#)]
28. Hyndman, R.; Koehler, A.B.; Ord, J.K.; Snyder, R.D. *Forecasting with Exponential Smoothing: The State Space Approach*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008.
29. Bogachev, V.I. *Measure Theory*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2007; Volume 1.