KNOWCON
Knowledge on Economics and Management
International Scientific Conference

Recent Advances in Business Analytics. Selected papers of the 2021 KNOWCON-NSAIS workshop on Business Analytics pp. 63–73

DOI: 10.15439/2021B2
ISSN 2300-5963 ACSIS, Vol. 29

Possible drivers of high performance of European mutual ESG funds - an fsQCA view on sustainable investing

Fanni Welling
School of Business and Management, LUT University,
Yliopistonkatu 34, 53851 Lappeenranta, Finland
Email: fanni.welling@student.lut.fi

Jan Stoklasa
School of Business and Management, LUT University
Yliopistonkatu 34, 53851 Lappeenranta, Finland, and
Palacky University Olomouc, Faculty of Arts
Department of Economic and Managerial Studies
Email: jan.stoklasa@lut.fi
Email: jan.stoklasa@upol.cz

***Abstract*—The paper applies the tools of fsQCA and their recent modifications by Stoklasa, Luukka and Talášek to analyze the possible drivers of high performance of European ESG funds. 429 mutual equity growth ESG funds from the European area are being analyzed. We focus mainly on the connection of Morningstar sustainability rating with the performance of the funds during 2018-2021 measured by Jensen's alpha and the Sharpe ratio. Other possible drivers of the success of these funds are also being explored. We identify the prevailing assumed relationships between funds' sustainability and other characteristics with their performance and formulate rules to be investigated using the fsQCA methodology. More specifically the possibility of high performance being associated with a high sustainability rating of the funds is explored in detail. Our results indicate that although the high performance cannot be clearly associated with the high sustainability rating of a fund, high sustainability rating seems to be preventing the low performance of the fund.**

I. Introduction

SUSTAINABILITY and responsibility are not only topical issues in business scientific literature and practice [1], [2], but these concepts are also potentially influencing the investment decision-making of individual investors. In this paper, we discuss three factors that might potentially influence the performance of mutual funds, namely the size of the fund, the length of its managers' tenure and its sustainability rating, show the relationships that have already been identified in the literature between these factors and the performance of the fund. In line with the usual approaches in the literature, the performance of the funds is measured using Jensen's alpha and the Sharpe ratio in this paper.

We then set the goal of validating the existence of the "prevailing" relationships on a chosen sample of 429 European growth funds in the 2018-2021 period. Given the fact that most studies (see the brief literature reviews for each feature further in the text) use statistical methods (regression etc.) to investigate the existence of relationships, it is reasonable to try to verify the (non)existence of the relationships between the chosen features and the fund performance using a different methodology.

We therefore apply the tools of the set-theoretic approach and its fuzzification, that are utilized in the frame of the fuzzy set qualitative comparative analysis (fsQCA) - namely we focus on the concepts of the consistency of the rules representing specific assumed relationships with the data and the coverage of these relationships by the available data [3], [4]. Given the recent advances in the methods for fsQCA focusing on the investigation of consistency and coverage of assumed relationships in the fuzzy context, we also apply the recently introduced fuzzified consistency and coverage measures and their alternatives [5], [6]. Another reason to reach for the set-theoretic methods is the fact that based on the definition of the rules (investigated relationships formulated as IF-THEN rules) we can postulate and verify the existence of non-linear relationships between the features of the funds and their performance. This has proven to be beneficial in the recent studies on strategic decision-making [7], [8].

Even though our focus is mainly on the possible relationship between sustainability (or sustainability ratings) of the funds and their performance, we include the other fund features too to be able to assess the performance of the fsQCA methods on the data. This way we will be able to interpret the results concerning sustainability in the context of fund size and manager tenure as well. Other potentially relevant features such as green approach to HR management [9], corporate social responsibility or company's reputation [1], [2] and others are left out of the scope of this paper.

II. Preliminaries

Let U be a nonempty set. A *fuzzy set* A on U is defined by a mapping $\mu_A : U \to [0,1]$, where μ_A is called a *membership function* of A (see e.g. [10], [11] for more details). The set of all fuzzy sets on U is denoted $\mathcal{F}(U)$. For simplicity, we can denote a fuzzy set and its membership function by the same

This work was supported by LUT research platform AMBI- Analytics-based management for business and manufacturing industry and partially also by the grant IGA_FF_2021_001 Barriers to the expansion of sustainable consumption.

symbol (that way the membership function of a fuzzy set A will be denoted $A(.)$). Let $A \in \mathcal{F}(U)$, then

- the *kernel* of A is a crisp set $\mathrm{Ker}(A) = \{x \in U \mid A(x) = 1\}$.
- the *support* of A is a crisp set $\mathrm{Supp}(A) = \{x \in U \mid A(x) > 0\}$.
- the *height* of A is $\mathrm{hgt}(A) = \sup\{A(x) \mid x \in U\}$
- the $\alpha-$*cut* of A is a crisp set $A^\alpha = \{x \in U \mid A(x) \geq \alpha\}$ for any $\alpha \in [0, 1]$.

A *negation of a fuzzy* set $A \in \mathcal{F}(U)$ is a fuzzy set $\neg A \in \mathcal{F}(U)$ such that for any $x \in U$ we have $\neg A(x) = 1 - A(x)$. Let A be a fuzzy set on $\mathbb{R}$, such that all the following conditions are met:

1) A is normal that is, $\mathrm{hgt}(A) = 1$,
2) A^α is a closed interval for all $\alpha \in (0, 1]$,
3) $\mathrm{Supp}(A)$ is bounded,

then A is called a *fuzzy number* on $\mathbb{R}$, denoted as $A \in \mathcal{F}_N(\mathbb{R})$. Each fuzzy number $B \in \mathcal{F}_N(\mathbb{R})$ can be represented by a quadruple of characteristic values $B \sim (b_1, b_2, b_3, b_4)$, where $b_1, ..., b_4 \in \mathbb{R}$, $b_1 \leq b_2 \leq b_3 \leq b_4$, and $[b_1, b_4] = \mathrm{Cl}(\mathrm{Supp}(B))$, $[b_2, b_3] = \{x \in \mathbb{R} \mid B(x) = 1\} = \mathrm{Ker}(B)$ and $B(x) = 0$ for all $x \in (-\infty, b_1] \cup [b_4, \infty)$. For a *triangular fuzzy number* we have $b_2 = b_3$ and the membership function is continuous, linear and strictly increasing between the points b_1 and b_2 and continuous, linear and strictly decreasing between b_3 and b_4. For a *trapezoidal fuzzy number* we assume the same, we just allow $b_2 \neq b_3$. If $[b_1, b_4] \subseteq [r, s]$ we call B a *fuzzy number on an interval* $[r, s]$. The set of all fuzzy numbers on an interval $[r, s]$ will be denoted $\mathcal{F}_N([r, s])$. In this paper, we will only consider these two types of fuzzy numbers to represent the linguistically defined values of the features under investigation.

As the main methodology chosen for this paper is the set-theoretic investigation of the consistency of the investigated rules with the data, we will need to introduce the basic (fuzzy) set-theoretic concepts of consistency and coverage as used in the fsQCA [12] and as recently generalized by Stoklasa et. al [5], [6]. We will be employing the revised fuzzification of the consistency and coverage measures [5], [6] as these have already proven useful in practical investigation of real-life relationships in business data [7]. Let us consider a set of observations $U = \{x_1, x_2, \ldots, x_n\}$. Let us consider a feature A and an indicator function $\chi^A : U \to \{0, 1\}$ such that $\chi^A(x_i) = 1$ if and only if x_i has the feature A and $\chi^A(x_i) = 0$ otherwise, for all $i = 1, \ldots, n$. Let us also consider a feature B with an analogous indicator function $\chi^B : U \to \{0, 1\}$. Let us also introduce a negation of the feature B representing the absence of the feature B (denoted B' and meaning "not B"), for which the indicator function is $\chi^{B'} : U \to \{0, 1\}$ such that $\chi^{B'}(x_i) = 1$ if and only if x_i does not have the feature B and $\chi^{B'}(x_i) = 0$ otherwise. In other words, we have $\chi^B(x_i) = 1 - \chi^{B'}(x_i)$ and as long as "possessing a feature" is considered as a crisp (binary) state, we have $\chi^B(x_i), \chi^{B'}(x_i) \in \{0, 1\}$. Now we assume that we need to investigate the assumption that *an observation having a feature A also implies it having the feature B as well*, or $A \Rightarrow B$ for short. Given the set of observations U and given $A \subseteq U, B \subseteq U$, we can assess the consistency [12] of such a crisp assumption with the data (its support by the data) computing the consistency of $A \Rightarrow B$:

$$\mathrm{Consistency}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\{\chi^A(x_i), \chi^B(x_i)\}}{\sum_{i=1}^{n} \chi^A(x_i)} = \frac{\mathrm{Card}(A \cap B)}{\mathrm{Card}(A)}, \quad (1)$$

where $\mathrm{Card}(A)$ represents the cardinality of the set A, i.e. the number of its elements, and $\cap$ is the standard set intersection, i.e. $\chi^{(A \cap B)}(x_i) = \min\{\chi^A(x_i), \chi^B(x_i)\}$. Note, that U is fully consistent with $A \Rightarrow B$ as long as $A \subseteq B$ (which implies that $A \cap B = A$), i.e. in this case $\mathrm{Consistency}(A \Rightarrow B) = 1$ and we can interpret this as the absence of counterexamples to $(A \Rightarrow B)$; obviously we need to assume that $\mathrm{Card}(A) \neq 0$. If the cardinality of A was zero, then there would be no observations that possess the feature A and it would make no sense to try to investigate the compatibility of the assumption $A \Rightarrow B$ with the given dataset. Analogously we can calculate a measure of "universality" of the assumption $A \Rightarrow B$ for the given set of observations U as the coverage of $A \Rightarrow B$ (assuming again that $\mathrm{Card}(B) \neq 0$):

$$\mathrm{Coverage}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\{\chi^A(x_i), \chi^B(x_i)\}}{\sum_{i=1}^{n} \chi^B(x_i)} = \frac{\mathrm{Card}(A \cap B)}{\mathrm{Card}(B)}. \quad (2)$$

Apparently $\mathrm{Coverage}(A \Rightarrow B) = 1$ if and only if $B \subseteq A$. In other words, both measures are based on subsethood. This means that the validity of the assumption that A leads to B is assessed based on the available data - if the set of observations having feature A is a subset of those observations that have the feature B, then having the feature A can be considered a sufficient condition for having the feature B too (see [12] or [5] for more details). If the possession of the feature can be understood in gradual and not binary terms, a fuzzification of the whole approach is necessary. We can still assume that the possession of the feature A by an element of U can be described by its membership to A, we just need to allow $A \in \mathcal{F}(U)$, that is we need to allow for A to be a fuzzy subset of U.

If we now assume that A and B are fuzzy sets ($A, B \in \mathcal{F}(U)$) and $\mu_A : U \to [0, 1]$ and $\mu_B : U \to [0, 1]$ are their respective membership functions, we need to introduce at least the fuzzy-set subsethood, fuzzy-set intersection operation and the notion of a cardinality of a fuzzy set to be able to generalize (1) and (2). The intersection of two fuzzy sets A and B on the same universe U is a fuzzy set $(A \cap B)$ on U with the membership function $\mu_{A \cap B} : U \to [0, 1]$ such that for any $x \in U$ we have $\mu_{A \cap B}(x) = \min\{\mu_A(x), \mu_B(x)\}$. A is a fuzzy subset of B (denoted $A \subseteq_F B$) if for all $x \in U$ it holds that $\mu_A(x) \leq \mu_B(x)$. The cardinality of a fuzzy set $A \in \mathcal{F}(U)$ is calculated as $\mathrm{Card}(A) = \sum_{x_i \in U} A(x_i)$ as long as U is a discrete set, and $\mathrm{Card}(A) = \int_{x_i \in U} A(x_i) dx$ as long as U is a continuous universe (e.g. a subinterval of the real axis). The direct fuzzification of (1) and (2) stemming from the subsethood interpretation of consistency and coverage can be expressed by the following formulas [12] (the fuzzified formulas will be denoted by the subscript F):

$$\text{Consistency}_{F_1}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\left(\mu_A(x_i), \mu_B(x_i)\right)}{\sum_{i=1}^{n} \mu_A(x_i)}, \tag{3}$$

$$\text{Coverage}_{F_1}(A \Rightarrow B) = \frac{\sum_{i=1}^{n} \min\left(\mu_A(x_i), \mu_B(x_i)\right)}{\sum_{i=1}^{n} \mu_B(x_i)}. \tag{4}$$

Stoklasa et al. [5] proposed a different fuzzification of (1) and (2) that deals with the fact that the transition to a gradual possession of a feature ultimately implies that a feature can be partially possessed and partially not possessed (in a nonzero degree) by the same observation at the same time. This results in the ambivalence of evidence in the set-theoretic investigation, as the same observation can now simultaneously support $A \Rightarrow B$ and $A \Rightarrow B'$ to some extent. Stoklasa et al. therefore suggested several alternative fuzzifications of formulas (1) and (2) - namely the F_2 fuzzification [5] represented by formulas (5) and (6) that removes that part of evidence that is ambivalent, F_3 fuzzification [5] that focuses of "pure support" of the investigated relationship by removing ambivalent evidence as well as reducing the evidence by the amount of available "pure" counterevidence represented by formulas (7) and (8). Finally, a modification of (7) and (8) was proposed in [6] that deals with the partial loss of information introduced to F_3 formulas by the use of the maximum operator. These F_4 fuzzifications are represented by formulas (9) and (10); note that the results of these formulas have a slightly different interpretation - for example if $\text{Consistency}_{F_4}(A \Rightarrow B) = 0.5$, then there is the same amount of "pure" evidence as there is counterevidence with regards to the investigated relationship, whereas if $\text{Consistency}_{F_4}(A \Rightarrow B) = 1$, then there is only "pure evidence" in its favor etc. It should be noted that Stoklasa et al. also proposed a completely different approach to the assessment of consistency and coverage of the investigated relationships [5] represented by the degree of (unconditional) support and degree of (unconditional) disproof, that are based on α-cuts of the fuzzy numbers used to represent the investigated values of the variables, namely it takes into account the amount of fulfillment of the outcome of the investigated rule. A more detailed discussion of the degrees of support/disproof is not necessary here, we therefore refer the interested readers to [5] and here we will simply calculate and discuss the values.

To make the description of the methods complete, we need to specify the measures applied to the assessment of the performance of the selected mutual equity growth ESG funds. The first measure applied in this paper is Jensen's alpha [13] which is calculated for a portfolio i using equation (11), where r_i is the return of the portfolio, β_i is the beta coefficient of the portfolio, r_m is the return of the market and r_f is the risk-free rate. From its construction, it is apparent that α_i is a risk-adjusted measure of portfolio performance that represents the excess returns of the portfolio above the expected level (derived through the capital asset pricing model (CAPM)). It is a benefit-type criterion of fund performance and positive values are interpreted as desirable as they represent situations when the portfolio under investigation outperforms the benchmark market portfolio.

$$\alpha_i = r_i - (\beta_i(r_m - r_f)) \tag{11}$$

Another performance measure applied in this paper is the Sharpe ratio [14]. This measure S_i reflects the returns of portfolio i per unit of risk and it is defined using (12), where r_i and r_f have the same interpretation as in Jensen's alpha and δ_i is the standard deviation of the i-th portfolio.

$$S_i = \frac{r_i - r_f}{\delta_i} \tag{12}$$

Unfortunately, Sharpe ratio's interpretability is limited when the information about the actual size of risk is not available or when a reference investment is not available. Higher values of this measure are preferred as they indicate better performance, however one can never be sure whether a high value of the ratio is obtained due to high excess returns, or due to low volatility of the portfolio. Sharpe ratio is therefore used as a secondary performance measure in this analysis.

III. Features of the Mutual Funds and Their Relationships with Fund Performance

In this part, we will briefly summarize the results of previous research on the possible links between the performance of mutual funds and their size, the length of their managers' tenure and their sustainability ratings. We do not claim the literature review in this aspect is complete, we mainly use the presented papers as a basis for the formulation of the assumed relationships to avoid data-mining bias.

A. *Relationship between mutual fund size and its performance*

Table I lists seven papers that focus on the relationship between the performance of the fund and its size. The analyzed periods do not cover the last 20 years, yet the most recent papers tend to agree on the existence of a negative relationship between the size of the fund and its performance. The only discovered relationships that can be considered positive are dealing with economies of scale and suggest that the larger the funds get, the lower the fees and thus the higher the potential returns for the investors (a simplified interpretation). Most of the research also relies on regression or other statistical methods. Based on the presented summary, we postulate the following potential relationship to be investigated: *If the fund size is large, then the risk-adjusted returns are low.* We will specify the meanings of "large" fund size and "low" risk-adjusted returns in the data section, where the meanings of these linguistic descriptions will be provided in terms of fuzzy numbers. In line with the recommendations by Stoklasa et al. [5], the opposite relationship *If the size of the fund is large, then its risk-adjusted returns are not low* will also be investigated to get a more complete picture.

$$\text{Consistency}_{F_2}(A \Rightarrow B) = \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_A(x_i), \mu_B(x_i), \mu_{B'}(x_i)))}{\sum_{i=1}^{n}\mu_A(x_i)} \tag{5}$$

$$\text{Coverage}_{F_2}(A \Rightarrow B) = \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_A(x_i), \mu_B(x_i), \mu_{A'}(x_i)))}{\sum_{i=1}^{n}\mu_B(x_i)} \tag{6}$$

$$\text{Consistency}_{F_3}(A \Rightarrow B) = \max\left\{0; \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_A(x_i), \mu_{B'}(x_i)))}{\sum_{i=1}^{n}\mu_A(x_i)}\right\} \tag{7}$$

$$\text{Coverage}_{F_3}(A \Rightarrow B) = \max\left\{0; \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_B(x_i), \mu_{A'}(x_i)))}{\sum_{i=1}^{n}\mu_B(x_i)}\right\} \tag{8}$$

$$\text{Consistency}_{F_4}(A \Rightarrow B) = \frac{1}{2}\left(1 + \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_A(x_i), \mu_{B'}(x_i)))}{\sum_{i=1}^{n}\mu_A(x_i)}\right) \tag{9}$$

$$\text{Coverage}_{F_4}(A \Rightarrow B) = \frac{1}{2}\left(1 + \frac{\sum_{i=1}^{n}(\min(\mu_A(x_i), \mu_B(x_i)) - \min(\mu_B(x_i), \mu_{A'}(x_i)))}{\sum_{i=1}^{n}\mu_B(x_i)}\right) \tag{10}$$

TABLE I

SUMMARY OF THE REVIEWED PAPERS DEALING WITH THE RELATIONSHIP OF THE SIZE OF THE FUND AND ITS PERFORMANCE.

Year	Paper, Authors	Declared objective(s)	Period	Data characteristics	Methodology	Results	Assumed effect[a]
2009	Chan, Faff, Gallagher and Looi [15]	To investigate if fund size affects performance. To identify the causes for the possible relation.	1998-2001 (40mths)	35 Australian equity funds	Regression analysis and simulation.	Fund size lowers performance, especially for funds with highly active trading approaches.	-
2008	Yan [16]	To examine the impact of liquidity and investment style on the relationship between fund size and fund performance.	1993-2002	1024 actively managed U.S. mutual funds.	Cross-sectional regression analysis and a portfolio approach. Performance measured with Alpha, CAPM, three- and four-factor models.	A negative relationship between fund size and fund performance. Liquidity is proposed as an important reason to cause this relation.	-
2004	Chen, Hong, Huang and Kubik [17]	To investigate if fund size affects fund performance.	1962-1999	3439 funds from the U.S.	Regression analysis. Performance measured with CAPM, three- and four-factor models.	A negative relationship between fund size and fund performance mainly caused by the lack of liquidity.	-
2001	Beckers and Vaughan [18]	To examine how fund size affects investment performance.	1996-1999	250 stocks from an Australian Index; Daily prices and trading volumes	Historical real-life simulation.	Bigger funds are less flexible in implementing their ideas and thus creating value-added is harder as the number of assets under management grow.	-
1997	Tufano and Sevick [19]	To research the relationship between fund board structure and fund fees. Also the relationship between fund size and fees is examined.	1991-1992 (12mths)	1587 U.S. open-end mutual funds.	Regression analysis.	Fund fees are inversely related to fund size, and thus larger funds have economies of scale.	+
1996	Golec [20]	To study if mutual fund manager's features affect fund fees, performance and risks. Also the effect of fund size is examined.	1988-1990	530 mutual funds; geographically not specified.	Regression analysis. Alpha and yield as performance measures.	Larger funds discover economies of scale. Large funds' fees are lower leading to larger yields.	+
1991	Perold and Salomon [21]	To detect the right amount of assets under management for financial maximization.	1982	Examples from [22] 1200 observations.	A mathematical analysis using a wealth-maximizing tradeoff. Alpha as performance measure.	The optimal fund size is when trading costs exceed the opportunity cost of not trading. A larger asset base than that leads to higher opportunity costs and lower returns.	-

[a] + indicates a positive relationship, - indicates a negative relationship, 0 indicates no relationship between the size of the fund and its performance; adapted from [23]

TABLE II
SUMMARY OF THE REVIEWED PAPERS DEALING WITH THE RELATIONSHIP OF THE LENGTH OF THE TENURE OF FUND MANAGER AND THE PERFORMANCE OF THE FUND.

Year	Paper, Authors	Declared objective(s)	Period	Data characteristics	Methodology	Results	Assumed effect[a]
2016	Kjetsaa and Kieff [24]	To explore the effect of manager tenure, expenses and turnover on blend fund performance.	2002-2012	559 blend funds; geographically not specified.	Regression analysis for three time horizons (3, 5 and 10 years). Returns as a performance measure.	There is a positive relation between manager tenure and mutual fund returns.	+
2006	Costa, Jakob and Porter [25]	To examine how market trends and fund managerial experience affect the ability to outperform the market.	1990-2001	1249 mutual equity funds from the U.S.	Regression analysis. Alpha from a four-factor model as a performance measure.	Longer-tenured managers do not outperform shorter tenured managers.	0
2004	Filbeck and Tompkins [26]	To investigate if there is a relation between manager tenure and risk-adjusted returns.	1990-2000	sample size or geographical area not specified.	Regression analysis. M-squared as a measure of risk-adjusted performance.	Longer-tenured managers outperformed the market more than shorter-tenured managers. Long-tenured managers were able to manage funds on lower expenses and thus more efficiently.	+
2002	Brooks and Tompkins [27]	To investigate the effect of mutual fund characteristics on mutual fund performance.	1989-1999	474 mutual funds; geographically not specified.	A two-tailed Z-test and regression analysis. M-squared as a measure of risk-adjusted performance.	A slight adverse relationship between manager tenure and risk-adjusted returns.	-
1999	Fortin, Michelson and Jordan-Wagner [28]	To research how manager tenure affects mutual fund performance across all investment classes.	1985-1995	800 bond and equity funds; geographically not specified.	Comparison of short-term and long-term fund managers' performance and regression analysis. Alpha as a performance measure.	Manager tenure does not affect mutual fund performance. There is an adverse relation between manager tenure and fund turnover.	0
1996	Golec [20]	To study if mutual fund manager's features affect fund fees, performance and risks. Also the effect of fund size is examined.	1988-1990	530 mutual funds; geographically not specified.	A three-stage least squares (3SLS) regression analysis. Yield and Jensen's Alpha as performance measures.	There is a positive connection between manager tenure and fund performance.	+
1996	Lemak and Satish [29]	To examine the differences in mutual fund performance and risk between longer-tenured mutual fund managers (>10 years) and shorter tenured managers (<10 years).	1984-1994	313 mutual funds; geographically not specified.	Comparison of short-term and long-term fund managers' performance. Regression analysis. Return as a performance measure.	Longer-tenured (10 years or more) fund managers performed better than shorter tenured managers.	+

[a] + indicates a positive relationship, - indicates a negative relationship, 0 indicates no relationship between the length of tenure and the performance of the fund; adapted from [23]

B. Relationship between the length of the tenure of mutual fund's manager and the performance of the fund

As can be seen in Table II, manager tenure and its effect on the performance of the mutual funds is a more actual topic with periods being analyzed stretching at least to 2012. Also in this context, the majority of the research is based on regression (statistical) models that in many cases involve the assumption of linearity of the relationship in one way or another. Also, the results of the research are a bit less consistent. We can find research that did not discover any sort of relationship between the length of manager tenure and fund performance, also some weak evidence of a negative-type of relationship can also be found; the prevailing result, however, seems to be one that confirm the existence of a positive relationship between the length of manager's tenure and fund performance. The positive relationship can be expressed by the manager's experience and ability to manage the fund more efficiently, while the negative relationship might be stemming from the inability of long-term managers to "think out of the box" and thus missing some opportunities.

Based on the presented summary of previous research, we consider the relationship *If fund manager's tenure is high, then the risk-adjusted returns of the fund are high* to be the one to validate on our data. Again, we will also investigate the validity of the opposite relationship *If fund manager's tenure is high, then the risk-adjusted returns of the fund are* <u>*not*</u> *high.* The definition of the fuzzy-number representation of high tenure will be provided further on in the data section.

C. The relationship between the sustainability rating of the fund and its performance

Out of all the variables, whose potential effect on fund performance is being studied in this paper, sustainability is definitely the one that has been receiving researchers' attention most recently (see Table III). From the conducted literature review it is obvious, that even though the topic is being currently researched, the findings are far from being unanimous. One main issue in the scientific investigation of the effect of sustainability on (or the relationship thereof with) other variables suffers from the multitude of possible approaches to sustainability and its definition. We can see the terms sustainable, responsible, green and many others being used interchangeably, and we can also frequently encounter the "environmental, social and governance" (ESG) label denoting those funds (companies) that either consider these factors in the composition of their investment portfolios or set explicit goals concerning these areas. In older literature mainly the predecessor of ESG - the corporate social responsibility (CRS) - can be found. Even though all these terms and concepts might share some goals or an ultimate vision, their definitions are not identical, the measures for the fulfillment of all the necessary criteria to use some of these labels are not widely available and there are also some potential methodological issues with the measurement of a "sustainability level" of a mutual fund or a company. Sustainability as a concept requires such behavior, goals and actions that allow for the continuous existence of all the elements of the system (all the stakeholders) or at least give a chance for "survival" to most. Even though this is a very simplified summary of the concept of sustainability, it helps us point out the key methodological issues connected with the concept: first of all sustainability is by definition a system issue - it is difficult to measure without the inputs concerning all the elements of the system, second it is a forward-looking concept meaning that its assessment needs to rely on predictions, and third there seem to be many ways to assess sustainability, most of which sooner or later degenerate to binary ones (sustainable/unsustainable, ESG/nonESG, etc.) or are at least interpreted as such.

There are, on the other hand, some indices for sustainability like the Morningstar Sustainability Ratings (MSR) [30] which allow for some graduality in the transition from non-sustainable to sustainable labels. It is also good to note that many ratings such as the one provided by Morningstar are intrinsically relative, i.e. they identify the "most sustainable" and the "least sustainable" units in the given set. Nothing guarantees that the most sustainable units are "sustainable enough" as well as nothing says that the least sustainable units are "not sustainable at all". It is also interesting to note that for a portfolio to obtain a Morningstar Sustainability Score, only 2/3 of its assets under management need to have the ESG risk rating. This means that the MSR might not reflect the full ESG risk and full information concerning the funds being assessed. It also considers the environmental, social and governance issues as proxies for sustainability, without an explicitly declared overall sustainability focus. Still, as evidenced also by the literature review conducted by us (see Table III), MSR is a frequently used proxy for fund sustainability.

Given the issues we have discussed above (which are just some of the issues connected with the measurement of something as complex and ill-defined as sustainability), it is not surprising that one can find research papers that do not find any relationship between funds' sustainability ratings and their performance, research that suggests the existence of a positive relationship between these two variables, but also research that points out the inability of sustainable (ESG) funds to outperform the market during non-crisis periods. Again, the prevalence of regression methods in the research is high, which only stresses the need for validation of these nonuniform findings by another approach. Given the results presented in Table III, we will further investigate the consistency of the following relationship with our data: *If the Morningstar Sustainability Rating of the fund is high, then the risk-adjusted returns are high*. Also, in this case, we will investigate the opposite relationship *If the Morningstar Sustainability Rating of the fund is high, then the risk-adjusted returns are not high.* Now that we know what relationships are expected based on the previous research, we can describe the dataset used in our analysis and also provide the fuzzy-number meanings of the linguistic terms used in the relationships to be investigated by the tools of fsQCA.

IV. DATA AND IMPLIED DEFINITIONS OF THE FUZZY-NUMBER MEANINGS OF HIGH/LOW VALUES OF THE FUND FEATURES

For our analysis, we have obtained a set of 429 mutual equity growth ESG funds from the European area from the Morningstar Mutual Fund Screener. Out of the over 31 000 mutual funds available in the database at the time of data retrieval (March 2021) we strived to get a compact sample by limiting our scope to

- "Europe Developed" or "Europe Developing" which limited the number of funds available for the analysis to 3583
- "Growth" funds ruling out funds that would be dividend-paying to simplify the performance assessment of the funds
- "Euro" as the currency to further facilitate the intercomparability of the funds and their performance
- at least three years old funds to ensure sufficient history of the analyzed funds; more specifically we required the funds to be in the database for the whole March 2018 - March 2021 period
- funds for which the MSR value is available
- equity funds; the reason for this is that other than equity funds were very infrequent in the resulting sample and their different characteristics might not be strong enough to have significant effect in the results, but might have biased the results for the equity funds.

TABLE III
SUMMARY OF THE REVIEWED PAPERS DEALING WITH THE RELATIONSHIP OF SUSTAINABILITY OF THE FUND (MEASURED IN VARIOUS WAYS) AND ITS PERFORMANCE.

Year	Paper, Authors	Declared objective(s)	Period	Data characteristics	Methodology	Results	Assumed effect[a]
2020	Steen, Moussawi and Gjolberg [31]	To analyze the relationship between the Morningstar Sustainability rating and fund performance	2014-2018	146 mutual funds domiciled in Norway.	Fama-French regression, geographical bias of the ratings considered. Sustainability measured with the MSR. Alpha as a performance measure.	Among categorized European funds (to avoid geographical bias) the performance improves in parallel with improving ESG ratings.	+
2019	Dolvin, Fulkerson and Krukover [32]	To investigate the effect of sustainable investing on investment performance.	2012-2016	1853 U.S. mutual funds.	Performance measured with Carhart alpha. Sustainability measured with the Morningstar Sustainability scores.	No difference in risk-adjusted returns between sustainable and conventional funds. However, sustainable funds limited to large-cap funds and thus can feature a higher risk and weaker diversification.	0
2016	Henke [33]	To examine the financial effect of screening ESG criteria on corporate bond fund portfolios.	2001-2014	103 socially responsible and 309 matched conventional bond mutual funds from the U.S. and Eurozone.	Regression analysis. Comparing socially responsible funds with their conventional pairs. Performance measured with risk-adjusted returns (a five-factor model). Sustainability is measured with ESG ratings based on information provided by the US Sustainable Investment Forum and the European Social Investment Forum.	Socially responsible bond mutual funds performed better than their conventional pairs annually.	+
2016	Nagy, Kassam and Lee [34]	To investigate if ESG factors of an investment affect investment performance.	2007-2015	global MSCI stock data.	Back-testing two global model portfolios that regard ESG criteria: "ESG tilt" and "ESG momentum." Alpha as a performance measure. MSCI ESG ratings as a sustainability measure.	Both tested portfolios that consider ESG criteria beat the global benchmark index MSCI World Index.	+
2014	Nofsinger and Varma [35]	To examine the performance of socially responsible funds during periods of market crisis and periods of non-crisis.	2000-2011	240 U.S. equity mutual funds and their 209 conventional pairs.	Regression analysis. CAPM, three-factor and four-factor models as performance measures.	Socially responsible mutual funds outperform their conventional pairs in periods of market crisis and underperform conventional funds during periods of non-crisis.	+/-
2005	Bello [36]	To examine the effects of socially responsible investing on portfolio diversification and fund performance.	1994-2001	42 socially responsible funds provided by Morningstar and 84 conventional funds from the U.S.	Regression analysis. Comparing socially responsible funds with their conventional pairs. Performance measured with Jensen's Alpha, Sharpe Ratio and excess standard deviation adjusted return.	There is no notable difference between the performance or diversification of socially responsible and conventional funds.	0
1993	Hamilton, Jo amd Statman [37]	To evaluate the financial effect of socially responsible investing in mutual fund performance.	1981-1990	32 socially responsible funds and 150 conventional funds.	Performance comparison between socially responsible and conventional funds. Jensen's Alpha as a performance measure. The selected funds were identified as socially responsible funds by their managers.	There is no practical difference between the performance of socially responsible and conventional funds.	0

[a] + indicates a positive relationship, - indicates a negative relationship, 0 indicates no relationship between the fund's sustainability rating and its performance; adapted from [23]

- funds having no missing values of the relevant variables (assets under management, manager tenure etc.) in the investigated period

After the selection of the dataset and ensuring that all the funds within do not have any missing values of the variables relevant for our research, the fuzzy numbers representing the meanings (denoted by the M operator) of "high", "middle" and "low" values of the variables were defined in the following way.

Fund performance measures values

For Jensen's alpha the prototype of the middle value representing "middle value" of alpha can be considered to be 0, which is the natural middle value of this variable. All values within the $(-3, 3)$ interval were considered at least partially fitting for the description "middle". These thresholds are set by the authors and can be modified if needed in future analyses. The idea of not setting the definition of "middle alpha" around the median of the data is that the alpha has a natural middle (neutral) point at zero. The minimum and maximum values of alpha were set relative to the available values of the funds. This resulted in:

- M("high alpha") $\sim (0, 3, 14.23, 14.23)$
- M("middle alpha") $\sim (-3, 0, 0, 3)$
- M("low alpha") $\sim (-12.44, -12.44, -3, 0)$

which implies

- M("not high alpha") $\sim (-12.44, -12.44, 0, 3)$
- M("not low alpha") $\sim (-3, 0, 14.23, 14.23)$.

For the Sharpe ratio, there is no natural minimum, middle or maximum value prototype. We have therefore identified the minimum, first, second and third quartile and the maximum value of the Sharpe ratios available in the given sample, which were -0.19, 0.25, 0.39, 0.63 and 1.58 respectively. We have used these values to define the meanings of "high", "middle" and "low" values of Sharpe ratio in the following way:

- M("high Sharpe ratio") $\sim (0.39, 0.63, 1.58, 1.58)$
- M("middle Sharpe ratio") $\sim (0.25, 0.39, 0.39, 0.63)$
- M("low Sharpe ratio") $\sim (-0.19, -0.19, 0.25, 0.39)$

which implies

- M("not high Sharpe ratio") $\sim (-0.19, -0.19, 0.39, 0.63)$
- M("not low Sharpe ratio") $\sim (0.25, 0.39, 1.58, 1.58)$.

It is clear that for variables without specific natural middle points, maxima or minima, the definitions of the meanings of the linguistic terms used in the investigated relationships need to be defined either relatively to the available values of the variables, or based on experience or expert knowledge.

Fund size values

Fund size was measured by assets under management (in millions of EUR). This variable has a natural minimum at 0, but no natural middle or maximum values. Therefore the first, second and third quartiles as well as the maximum value of this variable were determined: 78.39, 235.03, 664.91 and 7124.65 respectively. The meanings of "large", "middle" and "small" values of fund size were thus defined in the following way:

- M("large size") $\sim (235.03, 664.91, 7124.65, 7124.65)$
- M("middle size") $\sim (78.39, 235.03, 235.03, 664.91)$
- M("small size") $\sim (0, 0, 78.39, 235.03)$.

Manager tenure values

The length of manager tenure (measured in years) also has a natural minimum at 0, but no natural middle or maximum values. We have thus again decided to use the first, second and third quartiles as well as the maximum value of this variable, which were 3.58, 7.83, 12.08 and 23.58 respectively. The meanings of "long", "middle" and "short" values of manager tenure length were thus defined in the following way:

- M("long tenure") $\sim (7.83, 12.08, 23.58, 23.58)$
- M("middle tenure") $\sim (3.58, 7.83, 7.83, 12.08)$
- M("short tenure") $\sim (0, 0, 3.58, 7.83)$.

Sustainability rating values

The values of the MSR are always from the $\{1, 2, 3, 4, 5\}$ set, in other words, there are only five possible ratings to be assigned. As such the scale has a natural maximum, minimum and middle point which can be used for the definitions of the fuzzy-number meanings of the linguistic values used in the investigated relationships. Given the limited number of numerical values of this scale, we have decided to distinguish only between "low" and "high" sustainability defined in the following way:

- M("high sustainability") $\sim (2, 4, 5, 5)$
- M("low sustainability") $\sim (1, 1, 2, 4)$.

V. Results of the analysis using fsQCA methods

For the assumed relationships between each of the individual variables (fund size, manager tenure, fund sustainability rating) and fund performance (measured using Jensen's alpha and Sharpe ratio), we have calculated all four fuzzified consistency and coverage measures (3)-(10). To gain additional insights into the relationships between the variables, we have investigated not only the assumed relationships and their negations, but also relationships that lead to the outcome represented by the opposite linguistic term on the scale than the one that was postulated. In other words, we investigate (Long Tenure $\Rightarrow$ High risk-adjusted returns), (Long Tenure $\Rightarrow$ not High risk-adjusted returns), but also (Long Tenure $\Rightarrow$ Low risk-adjusted returns) and (Long Tenure $\Rightarrow$ not Low risk-

TABLE IV
SUMMARY OF THE RESULTS OBTAINED BY APPLYING THE F_1 - F_4 CONSISTENCY AND COVERAGE MEASURES ON THE INVESTIGATED RELATIONSHIPS.

A = Large Size, B = High Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.318	F1 consistency	0.709
F1 coverage	0.398	F1 coverage	0.355
F2 consistency	0.236	F2 consistency	0.626
F2 coverage	0.316	F2 coverage	0.290
F3 consistency	0.000	F3 consistency	0.390
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.305	F4 consistency	0.695
F4 coverage	0.381	F4 coverage	0.348
SUP1(A=>B) =	0.185	DISP1(A=>B) =	0.449
SUP0.9(A=>B) =	0.212	DISP0.9(A=>B) =	0.532
SUP0.8(A=>B) =	0.222	DISP0.8(A=>B) =	0.606
SUP0.7(A=>B) =	0.249	DISP0.7(A=>B) =	0.674
SUP0.6(A=>B) =	0.272	DISP0.6(A=>B) =	0.708
SUP0.5(A=>B) =	0.279	DISP0.5(A=>B) =	0.721
SUP0.4(A=>B) =	0.292	DISP0.4(A=>B) =	0.728
SUP0.3(A=>B) =	0.332	DISP0.3(A=>B) =	0.751
SUP0.2(A=>B) =	0.394	DISP0.2(A=>B) =	0.778
SUP0.1(A=>B) =	0.468	DISP0.1(A=>B) =	0.788
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Large Size, B = High Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.438	F1 consistency	0.590
F1 coverage	0.430	F1 coverage	0.332
F2 consistency	0.358	F2 consistency	0.510
F2 coverage	0.354	F2 coverage	0.266
F3 consistency	0.000	F3 consistency	0.152
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.424	F4 consistency	0.576
F4 coverage	0.417	F4 coverage	0.324
SUP1(A=>B) =	0.302	DISP1(A=>B) =	0.449
SUP0.9(A=>B) =	0.322	DISP0.9(A=>B) =	0.474
SUP0.8(A=>B) =	0.342	DISP0.8(A=>B) =	0.484
SUP0.7(A=>B) =	0.370	DISP0.7(A=>B) =	0.529
SUP0.6(A=>B) =	0.394	DISP0.6(A=>B) =	0.559
SUP0.5(A=>B) =	0.431	DISP0.5(A=>B) =	0.582
SUP0.4(A=>B) =	0.441	DISP0.4(A=>B) =	0.606
SUP0.3(A=>B) =	0.471	DISP0.3(A=>B) =	0.630
SUP0.2(A=>B) =	0.516	DISP0.2(A=>B) =	0.658
SUP0.1(A=>B) =	0.526	DISP0.1(A=>B) =	0.678
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Large Size, B = Low Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.293	F1 consistency	0.746
F1 coverage	0.307	F1 coverage	0.405
F2 consistency	0.196	F2 consistency	0.649
F2 coverage	0.229	F2 coverage	0.333
F3 consistency	0.000	F3 consistency	0.454
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.273	F4 consistency	0.727
F4 coverage	0.287	F4 coverage	0.394
SUP1(A=>B) =	0.132	DISP1(A=>B) =	0.551
SUP0.9(A=>B) =	0.149	DISP0.9(A=>B) =	0.591
SUP0.8(A=>B) =	0.170	DISP0.8(A=>B) =	0.638
SUP0.7(A=>B) =	0.207	DISP0.7(A=>B) =	0.679
SUP0.6(A=>B) =	0.244	DISP0.6(A=>B) =	0.703
SUP0.5(A=>B) =	0.281	DISP0.5(A=>B) =	0.719
SUP0.4(A=>B) =	0.297	DISP0.4(A=>B) =	0.762
SUP0.3(A=>B) =	0.321	DISP0.3(A=>B) =	0.793
SUP0.2(A=>B) =	0.362	DISP0.2(A=>B) =	0.830
SUP0.1(A=>B) =	0.409	DISP0.1(A=>B) =	0.851
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Large Size, B = Low Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.316	F1 consistency	0.717
F1 coverage	0.314	F1 coverage	0.401
F2 consistency	0.230	F2 consistency	0.631
F2 coverage	0.243	F2 coverage	0.329
F3 consistency	0.000	F3 consistency	0.401
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.300	F4 consistency	0.700
F4 coverage	0.297	F4 coverage	0.392
SUP1(A=>B) =	0.192	DISP1(A=>B) =	0.565
SUP0.9(A=>B) =	0.202	DISP0.9(A=>B) =	0.578
SUP0.8(A=>B) =	0.210	DISP0.8(A=>B) =	0.609
SUP0.7(A=>B) =	0.235	DISP0.7(A=>B) =	0.643
SUP0.6(A=>B) =	0.257	DISP0.6(A=>B) =	0.681
SUP0.5(A=>B) =	0.294	DISP0.5(A=>B) =	0.715
SUP0.4(A=>B) =	0.319	DISP0.4(A=>B) =	0.743
SUP0.3(A=>B) =	0.357	DISP0.3(A=>B) =	0.765
SUP0.2(A=>B) =	0.391	DISP0.2(A=>B) =	0.790
SUP0.1(A=>B) =	0.422	DISP0.1(A=>B) =	0.798
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Long Tenure, B = High Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.288	F1 consistency	0.746
F1 coverage	0.381	F1 coverage	0.395
F2 consistency	0.194	F2 consistency	0.652
F2 coverage	0.285	F2 coverage	0.331
F3 consistency	0.000	F3 consistency	0.458
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.271	F4 consistency	0.729
F4 coverage	0.359	F4 coverage	0.387
SUP1(A=>B) =	0.133	DISP1(A=>B) =	0.501
SUP0.9(A=>B) =	0.166	DISP0.9(A=>B) =	0.567
SUP0.8(A=>B) =	0.178	DISP0.8(A=>B) =	0.627
SUP0.7(A=>B) =	0.197	DISP0.7(A=>B) =	0.686
SUP0.6(A=>B) =	0.236	DISP0.6(A=>B) =	0.726
SUP0.5(A=>B) =	0.246	DISP0.5(A=>B) =	0.754
SUP0.4(A=>B) =	0.274	DISP0.4(A=>B) =	0.764
SUP0.3(A=>B) =	0.320	DISP0.3(A=>B) =	0.803
SUP0.2(A=>B) =	0.373	DISP0.2(A=>B) =	0.822
SUP0.1(A=>B) =	0.433	DISP0.1(A=>B) =	0.840
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Long Tenure, B = High Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.355	F1 consistency	0.670
F1 coverage	0.369	F1 coverage	0.399
F2 consistency	0.278	F2 consistency	0.593
F2 coverage	0.293	F2 coverage	0.332
F3 consistency	0.000	F3 consistency	0.315
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.343	F4 consistency	0.657
F4 coverage	0.356	F4 coverage	0.391
SUP1(A=>B) =	0.217	DISP1(A=>B) =	0.513
SUP0.9(A=>B) =	0.237	DISP0.9(A=>B) =	0.538
SUP0.8(A=>B) =	0.252	DISP0.8(A=>B) =	0.570
SUP0.7(A=>B) =	0.299	DISP0.7(A=>B) =	0.613
SUP0.6(A=>B) =	0.320	DISP0.6(A=>B) =	0.637
SUP0.5(A=>B) =	0.340	DISP0.5(A=>B) =	0.662
SUP0.4(A=>B) =	0.363	DISP0.4(A=>B) =	0.680
SUP0.3(A=>B) =	0.387	DISP0.3(A=>B) =	0.701
SUP0.2(A=>B) =	0.430	DISP0.2(A=>B) =	0.748
SUP0.1(A=>B) =	0.462	DISP0.1(A=>B) =	0.763
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Long Tenure, B = Low Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.346	F1 consistency	0.680
F1 coverage	0.385	F1 coverage	0.390
F2 consistency	0.252	F2 consistency	0.585
F2 coverage	0.311	F2 coverage	0.321
F3 consistency	0.000	F3 consistency	0.334
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.333	F4 consistency	0.667
F4 coverage	0.370	F4 coverage	0.383
SUP1(A=>B) =	0.189	DISP1(A=>B) =	0.499
SUP0.9(A=>B) =	0.207	DISP0.9(A=>B) =	0.534
SUP0.8(A=>B) =	0.233	DISP0.8(A=>B) =	0.555
SUP0.7(A=>B) =	0.265	DISP0.7(A=>B) =	0.609
SUP0.6(A=>B) =	0.297	DISP0.6(A=>B) =	0.658
SUP0.5(A=>B) =	0.321	DISP0.5(A=>B) =	0.679
SUP0.4(A=>B) =	0.342	DISP0.4(A=>B) =	0.710
SUP0.3(A=>B) =	0.398	DISP0.3(A=>B) =	0.735
SUP0.2(A=>B) =	0.445	DISP0.2(A=>B) =	0.767
SUP0.1(A=>B) =	0.466	DISP0.1(A=>B) =	0.793
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = Long Tenure, B = Low Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.391	F1 consistency	0.644
F1 coverage	0.410	F1 coverage	0.382
F2 consistency	0.304	F2 consistency	0.558
F2 coverage	0.337	F2 coverage	0.307
F3 consistency	0.000	F3 consistency	0.253
F3 coverage	0.000	F3 coverage	0.000
F4 consistency	0.373	F4 consistency	0.627
F4 coverage	0.391	F4 coverage	0.371
SUP1(A=>B) =	0.253	DISP1(A=>B) =	0.505
SUP0.9(A=>B) =	0.266	DISP0.9(A=>B) =	0.509
SUP0.8(A=>B) =	0.286	DISP0.8(A=>B) =	0.536
SUP0.7(A=>B) =	0.326	DISP0.7(A=>B) =	0.575
SUP0.6(A=>B) =	0.340	DISP0.6(A=>B) =	0.594
SUP0.5(A=>B) =	0.376	DISP0.5(A=>B) =	0.642
SUP0.4(A=>B) =	0.406	DISP0.4(A=>B) =	0.660
SUP0.3(A=>B) =	0.425	DISP0.3(A=>B) =	0.674
SUP0.2(A=>B) =	0.464	DISP0.2(A=>B) =	0.714
SUP0.1(A=>B) =	0.491	DISP0.1(A=>B) =	0.734
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = High Sustainability, B = High Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.357	F1 consistency	0.698
F1 coverage	0.690	F1 coverage	0.541
F2 consistency	0.263	F2 consistency	0.604
F2 coverage	0.464	F2 coverage	0.304
F3 consistency	0.000	F3 consistency	0.341
F3 coverage	0.272	F3 coverage	0.038
F4 consistency	0.329	F4 consistency	0.671
F4 coverage	0.636	F4 coverage	0.519
SUP1(A=>B) =	0.203	DISP1(A=>B) =	0.483
SUP0.9(A=>B) =	0.230	DISP0.9(A=>B) =	0.530
SUP0.8(A=>B) =	0.247	DISP0.8(A=>B) =	0.589
SUP0.7(A=>B) =	0.270	DISP0.7(A=>B) =	0.633
SUP0.6(A=>B) =	0.302	DISP0.6(A=>B) =	0.669
SUP0.5(A=>B) =	0.314	DISP0.5(A=>B) =	0.686
SUP0.4(A=>B) =	0.331	DISP0.4(A=>B) =	0.698
SUP0.3(A=>B) =	0.371	DISP0.3(A=>B) =	0.730
SUP0.2(A=>B) =	0.411	DISP0.2(A=>B) =	0.753
SUP0.1(A=>B) =	0.470	DISP0.1(A=>B) =	0.774
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = High Sustainability, B = High Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.473	F1 consistency	0.587
F1 coverage	0.718	F1 coverage	0.510
F2 consistency	0.378	F2 consistency	0.493
F2 coverage	0.498	F2 coverage	0.265
F3 consistency	0.000	F3 consistency	0.114
F3 coverage	0.344	F3 coverage	0.000
F4 consistency	0.443	F4 consistency	0.557
F4 coverage	0.672	F4 coverage	0.484
SUP1(A=>B) =	0.323	DISP1(A=>B) =	0.441
SUP0.9(A=>B) =	0.340	DISP0.9(A=>B) =	0.460
SUP0.8(A=>B) =	0.357	DISP0.8(A=>B) =	0.481
SUP0.7(A=>B) =	0.390	DISP0.7(A=>B) =	0.502
SUP0.6(A=>B) =	0.418	DISP0.6(A=>B) =	0.523
SUP0.5(A=>B) =	0.456	DISP0.5(A=>B) =	0.553
SUP0.4(A=>B) =	0.477	DISP0.4(A=>B) =	0.582
SUP0.3(A=>B) =	0.498	DISP0.3(A=>B) =	0.610
SUP0.2(A=>B) =	0.519	DISP0.2(A=>B) =	0.643
SUP0.1(A=>B) =	0.540	DISP0.1(A=>B) =	0.660
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = High Sustainability, B = Low Jensen's alpha

	A=>B		A=>notB
F1 consistency	0.335	F1 consistency	0.727
F1 coverage	0.543	F1 coverage	0.609
F2 consistency	0.225	F2 consistency	0.616
F2 coverage	0.289	F2 coverage	0.381
F3 consistency	0.000	F3 consistency	0.392
F3 coverage	0.000	F3 coverage	0.166
F4 consistency	0.304	F4 consistency	0.696
F4 coverage	0.493	F4 coverage	0.583
SUP1(A=>B) =	0.173	DISP1(A=>B) =	0.521
SUP0.9(A=>B) =	0.179	DISP0.9(A=>B) =	0.559
SUP0.8(A=>B) =	0.194	DISP0.8(A=>B) =	0.589
SUP0.7(A=>B) =	0.234	DISP0.7(A=>B) =	0.639
SUP0.6(A=>B) =	0.268	DISP0.6(A=>B) =	0.688
SUP0.5(A=>B) =	0.297	DISP0.5(A=>B) =	0.703
SUP0.4(A=>B) =	0.312	DISP0.4(A=>B) =	0.732
SUP0.3(A=>B) =	0.365	DISP0.3(A=>B) =	0.766
SUP0.2(A=>B) =	0.411	DISP0.2(A=>B) =	0.806
SUP0.1(A=>B) =	0.441	DISP0.1(A=>B) =	0.821
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

A = High Sustainability, B = Low Sharpe ratio

	A=>B		A=>notB
F1 consistency	0.331	F1 consistency	0.725
F1 coverage	0.507	F1 coverage	0.627
F2 consistency	0.240	F2 consistency	0.634
F2 coverage	0.283	F2 coverage	0.387
F3 consistency	0.000	F3 consistency	0.394
F3 coverage	0.000	F3 coverage	0.205
F4 consistency	0.303	F4 consistency	0.697
F4 coverage	0.464	F4 coverage	0.603
SUP1(A=>B) =	0.211	DISP1(A=>B) =	0.586
SUP0.9(A=>B) =	0.213	DISP0.9(A=>B) =	0.599
SUP0.8(A=>B) =	0.222	DISP0.8(A=>B) =	0.614
SUP0.7(A=>B) =	0.249	DISP0.7(A=>B) =	0.639
SUP0.6(A=>B) =	0.264	DISP0.6(A=>B) =	0.677
SUP0.5(A=>B) =	0.304	DISP0.5(A=>B) =	0.711
SUP0.4(A=>B) =	0.323	DISP0.4(A=>B) =	0.736
SUP0.3(A=>B) =	0.361	DISP0.3(A=>B) =	0.751
SUP0.2(A=>B) =	0.386	DISP0.2(A=>B) =	0.778
SUP0.1(A=>B) =	0.401	DISP0.1(A=>B) =	0.787
SUP0.0(A=>B) =	1.000	DISP0.0(A=>B) =	1.000
alpha-SUP =	1.000	alpha-DISP =	1.000

adjusted returns) and analogously for the other two features and the funds' performance.

Let us start with the relationship of the size of the fund and its performance. The original assumed relationship was that a large fund size should result in low risk-adjusted returns. The results for this relationship are available in the last two subtables in the top row of subtables in Table IV. We can clearly see that the consistency of the "large fund size resulting in low fund performance" (both measured by Jensen's alpha and Sharpe ratio) is much lower than the consistency of "large fund size resulting in not low performance". The values of F_1 consistencies of "not low performance" resulting from "large fund size" being higher than F_1 consistencies of "low performance" resulting from "large fund size", as well as the values of F_4 consistencies being higher than 0.5 (0.727 for alpha, 0.700 for Sharpe) suggest that there is much more evidence in favor of the *If the size of the fund is large, then its risk-adjusted returns are <u>not low</u>* relationship in the data than there is for the originally assumed one. The originally assumed relationship even does not have any pure non-ambivalent excess evidence in its favor meaning that there is no evidence in its favor as defined by the F_3 consistency measure (both values of F_3 consistency are zero for alpha and for Sharpe). By the same logic, looking at the first two subtables in the top row of Table IV, we can see that there is also no pure non-ambivalent excess evidence for high size resulting in high performance of the funds. Given the fact that all values of unconditional support and disproof are nonzero in the first row of subtables in Table IV and that the largest values of the unconditional support/disproof are for $\mathrm{DISP}_1(A \Rightarrow B)$ in the Large size $\rightarrow$ Low Sharpe ratio (0.565) and in the Large size $\rightarrow$ Low Jensen's alpha (0.551) with comparatively lower values of the unconditional support $\mathrm{SUP}_1(A \Rightarrow B)$, we can conclude that Large fund size being related to not low performance seems to be the most plausible of the investigated relationships. Note, that "not low performance" covers the "middle or high" performance in this case. The F_3 consistency of "Large fund size $\rightarrow$ not High Sharpe ratio" being rather low (0.152) prevents us from claiming that large fund size would be related with not high performance of the fund in general, though. We, however, do not see any clear support for the claim that the large funds are high-performing either. Still large fund size seems to be preventing low performance.

As far as the relationship between manager tenure and fund performance is concerned, we need to look at the middle row of subtables in Table IV. By the same logic applied here, we can see that the most viable relationship that can be found in the data is *If fund manager's tenure is high, then the risk-adjusted returns of the fund are <u>not high</u>*. Again, here "not high" covers "middle or low". It is, however, true to say that the relationship *If fund manager's tenure is high, then the risk-adjusted returns of the fund are <u>not low</u>* has a similar support by the data. Overall, we can see that high manager tenure does not seem to guarantee high performance and it also does not guarantee low performance of the fund.

Now we can focus on the bottom row of subtables in Table IV that investigates the relationships between the sustainability rating of the funds and their performance. Applying the same logic in this case, we can clearly see that the most supported relationship in the data is the one of "High MSR $\rightarrow$ not Low performance of the fund" both measured by Sharpe ratio and by Jensen's alpha. If we have a close look at the values, the F_1 consistencies and coverages are rather high for them, the F_3 consistencies are nonzero and reasonably high implying that there is pure non-ambivalent excess evidence for these relationships and there is also nonzero F_3 coverage for these relationships. These relationships are the only ones (except for "High MSR $\rightarrow$ not High Jensen's alpha") with nonzero F_3 consistency and coverage, but for the "not Low performance of the fund" the values of F_3 consistencies and coverages are such that one can see clear evidence in favor of the given relationship in the data. Given all this, we can conclude that the data supports the relationship *If the Morningstar sustainability rating of the fund is high, then the risk-adjusted returns of the fund are <u>not low</u>.* There is not enough evidence to conclusively prove the validity of the claims that high sustainability ratings would be related with high fund performance. The evidence in favor of claiming that high sustainability is related with not high fund performance is inconclusive.

VI. Conclusions

In this paper, we have analyzed the relationship between three selected characteristics of mutual funds, namely their size, length of their manager's tenure and their sustainability ratings, and the performance of these funds. The analysis was carried out on a sample of European growth mutual funds using the fsQCA tools, mainly the recently proposed fuzzified versions of consistencies and coverages.

Overall the strongest relationship found in the data can be expressed in general terms by the statement *If the sustainability rating of the fund is high, then its performance is not low.* This is well in line with the previous research that suggests that sustainable/responsible funds might overperform the non-sustainable ones in crises periods, but at the same time there seems to be evidence that they might underperform during calmer times (see Table III and its discussion). Our findings suggest, on the given sample and under the given definitions of the variables, that although the high sustainability rating does not guarantee the high performance of the fund, it seems to indicate that low performance of the fund is not to be expected.

There are several ways in which to continue this research. First of all this paper focused on the drivers of high performance of the European growth mutual funds and thus the potential reasons for low performance etc. were not analyzed. An analogous analysis can be performed with the intention of identifying potential sources of low performance for these funds even using the same dataset. We have also analyzed only isolated effects of single features on the performance. The fsQCA methodology allows for the investigation of combined effects (for example of the type "IF sustainability rating is High and manager tenure is Not low, THEN the performance of the fund is High"). These combined effects were left out of the scope of this paper and can constitute a research direction that sheds more light on the drivers of the performance of mutual funds.

VII. Acknowledgement

This paper is an extended summary of the analysis and results obtained in the thesis by Fanni Welling [23].

References

[1] J. Tomar, S. Agarwal, and K. R. Chaturvedi, "Relationship between Employer Branding and Corporate Social Responsibility," in S. Agarwal, D. N. Burrell, and V. K. Solanki (eds.) *Proceedings of the 2020 International Conference on Research in Management & Technovation*, *ACSIS*, vol. 24, pp. 123-127, 2020, doi: http://dx.doi.org/10.15439/2020KM19

[2] H. D. Hai, and C. N. Huu, "The Effect of Corporate Social Responsibility Factor on the Sustainable Development of Industrial SMEs: A Case Study in Hanoi-Vietnam," in S. Agarwal, D. N. Burrell, and V. K. Solanki (eds.) *Proceedings of the 2020 International Conference on Research in Management & Technovation*, *ACSIS*, vol. 24, pp. 57-62, 2020, doi: http://dx.doi.org/10.15439/2020KM23

[3] P. C. Fiss, "Building better causal theories: A fuzzy set approach to typologies in organization research," *Acad. Manag. J.*, vol. 54, no. 2, pp. 393-420, 2011.

[4] P. C. Fiss, "A Set-Theoretical Approach to Organizational Configurations," *Acad. Manag. Rev.*, vol. 32, no. 4, pp. 1180-1198, 2007.

[5] J. Stoklasa, P. Luukka, and T. Talášek, "Set-theoretic methodology using fuzzy sets in rule extraction and validation - consistency and coverage revisited," *Inf. Sci.*, vol. 412-413, pp. 154-173, 2017, doi: 10.1016/j.ins.2017.05.042

[6] J. Stoklasa, T. Talášek, and P. Luukka, "On consistency and coverage measures in the fuzzified set-theoretic approach for social sciences: dealing with ambivalent evidence in the data," in *Proceedings of the 36th International Conference on Mathematical Methods in Economics*, 2018, pp. 521-526.

[7] M. M. Kumbure, A. Tarkiainen, P. Luukka, J. Stoklasa, and A. Jantunen, "Relation between managerial cognition and industrial performance : An assessment with strategic cognitive maps using fuzzy-set qualitative comparative analysis," *J. Bus. Res.*, vol. 114, no. June 2020, pp. 160-172, 2020, doi: 10.1016/j.jbusres.2020.04.001

[8] M. M. Kumbure, P. Luukka, A. Tarkiainen, J. Stoklasa and A. Jantunen, "An investigation of hidden shared linkages among perceived causal relationships in cognitive maps," in P. Luukka and J. Stoklasa (eds.), *Intelligent Systems and Applications in Business and Finance,* Springer, in press.

[9] S. Tomer, and G. Rana, "Green Human Resource Management: A Conceptual Study," in S. Agarwal, D. N. Burrell, and V. K. Solanki (eds.) *Proceedings of the 2020 International Conference on Research in Management & Technovation*, *ACSIS*, vol. 24, pp. 147-150, 2020, doi: http://dx.doi.org/10.15439/2020KM10

[10] L. A. Zadeh, "Fuzzy sets," *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.

[11] G. J. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic: Theory and Aplications*. New Jersey: Prentice Hall, 1995.

[12] C. C. Ragin, *Redesigning Social Inquiry: Fuzzy sets and Beyond.* Chicago: University of Chicago Press, 2008.

[13] M. C. Jensen, "The Performance of Mutual Funds in the Period 1945-1964", *J. Financ.*, vol. 23, no. 2, pp. 389-416, 1968.

[14] W. F. Sharpe, "Mutual Fund Performance", *J. Bus. (Chicago, Ill.)*, vol. 39, no. 1, pp. 119-138, 1966.

[15] H. W. H. Chan, R. W. Faff, D. R. Gallagher and A. Looi, "Fund Size, Transaction Costs and Performance: Size Matters", *Aust. J. Manage.*, vol. 34, no. 1, pp. 73-96, 2009.

[16] X. Yan, "Liquidity, Investment Style, and the Relation between Fund Size and Fund Performance", *J. Financ. Quant. Anal.*, vol. 43, no. 3, pp. 741-767, 2008.

[17] J. Chen, H. Hong, M. Huang and J. D. Kubik, "Does Fund Size Erode Mutual Fund Performance? The Role of Liquidity and Organization", *Am. Econ. Rev.*, vol. 94, no. 5, pp. 1276-1302, 2004.

[18] S. E. Beckers and G. Vaughan, "Small Is Beautiful", *J. Portfolio Manage.*, vol. 27, no. 4, pp. 9-17, 2001.

[19] P. Tufano and M. Sevick, "Board structure and fee-setting in the U.S. mutual fund industry", *J. Financ. Econ.*, vol. 46, no. 3, pp. 321-355, 1997.

[20] J. Golec, "The effects of mutual fund managers' characteristics on their portfolio performance, risk and fees", *Financ. Serv. Rev. (Greenwich, Conn.)*, vol. 5, no. 2, pp. 133-147, 1996.

[21] A. F. Perold and R. S. Salomon, "The Right Amount of Assets under Management", *Financ. Analysts J.*, vol. 47, no. 3, pp. 31-39, 1991.

[22] T. F. Loeb. "Trading Cost: The Critical Link between Investment Information and Results." *Financ. Analysts J.*, vol. 39, no. 3, pp. 39-44, 1983.

[23] F. Welling, "The role of sustainability, manager tenure, and fund size in European mutual growth fund performance," BSc thesis, LUT University, 2021.

[24] R. Kjetsaa and M. Kieff, "Impact of Expenses, Turnover and Manager Tenure on Blend Fund Performance", *J. Bus. Account.*, vol. 9, no. 1, pp. 99-115, 2016.

[25] B. A. Costa, K. Jakob and G. E. Porter, "Mutual Fund Performance and Changing Market Trends 1990-2001: Does Manager Experience Matter?", *J. Investing*, vol. 15, no. 2, pp. 79-86, 2006.

[26] G. Filbeck and D. L. Tompkins, "Management Tenure and Risk-Adjusted Performance of Mutual Funds", *J. Investing*, vol. 13, no. 2, pp. 72-80, 2004.

[27] M. Brooks and D. L. Tompkins, "Mutual funds' risk adjusted performance", *J. Commercial Bank. Financ.*, Vol. 1, pp. 111-121, 2002.

[28] R. Fortin, S. Michelson and J. Jordan-Wagner, "Does mutual fund manager tenure matter?", *J. Financ. Plan.*, vol. 12, no. 7, pp. 72-79, 1999.

[29] D. Lemak and P. Satish, "Mutual Fund Performance and Managers' Terms of Service: Are There Performance Differences?" *J. Investing*, Winter 1996, pp. 59-63, 1996.

[30] Morningstar Research, "Morningstar Sustainability Rating Methodology," *Morningstar*, 2019. [Web document] [Accessed 20.3.2021]. Available: https://www.morningstar.com/content/dam/marketing/shared/research/methodology/744156_Morningstar_Sustainability_Rating_for_Funds_Methodology.pdf

[31] M. Steen, J. T. Moussawi and O. Gjolberg, "Is there a relationship between Morningstar's ESG ratings and mutual fund performance?", *J. Sustain. Financ. Invest.*, vol. 10, no. 4, pp. 349-370, 2020.

[32] S. Dolvin, J. Fulkerson and A. Krukover, "Do "Good Guys" Finish Last? The Relationship between Morningstar Sustainability Ratings and Mutual Fund Performance", *J. Investing*, vol. 28, no. 2, pp. 77-91, 2019.

[33] H. Henke, "The effect of social screening on bond mutual fund performance", *J. Bank. Financ.*, vol. 67, pp. 69-84, 2016.

[34] Z. Nagy, A. Kassam and L. Lee, "Can ESG Add Alpha? An Analysis of ESG Tilt and Momentum Strategies", *J. Investing*, vol. 25, no. 2, pp. 113-124, 2016.

[35] J. Nofsinger and A. Varma, "Socially responsible funds and market crises", *J. Bank. Financ.*, vol. 48, pp. 180-193, 2014.

[36] Z. Y. Bello, "Socially Responsible Investing and Portfolio Diversification", *J. Financ. Res.*, vol. 28, no. 1, pp. 41-57, 2005.

[37] S. Hamilton, H. Jo and M. Statman, "Doing Well While Doing Good? The Investment Performance of Socially Responsible Mutual Funds", *Financ. Analysts J.*, vol. 49, no. 6, pp. 62-66, 1993.