

**CHARLES UNIVERSITY**  
**FACULTY OF SOCIAL SCIENCES**

Institute of Economic Studies

**Do mutual funds offered in Czech  
Republic add value to investors?**

Master's thesis

Author: Bc. Jiří Nosek

Study program: Economics and Finance

Supervisor: Mgr. Martin Hronec

Year of defense: 2022

## **Declaration of Authorship**

The author hereby declares that he or she compiled this thesis independently, using only the listed resources and literature, and the thesis has not been used to obtain any other academic title.

The author grants to Charles University permission to reproduce and to distribute copies of this thesis in whole or in part and agrees with the thesis being used for study and scientific purposes.

Prague, December 29, 2021

Jiri Nosek

## Abstract

We estimate the proportions of skilled, unskilled, and zero-alpha funds prevalent in the mutual Funds population easily accessible by Czech Investors. We estimate alphas from a regression against a concise set of Exchange Traded Funds and control for luck using False Discovery rate. We design a straightforward ETF selection algorithm and find that if investors adhere to simple diversification rules, they can outperform a large proportion of mutual funds. We further document a negative relationship between the performance of mutual funds and its Total Expense ratio, suggesting that portfolio managers are on average unable to compensate their costs with better performance.

**JEL Classification** C12, C20, G12, G23

**Keywords** Mutual Funds, Exchange Traded Funds, Performance evaluation

**Title** Do mutual funds offered in Czech Republic add value to investors?

## Abstrakt

Naším cílem je zjistit, jak velká část fondů je schopná generovat nadprůměrné výnosy a jak velká část nadprůměrné zhodnocení schopná přinést není. Výkonnost podílových fondů reprezentujeme pomocí *alphy* odhadnuté z regrese vůči několika vybraným ETF. Výsledky dále kontrolujeme o efekt štěstí pomocí pokročilé statistické metodologie zvané Míra falešných objevů (False Discovery Rate). Navrhli jsem přímočarý algoritmus vybírající ETF a zjistili, že pokud se investoři budou držet jednoduchých pravidel diverzifikace, tak mohou dosáhnout vyššího výnosu než velká část podílových fondů. Dále jsme našli negativní vztah mezi výkonností fondu a celkového poměru nákladů který naznačuje že manažeři portfolií nejsou schopni vykompenzovat jejich náklady vyšším výnosem jejich portfolia.

**Klasifikace JEL** C12, C20, G12, G23

**Klíčová slova** Podílové fondy, ETFs, Hodnocení výkonnosti

**Název práce** Přinášejí podílové fondy nabízené v České republice hodnotu svým investorům?

## Acknowledgments

I would like to express my sincere gratitude to my supervisor Mgr. Martin Hronec for his invaluable advice, useful comments, and willingness to help during the whole process of thesis writing.

Typeset in L<sup>A</sup>T<sub>E</sub>X using the IES Thesis Template.

## Bibliographic Record

Nosek, Jiří: *Do mutual funds offered in Czech Republic add value to investors?*. Master's thesis. Charles University, Faculty of Social Sciences, Institute of Economic Studies, Prague. 2022, pages 70. Advisor: Mgr. Martin Hronec

# Contents

List of Tables	vii
List of Figures	ix
Acronyms	x
Thesis Proposal	xi
<b>1 Introduction</b>	<b>1</b>
<b>2 Literature review</b>	<b>4</b>
<b>3 Theory Recap</b>	<b>8</b>
3.1 Investment Vehicles . . . . .	8
3.1.1 Summary of Mutual Fund's history . . . . .	9
3.1.2 Exchange Traded Funds (& UCITS ETFs) . . . . .	10
3.2 Standard performance evaluation metrics . . . . .	11
3.2.1 Sharpe Ratio . . . . .	11
3.2.2 Treynor Ratio . . . . .	13
3.2.3 Jensen's Alpha . . . . .	14
3.3 Asset pricing based performance evaluation metrics . . . . .	15
3.3.1 Arbitrage pricing Theory . . . . .	16
3.3.2 Fama-French factor models . . . . .	16
3.3.3 Multi-index models . . . . .	18
<b>4 Methodology</b>	<b>19</b>
4.1 Mutual funds to ETFs regression . . . . .	19
4.1.1 ETF selection algorithms . . . . .	20
4.1.2 Alpha adjustment procedure . . . . .	23
4.2 Correction for lucky and unlucky funds . . . . .	24

---

4.2.1	Problem of Multiple comparisons . . . . .	25
4.2.2	False Discovery Rate (FDR) . . . . .	25
4.2.3	FDR Simulation Example . . . . .	29
<b>5</b>	<b>Data</b>	<b>32</b>
5.1	Characteristics of Czech Mutual Funds Universe . . . . .	35
<b>6</b>	<b>Results</b>	<b>39</b>
6.1	Analysis of Skilled funds proportions . . . . .	39
6.2	Relationship between performance and cost . . . . .	45
<b>7</b>	<b>Conclusion</b>	<b>52</b>
	<b>Bibliography</b>	<b>56</b>

# List of Tables

4.1	Simulation results - This table summarises the proportions of zero-alpha, skilled and unskilled funds ( $\pi_-$ , $\pi_0$ , $\pi_+$ ). It shows the true proportions which were used to generate data. Then it displays the estimated proportions using simple counts and FDR methodology. . . . .	31
5.1	The table shows descriptive statistics of the final dataset. Panel A corresponds to our sample of mutual funds, Panel B to a sample of ETFs. Statistics for average monthly return, ex-post Sharpe Ratio and the length of our time series are shown in both panels. For each variable, we present Mean, Standard Deviation, Minimum, Maximum, Median and 5th (p05) and 95th (p95) percentiles. . . . .	34
6.1	The table shows the proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds estimated on alphas from a regression where Factor models or individual factors were used as regressors. The first part of the table displays results of using three and four-factor models, the second part shows results with individual factors where Mkt-RF is the Market Factor, HML (High-Minus-Low) refers to the Value factor, SMB (Small-Minus-Big) corresponds to the Size factor, and MOM is a Momentum Factor. Data for these factors were obtained from Kenneth French website. . . . .	40

6.2	The table summarizes the results of the Random Selection algorithm, which randomly selects a given number of ETFs (from 1 to 10) and uses them to estimate the overall skill proportions. The procedure is repeated 500 times for each number of ETFs, and the average proportions calculated over all runs are stored. These average proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds are shown in this table. . . . .	41
6.3	This table displays proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds estimated on alphas from a regression where well-known and diversified ETFs were used. These ETFs approximate indices from various regions all around the world and can be thought of as a representation of average returns in these regions. MSCI World approximates the Market index, S&P 500 is the proxy for the United States market, FTSE 100 reflects the European market, and MSCI EM Asia proxies financial markets of Asia. . . . .	43
6.4	Results from regression of fund's alpha against fund's Total Expense Ratio (Equation 6.3). Note that in this case alphas were estimated in a regression against MSCI World. . . . .	46
6.5	Results from regression of fund's alpha against fund's Total Expense Ratio. Note that alphas were estimated in a regression against the 3-Factor model in this case. . . . .	49
6.6	Results from regression of fund's alpha against fund's Total Expense Ratio (TER), Assets Under Management (AUM) and its Standard Deviation (STD). Note that in this case alphas were estimated in a regression against MSCI World. . . . .	50

# List of Figures

5.1	Number of Mutual Funds (grey line) and Exchange Traded Funds (blue line) from January 1987 to December 2020. . . . .	33
5.2	The figure shows histograms of annualized Sharpe ratio and annualized Treynor ratio calculated from monthly data. . . . .	36
5.3	Densities of Total Expense Ratios (TER) of mutual funds investing in Stocks (blue), bonds (orange), and mixed assets (green). TER of ETFs is shown in red colour. . . . .	37
6.1	Figure depicts the residuals from the regression of fund's alphas against Total Expense Ratio (TER). The upper sub-figure is from the regression on Stock mutual funds exclusively, while the bottom sub-figure is from regression on Bond funds. The original value of TER is depicted on the x-axis, and the y-axis shows the value of a respective error term. The blacked dotted line shows the value of 0 across the whole plot and the blue curve is a non-parametric lowess model. In an ideal case the blue line should overlay the dotted line. . . . .	48

# Acronyms

**MF** Mutual Fund

**ETF** Exchange Traded Fund

**UCITS** Undertakings for the Collective Investment of Transferable Securities

**PRIIPs** Packaged Retail Investment and Insurance Products

**KIID** Key Investors Information Document

**TER** Total Expense Ratio

**NAV** Net Asset Value

**AUM** Assets Under Management

**CAPM** Capital Asset Pricing Model

**APT** Arbitrage Pricing Theory

**HML** High-Minus-Low Factor

**SMB** Small-Minus-Big Factor

**MOM** Momentum Factor

**DJIA** Dow Jones Industrial Average

**US** United States

**EU** European Union

# Master's Thesis Proposal

---

<b>Author</b>	Bc. Jiří Nosek
<b>Supervisor</b>	Mgr. Martin Hronec
<b>Proposed topic</b>	Do mutual funds offered in Czech Republic add value to investors?

---

**Motivation** Historically, open-end mutual funds represented a fair value for money proposition. For a reasonable fee, they enabled average investors to hold a large and diversified portfolio that would otherwise be too costly to be constructed by themselves individually. Even then, many academics raised the question that for most investors holding the market is likely the best investment strategy, see Cowles (1933). But at that time holding the market was not an option.

However, since then the situation has changed dramatically. The significant improvements in computer technology led to the introduction of digital trading systems that together with increased market liquidity resulted in a significant decrease in trading costs. And with the introduction of low-cost index funds and ETFs now anybody has the affordable opportunity to invest in the broad market of his/her choice. Hence suddenly, the original reason for providing access to a diversified portfolio that justified the presence of actively managed mutual funds disappeared. The existence of these funds, which are now considered expensive, makes sense only if they provide market-beating returns that are large enough to compensate investors for the extra cost in the form of the high fees. There are several studies Sharpe (1966), Fama French (2010) indicating that in the US mutual funds are unable to deliver on their promise to reliably outperform their benchmarks. But even though there is a large body of literature written on the performance of US mutual funds, similar studies for the European region are much scarcer and Otten Bams (2002) even reported that mutual funds on aggregate outperformed in 4 out of 5 studied European countries. However, according to our knowledge, there is a gap for comprehensive research of mutual fund performance in the Czech Republic that we attempt to fulfill with this study.

Further, as documented by Cavalcante Filho *et al.* (2021), the performance of

mutual funds might be replicable to a sufficient degree with a portfolio of low-cost ETFs which would provide similar exposures and similar performance, but importantly, for a much lower price. This approach may possibly provide Czech investors with an affordable alternative to widely used mutual funds.

## Hypotheses

Hypothesis #1: Do mutual funds in Czech Republic add value to the investors?

Hypothesis #2: : Is there a relationship between the size of the fee and fund's performance?

Hypothesis #3: Can the performance of these funds be replicated using low-cost ETFs?

**Methodology** We plan to construct a dataset that would approximately reflect the opportunity set that faces investors in the Czech Republic when considering investment into mutual funds. To achieve this goal, we attempt to collect ISINs of mutual funds offered by major providers. For these ISINs, we obtain relevant information from the Refinitive Lipper Fund Research Database which is available to the IES students. If the respective fund is not present in the Lipper database we collect the information manually directly from the provider.

Using this dataset, we evaluate the performance of Czech mutual funds with the aspiration to determine whether the fund managers do have an investment skill that materializes as abnormal returns which are high enough to compensate for the fees paid by fund investors. We are particularly interested in the fund's alphas, as described in Jensen (1968) which should represent the added value of the active management. Particularly in the case of replicating the fund's performance with a portfolio of ETFs the ability of active management to outperform the passive alternative could be analyzed with the following regression.

$$r_{i,t} = \alpha_i + \beta_i ETF_t + \epsilon_{i,t} \quad (1)$$

Where  $t \in \{1, \dots, T\}$ ,  $ETF_t$  is a vector of  $P$  ETFs and  $\beta_i$  is the vector of weights to respective ETFs. The statistical significance of  $\alpha$  would be then the indication of skilled management in case of  $\alpha > 0$  and of unskilled management when  $\alpha < 0$ .

Finally, we evaluate alpha in a more theoretical setting when controlling for several risk factors as in Fama French (2010). Also, we want to rate the performance of mutual funds by the ranking system described in Treynor (2007) which is based on the fund's characteristic line and portfolio choice theory. Ultimately, the relationship between the management fee of mutual funds and their performance would be analyzed by standard econometric methods.

**Expected Contribution** Mutual funds are still one of the most frequently used investment instruments by Czech investors despite having cheaper alternatives in the form of low-cost index funds and ETFs readily available. The results of this study may then have normative implications on the selection process of investment instruments of Czech investors.

## Outline

1. Introduction: The goal of this section will be to position the thesis into a broader context.
2. Literature review: I will briefly describe the most important literature which addresses the analysis of mutual fund performance.
3. Theory: In this section, I will introduce the main theoretical concepts which are needed to understand the methodology part.
4. Methodology: The detailed description of the methods used for evaluating mutual fund performance.
5. Data: I will provide a description of the data collection process and its main characteristics.
6. Empirical results: In this part, I will describe the results that I obtained for the ability of mutual fund managers to outperform the market and the degree to which these returns can be approximated by the portfolios of ETFs.
7. Conclusion: A summary of the results and the practical implication for the Czech investors will be provided in this section.

## Core bibliography

Cowles 3rd, A. (1933). Can stock market forecasters forecast?. *Econometrica: Journal of the Econometric Society*, 309-324.

Sharpe, W. F. (1966). Mutual fund performance. *The Journal of business*, 39(1), 119-138.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of finance*, 23(2), 389-416.

Otten, Rogér Bams, Dennis. (2002). European Mutual Fund Performance. *European Financial Management*. 8. 75 - 101. 10.1111/1468-036X.00177.

Treynor, J. L. (2007). Treynor On Institutional Investing. In *Treynor on Institutional Investing* (1st ed., pp. 65–87). Wiley.

FAMA, E.F. and FRENCH, K.R. (2010), Luck versus Skill in the Cross-Section of Mutual Fund Returns. *The Journal of Finance*, 65: 1915-1947

Junior, E. C., Fernando Moraes, R. D. L. (2020). Unskilled Fund Managers: Replicating Active Fund Performance With Few ETFs (No. 2020\_14). University of São Paulo (FEA-USP).

# Chapter 1

## Introduction

Despite the relative consensus of academic literature that mutual funds fail to beat the market, they are still among the most popular forms of investments in the Czech Republic. This sparks an important question - are Czech mutual funds (implicitly their managers) exceptionally good in capturing excess returns? Or rather, are Czech investors just unaware of the suboptimality of their investments?

In this thesis, we conduct the performance evaluation of mutual funds available to Czech investors. We use a regression setting to examine their performance side-by-side with Exchange Traded Funds. Our primary contributions are at least twofold. Firstly, we find that mutual funds offered in the Czech Republic are no different from their international counterparts and struggle to beat the cost-efficiency of passive investing. We conclude that their popularity is not justified from the investment perspective and has a different cause.<sup>1</sup> Secondly, as the viewpoint of retail investors is of our prime interest, we investigate whether these investors can leverage Exchange Traded Funds to outperform mutual funds with only limited knowledge. For that purpose, we devised a Naive ETF selection algorithm which demonstrates that if investors adhere to simple principles of diversification, they can perform significantly better than the large proportion of mutual funds. Lastly, we investigate the relationship between the fund's performance and the Total Expense Ratio. Economic theory would suggest a positive relationship while the arithmetic of active investing described in Sharpe (1991) predicts a negative link. We document a negative

---

<sup>1</sup>We speculate that the source of such popularity may be a legacy of the historical development of the Czech personal finance market driven by the many financial advisory companies.

relationship, thus providing further evidence for Sharpe's theory.

Throughout this thesis, we assume that a population of all mutual funds can be divided into three distinct groups based on their performance and that the performance can be captured by a fund's *alpha* estimated from the following regression model against a given set of ETFs.

$$r_{i,t} = \alpha_i + \beta_i \mathbf{ETF}_t + \epsilon_{i,t}$$

Where  $t = 1, \dots, T$ ,  $\mathbf{ETF}_t$  is a vector of  $P$  ETFs and  $\beta_i$  is the vector of weights to respective ETFs. The sign of alpha is then what separates the three groups. The negative alpha ( $\alpha < 0$ ) determines the "Unskilled funds", positive alpha ( $\alpha > 0$ ) is an characteristic of "Skilled funds" and  $\alpha = 0$  determines "Zero-alpha funds". Further, to give the mutual funds the benefit of the doubt, we consider all funds to be zero-alpha until the alpha is statistically different from zero. However, as the hypothesis about fund's alphas is evaluated for a large number of mutual funds simultaneously, it leads to multiple hypotheses test setting, and some approach to properly account for false-positive observations is necessary. Several statistical approaches were developed to address the problem of multiple comparisons. Classical methods dealing with Family-Wise-Error-Rate were devised by Bonferroni (1936) or Šidák (1967). However, these approaches are overly strict and might not be accurate enough. We hence employ the False Discovery Rate methodology devised in Barras *et al.* (2010) which is designed to control for luck when evaluating mutual fund performance. Our main goal is to estimate true proportions of zero, positive and negative alpha funds in the population. This result is important because it reveals the average odds investors face when choosing a mutual fund "at random"; hence, it is of their prime interest.

If we think about the evaluation of mutual fund performance as a horse race between the active (mutual funds) and passive (Exchange Traded Funds) investment approaches. We could pit the realized performance of a mutual fund against a sparse set of ETFs to which investors could have placed money instead. However, the question is, which ETFs to choose? It turns out that the answer to this question matters a lot because the results depend significantly on this choice, as shown in Cavalcante Filho *et al.* (2021). In this thesis, we analyze three main approaches. Firstly, We implement a Random

Selection algorithm that randomly chooses a set of ETFs and estimates the skilled proportions. With a bootstrap procedure on top of that, we can examine the average outperforming capacity of ETFs. Secondly, we propose a straightforward adjustment to this Random selection procedure that enforces simple diversification rules, which helps us answer a question to what extent retail investors are capable of outperforming Mutual Funds. Thirdly, we check what a sophisticated investor can achieve with a guided selection algorithm. In addition, we examine the skill proportions against ETFs that track well-known stock indices, and we also apply the standard factor regression.

The literature on return predictability is quite extensive, Fama (1970) proposed the Efficient Market Hypothesis, which implies that beating the market should be impossible. Subsequently, other academics found some predictability in the stock returns can be exploited in order to generate market-beating returns, see Campbell & Shiller (1988) or Jegadeesh & Titman (1993). Still, this predictability seems to be unstable and works poorly out-of-sample as shown in Welch & Goyal (2008). McLean & Pontiff (2016) points out that the discovered anomalies tend to be traded away following their publication because sophisticated funds attempt to exploit them. So the current state of the literature on return predictability is that it could be theoretically possible to beat the market, but it is very difficult. It then leads to a question of what percentage of mutual funds is capable of fulfilling their promise of outperforming market averages, which we analyze in this thesis.

The thesis is structured as follows. In the following chapter, we present a summary of academic literature on the topic of Mutual funds performance. Chapter 3 introduces the main theoretical concepts that this thesis is built upon. Chapter 4 provides a comprehensive description of the methodology which we apply to our dataset. This dataset is detailed in Chapter 5. Then the summary of the results is provided in Chapter 6 and the conclusion is given in the final chapter.

# Chapter 2

## Literature review

The beginnings of the literature on performance evaluation of investment vehicles date back to the 1930s. The first comprehensive analysis of investment professionals was performed in Cowles (1933). Proper risk-adjusted metrics were not yet discovered at that time, so the analysis revolves around the comparisons of average return with a Dow Jones Industrial Average index serving as a benchmark. The results suggest that about 2/3 of financial specialists performed worse than the benchmark, while only 1/3 outperformed. However, while comparing the average return of a mutual fund to a broad benchmark provides some perspective, one crucial element is still ignored – the risk. Several decades later, Markowitz (1952), gave birth to modern portfolio theory when presenting his solution to the mean-variance optimization problem of portfolio choice. He pointed out a strong relationship between risk and return and that the performance of investments should be evaluated side by side with risk undertaken. The introduction of risk opened a new chapter of performance evaluation and led to the development of performance metrics that are still frequently used. Sharpe (1966) presented the measure linking reward to variability (nowadays called Sharpe ratio). This ratio, together with the Treynor ratio (described in Treynor (1965)), was used to evaluate the performance of 34 mutual funds over a period from 1954 to 1963. The results suggested that more than two-thirds of funds underperformed the Dow Jones Industrial Average during that period. Additionally, the results showed that the average Sharpe ratio of the funds was smaller than the Sharpe ratio of DJIA by a considerable margin hence indicating that mutual funds offer less return per unit of risk. Further, Jensen (1968) introduced the innovative concept of alpha. It is a regression-based approach that enables the evaluation of the forecasting

ability of funds management by controlling for other risk factors. Its original time-series form is the following:

$$R_{i,t} - R_{f,t} = \alpha_i + \beta_i [R_{M,t} - R_{f,t}] + e_{i,t}$$

where  $E[R_j]$  is the expected return of asset  $j$ ,  $R_F$  is the risk-free rate of return for matching holding period and  $E[R_M]$  is the return of the market, and  $\alpha$  is a variable of interest that represents the predictive ability of the fund's manager. More details can be found in Section 3.2.3. He tested it on 115 surviving mutual funds from 1945 to 1964 and found that 76 funds were unable to beat the market and that only 39 funds outperformed. This again provides about a 1-to-2 win-loss ratio. He finally concluded that, in general, the average fund managers failed to generate alpha for their investors.

In Jensen (1968) only the market factor is used to control for the risk when calculating alpha. However, during the 1980s and 1990s, further advances in performance analysis were made, and prominently more factors explaining asset returns were discovered. These findings lead to the introduction of the 3-factor model in Fama & French (1993) as

$$R_{i,t} - R_{f,t} = \alpha_i + b_{1,i}MKT + b_{2,i}SMB_t + b_{3,i}HML_t + e_{i,t}$$

where  $R_{i,t}$  are returns of asset  $i$ ,  $R_{f,t}$  are returns of risk-free asset,  $\alpha_i$  is an indication of fund's management skill,  $MKT$  is a standard market factor and  $SMB$  (small minus big) and  $HML$  (high minus low) are the new proposed risk factors. Lastly,  $b_{k,i}$  for  $k = 1, 2, 3$  are respective factor loadings. For detailed description refer to the Section 3.3.2. This model later became the industry standard for performance evaluation. Fama & French (2010) utilize this model to analyze the set of more than 3000 mutual funds over 23 years. They found that about one-third of the funds outperformed (with the average excess return being 0.6%), and two-thirds underperformed the three-factor model (with average excess return being -1.2%). These results provide us with the odds of picking a mutual fund at random which leaves us with the expectation to underperform the three-factor model on average by 0.6%<sup>1</sup>. However, theoretical factors have the drawback of not including the real cost of constructing these portfolios; hence the comparison with mutual funds, which are net of fees, might not be completely fair. This issue was addressed in Cavalcante Filho

---

<sup>1</sup> $Expectation = \frac{2}{3} * -1.2 + \frac{1}{3} * 0.6 = -0.6$

*et al.* (2021) by replacing theoretical factors with ETFs, which have transaction costs included in the price. They conduct the following regression analysis

$$r_{i,t} = \alpha_i + \beta_i \mathbf{ETF}_t + \epsilon_{i,t} \quad (2.1)$$

Where  $t = 1, \dots, T$ ,  $\mathbf{ETF}_t$  is a vector of  $P$  ETFs and  $\beta_i$  is the vector of weights to respective ETFs. Conducting the hypotheses test  $H_0 : \alpha = 0$  against  $H_A : \alpha \neq 0$ , the statistical significance of  $\alpha$  would then indicate skilled management in the case of  $\alpha > 0$  and unskilled management when  $\alpha < 0$ . Insignificant alpha would mean that the null hypothesis is true and that the fund is considered to have a zero alpha. This means that the fund's management is just skilful enough to cover their fees. The author's results suggest that simple random selection of ETFs is not well suited to outperform mutual funds. However, with sophisticated selection methods, they estimate that 95% of fund managers are unable to generate alpha for their investors. Additionally, many academics realized that when evaluating the performance of mutual funds, there is a role of luck inherently involved. The alpha then provides only an indication of skilled management, not its direct evidence. It is because funds can be both lucky and unlucky, and the alpha is not designed to capture this. This has serious consequences, especially in a Multiple hypothesis test setting. There are several statistical methodologies designed to address this issue. Bonferroni (1936) proposed a stringent correction - dividing critical value  $\alpha$  by a number of tests conducted which ensures that Family-Wise-Error-Rate will be below  $\alpha$ . Benjamini & Hochberg (1995) devised another and less strict approach to control for false positives, which is nowadays called False Discovery Rate. This methodology is based on the knowledge of how p-values behave when the null hypothesis is False (in contrast to when it is true). An extension oriented to finance literature was described in Barras *et al.* (2010) where False Discovery Rate (FDR) methodology is utilized to control for luck in estimating mutual fund's alphas. Andrikogiannopoulou & Papakonstantinou (2019) points out a few limitations of the FDR procedure. Specifically, they show that the estimator of zero-alpha funds can be biased when the number of observations of individual funds is small. Barras *et al.* (2019) responded by acknowledging the bias in the aforementioned condition but showed that in larger datasets, the bias is of minor importance.

Despite the exhaustive literature written on the performance of US mu-

tual funds, comparative studies for the European region are much scarcer and sometimes of conflicting results. Otten & Bams (2002) reported that the European mutual funds, and primarily funds investing in small caps, are capable of adding value to their investors, as indicated by their positive after cost alphas. Authors theorize a causal connection with the relatively small size of the European mutual fund industry compared to the stock market (around 13% in 1998). They revisited this analysis almost a decade later and in Otten & Thevissen (2011) and reported that with the size of the mutual fund's market almost doubling, the ability of mutual fund managers to outperform has significantly diminished. Moreover, according to our knowledge, there is a gap for comprehensive research of mutual fund performance in the Czech Republic. This brings us to where our thesis lines up with the current literature. We base our analysis on the most recent research papers – we benchmark the performance of mutual funds to ETFs similarly to Cavalcante Filho *et al.* (2021), and we control for lucky and unlucky funds as Barras *et al.* (2010). However, we identify two main distinctive features that separate this work from what has already been done. Firstly, we evaluate the outperforming potential of *naive selection* strategies conducted by retail investors. Hence we try to find the answer if investors are better off with this approach or the mutual funds on aggregate provide enough value for investors that they should rather entrust their money to them. Secondly, we analyze the mutual funds provided in the Czech Republic, which offers a unique perspective because European markets were overlooked in performance analysis studies.

In the following chapter, we describe the theoretical concepts that are useful for understanding the primary substance of this thesis. After that, we move to the core chapters covering the Methodology, Data description, and Results.

# Chapter 3

## Theory Recap

### 3.1 Investment Vehicles

In this thesis, we mainly deal with two types of Investment vehicles – Mutual Funds and Exchange Traded Funds. These instruments are often mixed by unacquainted investors. While having many similarities, they also differ in a few critical aspects. Mutual Funds are the typical example of an actively managed investment vehicle, while Exchange Traded Funds are nowadays synonym with passive investing. Hopefully, we shed more light and distinguish these instruments in this section.

#### **Mutual Funds**

A mutual fund is an investment vehicle that pools money from many investors, which is then further invested by a professional fund manager. The fund manager then uses this pooled capital to purchase securities like stocks, bonds, commodities, or any combination of these. Every mutual fund has its own investment objectives stated in its prospectus and strives to achieve them. One of the indisputable benefits of mutual funds is convenience. Investors can gain exposure to a diversified portfolio by buying shares of a mutual fund. The minimal necessary investments are very small, and practically anybody can invest in them starting at about 500 CZK per month. Another advantage that is often overlooked is liquidity. Mutual funds stand ready every day to redeem investors shares for the current Net Asset Value (NAV); hence investors can convert their assets to cash with relative ease. On the other hand, a remarkable disadvantage of mutual funds is the cost associated with them. Since they are led by professional management, mutual funds need to charge for operating

costs, fund managers salaries, distribution costs, etc. Depending on the fund, these charges can be significant. This makes them usually several times more expensive than Exchange Traded Funds (ETFs).

Mutual Funds play an important role in the financial market of the Czech Republic. They are among the most popular investments, and substantial capital flows into them every year. There are several reasons for this. Firstly, investments in mutual funds are heavily pushed forward by financial advisors motivated by generous brokerage fees. Secondly, there was a relative lack of sound investing options promoted to retail investors for a long time. There is a system of pension funds heavily subsidised by the government. However, due to a strict limitations<sup>1</sup>, these funds are overly conservative and deliver dismal returns<sup>2</sup>. Nevertheless, the pool of possibilities for retail investors has grown considerably in recent years. Among other things, Robo-advisory companies offer seamless investments into portfolios that reflect the client's risk profile. But importantly, Exchange Traded Funds (ETFs) are now easily accessible on most brokerage platforms. Some even have their web pages or investment platforms in the Czech language. We firstly summarise the history, motivations and purpose of Mutual Funds, and then we take a closer look at ETFs in the next section.

### 3.1.1 Summary of Mutual Fund's history

Modern portfolio theory teaches us that diversification can reduce the expected risk without altering the expected return. When holding a diversified portfolio, specific risks of individual constituents get vastly diversified away, and mainly the aggregate risk of the whole market remains. Initially, providing access to these widely diversified portfolios was a primary added value of mutual funds. Investors paid fees to these funds, and in exchange for that, they received exposure to a portfolio with less overall risk than what they would otherwise be able to achieve by themselves. It was a fair value-for-money service, and both parties were satisfied. Nevertheless, some academics started noticing that apart from offering exposure to a diversified portfolio, managers of mutual funds might not have much else to offer and suggested that holding the market could be the most appropriate investment strategy for most investors, see Cowles (1933).

---

<sup>1</sup>For example so-called transformed funds must guarantee non-negative return every year.

<sup>2</sup>As reported in OECD (2021), the 15-year average annual real investment rate of return of retirement savings plans was (-0.4%) in the Czech Republic.

However, holding the market was not possible at that time. It changed during the 1970s when the first index funds tracking the broad market emerged. Their introduction represented an inflexion point in mutual fund history and forced classical mutual funds to redefine their original purpose. Providing access to a diversified portfolio was suddenly not enough because more cost-efficient alternatives in the form of index funds now existed, so the rhetoric had to change. It shifted from offering well-diversified portfolios to beating the market by generating above-average returns. Mutual Funds pushed forward the paradigm that the average is just not enough. They asked - "Who wants to be operated on by an average surgeon or be advised by an average lawyer?". Not surprisingly, this question was extensively analyzed, and as detailed in the literature review section, the evidence suggests that consistently delivering above-average returns is extremely hard.

### 3.1.2 Exchange Traded Funds (& UCITS ETFs)

Exchange Traded Funds, or shortly, ETFs are baskets of securities that could be bought or sold directly on a stock exchange via a brokerage firm. Similar to mutual funds, they provide investors with diversification – investors buy a single instrument to gain exposure to dozens of various securities. However, unlike mutual funds, they are not managed actively. Entirely on the contrary, they are passively managed. They follow a given protocol what to buy and what to sell. A prime example is index tracking, where the ETF's strategy is to replicate a given index like the S&P 500 (or any other). Hence they have much lower fees because they do not need to provide for their managers and analysts. Furthermore, these reduced costs are one of the main reasons why ETFs are considered to provide excellent value-for-money propositions on financial markets and why they are becoming so popular among retail investors.

The cradle of ETFs is the United States. It is the largest ETF market in the world, and it is also the most dynamic and competitive. Hence it is no surprise that modern and progressive trends like the cryptocurrency ETFs are firstly introduced there. The European Union takes a more protective view than the United States on safeguarding the retail investors. So based on PRIIPs<sup>3</sup> regulation, it requires all investment vehicles to produce a Key Investors Information

---

<sup>3</sup>PRIIPs is short for Packaged Retail Investment and Insurance Products.

Document (KIID) that provides investors with relevant information to evaluate costs, risks, and rewards of various investment products. Many US ETFs do not comply with this regulation, so they cannot be offered in EU markets to retail investors<sup>4</sup>. So for an average European investor, it is safe to assume that the set of ETFs from which he/she could choose is limited to a UCITS (Undertakings for the Collective Investment of Transferable Securities) ETFs. These ETFs are domiciled in Europe and fully conform to a regulation of the European Union. The UCITS regulation represents a set of safety standards for protecting investors from unsuitable investment instruments. It levies rules to ensure diversification – there can be no security in a fund’s portfolio in which market value exceeds 20% of the fund’s Net Asset Value. Also, UCITS ETF assets must be detached from the ETF provider and overlooked by an independent custodian. These measures protect investor’s assets in case of ETF issuer runs into financial difficulties.

With this, we end the section describing Investment vehicles, and we follow with the description of various metrics that are used for performance evaluation.

## 3.2 Standard performance evaluation metrics

In the previous section, we described the main characteristics of Mutual Funds and Exchange Traded Funds. Analysis of mutual fund performance was a leading research topic in finance literature in the 1960s. Academics searched intensively for a concise measure that could determine whether a fund’s management is skillful and able to generate superior returns. At that time, several breakthroughs in investigating risk-adjusted fund returns were made. In this section, we provide a summary of the most important ones, which are in their original form (or in the form of improved extensions) used by both academics and practitioners until today. Sharpe ratio, Treynor ratio and Jensen’s alpha are described in Sections 3.2.1, 3.2.2 and 3.2.3.

### 3.2.1 Sharpe Ratio

The Sharpe ratio, originally called reward-to-variability ratio in Sharpe (1966) is a risk-adjusted performance measure that expresses how much excess return

---

<sup>4</sup>Some brokers enable to trade these US ETFs to so-called qualified investors. These are typically investors with investment account larger than 500 000 EUR.

the portfolio is expected to generate per unit of total risk represented by a standard deviation. Mathematically,

$$S_p = \frac{E[R_p] - R_f}{\sigma_p}$$

where  $S_p$  is a Sharpe ratio of a portfolio  $p$ ,  $E[R_p]$  is its expected return,  $\sigma_p$  is standard deviation of its returns and  $R_f$  is a risk-free rate. Note that the formula is derived from the slope of a Capital Market Line.

Since the Sharpe ratio takes total risk into account, it is suitable to evaluate the performance of portfolios which are not entirely diversified. Also, when considering investment into a single asset without forming a larger portfolio, this property makes it an appropriate metric. Since it was published, it has become an industry standard to evaluate the performance of investment vehicles. The ratio is based solely on a portfolio theory and not the CAPM as Treynor and Jensen metrics. Hence it is not relying on the relationship with unobservable and hard to replicate market portfolio and is exempt from the criticism in Roll (1977).

During the following years after publishing, several variations of the Sharpe ratio were developed. They all are well summarised by their original author in Sharpe (1994). This paper takes the calculation methodology for the ex-post Sharpe ratio, which we utilize in this thesis. The exact procedure is described below.

For a mutual fund  $i$  consider a sequence of monthly returns  $R_{i,1}, R_{i,2}, \dots, R_{i,T}$  and a sequence of monthly risk-free rates for similar periods  $R_{f,1}, R_{f,2}, \dots, R_{f,T}$  then the Sharpe Ratio is calculated in the following way. Firstly, we calculate the fund's excess return over risk-free rate as

$$r_t = R_{i,t} - R_{f,t}$$

Then we calculate the average monthly excess return of a mutual fund  $i$  as

$$\bar{X} = \sum_{t=1}^T r_t$$

Next we calculate the standard deviation of monthly excess returns as

$$s = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \bar{X})^2}$$

Finally, the annualized Sharpe ratio of a mutual fund  $m$  equals

$$S_m = \frac{\bar{X}}{s} \sqrt{12}$$

Where the multiplication with  $\sqrt{12}$  handles the annualization under the assumption that the returns are independent.

### 3.2.2 Treynor Ratio

The Treynor ratio is a metric that captures the relationship between the excess return of a portfolio and its systematic risk. It tells us how much excess return the portfolio is expected to generate per unit of systematic risk, and it was defined in Treynor (1965) as

$$T_p = \frac{E[R_p] - R_f}{\beta_p}$$

where  $T_p$  is a Treynor ratio of a portfolio  $p$ ,  $E[R_p]$  is its expected return,  $\beta_p$  its market beta and  $R_f$  is a risk-free rate. This ratio arises directly from CAPM by dividing both sides with  $\beta_p$ .

Unlike the Sharpe Ratio, which considers the overall risk of a portfolio, the Treynor ratio focuses only on a systematic part of a risk expressed by market beta. The systematic risk cannot be diversified away, and it is the only risk that matters in a well-diversified portfolio. Hence particular usefulness of this ratio is highlighted when evaluating the performance of a portfolio (e.g. mutual fund), which is only a constituent of a larger portfolio of particular investors investments. However, the main drawback of this ratio is that the ratio is highly dependent on the choice of a suitable market proxy. Furthermore, as pointed by Roll (1977) the results can vary widely.

In this thesis we use annualized Treynor ratio which we calculate in the following way. Denote  $\mathbf{r}_i = [r_{i,1}, \dots, r_{i,T}]$  monthly excess returns of mutual fund  $i$ . Then denote  $\mathbf{r}_M = [r_{M,1}, \dots, r_{M,T}]$  to be monthly excess returns of a market proxy. In our case we use Market returns are taken from Kenneth French website. We then calculate beta of a mutual fund  $m$  as

$$\beta_m = \frac{\text{cov}(\mathbf{r}_i, \mathbf{r}_M)}{\text{var}(\mathbf{r}_M)}$$

Now the annualized Treynor ratio is calculated as

$$T_p = \frac{\bar{X} * 12}{\beta_m}$$

where  $\bar{X}$  is the average of monthly excess returns,  $\beta$  is market beta and multiplication by 12 ensures annualization.

### 3.2.3 Jensen's Alpha

The theoretical backbone of Jensen's alpha starts with Capital Asset Pricing Model (CAPM) of Sharpe (1964), Treynor (1965), Lintner (1965) and Mossin (1966). Under standard CAPM assumptions the model for one-period expected return on asset  $j$  is

$$E[R_j] - R_f = \beta_j [E[R_M] - R_f]$$

where  $E[R_j]$  is the expected return of asset  $j$ ,  $R_f$  is the risk-free rate of return for matching holding period, and  $E[R_M]$  is the return of the market. However, for empirical studies, this model has one major drawback. It is described in expectations that are inherently unobservable. But derived in Jensen (1968) it can be rewritten to an empirically valid time-series model

$$R_{j,t} - R_{f,t} = \beta_j [R_{M,t} - R_{f,t}] + e_{j,t}$$

Now, if the manager has a superior forecasting skills he/she will systematically select securities that ultimately realize  $e_{j,t} > 0$ . Such a portfolio will generate higher returns than expected by its risk premium. To account for the effect of forecasting ability constant can be added to a previous model, and we arrive to

$$R_{j,t} - R_{f,t} = \alpha_j + \beta_j [R_{M,t} - R_{f,t}] + e_{j,t}$$

where  $\alpha$  is a variable of interest that represents the predictive ability of the fund's manager. A positive alpha indicates that the manager is making unusually profitable investments within his portfolio and hence earning higher returns than we could expect given the level of riskiness of his portfolio. On

the other end, if the alpha is negative, it means that the manager's portfolio is earning less than it is supposed to, given its level of risk. Since the returns of mutual funds are reported net of fees, negative alpha can be easily obtained due to high fees. If the manager is earning average market returns, the alpha will be negative and about the size of the fund fees. Zero alpha means that the manager's portfolio selection skills are good enough to pay for fund fees. Importantly, this model can be easily estimated by a standard Least Squares regression model.

The true alpha of a mutual fund is unobservable and has to be estimated as is described above. However, when we test hypotheses simultaneously for a large group of funds, we conduct a multiple hypotheses test. Proper statistical methodology should be implemented to account for alphas that end up significant by chance alone (meaning their true alpha is zero). We adjust for this bias with the False Discovery rate approach described in Barras *et al.* (2010). Details of this correction are described in section 4.2.2.

In the next section, we look at more complex approaches that build on the idea of Jensen's alpha. Then, we describe the Arbitrage Pricing Theory, go through the prominent Fama-French Factor models, and end the theoretical part with a concept of Multi-index models.

### 3.3 Asset pricing based performance evaluation metrics

In Section 3.2 we described some performance evaluation metrics which are based on CAPM. It is essential to realize that the model was built on a set of very strong assumptions which might be too far from reality. Some of the assumptions were criticized by other academics. Most importantly, Roll (1977) points out that the CAPM model is not empirically testable since the market index is infeasible to construct. Another particularly restrictive assumption comes from the notion that the market portfolio is efficient in the mean-variance framework. Such efficiency can be obtained only in the case of normally distributed returns, which might not be the case in reality. Naturally, the violations of model assumptions have an effect on the performance measures, whose accuracy might then deteriorate. Search for a less restrictive

theory resulted in Arbitrage Pricing theory which we describe in the upcoming section 3.3.1. Fama-French then followed with the introduction of several empirical factor models that we describe in Section 3.3.2. Finally, we end this chapter with a short description of a multi-index model in Section 3.3.3.

### 3.3.1 Arbitrage pricing Theory

The search for a more general model with less restrictive assumptions resulted in a stream of literature focusing on what are collectively called factor models. The seminal model was presented in Ross (1976) as Arbitrage Pricing Theory (APT). The APT model attempts to explain the returns of a particular asset with several common factors. However, instead of a properly specified market factor, it uses  $K$  undefined factors. The APT states that there is a linear relationship between realized returns of the asset and the  $K$  common factors in a form

$$R_{i,t} = \alpha_i + \sum_{k=1}^K \beta_{i,k} F_{k,t} + \epsilon_{i,t}$$

where,  $\alpha_i$  is an intercept of asset  $i$  in time  $t$ ,  $F_{k,t}$  is a factor  $k$  in time  $t$ ,  $\beta_{i,k}$  is a loading of factor  $k$  with respect to asset  $i$  and  $\epsilon_{i,t}$  represents an idiosyncratic risk. The assumption that the investors are risk averse stays but the assumptions about Normally distributed returns and quadratic utility are levied.

APT is not specifying which factors are important, but academics have identified several factors which seem empirically promising. The most well-regarded and the most often used are 3-Factor and 4-Factor models, which we define in the following section. These models can serve as a foundation for a comprehensive performance evaluation.

### 3.3.2 Fama-French factor models

Fama and French have carried out several empirical studies to identify the fundamental factors that explain cross-section of asset returns and that can complement market factor. Initially, they highlighted two important factors that characterize a company's risk: the book-to-market ratio and its size measured by its market capitalization.

**Three-factor model** Eugene Fama conducted extensive research on the behavior of asset prices, and together with Kenneth French, they examined various factors that could potentially explain asset returns. In Fama & French (1992) they distilled their research and proposed three-factor model

$$E(R_i) - R_f = b_{i,1}[E(R_M) - R_f] + b_{i,2}E(SMB) + b_{i,3}E(HML)$$

where  $E(R_i)$  is the expected return of asset  $i$ ,  $E(R_m)$  is the expected return of a market portfolio,  $R_f$  is a riskfree rate, and SMB (small minus big) and HML (high minus low) are the newly proposed risk factors. SMB represents a difference of returns of a small-cap and large-cap portfolio, and HML denotes the difference between returns of high book-to-market ratio and low book-to-market ratio. Lastly,  $b_{i,k}$  for  $k = 1, 2, 3$  are respective factor loading which are estimated from the following equation

$$R_{i,t} - R_{f,t} = b_{i,1}[R_{M,t} - R_{f,t}] + b_{i,2}SMB_t + b_{i,3}HML_t$$

The view of Fama and French is that financial markets are indeed quite efficient, but that the market factor alone is not enough to explain all the risk on its own. They consider the three-factor model to be better suited to explain stock returns. However, they acknowledge that the list of factors might not be complete and other risk factors might also matter.

**Four-factor model** This extension of Fama and French's three-factor model was proposed in Carhart (1997) by appending momentum factor, which is based on the momentum anomaly found in Jegadeesh & Titman (1993). The model in a form that could be estimated from data has a form

$$E(R_i) - R_f = b_{i,1}[E(R_M) - R_f] + b_{i,2}E(SMB) + b_{i,3}E(HML) + b_{i,4}E(MOM)$$

The model uses a similar notation as the three-factor model only  $MOM$  is added, which denotes the difference between returns on best performing and worst performing portfolios from the previous year. By including the Momentum factor, the model becomes even better suited to explain asset returns and serves as a benchmark in modern performance evaluation studies.

### 3.3.3 Multi-index models

In the Arbitrage Pricing Theory models, returns of an asset are explained with theoretical factors. For multi-index modeling, factors are replaced with indexes. Build on the assumption that a set of  $K$  indexes can explain the variation of asset returns. The general model can be represented as

$$R_{i,t} = a_i + \sum_{k=1}^K b_{i,k} I_{k,t} + \epsilon_{i,t}$$

where  $I_{k,t}$  is a return of index  $k$  during period  $t$ . In this thesis, we use ETFs available in the Czech Republic as indexes by which we try to explain the performance of mutual funds.

Having established the theoretical background of this thesis, we can move to a Methodology section where we describe the ETF selection algorithms and how we control for luck using the False Discovery Rate approach.

# Chapter 4

## Methodology

From the computational perspective, the methodology of this thesis can be split into two parts. The first part deals with the sole estimation process of the fund's alphas. The second part is concerned with the correction for the luck of obtained alphas. The former, described in the following section, is conducted through the linear regression of the fund's returns on a certain number of ETFs selected based on various procedures. Random Selection, Naive Selection, and Proposed Selection algorithms are introduced for the ETF selection procedure. The latter adjusts for luck through False Discovery rate adjustment and is detailed in Section 4.2.

### 4.1 Mutual funds to ETFs regression

We estimate the mutual fund's alpha by fitting the following regression model

$$r_{i,t} = \alpha_i + \beta_i \mathbf{ETF}_t + \epsilon_{i,t} \quad (4.1)$$

Where  $t = 1, \dots, T$ ,  $\mathbf{ETF}_t$  is a vector of  $P$  ETFs and  $\beta_i$  is the vector of weights to respective ETFs. Conducting the hypotheses test  $H_0 : \alpha = 0$  against  $H_A : \alpha \neq 0$ , the statistical significance of  $\alpha$  would then be the indication of skilled management in case of  $\alpha > 0$  and of unskilled management when  $\alpha < 0$ . Insignificant alpha would mean that the null hypothesis is true and that the fund is considered to have a zero alpha. This means that the fund's management is just skillful enough to recover their fees.

We approach the evaluation of the mutual fund performance as a horse race between the active management of a mutual fund and a passive alternative of

ETFs. We analyze the realized performance of a mutual fund side-by-side with a sparse set of ETFs to which investors could invest otherwise. However, as shown in Cavalcante Filho *et al.* (2021), it matters how you choose the set of ETFs. In this section, we describe several algorithms designed for this task. First, we introduce the Random selection algorithm, which selects ETFs at random. It represents a proxy for the average outperforming capacity of ETFs. Next, we present our extension of Cavalcante's Random Selection Algorithm that adheres to simple diversification principles. We call this procedure the Naive Selection algorithm. Finally, we outline the Proposed ETF selection procedure, which chooses the ETF set based on their ability to outperform mutual funds.

### 4.1.1 ETF selection algorithms

#### Random ETF selection

Imagine an investor who chooses to invest in a few ETFs (which she selects completely arbitrarily) instead of a mutual fund. Will she be better or worse off? We attempt to find the answer with a simulation of this scenario by selecting the  $\mathbf{ETF}_t$  at random and then measuring its outperforming capacity. If this process is repeated multiple times, we obtain the average ability of a randomly selected set of ETFs to outperform mutual funds. We follow the original design of the random ETF selection algorithm, which was proposed in Cavalcante Filho *et al.* (2021).

As  $\mathbf{ETF}_t$  is a  $(P \times 1)$  vector, the procedure starts by manually defining the maximum number of ETFs that we want to include,  $P^*$ . As we want to keep the dimension of  $\mathbf{ETF}_t$  low, we set the threshold to  $P^* = 10$ . Then for each  $P = 1, \dots, P^*$  we randomly select a  $b$  subsample from  $\mathbf{ETF}_t$  denoted as  $\mathbf{ETF}_{(P),(b),t}$ . Next, for each mutual fund  $i = 1, \dots, M$  we estimate the following model

$$r_{i,t} = \alpha_i + \beta_i \mathbf{ETF}_{(P),(b),t} + \epsilon_{i,t} \quad (4.2)$$

We store  $R_{(P),(b),i}^2$ ,  $\hat{\alpha}_{(P),(b),i}$  for each  $i = 1, \dots, M$  and  $\hat{\boldsymbol{\pi}}_{(P)(b)}$ .

This procedure is repeated  $B = 500$  times, so  $b = 1, \dots, B$  and finally, the following statistics are computed

$$R_{(P)}^2 = \frac{1}{BM} \sum_{b=1}^B \sum_{i=1}^M R_{(P),(b),i}^2 \quad (4.3)$$

$$\hat{\alpha}_{(P)} = \frac{1}{BM} \sum_{b=1}^B \sum_{i=1}^M \hat{\alpha}_{(P),(b),i} \quad (4.4)$$

$$\hat{\pi}_{(P)} = \frac{1}{B} \sum_{b=1}^B \hat{\pi}_{(P),(b)} \quad (4.5)$$

So for a given  $P$  of the  $\mathbf{ETF}_t$ , average adjusted- $R^2$ , average alpha, and average fund category proportion are calculated.

### Naive ETF selection

A Naive ETF selection algorithm is our extension of the Random selection algorithm described in Cavalcante Filho *et al.* (2021). The idea behind it is that average investors do not choose ETFs at random. They instead follow simple guidelines and rules of thumb that can be found on various investment websites or popular books about investing. The core idea is that the investor should choose a set of ETFs that diversifies the risks across various asset classes, geographic regions, industries, and investment styles. We enforce the asset class diversification by requiring the set of ETFs to consist always of at least one ETF and at least one bond.<sup>1</sup> Furthermore, we ensure that the stock ETFs are diversified across regions by requiring each of the following four regions (Global, North America, Europe, Asia Pacific) to be present. The diversification across investment styles and industries is not explicitly enforced, but as ETFs are chosen at random within the described constraints, it is also marginally reflected.

The algorithm begins by randomly selecting a subsample  $b$  of 5 ETFs<sup>2</sup> lim-

<sup>1</sup>Hence the composition of a 5 assets portfolio varies between 4 bond ETFs combined with 1 stock ETF to 1 Bond ETF combined 4 stock ETFs.

<sup>2</sup>The choice of 5 ETFs is somewhat arbitrary. However, in our view, five instruments provide large enough diversification while at the same time it is a small and manageable number.

ited by the constraints described above. Next, for each mutual fund  $i = 1, \dots, M$  the following model is estimated

$$r_{i,t} = \alpha_i + \beta_i \mathbf{ETF}_{(b),t} + \epsilon_{i,t} \quad (4.6)$$

From this regression  $R_{(b),i}^2$ ,  $\hat{\alpha}_{(b),i}$  is stored for each  $i$ . After going over all  $i = 1, \dots, M$  skilled proportions  $\hat{\pi}_{(b)}$  are estimated with FDR methodology detailed in Section 4.2.

This procedure is repeated  $B = 500$  times, so  $b = 1, \dots, B$ , and ultimately the algorithm returns average adjusted  $R^2$ , average estimated alpha, and average skilled proportions. Mathematically,

$$R_{(P)}^2 = \frac{1}{BM} \sum_{b=1}^B \sum_{i=1}^M R_{(b),i}^2 \quad (4.7)$$

$$\hat{\alpha}_{(P)} = \frac{1}{BM} \sum_{b=1}^B \sum_{i=1}^M \hat{\alpha}_{(b),i} \quad (4.8)$$

$$\hat{\pi}_{(P)} = \frac{1}{B} \sum_{b=1}^B \hat{\pi}_{(b)} \quad (4.9)$$

### Proposed ETF selection

This ETF selection algorithm was also introduced in Cavalcante Filho *et al.* (2021), and it selects ETFs by evaluating their ability to outperform mutual funds (low  $\hat{\pi}^+$ ) and to explain fund return variability (high average adjusted- $R^2$ ). The algorithm is divided into two steps where the second step runs recursively till the finishing conditions are met.

Lets begin by defining  $E$  as the overall number of ETFs in our dataset so  $\mathbf{ETF}_t = (ETF_{(1),t}, \dots, ETF_{(E),t})$ . The first step of our algorithm starts by estimating for each  $e = 1, \dots, E$  over all mutual funds  $i = 1, \dots, M$  following regression model:

$$r_{i,t} = \alpha_i + \beta_i ETF_{(e),t} + \epsilon_{i,t} \quad (4.10)$$

For each  $e$ , we calculate the average adjusted- $R^2$  ( $R_{(e)}^2$ )<sup>3</sup>, and the estimated

---

<sup>3</sup>For a given  $e$ , the average adjusted- $R^2$  over all mutual funds is calculated as  $R_{(e)}^2 = \frac{\sum_{i=1}^M R_{(e),i}^2}{M}$

skilled fund proportion  $(\hat{\pi}_{(e),+})$ . We follow by selecting a set of ETFs that achieved  $R_{(e)}^2$  higher than a given threshold  $R^{2(*)}$ .

$$\mathbf{ETF}_{1,t}^{R^2} = \left\{ e : R_{(e)}^2 \geq R^{2(*)} \right\} \quad (4.11)$$

We use  $R^{2(*)} = 0.5$ . The first ETF is then selected as

$$ETF_{(e_1),t} = \left\{ e : e = \underset{e \in \mathbf{ETF}_{1,t}^{R^2}}{\operatorname{argmin}} \hat{\pi}_{(e),+} \right\} \quad (4.12)$$

This concludes the first step of the algorithm. In the second step, we estimate

$$r_{i,t} = \alpha_i + \beta_{i,e_1} ETF_{(e_1),t} + \beta_i ETF_{(e),t} + \epsilon_{i,t} \quad (4.13)$$

where  $e = (1, \dots, E) \setminus \{e_1\}$ . Then, the pair of ETF sets that produce a higher adjusted  $R^2$  than that of previous  $ETF_{e_1,t}$  is chosen.<sup>4</sup> Mathematically,

$$ETF_{2,t}^{R^2} = \left\{ e : R_{(e)}^2 \geq R_{(e_1)}^2 \right\} \quad (4.14)$$

Next, the second ETF is chosen as

$$ETF_{(e_2),t} = \left\{ e : e = \underset{e \in ETF_{2,t}^{R^2}}{\operatorname{argmin}} \hat{\pi}_{(e),+} \wedge \hat{\pi}_{(e),+} < \hat{\pi}_{(e_1),+} \right\} \quad (4.15)$$

Now, the step two is recursively repeated until  $\mathbf{ETF}_{s,t}^{R^2} = \emptyset$  or  $ETF_{(e_s),t} = \emptyset$ , where  $s$  is the number of steps. Hence we arrive to a final selection  $\mathbf{ETF}_{S,t} = \bigcup_{s=1}^S \{ETF_{(e_s),t}\}$ .

### 4.1.2 Alpha adjustment procedure

For each mutual fund  $i$ ,  $i \in \{1, \dots, M\}$ , we estimate the multiple regression model 4.1 for which we choose a matrix of ETFs based on the algorithms described in Section 4.1.1. The t-statistic for a hypothesis test  $H_0 : \alpha_i = 0$  against  $H_A : \alpha_i \neq 0$  is calculated by  $\hat{t}_i = \frac{\hat{\alpha}_i}{\hat{\sigma}_{\alpha_i}}$ , where  $\hat{\sigma}_{\alpha_i}$  is a HAC estimator described in Newey & West (1987). Determination of p-value using normal distribution cannot be used because as stressed in Kosowski *et al.* (2006), the cross-section of mutual fund alphas is not normally distributed as a result of heterogeneous risk-taking by mutual funds managers combined with non-normalities in individual fund alpha distributions. The aforementioned paper also introduces a bootstrap procedure to estimate the correct p-value. We follow this approach

<sup>4</sup>During recursion of step two, this is replaced by the set of previously chosen ETFs

in this thesis.

The process starts with estimating 4.1 and storing  $\{\hat{\alpha}_i, \hat{\beta}_i, \hat{t}_{\alpha_i}, \hat{\epsilon}_{i,t}\}$  where  $\hat{\epsilon}_{i,t}$  are estimated residuals. Next, we create  $B = 1000$  bootstrap samples of residuals  $\epsilon_{i,t}^b$  and create pseudo excess return time series as

$$\{r_{i,t}^b = \beta_i \mathbf{ETF}_{i,t} + \epsilon_{i,t}^b\}_{b=1}^B \quad (4.16)$$

Note that this manufactured series has a zero alpha by construction. However, when we fit this "bootstrapped" regression, we may receive positive (negative) alpha simply by the variation in the residual bootstrap sample - during bootstrap, we may end up choosing an unusually high number of positive (negative) residuals by chance alone. After storing t-statistics from these "bootstrapped" regressions, we can estimate the p-value of our original t-statistic as

$$\hat{p}_i = 2\min\left(\frac{1}{B} \sum_{q=1}^B \mathbb{1}(\hat{t}_i^b > \hat{t}_i), \frac{1}{B} \sum_{q=1}^B \mathbb{1}(\hat{t}_i^b < \hat{t}_i)\right) \quad (4.17)$$

where  $\mathbb{1}(\hat{t}_i^b > \hat{t}_i)$  is an indicator function which equals 1 if  $\hat{t}_i^b > \hat{t}_i$  and 0 otherwise.

## 4.2 Correction for lucky and unlucky funds

In the previous section, we described various selection algorithms based on which we choose the set of ETFs, which we then use in equation 4.1 to estimate the fund's alpha. However, if we repeat the procedure for all funds in our sample and then simultaneously test hypotheses about their alphas, we conduct multiple hypothesis testing. In such a situation, a straightforward count of significant alphas does not properly account for luck in such a multiple test setting – there will be many funds with significant estimated alphas by a mere chance. This issue is known as the problem of multiple comparisons and is described in more detail in section 4.2.1. One of the main pillars of this thesis is to uncover the true proportions of zero, positive and negative alpha funds prevalent in the mutual fund's investment universe attainable by Czech investors. Therefore we decided to address this issue by using the False Discovery rate procedure developed by Barras *et al.* (2010) which we describe in the section 4.2.2. We end this part with Section 4.2.3, where we demonstrate the usefulness of the False Discovery Rate approach in a simulation.

### 4.2.1 Problem of Multiple comparisons

Suppose that a hypothesis at 5% significance level about 100 parameters has to be conducted. For the sake of argument, we can assume that the null hypothesis is true for all of them. Hence we conduct the following series of hypothesis tests

$$\begin{aligned} H_{0,1} : \theta_1 = 0, H_{A,1} : \theta_1 \neq 0 \\ \vdots \\ H_{0,100} : \theta_{100} = 0, H_{A,100} : \theta_{100} \neq 0 \end{aligned}$$

Even though we know that the null hypothesis holds, provided the 5% significance level, we would expect to observe about five significant rejections by chance alone<sup>5</sup>. These particular cases, where we reject the null hypothesis even though it is true, are called false positives or Type I errors, and in multiple hypothesis test setting, we should be aware of them and try to account for them properly. Several techniques were developed to account for the inflation caused by false positives. Early methods focused on adjusting the significance level based on the number of comparisons made. These methods, with family-wise error rate, put upfront, are well described in Hochberg & Tamhane (1987). Modern approaches adjust the number of significant rejections with the number of expected false positive results. The frequently used method is called False Discovery Rate and was pioneered in Benjamini & Hochberg (1995). Our approach follows Barras *et al.* (2010) and is described in detail in the Section 4.2.2 which follows.

### 4.2.2 False Discovery Rate (FDR)

Our methodology strictly follows Barras *et al.* (2010). We start with the table, which defines all important variables. This should be helpful for orientation in further text.

---

<sup>5</sup>If we assume independence of individual tests, the probability that there would be at least one "incorrect" rejection is approximately 99.4 % ( $1 - 0.95^{100} = 0.9941$ ).

## Core Variables

$F_\gamma^+$	Proportion of false positive "lucky" funds
$F_\gamma^-$	Proportion of false negative "unlucky" funds
$S_\gamma^+$	Proportion of observed significantly positive funds
$S_\gamma^-$	Proportion of observed significantly negative funds
$\pi_0$	True proportion of zero alpha funds
$\pi_+$	True proportion of positive alpha funds
$\pi_-$	True proportion of negative alpha funds

## Further Parameters

$\gamma$	Significance level
$\lambda$	Threshold to separate zero-alpha p-values

The main idea of the FDR approach is to adjust the observed significant results,  $S_\gamma^+$  and  $S_\gamma^-$ , which we obtain in our sample for the number of false-positive results,  $F_\gamma^+$  and  $F_\gamma^-$ , that are expected to occur by chance alone. Consider that our sample of mutual funds can be split into three distinctive groups – zero-alpha, positive-alpha, and negative-alpha funds, denoted as  $\pi_0, \pi_+, \pi_-$ . Then for a given level of significance,  $\gamma$ , the expected proportion of false-positive results can be calculated as

$$E(F_\gamma^+) = \pi_0 \frac{\gamma}{2} \quad (4.18)$$

This is the proportion of funds that we expect will end up with significantly positive alpha just by chance. Given p-values from a hypothesis test  $H_i^0 : \alpha_i = 0$  against  $H_i^A : \alpha_i \neq 0$  obtained as described in Section 4.1.2, we can calculate the observed proportion of significant positive alpha funds as

$$\hat{S}_\gamma^+ = \frac{\sum_{i \in I_{\hat{\alpha}_+}} \mathbb{1}(\hat{p}_i < \frac{\gamma}{2})}{M} \quad (4.19)$$

where  $\mathbb{1}(\hat{p}_i < \frac{\gamma}{2})$  is an indicator function that yields 1 if  $p_i < \frac{\gamma}{2}$  and zero otherwise,  $i \in I_{\hat{\alpha}_+} = \{i : \alpha_i > 0\}$  signifies that only positive alphas are considered. Now we can estimate the expected proportion of skilled funds prevalent in the population as

$$E(\pi_+) = E(S_\gamma^+) - E(F_\gamma^+) \quad (4.20)$$

So with the estimated proportion of zero-alpha funds,  $\hat{\pi}_0$ , and an optimal significance level  $\gamma^*$ , the expected proportion of positive-alpha funds can be

calculated as

$$\hat{\pi}_+ = \hat{S}_\gamma^+ - \hat{\pi}_0 \frac{\gamma^*}{2} \quad (4.21)$$

Ultimately, the proportion of negative-alpha funds is calculated leveraging the fact that the sum of zero-alpha, positive-alpha, and negative-alpha funds has to be equal to 1. Hence,

$$\hat{\pi}_- = 1 - \hat{\pi}_0 - \hat{\pi}_+ \quad (4.22)$$

The procedure described above started with estimating  $\hat{\pi}_+$  first and then calculating  $\hat{\pi}_-$ . However, the whole process can be reversed and  $\hat{\pi}_-$  can be calculated first by equation 4.21 and then calculate  $\hat{\pi}_+$  by equation 4.22. The direction from which to start the calculation process should be decided based on the number of elements in sets  $I_{\hat{\alpha}_+}$  and  $I_{\hat{\alpha}_-}$  if  $I_{\hat{\alpha}_+}$  has more elements than  $I_{\hat{\alpha}_-}$  we estimate proportions by estimating  $\hat{\pi}_+$  first. Otherwise, the opposite approach is taken.

#### Estimating procedure for $\pi_0$

The estimation process of  $\pi_0$  is built on the knowledge about the behavior of the p-value when the Null hypothesis is true. In that setting, the p-value on a continuous test static is uniformly distributed over the interval  $[0,1]$ . Mathematically speaking, if the null hypothesis is true, it holds that

$$H_{0,i} : \alpha_i = 0 \Rightarrow p_i \sim U(0, 1) \quad (4.23)$$

However, in the scenario when the alternative hypothesis is true for a certain proportion of tests, the resulting p-value distribution will have more mass on the left tail because, in such a setting, there is a much higher likelihood that the p-values of these observations will be small. However, we will assume that p-values that are larger than a certain threshold  $\lambda^*$  can only arise from funds with zero-alpha. Therefore, taking into consideration the p-value empirical distribution, we can denote that the area,  $W(\lambda_*)$  described by

$$W(\lambda^*) = \frac{\sum_{i=1}^M \mathbb{1}(\hat{p}_i > \lambda_*)}{M} \quad (4.24)$$

becomes the section of the uniform distribution of p-values originating from zero-alpha funds on the interval  $[\lambda^*, 1]$ . Extrapolating this section to the whole interval  $[0, 1]$  gives us the estimator of  $\pi_0$ . Hence,

$$\hat{\pi}_0(\lambda^*) = \frac{W(\lambda^*)}{M} \frac{1}{1 - \lambda^*} \quad (4.25)$$

where the optimal threshold,  $\lambda^*$ , is determined by the bootstrapping method described in the next section.

#### Estimating procedure for $\lambda^*$

The optimal lambda,  $\lambda^*$ , is estimated by bootstrap method described in Barras *et al.* (2010). This approach chooses  $\lambda$  from the data by minimizing the Mean Squared Error (MSE) of  $\hat{\pi}_0(\lambda)$ . The procedure starts with creating a grid of lambdas  $\lambda = (\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_K)$  which we set in line with Barras *et al.* (2010) to  $(0.3, 0.35, 0.4, \dots, 0.7)$ . Then for each value  $\lambda_k \in \lambda$  we estimate the proportion of zero-alpha funds by equation 4.25 as  $\hat{\pi}_0(\lambda_k) = \frac{\sum_{i=1}^M \mathbb{1}(\hat{p}_i > \lambda_k)}{M} \frac{1}{1 - \lambda_k}$ . Then we form  $B$  p-value bootstrap samples<sup>6</sup> on which we estimate  $\hat{\pi}_0^b(\lambda_k)$ . We denote these as  $\{\hat{\pi}_0^b(\lambda_k)\}_{b=1}^B$ . Lastly, we calculate the estimate of Mean Square Error for each  $\lambda_k \in \lambda$  as follows:

$$\widehat{MSE}(\lambda_k) = \frac{1}{B} \sum_{b=1}^B \left( \hat{\pi}_0^b(\lambda_k) - \min_{\lambda} \hat{\pi}_0(\lambda) \right)^2 \quad (4.26)$$

The optimal lambda is then chosen such that

$$\lambda^* = \operatorname{argmin}_{\lambda} \widehat{MSE}(\lambda) \quad (4.27)$$

#### Estimating procedure for $\gamma^*$

Similarly as we did for the estimation of  $\lambda^*$ , we start by creating a grid of candidate gammas  $\gamma = (\gamma_1, \dots, \gamma_K)$ . To conform with Cavalcante Filho *et al.* (2021), we set the range to  $(0.05, 0.10, \dots, 0.50)$ . Then, for each  $\gamma_k \in \gamma$  we estimate the proportion of positive alpha funds using equation 4.21 as  $\hat{\pi}_+(\gamma_k) = S_{\gamma_k}^+ - \hat{\pi}_0(\lambda^*) \frac{\gamma_k}{2}$ . Then we form  $B$  bootstrap samples<sup>7</sup> on which we estimate

<sup>6</sup>Similarly to Barras *et al.* (2010), we set the  $B = 1000$

<sup>7</sup>Similarly to Barras *et al.* (2010), we set the  $B = 1000$

$\pi_+^b(\gamma_k)$ . We denote these as  $\{\pi_+^b(\gamma_k)\}_{b=1}^B$ . Lastly, we calculate the estimate of Mean Square Error for each  $\lambda_k \in \lambda$  as follows:

$$\widehat{MSE}^+(\gamma_k) = \frac{1}{B} \sum_{b=1}^B \left( \hat{\pi}_+^b(\gamma_k) - \max_{\gamma} \hat{\pi}_+(\gamma) \right)^2 \quad (4.28)$$

The  $\gamma^+$  is then determined as

$$\gamma^+ = \operatorname{argmin}_{\gamma} \widehat{MSE}^+(\gamma) \quad (4.29)$$

The same procedure is used to estimate  $\gamma^- = \operatorname{argmin}_{\gamma} \widehat{MSE}^-(\gamma)$ . Then the optimal gamma,  $\gamma^*$ , is set based on which MSE of the two is smaller. Hence if  $\min_{\gamma} \widehat{MSE}^-(\gamma) < \min_{\gamma} \widehat{MSE}^+(\gamma)$  we set  $\gamma^* = \gamma^-$ .

### 4.2.3 FDR Simulation Example

In this section, we would like to demonstrate that in a Multiple hypothesis setting, the simple count of significant test statistics fails to control for the false positives, while the FDR approach is able to adjust for them successfully and provides much better estimates. Also, we want to illuminate the estimation process of parameters  $\boldsymbol{\pi} = (\pi_-, \pi_0, \pi_+)$  which correspond to proportions of unskilled funds, zero-alpha funds, and skilled funds.

We generated 16000 alphas of zero-alpha funds, 3000 alphas of negative alpha funds, and 1000 alphas of positive alpha funds. Hence, the true fund proportions are  $\boldsymbol{\pi} = (0.15, 0.8, 0.5)$  by construction. We follow by conducting hypothesis test  $H_i^0 : \alpha_i = 0$  against  $H_i^A : \alpha_i \neq 0$  for each  $\alpha_i$  where  $i = 1, \dots, 20000$ . Nevertheless, since the alphas were generated from Normal distributions, there is still some randomness involved, and there will be some false positives. They correspond to funds that appear to have significant positive/negative alpha, even though they come from zero alpha distribution. As mentioned before, these funds can be considered to be lucky/unlucky. So as outlined in the previous section, we estimate respective p-values  $\hat{p}_i$ . Hence now we have a whole distribution of p-values, which we can see in Figure 4.1. Based on the estimation procedure of  $\lambda^*$  described in Section 4.2.2 we find that optimal  $\lambda^*$  for our simulated dataset is 0.30.

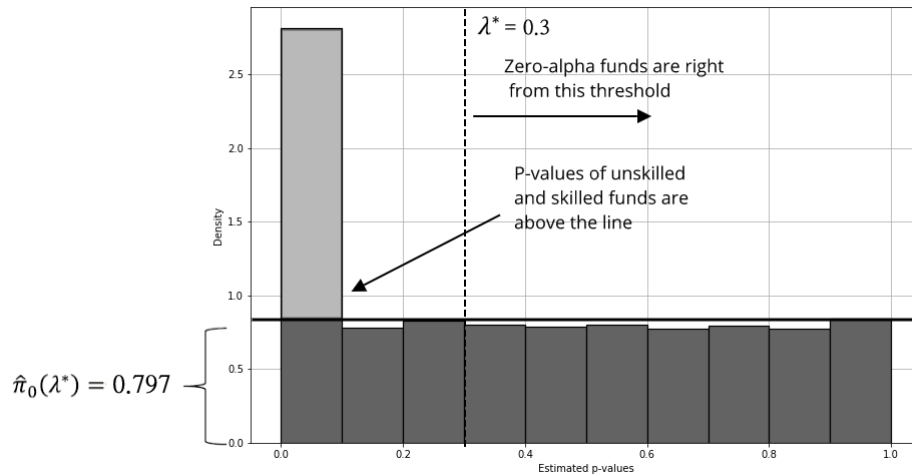


Figure 4.1: Figure shows density of p-values from 20000 simulated funds. We generated 16000 alphas from zero-alpha distribution, 3000 from negative alpha distribution and 1000 from positive alpha distribution. The dark grey colour depicts the region of p-values of zero-alpha funds. The light grey colour corresponds to statistically significant skilled/unskilled funds.

All p-values such  $\hat{p}_i > \lambda^*$  are assumed to come from zero-alpha distribution. Extrapolation of this area to the left of  $\lambda^*$  threshold, as described in equation 4.25, gives us the estimate of the proportion of zero-alpha funds in the whole population. In our case, the estimate is 0.797 (The true proportion is 0.8). The light grey area then represents the proportion of either skilled or unskilled funds.

So using the FDR methodology described in Section 4.2.2, we receive the following estimates for our simulation  $\pi^{FDR} = (0.155, 0.797, 0.048)$ . The simple count of significant alphas without any correction results in  $\pi^{count} = (0.170, 0.7658, 0.072)$ . Since we do not control for the false positives, they inflate the observed proportions of significant alphas and highlight the need for using FDR methodology since the estimates are much closer to the actual proportions used to generate the data. The results are summarized in Table 4.1.

	$\pi_-$	$\pi_0$	$\pi_+$
True proportions	0.15	0.8	0.5
Simple Count	0.172	0.757	0.071
FDR approach	0.155	0.797	0.048

**Table 4.1:** Simulation results - This table summarises the proportions of zero-alpha, skilled and unskilled funds ( $\pi_-$ ,  $\pi_0$ ,  $\pi_+$ ). It shows the true proportions which were used to generate data. Then it displays the estimated proportions using simple counts and FDR methodology.

With this demonstration of usefulness of False Discovery Rate methodology in estimating proportions of skilled, zero-alpha and unskilled funds in multiple hypothesis setting we conclude the methodology section. In the next section, we provide detailed description of our Mutual Funds and Exchange Traded Funds datasets.

# Chapter 5

## Data

As the viewpoint of Czech investors is of major interest in this thesis, we construct our universe with the best effort to reflect the opportunity set they face. To achieve this, we collect ISINs from all major providers of mutual funds in the Czech republic. To obtain information about each ISIN, we leverage the Refinitive Lipper Fund Research Database available to the IES students. If some fund is not present in the Lipper database, we get the information directly from the provider. Data is gathered for all major banks - ČSOB, Česká Spořitelna, Generali, Raiffeisen and from other major providers - CONSEQ and Amundi. Together totalling 1484 mutual funds covering a variety of asset classes, sizes, styles and geographic focus. However, pricing data on Refinitive Lipper Fund Research Database is not available for all funds. So after trying to download pricing data for all funds, we end up with 1381 funds with at least one NAV record. Furthermore, we only keep funds with at least 30 monthly observations for our analysis. Funds with fewer observations are ignored. Although this dataset can suffer from a survivorship bias to some degree, we believe that it still serves as a high-quality approximation of the whole population.

Dataset of UCITS ETFs was obtained from iShares<sup>1</sup> and it contains 521 UCITS ETFs. We decided to stick to only this provider due to the high computational complexity of implemented methods. Furthermore, since iShares is one of the largest ETF providers in the world, it has wide-ranging coverage of ETFs, and we believe that it resembles a good approximation of the overall ETF universe. From the 521 ETFs in our raw dataset, only 512 have pricing data available on Thompson Reuters Eikon. We also apply the same limit of

---

<sup>1</sup>iShares is one of the world's largest providers of ETFs. It is a subsidiary of BlackRock – The world's largest asset management company.

at least 30 observations. We show the evolution of the number of both mutual funds and ETFs in the Figure 5.1.

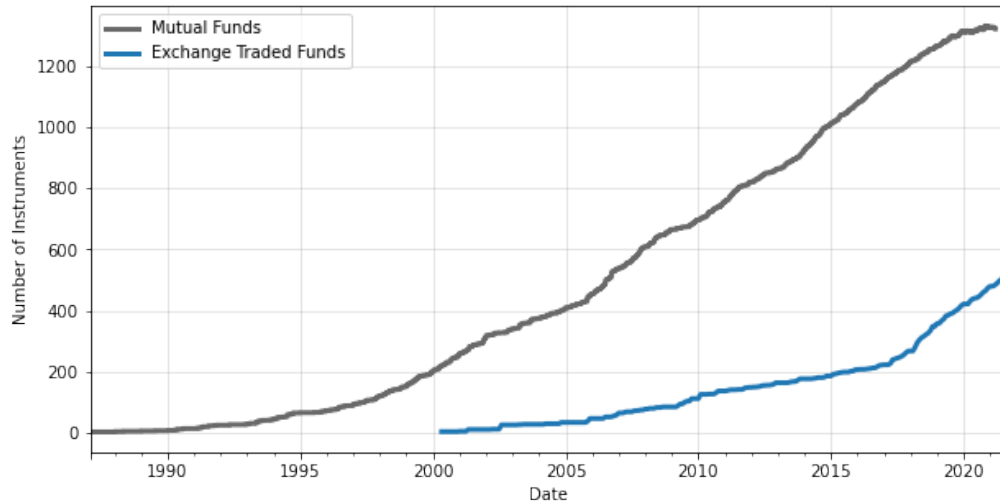


Figure 5.1: Number of Mutual Funds (grey line) and Exchange Traded Funds (blue line) from January 1987 to December 2020.

The figure shows that the mutual fund's data spans from January 1987 to December 2020, totalling 1322 mutual funds as of December 2020. It also clearly demonstrates the stable increase in the number of funds available on the market. As ETFs are more recent instruments than mutual funds, their pricing data starts in May 2000 and ends in December 2020. Similarly to mutual funds, the increasing trend in the number of ETFs is also apparent. On top of that, it seems to be accelerating from the year 2018. As we can see, the variety of choices that investors currently have is greater than ever. Since we will run the regressions of Mutual funds to ETFs, we need both datasets to overlay. So after limiting both datasets to a common window from May 2000 to December 2020, we receive our final dataset with 1241 Mutual funds and 315 ETFs. the length of our dataset is

In Table 5.1 we present the sample statistics for both mutual funds and ETFs dataset.

<b>Panel A: Mutual funds (MFs)</b>							
	Mean	Std.	Min	p05	Median	p95	Max
Avg. return (%p.m.)	0.39%	0.35%	-0.91%	-0.03%	0.33%	0.99%	2.63%
Sharpe ratio (ex-post)	0.13	0.12	-0.45	-0.01	0.11	0.30	1.55
Time Series Length	150	72	30	42	153	248	248
Total number of Funds: 1241							
<b>Panel B: Exchange Traded Funds (ETFs)</b>							
	Mean	Std.	Min	p05	Median	p95	Max
Avg. return (%p.m.)	0.45%	0.41%	-0.38%	-0.10%	0.38%	1.19%	2.09%
Sharpe ratio (ex-post)	0.13	0.11	-0.24	-0.02	0.13	0.31	0.53
Time Series Length	105	62	30	32	98	221	248
Total number of ETFs: 315							

**Table 5.1:** The table shows descriptive statistics of the final dataset. Panel A corresponds to our sample of mutual funds, Panel B to a sample of ETFs. Statistics for average monthly return, ex-post Sharpe Ratio and the length of our time series are shown in both panels. For each variable, we present Mean, Standard Deviation, Minimum, Maximum, Median and 5th (p05) and 95th (p95) percentiles.

The table shows descriptive statistics of average monthly return, Ex-post Sharpe ratio, and the length of a time series for both our samples. We can see that the average monthly returns of MFs and ETFs are 0.39% and 0.45%, respectively. They are both positive, and the average return of ETFs is slightly higher in comparison to Mutual funds. Additionally, the average length of a mutual fund time series is 150 months (12.5 years) against 105 months (8.75 years), reflecting that ETF is a relatively new instrument compared to Mutual Funds. Nevertheless, this period is long enough from both econometrical and economic perspectives. The average Sharpe ratio is similar for both MFs and ETFs and equals 0.13. Further, the maximum Sharpe ratio of a mutual fund is about three times higher than the maximum Sharpe ratio of an ETF. The former being 1.55 and the latter 0.53. However, it is fair to note that the minimum Sharpe ratio of a mutual fund is -0.45, which is about twice as large as the minimum Sharpe ratio for an ETF which equals -0.24. These statistics align with our intuition about the strategies employed by mutual funds and ETFs. Mutual Funds are usually actively managed, and their managers aim to earn above-average returns for their investors. They typically undertake more risks, and sometimes it pays off (exceptionally high Sharpe ratio), and sometimes it does not (exceptionally low Sharpe ratio). On the other hand, ETFs are

passively managed, and they track a specific basket of securities. Sometimes given basket is favoured by the market (high Sharpe ratio), and other times it is not (low Sharpe ratio), but in general, the swings are much smaller because of the less risk undertaken.

In the following section we provide a detailed description of the mutual funds universe available to Czech investors.

## 5.1 Characteristics of Czech Mutual Funds Universe

To shed even more light on the variety of our mutual fund dataset, we provide a few insights that characterise the mutual fund's investment universe that is easily obtainable for Czech investors.

Starting with currencies in which funds are denominated, approximately 42% of funds have a denomination in EUR, 29% in CZK and 27% in USD. Other denominations, namely CHF, JPY, PLN, AUD, GBP and NOK, have less than 1% of funds. Another characteristic that is of extreme importance is fund type. This variable specifies the asset class in which the fund is investing. A little less than half, 48%, of the funds invest mainly in stocks. 25% holds debt securities primarily. 17% are of a "mixed type", which means that they invest in both aforementioned asset classes – bonds and equities. Remaining asset classes are alternatives (5%), short-term (2%), real estate (2%), commodities (1%). The size of the funds is also very heterogeneous. Our database contains very small funds with several hundred thousand euros in AUM but also humongous funds with billions of euros under management.

Moving to general characteristics of performance, histograms of ex-post Sharpe ratios and Treynor ratios are shown in Figure 5.2.



Figure 5.2: The figure shows histograms of annualized Sharpe ratio and annualized Treynor ratio calculated from monthly data.

The annualized Sharpe ratio of the market factor is 0.44, while the average value for all funds in our sample is only 0.4. This indicates that the market is slightly more generous in rewarding investors for a unit of total risk. Furthermore, only 35% of funds have a higher Sharpe ratio than the market. The average annualized Treynor ratio of the market factor is 0.0685, while the average value among mutual funds is 0.06466. Both values are quite similar this time, and slightly less than 45% of mutual funds have a higher Treynor ratio than the market.

Services of mutual funds are not free, and there are a few costs associated with them. Firstly, there is an entry fee. Usually, it is used to pay for the sales and distribution expenses. Then there is the management fee and the performance fee. These fees are paid to the investment company which owns the respective fund, and it serves to cover the costs of operating the fund - salaries to employees, licensing costs and others. Sometimes there is also an exit fee, but it is zero for the majority of funds. For a few funds, it is zero if investors redeem their money only after a certain time, usually several years. However, the measure that best reflects the actual cost of a fund is called the Total expense ratio (TER). It is a measure of the total costs associated with managing and operating a mutual fund and it is calculated by the following formula

$$TER \% = \frac{\text{Total operating expenses in accounting currency}}{\text{Average net assets in accounting currency}}$$

In plain English, this ratio equals the sum of all fees and incidental costs

charged to the mutual fund assets on an ongoing basis taken retrospectively as a percentage of the net assets. For example, if a mutual fund A has a TER of 1.55%. It means that approximately 1.55% of assets were used to cover the expenses of the operation during the previous year, and importantly, it directly affects the investors return. If fund A generated a gross return of 8% during the last year, the net return to the investor would only be about 6.45 %.

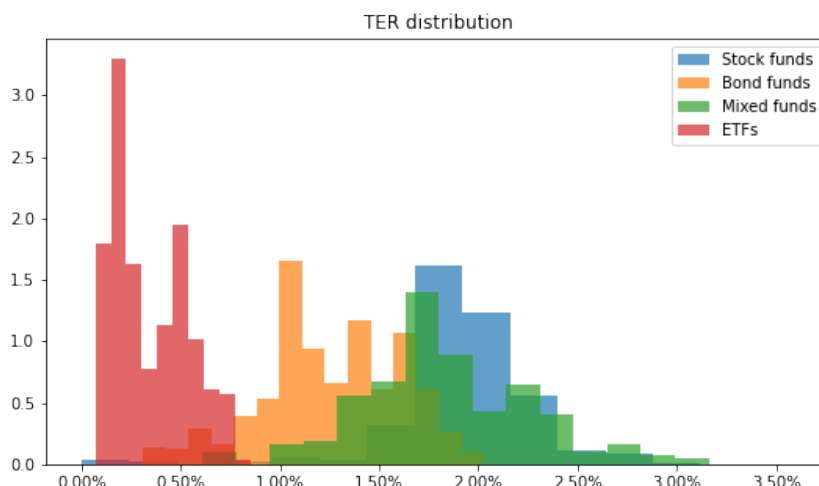


Figure 5.3: Densities of Total Expense Ratios (TER) of mutual funds investing in Stocks (blue), bonds (orange), and mixed assets (green). TER of ETFs is shown in red colour.

In figure 5.3, there is a histogram of the Total Expense Ratio for the three main mutual fund asset class types. As we can see from the plot TER of bond funds is generally smaller than the TER of stock or mixed funds. The average TER of bond funds is 1.23%, while the average for mixed and stock funds is 1.82% and 1.91%, respectively. As bond funds have, on average, smaller returns than stock funds, the investment companies cannot ask for such high fees as stock funds because it would destroy the investor's returns. To provide a better comparison between fees of mutual funds and of ETFs, we also include the TER of ETFs. As we can see from the plot, it is generally much smaller than the TER of mutual funds. The average TER of ETFs from our sample is 0.34% which is more than five times smaller than the average TER of stock or mixed mutual funds and more than three times smaller than bond mutual funds. This is really an enormous difference in the price that investors pay for their instruments.

---

There is one particularly interesting question regarding the fee size. One might ask if there is a relationship between the fund skill and its fee. Is it that the skilled fund managers ask for more money for their service, but they generate enough excess returns that the investor still benefits? Or, on the contrary, is the fee on average a burden for investors which drags down their overall performance? We investigate this question in Section 6.2. However, we first examine the proportions of skilled, zero-alpha and unskilled funds in the following section.

# Chapter 6

## Results

In this chapter, we would like to present our results. This chapter consists of two main parts. The first section elaborates on the central question of this thesis - whether mutual funds on aggregate add value to the investors. The second section is devoted to analyzing the relationship between a fund's performance and its Total Expense Ratio.

### 6.1 Analysis of Skilled funds proportions

The question we seek to answer is whether the mutual funds offered in the Czech republic on average add value to their investors. It means that we want to draw a conclusion about the population of mutual funds as a whole. To accomplish this task, we work with an extensive dataset consisting of most mutual funds available to Czech investors. We estimate each fund's alpha given a set of regressors and subsequently estimate the overall proportions of funds in each "skill group". The conclusion is then driven by the magnitudes of proportions in each skill group. We consider that mutual funds successfully deliver value to their investors if there is a large-enough proportion with significantly positive alpha from a regression against a set of ETFs. On the other hand, we conclude that mutual funds fail to bring value to their investors if there is a large-enough share with significantly negative alpha. Lastly, the scenario when there are not many funds with significantly positive or negative alphas would suggest that the majority of funds are "Zero-alpha", which means that they are comparable to the ETFs and that there is not a clear winner or loser.

The estimation process of alpha is conditional on the given set of regressors.

The choice of specific regressors enables us to draw conclusions on several levels. We start by using factor models, and then we continue by applying various approaches of ETF selection.

By applying factor models, we obtain a performance benchmark for our other results and a comparison with other authors' findings. We find that unskilled funds outnumber the skilled funds for both 3-Factor and 4-Factor models. The frequency of unskilled funds is almost ten times greater than that of skilled funds. This predominance of unskilled funds is in line with the results well documented throughout the literature (conducted mainly on US data). The results for both models are almost identical, and the proportions are approximately 36%, 60%, 3.8% for unskilled, zero-alpha and skilled funds, respectively. Cavalcante Filho *et al.* (2021) obtained 37.86%, 60.95%, 1.19%, further suggesting that the European mutual fund market seems to be no different from the United States. The complete results with using factors as regressors are summarized in table 6.1.

<b>Models</b>	$\pi_-$	$\pi_0$	$\pi_+$
4 Factor Model	36.77 %	59.48 %	3.75 %
3 Factor Model	35.46 %	60.75 %	3.80 %
MKT-RF	38.14 %	57.90 %	3.95 %
HML	0.00 %	78.53 %	22.14 %
SMB	0.27 %	31.16 %	68.57 %
MOM	0.00 %	52.38 %	47.62 %

**Table 6.1:** The table shows the proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds estimated on alphas from a regression where Factor models or individual factors were used as regressors. The first part of the table displays results of using three and four-factor models, the second part shows results with individual factors where Mkt-RF is the Market Factor, HML (High-Minus-Low) refers to the Value factor, SMB (Small-Minus-Big) corresponds to the Size factor, and MOM is a Momentum Factor. Data for these factors were obtained from Kenneth French website.

By investigating individual factors, we can see that the Market factor (Mkt-RF) is the most important. However, one unanticipated finding was that other factors alone are not sufficient to explain mutual funds returns. The reason for this is not clear, but it may be related to the unique risk exposures that are obtained by these long-short factors. These factors might be very useful in

combination with the Market Factor because of the unique exposures that they capture. However, in isolation, it seems that the absence of the Market Factor leaves so much variation unexplained that they simply fail to explain returns of mutual funds.

To sum it up, these results suggest that mutual funds fail to bring value to investors when compared to conventional risk factors. Importantly, we highlight that the performance of mutual funds can be replicated by having a particular tilt towards a well-regarded Market factor. Now, we want to investigate whether similar results can not be accomplished with tradable instruments instead of theoretical factors. We examine this question by employing various ETF selection algorithms. We start by evaluating the performance with Random Selection of ETFs, which is described in Section 4.1.1. In this case, no guided selection methodology is employed, and the ETFs are simply sampled at random. After conducting 500 bootstrap iterations, final average proportions are summarized in Figure 6.2.

Number of ETFs	$\pi_-$	$\pi_0$	$\pi_+$
1	0.00 %	71.46 %	28.54 %
2	4.64 %	71.83 %	23.51 %
3	7.68 %	76.13 %	16.19 %
4	7.99 %	75.57 %	16.44 %
5	8.93 %	76.73 %	14.34 %
6	6.95 %	81.84 %	11.21 %
7	9.14 %	81.24 %	9.62 %
8	5.51 %	83.38 %	11.12 %
9	5.30 %	80.51 %	14.19 %
10	6.47 %	79.86 %	13.68 %

Table 6.2: The table summarizes the results of the Random Selection algorithm, which randomly selects a given number of ETFs (from 1 to 10) and uses them to estimate the overall skill proportions. The procedure is repeated 500 times for each number of ETFs, and the average proportions calculated over all runs are stored. These average proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds are shown in this table.

The results show that the random selection of ETFs is unfit to outperform mutual funds. The vast majority of the funds are zero-alpha no matter how many ETFs are chosen. Furthermore, the proportion of skilled funds usually exceeds the proportion of unskilled funds. If we take averages over all runs,

we obtain that the average proportion of zero-alpha funds is around 77% and of skilled and unskilled funds about 16% and 6%, respectively. This is interesting because there are almost three times more skilled funds than unskilled ones when benchmarking against the random selection of ETFs. It is an indication that there is some skill present in the mutual fund industry because the random selection of ETFs clearly underperforms. Although these results are in line with Cavalcante Filho *et al.* (2021) they represent quite a surprise against the conventional academic belief that passive investing is superior to active investing. However, before drawing more conclusions, we have to pause and think about the context of these results. The procedure simply chooses ETFs at random. However, not all ETFs are equal at exposures they offer, and not all exposures are convenient to outperform mutual funds. So we theorize that these results are influenced by the fact that, unlike mutual funds, the primary purpose of ETFs is not to generate above-market returns. The purpose of ETFs is more shifted towards providing low-cost exposure to a specific sector, region, or risk in general. This may cause many ETFs not to be adequately diversified from a broad market perspective and offer a suboptimal risk-reward property. So while there are many universal ETFs like MSCI World or Global bonds ETFs, which are well equipped to benchmark mutual funds, there is also a large number of niche ETFs that provide unique exposures<sup>1</sup> but are unfit to beat mutual funds.

We explore this idea further by looking at the skill proportions obtained from a few different ETFs. If we propose that the results from a random selection of ETFs are driven down by too specific and, importantly, not diversified enough ETFs, we should see that general, properly diversified ETFs are significantly better at explaining mutual fund returns and lead to a higher proportion of unskilled funds (and a lower proportion of skilled funds). We show the results for major indexes across the globe in Table 6.3.

---

<sup>1</sup>Like MSCI South Africa UCITS ETF or MSCI EM Islamic UCITS ETF.

ETF	$\pi_-$	$\pi_0$	$\pi_+$
(1.) MSCI World	32.10 %	59.79 %	8.11 %
(2.) S&P 500	52.07 %	45.46 %	2.57 %
(3.) FTSE 100	1.55 %	79.7 %	18.75 %
(4.) MSCI EM Asia	6.99 %	71.31 %	21.70 %
(2.) - (4.) Combined	26.15 %	71.45 %	2.40 %

**Table 6.3:** This table displays proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) funds estimated on alphas from a regression where well-known and diversified ETFs were used. These ETFs approximate indices from various regions all around the world and can be thought of as a representation of average returns in these regions. MSCI World approximates the Market index, S&P 500 is the proxy for the United States market, FTSE 100 reflects the European market, and MSCI EM Asia proxies financial markets of Asia.

Although there are considerable differences in skill proportions between the ETFs from distinct regions, one common pattern can be derived – the broad diversification pays off. If the MSCI world<sup>2</sup> is used as a single regressor, the proportion of unskilled funds rises to 32.10 %, and the proportion of skilled funds drops to 8.11 %. Now the ratio completely shifted, and the prevalence of unskilled funds is about four times higher than skilled funds. Furthermore, there is a notable similarity between the results of the Mkt-RF factor and the MSCI world. The former having proportions  $\pi_-$ ,  $\pi_0$ ,  $\pi_+$  as 38.14 %, 57.90 %, 3.95 % and latter 32.10 %, 59.79 %, 8.11 %. These differences might be partly explained by the absence of transaction costs in the case of the Market factor.

If we take a look at the individual ETFs from specific regions, we can see that the results are proportional to how the market recently favoured the particular regions. The skill proportions obtained with S&P 500 are the most extreme. Only 2.57 % of funds can be considered skilled, while an alarming 52 % are estimated to be unskilled. This is clearly driven by the staggering uptrend that the stock market of the United States experienced in the last ten years. On the other hand, ETFs from not so fortunate markets performed poorly. FTSE 100 (European equities) and MSCI EM Asia (Asian equities) performed similarly or even worse than the random selection. With FTSE 100,

<sup>2</sup>MSCI World comprises 1559 holdings from 23 countries covering 85% of the listed equities in each country. Generally speaking, it is one of the most diversified ETFs available and can be thought of as an approximation of "The Market index/factor".

the estimated proportion of unskilled funds was 1.55% (18.75 % of skilled), and with MSCI EM Asia, it was 6.99% (21.70 % of skilled). So as we can see, regional diversification matters a lot. And because it is inherently difficult (or even impossible) to predict which region will outperform others in the future, it is probably safest to hold all of them. For example, when combining the three previously mentioned "localized" ETFs, we obtain proportions quite similar to what was reached with MSCI World.

This leads us to a Naive Selection algorithm that we designed to enforce simple diversification rules. We want to investigate whether a few simple constraints for choosing ETFs can be enough to give average investors an edge over mutual funds. The procedure runs 500 times and always selects five ETFs to calculate the skill proportions. However, it enforces limits to ensure that ETFs are diversified enough. It requires that at least one Bond ETF and one Stock ETF are always present. Above that, it requires the portfolio to be diversified across several geographical regions. Further details are provided in Section 4.1.1. The average proportions  $\pi_-$ ,  $\pi_0$ ,  $\pi_+$  estimated using Naive Selection algorithm are 24.82% , 68.58% and 6.26%. This is a very interesting result because it shows that almost one-quarter of mutual funds perform significantly worse than this naive diversification approach and that only about 6% of the funds are able to outperform it. This is an essential result because it suggests that if investors adhere to simple diversification rules, they can achieve better results than about 25% of mutual funds. The importance of this result stands out in the context of the Robo-advisory companies that recently gained considerable popularity in the Czech Republic. These companies promote passive investing in a transparent and easy-to-use way and are also cheaper than conventional mutual funds.<sup>3</sup> However, as we demonstrated with the Naive Selection algorithm, only a modest diversification is enough to outperform a large portion of mutual funds. Moreover, retail investors can achieve this diversification with only minimal effort and importantly, compared to robo-advisors, they can cut the average cost three times to only about 0.3%.

Finally, we show the results of the Proposed Selection Algorithm. This approach chooses ETFs based on their ability to explain fund's returns (high adjusted  $R^2$ ) and its ability to outperform hence causing high proportions

---

<sup>3</sup>The average management fee of these robo-advisors is around 1% which is almost half of the average management fee of a mutual fund which is about 1.9%.

of unskilled funds (high  $\pi_-$ ). Further details of the selection procedure are described in Section 4.1.1. After applying the algorithm on our data, three ETFs with the following isins were chosen: IE00B3WJKG14, IE00B428Z604, IE00BDFJYM28. The first one tracks the information technology sector of the SP 500, the second one invests in a full range of maturities of Spain government bonds, and the third holds a blend of bond types (mainly government and corporate bonds) from all around the world. When applying these ETFs, the proportions of unskilled ( $\pi_-$ ), zero-alpha ( $\pi_0$ ) and skilled ( $\pi_+$ ) are 85.60% , 13.7%, 0.71% respectively. With this methodology, only less than 1% of mutual funds can be considered skilled.

To wrap it up, we showed that random selection of ETFs is unfit to outperform mutual funds. However, as we demonstrated with the Naive Selection approach, enforcing simple diversification rules is enough to outperform a quarter of all mutual funds and that only about 6% of funds can be considered skilful against this approach. Nevertheless, holding the single diversified ETF, such as MSCI World, can be the best investment strategy for retail investors due to its capacity to outperform mutual funds combined with the simplicity of holding a single instrument.

In the next section, we examine the one thing that is often blamed as the cause of the mutual funds' underperformance – their high costs.

## 6.2 Relationship between performance and cost

In this section, we would like to present a concise analysis of the relationship between the alpha of a fund (a proxy for managers performance) and the Total Expense Ratio (a proxy for compensation of a fund's management). As portfolio management is the essential service provided by a funds management and TER reflects the "price" that is paid for that service, standard economic logic implies that in a well-functioning market, agents might be willing to pay a higher price for service only if the provided service is more valuable. Hence it follows that we should find a positive relationship between the fund's estimated alpha and its TER.

In order to investigate this question, we estimate the following regression model for every Mutual fund  $i = 1, \dots, M$ .

$$r_{i,t} = \alpha_i + \beta_i r_t^{MSCIWorld} + \epsilon_{i,t} \quad (6.1)$$

where  $r_{i,t}$  is a return of mutual fund  $i$  in time  $t$ ,  $\alpha_i$  is the estimated alpha,  $r_t^{MSCIWorld}$  is a return of MSCI World ETF at time  $t$  and  $\epsilon_{i,t}$  is an error term. We store the estimated alphas for each mutual fund, and then we find the corresponding Total Expense ratio of a respective fund. Now we can estimate the cross-sectional regression model

$$\hat{\alpha}_i = \gamma + \tilde{\beta} TER_i + \tilde{\epsilon}_i \quad (6.2)$$

where  $\hat{\alpha}_i$  is the estimated alpha from a regression against the MSCI World,  $\gamma$  is an intercept,  $TER_i$  is a Total Expense Ratio of a fund  $i$ ,  $\tilde{\beta}$  is the coefficient of interest which describes the relationship between fund's alpha and its TER, and  $\tilde{\epsilon}_i$  is an error term.

To check whether there is any difference between the Mutual Funds that invest exclusively in Stocks and the funds that invest only in bonds, we split the funds into two respective groups and estimate the regressions separately. Finally, the results are summarized in Table 6.4.

Panel A: Regression results for Stock Mutual Funds						
	coef	std err	z	P>  z	[0.025	0.975]
<b>TER</b>	-0.0013	0.000	-3.320	0.001	-0.002	-0.001
<b>const</b>	0.0006	0.001	0.880	0.379	-0.001	0.002

Panel B: Regression results for Bond Mutual Funds						
	coef	std err	z	P>  z	[0.025	0.975]
<b>TER</b>	-0.0019	0.000	-4.917	0.000	-0.003	-0.001
<b>const</b>	0.0029	0.001	5.643	0.000	0.002	0.004

Table 6.4: Results from regression of fund's alpha against fund's Total Expense Ratio (Equation 6.3). Note that in this case alphas were estimated in a regression against MSCI World.

Contrary to the original proposition, we find that the relationship between the fund's alpha and its Total Expense Ratio is negative for both asset classes, suggesting that funds that charge higher fees deliver worse risk-adjusted per-

formance to their investors. As we can see from the table, both coefficients are negative and statistically significant (at 1% significance for Stock funds and  $< 1\%$  for Bond funds). Residuals from these regressions are relatively well-behaved, as can be seen in the residual plot in the Figure 6.1. Their Expected value seems to be zero for the whole range of obtained TERs with the exception of the tails. Especially for the low values of TER in the case of bond funds, the regression seems to be slightly underfitting. However, a lack of observations in both tails can influence this result. Furthermore, the variance also appears to be relatively constant for most values of TER, so the potential issue of heteroscedasticity seems to be relatively small.

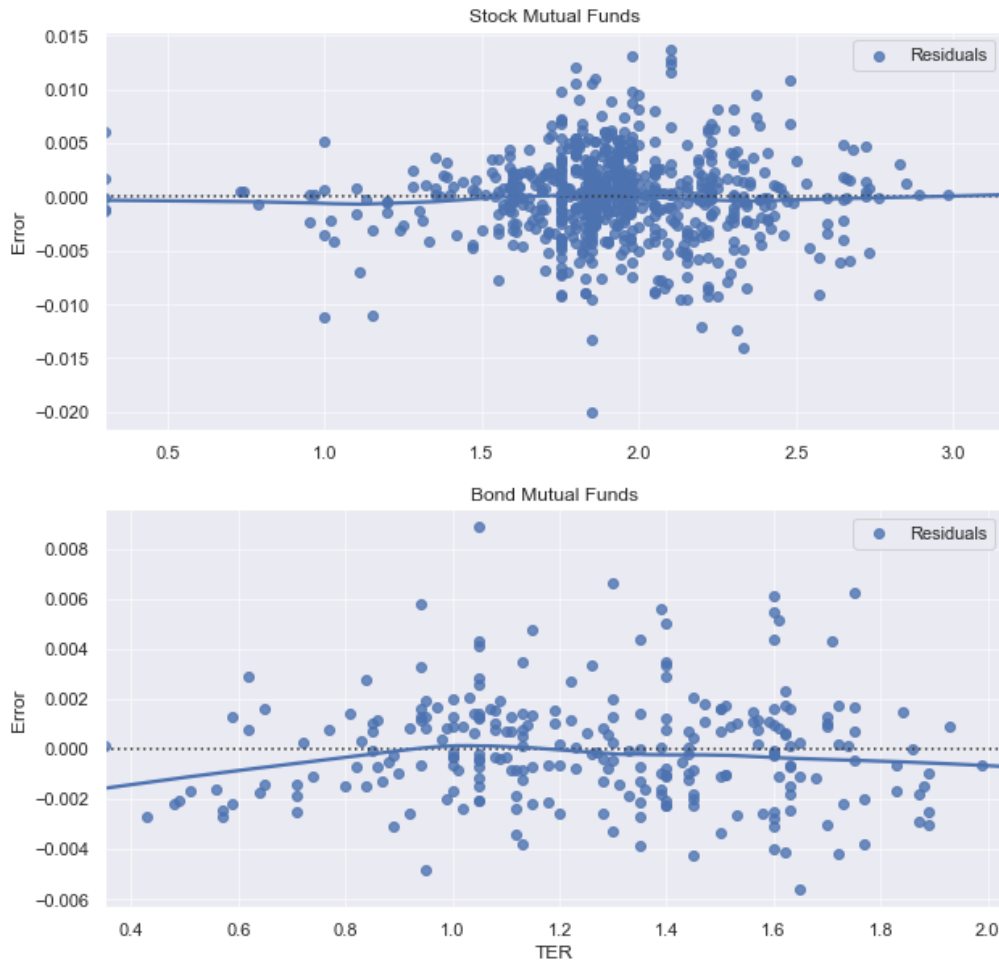


Figure 6.1: Figure depicts the residuals from the regression of fund's alphas against Total Expense Ratio (TER). The upper sub-figure is from the regression on Stock mutual funds exclusively, while the bottom sub-figure is from regression on Bond funds. The original value of TER is depicted on the x-axis, and the y-axis shows the value of a respective error term. The blacked dotted line shows the value of 0 across the whole plot and the blue curve is a non-parametric lowess model. In an ideal case the blue line should overlay the dotted line.

To check the robustness of this result, we replicated this methodology by applying factors instead of the MSCI World Exchange Traded Fund. Alphas from factors regression are standard performance metrics, so they represent an ideal candidate to broaden our analysis. The results are summarized in Table 6.4.

<b>Panel A: Regression results for Stock Mutual Funds</b>						
	<b>coef</b>	<b>std err</b>	<b>z</b>	<b>P &gt;  z </b>	<b>[0.025</b>	<b>0.975]</b>
<b>TER</b>	-0.0008	0.000	-2.945	0.003	-0.001	-0.000
<b>const</b>	-0.0010	0.001	-1.832	0.067	-0.002	6.68e-05

<b>Panel B: Regression results for Bond Mutual Funds</b>						
	<b>coef</b>	<b>std err</b>	<b>z</b>	<b>P &gt;  z </b>	<b>[0.025</b>	<b>0.975]</b>
<b>TER</b>	-0.0013	0.000	-4.290	0.000	-0.002	-0.001
<b>const</b>	0.0020	0.000	5.092	0.000	0.001	0.003

Table 6.5: Results from regression of fund's alpha against fund's Total Expense Ratio. Note that alphas were estimated in a regression against the 3-Factor model in this case.

We can see that the results are very similar. Even though the values of both coefficients are a little bit weaker than in the previous case, both coefficients are still statistically significant at 5% significance. This provides further evidence that these results are relatively robust and that the relationship between the fund's performance and its Total Expense Ratio is indeed negative.

Lastly, to address the possible problem of endogeneity, we extend the equation 6.3 with two more variables that may be hidden in the error term and have an effect on the performance of the fund. Hence we include the fund's Asset Under Management (AUM) and standard deviation of the fund's returns. The equation we estimate then changes to

$$\hat{\alpha}_i = \gamma + \tilde{\beta}_1 TER_i + \tilde{\beta}_2 AUM_i + \tilde{\beta}_3 STD_i + \tilde{\epsilon}_i \quad (6.3)$$

And after estimating it, we obtained the following results.

	coef	std err	z	P>  z	[0.025	0.975]
<b>TER</b>	-0.0007	0.000	-1.765	0.078	-0.001	7.43e-05
<b>AUM</b>	4.549e-15	2.71e-14	0.168	0.867	-4.86e-14	5.77e-14
<b>STD</b>	-0.0518	0.011	-4.915	0.000	-0.072	-0.031
<b>const</b>	0.0019	0.001	2.418	0.016	0.000	0.003

---

	coef	std err	z	P>  z	[0.025	0.975]
<b>TER</b>	-0.0025	0.001	-4.785	0.000	-0.004	-0.001
<b>AUM</b>	-1.581e-14	3.68e-15	-4.299	0.000	-2.3e-14	-8.6e-15
<b>STD</b>	0.0322	0.021	1.545	0.122	-0.009	0.073
<b>const</b>	0.0031	0.001	5.945	0.000	0.002	0.004

Table 6.6: Results from regression of fund's alpha against fund's Total Expense Ratio (TER), Assets Under Management (AUM) and its Standard Deviation (STD). Note that in this case alphas were estimated in a regression against MSCI World.

When we added two more variables, we obtained mixed results for TER, and the importance of AUM and STD also seems to be unclear. Assets Under Management appears to be relevant for bond funds but are statistically insignificant for stock funds. Conversely, for stock funds, the standard deviation is statistically significant at less than 1% significance, while for bond funds, it is insignificant. Nevertheless, the TER coefficient in the case of stock mutual funds is still negative, but it became insignificant at a 5% significance level. For Bond mutual funds, the coefficient remained negative and significant at even 1% significance.

Although the value of the TER coefficient is not always stable, it remains negative, and most importantly, the statistical significance holds for the majority of fitted regression. Based on the original paradigm that we stated at the beginning of this section, this result can be puzzling. However, based on the arithmetic of active management described in Sharpe (1991), it should not be a surprise. Sharpe shows that the before costs return on the average actively managed dollar have to be the same as the return on the average passively managed dollar. However, this holds only before fees. When the costs are included, the return on average actively managed dollar will be smaller than the return on average passively managed dollar because active managers incur greater costs. Note that this discussion is about the "average dollar", it does not rule out the possibility of exceptionally good funds. However, with this

---

paradigm, the higher the fee, the smaller the average return is generated for investors, which is exactly what we can see in the data. Furthermore, the relationship seems to be stronger for bonds. It means that what regards bond funds, investors should be especially aware of the fee size. Nevertheless, this implication is similar to both stock and bond funds - retail investors should try to keep the costs at a minimum.

# Chapter 7

## Conclusion

Despite the evidence from the United States indicating that mutual funds represent suboptimal investment vehicles, they are still among the most popular investments in the Czech Republic. In this thesis, we examine whether this popularity can be explained by the abnormal ability of local managers to outperform the market. Based on the alpha from a regression against a concise set of ETFs, we estimate the overall proportions of skilled (significantly positive alpha), zero-alpha (insignificant alpha) and unskilled (significantly negative alpha) that are prevalent in the population. We find that the random selection of ETFs is unfit to outperform the mutual funds on average. However, we introduced a straightforward extension that enforces simple diversification rules, and with this selection algorithm, we estimated that the proportion of skilled funds is only 6.26% while the proportion of unskilled funds is almost four times larger at 24.82%. We further showed that with the proposed selection algorithm, only a minimal fraction of mutual funds could be considered skilled. So our results suggest the following. Firstly, Exchange Traded Funds can be used to outperform a large proportion of mutual funds. And secondly, country diversification appears to be a key factor for constructing a portfolio of ETFs that consistently outperforms. These two findings have normative implications for retail investors because such etf selection approach can be easily implemented and, on average, beats the mutual funds. Moreover, we argue that if it is possible to outperform mutual funds with a set of passive and less expensive Exchange Traded Funds, a more predictable stream of returns can be achieved since it is not susceptible to shifts in mutual funds investing strategy or possible excessive risk-taking by a funds management.

---

Lastly, we investigated the relationship between the fund's performance and its cost. As a proxy for the fund's performance, we used estimated alpha, and the cost was proxied with the Total Expense Ratio (TER). We document a negative relationship between the two suggesting that investors should be highly cautious of the costs associated with their investments and try to minimize them.

# Bibliography

- ANDRIKOIANNPOULOU, A. & F. PAKONSTANTINO (2019): “Reassessing false discoveries in mutual fund performance: Skill, luck, or lack of power?” *The Journal of Finance* **74(5)**: pp. 2667–2688.
- BARRAS, L., O. SAILLET, & R. WERMERS (2010): “False discoveries in mutual fund performance: Measuring luck in estimated alphas.” *The journal of finance* **65(1)**: pp. 179–216.
- BARRAS, L., O. SAILLET, & R. WERMERS (2019): “Reassessing false discoveries in mutual fund performance: skill, luck, or lack of power? a reply.” *Journal of Finance, Forthcoming, Swiss Finance Institute Research Paper (19-61)*.
- BENJAMINI, Y. & Y. HOCHBERG (1995): “Controlling the false discovery rate: a practical and powerful approach to multiple testing.” *Journal of the Royal statistical society: series B (Methodological)* **57(1)**: pp. 289–300.
- BONFERRONI (1936): “Bonferroni, ce, 1936. teoria statistica delle classi e calcolo delle probabilita, pubblicazioni del r istituto superiore di scienze economiche e commerciali di firenze, 8: 3- 62. brenning, a., 2009. benchmarking classifiers to optimally integrate terrain analysis and multispectral remote sensing in automatic.” *Environment* **37(1)**: pp. 35–46.
- CAMPBELL, J. Y. & R. J. SHILLER (1988): “The dividend-price ratio and expectations of future dividends and discount factors.” *The Review of Financial Studies* **1(3)**: pp. 195–228.
- CARHART, M. M. (1997): “On persistence in mutual fund performance.” *The Journal of finance* **52(1)**: pp. 57–82.
- CAVALCANTE FILHO, E., F. MORAES, & R. DE-LOSSO (2021): “Unskilled

- fund managers: replicating active fund performance with few etfs.” *Available at SSRN 3634701* .
- COWLES, A. (1933): “Can stock market forecasters forecast?” *Econometrica* **1(3)**: pp. 309–324.
- FAMA, E. F. (1970): “Efficient capital markets: A review of theory and empirical work.” *The Journal of Finance* **25(2)**: pp. 383–417.
- FAMA, E. F. & K. R. FRENCH (1992): “The cross-section of expected stock returns.” *The Journal of Finance* **47(2)**: pp. 427–465.
- FAMA, E. F. & K. R. FRENCH (1993): “Common risk factors in the returns on stocks and bonds.” *Journal of Financial Economics* **33(1)**: pp. 3–56.
- FAMA, E. F. & K. R. FRENCH (2010): “Luck versus skill in the cross-section of mutual fund returns.” *The Journal of Finance* **65(5)**: pp. 1915–1947.
- HOCHBERG, Y. & A. C. TAMHANE (1987): *Multiple comparison procedures*. John Wiley & Sons, Inc.
- JEGADEESH, N. & S. TITMAN (1993): “Returns to buying winners and selling losers: Implications for stock market efficiency.” *The Journal of finance* **48(1)**: pp. 65–91.
- JENSEN, M. C. (1968): “The performance of mutual funds in the period 1945–1964.” *The Journal of finance* **23(2)**: pp. 389–416.
- KOSOWSKI, R., A. TIMMERMAN, R. WERMERS, & H. WHITE (2006): “Can mutual fund “stars” really pick stocks? new evidence from a bootstrap analysis.” *The Journal of finance* **61(6)**: pp. 2551–2595.
- LINTNER, J. (1965): “The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets.” *The Review of Economics and Statistics* **47(1)**: pp. 13–37.
- MARKOWITZ, H. (1952): “Portfolio selection.” *The Journal of Finance* **7(1)**: pp. 77–91.
- MCLEAN, R. D. & J. PONTIFF (2016): “Does academic research destroy stock return predictability?” *The Journal of Finance* **71(1)**: pp. 5–32.

- MOSSIN, J. (1966): "Equilibrium in a capital asset market." *Econometrica* **34(4)**: pp. 768–783.
- NEWKEY, W. K. & K. D. WEST (1987): "Hypothesis testing with efficient method of moments estimation." *International Economic Review* pp. 777–787.
- OECD (2021): *Pensions at a Glance 2021*.
- OTTEN, R. & D. BAMS (2002): "European mutual fund performance." *European financial management* **8(1)**: pp. 75–101.
- OTTEN, R. & K. THEVISSSEN (2011): "Does industry size matter? revisiting european mutual fund performance." *Revisiting European Mutual Fund Performance (January 15, 2011)* .
- ROLL, R. (1977): "A critique of the asset pricing theory's tests part i: On past and potential testability of the theory." *Journal of Financial Economics* **4(2)**: pp. 129–176.
- ROSS, S. A. (1976): "The arbitrage theory of capital asset pricing." *Journal of Economic Theory* **13(3)**: pp. 341–360.
- SHARPE, W. F. (1964): "Capital asset prices: A theory of market equilibrium under conditions of risk." *The journal of finance* **19(3)**: pp. 425–442.
- SHARPE, W. F. (1966): "Mutual fund performance." *The Journal of business* **39(1)**: pp. 119–138.
- SHARPE, W. F. (1991): "The arithmetic of active management." *Financial Analysts Journal* **47(1)**: pp. 7–9.
- SHARPE, W. F. (1994): "The sharpe ratio." *Journal of portfolio management* **21(1)**: pp. 49–58.
- ŠIDÁK, Z. (1967): "Rectangular confidence regions for the means of multivariate normal distributions." *Journal of the American Statistical Association* **62(318)**: pp. 626–633.
- TREYNOR, J. (1965): "How to rate management of investment funds." .
- WELCH, I. & A. GOYAL (2008): "A comprehensive look at the empirical performance of equity premium prediction." *The Review of Financial Studies* **21(4)**: pp. 1455–1508.