UNIVERSITY OF WARSAW

FACULTY OF ECONOMIC SCIENCES

Working Papers

No. 17/2023 (424)

OPTIMAL MARKOWITZ PORTFOLIO USING RETURNS FORECASTED WITH TIME SERIES AND MACHINE LEARNING MODELS

Damian Ślusarczyk

Robert Ślepaczuk

Warsaw 2023

UNIVERSITY OF WARSAW
FACULTY OF ECONOMIC SCIENCES
WORKING PAPERS

Optimal Markowitz Portfolio Using Returns Forecasted with Time Series and Machine Learning Models

Damian Ślusarczyk[a], Robert Ślepaczuk[b]

[a] *University of Warsaw, Faculty of Economic Sciences*
[b] *University of Warsaw, Quantitative Finance Research Group, Department of Quantitative Finance, Faculty of Economic Sciences*
Corresponding author: rslepaczuk@wne.uw.edu.pl

Abstract: We aim to answer the question of whether using forecasted stock returns based on machine learning and time series models in a mean-variance portfolio framework yields better results than relying on historical returns. Nevertheless, the problem of the efficient stock selection has been tested for more than 50 years, the issue of adequate construction of mean-variance portfolio framework and incorporating forecasts of returns in it has not been solved yet. Stock returns portfolios were created using 'raw' historical returns and forecasted return based on ARIMA-GARCH and the XGBoost models. Two optimization problems were concerned: global maximum information ratio and global mini-mum variance. Then strategies were compared with two benchmarks – an equally weighted portfolio and buy and hold on the DJIA index. Strategies were tested on Dow Jones Industrial Average stocks in the period from 2007-01-01 to 2022-12-31 and daily data was used. The main portfolio performance metrics were information ratio* and information ratio**. The results showed that using forecasted returns we can enhance our portfolio selection based on Markowitz framework, but it is not a universal solution, and we have to control all the parameters and hyperparameters of selected models.

Keywords: Algorithmic Investment Strategies, Markowitz framework, portfolio optimization, forecasting, ARIMA, GARCH, XGBoost, minimum variance

JEL codes: C4, C14, C45, C53, C58, G13

Note: This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

INTRODUCTION

Information is a crucial factor in trading. In order to ensure successful investment outcomes, investors need to have good knowledge of interdependencies in financial markets and the main factors affecting the prices of securities. But not only the quantity and the quality of information is important but also how one uses it. In this study, we limit the information we use to create investment strategies solely to historical stock prices. At the same time, we check if using this data as an input for both classic time-series models and modern machine learning models yields better investment outcomes.

The main framework in the study is the mean-variance portfolio framework first introduced by Henry Markowitz (1952). We verify the efficiency of the enhanced Markowitz's framework by testing three groups of strategies against simple benchmarks – an equally weighted portfolio and buy and hold on the DJIA index. These groups are strategies using 'raw' historical returns, strategies using the ARIMA-GARCH forecasts, and strategies using the XGBoost forecasts. Comparing these three groups with one another and with the benchmarks, we get some insights about, whether it is beneficial to further transform the return series (in this case, to forecast future returns).

The main aim of this study is synthesised in two hypotheses:
H1: *The strategies based on forecasted stock returns outperform (having higher values of information ratio**) the strategies based on historical stock returns.*
H2: *The strategies based on forecasted stock returns outperform (having higher values of information ratio**) an equally weighted portfolio and buy and hold on the equity index.*
Moreover, we ask the following research questions:
RQ1: *Which portfolio optimization method will perform better in terms of information ratio**?*
RQ2: *Which forecasting model performs better in the framework in terms of information ratio**?*
RQ3: *Are the results sensitive to the number of assets, the estimation and the rebalancing windows' lengths, the transaction costs, and the forecasting models' parameters?*

To verify hypotheses and answer research questions, we conducted an empirical study simulating the results of many different strategies in the period from January 1, 2007, to December 31, 2022. We used nearly all stocks that have been components of the Dow Jones Industrial Average during the discussed period. All of the calculations were performed using R and Python. The time needed to perform all the computations is estimated to be about 32 hours and over 200 000 models were fitted. The ARIMA-GARCH forecasts were performed using the `rugarch` R package (Ghalanos, 2022), and the XGBoost forecasts were performed using `XGBoost` R package (Chen and Guestrin, 2016). Portfolios' objective functions were optimized in Python using `scipy` package.

Nowadays huge amount of data is available. It is easy to fall into the trap of using much information without maximizing the utility of it. In our study, we limit the amount of information we use and create investment strategies attempting to use the restricted dataset most efficiently. Our main contribution to the portfolio optimization topic consists of presenting an approach that does not require expensive or difficult to obtain data and makes the most use only of historical stock returns. Moreover, we check whether utilizing the proposed approach can be beneficial compared to classical mean-variance portfolio, equally weighted portfolio and buy and hold on an equity index.

The structure of the thesis is as follows. In the first chapter, we present the theoretical background and describe recent studies concerning the discussed topics. The first chapter is divided into three parts: a part explaining the mean-variance framework with its modern applications, a part concerning the ARIMA-GARCH model, and a part describing the XGBoost model. In the second chapter, we describe the methodology, utilized strategies and data. In the third chapter, we show and describe the results of the base case scenario. In the last chapter, we perform an extensive sensitivity analysis of the results we obtained. In the end, we derive conclusions and discuss the results.

I. THEORETICAL BACKGROUND AND LITERATURE REVIEW

1.1. Markowitz model

In 1952 Markowitz laid the grounds for the Modern Portfolio Theory presenting an approach to construct a portfolio of different assets. Over 70 years later, his ideas are still alive in portfolio management. Many researchers have extended the theory, but its conceptual basics remained unchanged. Regardless, the most important problem with the classic mean-variance framework is that it relies on historical stock prices. They may be a good indicator of stock performance only when the stock is proven to perform consistently under various political and economic circumstances (Fabozzi et al., 2002), which is hardly the case. That is why researchers tried different approaches relying on stock price forecasting.

For example, Chen et al. (2021) proposed a strategy consisting of two components: stock prediction using XGBoost with IFA (Improved Firefly Algorithm) and portfolio optimization using the mean-variance framework. First, the authors forecast stock returns using XGBoost and optimize its hyperparameters using IFA, then they choose best-performing stocks and apply a mean-variance framework to allocate weights for the assets. The authors compared the proposed approach to a number of benchmarks, including the traditional mean-variance portfolio (based on historical returns), other machine learning models combined with a mean-variance portfolio or a number of equally weighted portfolios. Chen et al. tested this approach on 24 randomly selected Shanghai Stock Exchange 50 (SSE50) stocks, and their proposed strategy outperformed the others between Q4 2017 and Q4 2019. It is worth noting that the forecasts have been performed only for the one period ahead.

Ma et al. (2021) forecasted one day ahead stock returns using machine learning and traditional forecasting models and then used forecasts to form mean-variance and omega ratio portfolios. Used models include DMLP (deep multilayer perceptron), LSTM (long short term memory) neural network, CNN (convolutional neural network), random forest, SVR (support vector regression) and traditional ARIMA. Each forecasting model was combined with both mean-variance and omega ratio portfolio approaches. The strategies were tested on 49 China Securities 100 index components. The results showed that between January 2012 and December 2015, both approaches utilizing random forests outperformed other strategies. Due to daily rebalancing and high turnover rates, the best strategies' returns were decreased by half by transaction fees. Despite that, the authors recommended using the proposed approach with random forests for portfolio creation.

It should not be shocking that even in recent years, an over-70-year-old approach was used to optimize the portfolio. High returns and low risk are one of the most crucial factors for investors. The approach proposed by Markowitz (1952) combined with different optimization problems can create a portfolio with a high chance of fulfilling investors' expectations. It is worth noting that higher moments of return distributions also can play a significant role in the portfolio selection. Harvey et al. (2010) and Chen and Zhou (2018) showed that incorporating skewness and kurtosis in portfolio optimization can improve the results.

1.2. ARIMA-GARCH

Box and Jenkins (1976) introduced their model for time-series data called ARIMA – **A**uto-**R**egressive **I**ntegrated **M**oving **A**verage. In this model variable X in period t depends on previous values of X and previous values of residuals ε. The main problem of such a model is the assumption of homoscedasticity. Unfortunately, for financial time series, it is not often the case. They tend to be heteroskedastic, and periods of higher variance tend to cluster. This problem were addressed by Engle (1982). In his study, he presented the ARCH model – **A**uto-**R**egressive **C**onditional **H**eteroskedasticity. Moreover, he proposed a test, to check whether time series show ARCH effects. Due to the fact, that the ARCH model tends to have relatively high orders which require estimating multiple parameters, Bollerslev (1986) proposed the GARCH model – **G**eneralized **A**uto-**R**egressive **C**onditional **H**eteroskedasticity. It builds on the classic ARCH model adding lagged variances. GARCH models tend to be more parsimonious. When we use the ARIMA model as the mean equation with a GARCH model we obtain the ensemble ARIMA-GARCH model with the final equational form of:

$$(1-\sum_{i=1}^{p}\varphi_i L^i)(1-L)^d X_t = \mu + (1+\sum_{i=1}^{q}\theta_i L^i)u_t, \quad u_t \sim N(0, \sigma_t^2) \tag{1}$$

$$(1-\sum_{j=1}^{b}\beta_i L^i)\sigma_t^2 = \alpha_0 + (\sum_{i=1}^{a}\alpha_i L^i)u_t^2 \tag{2}$$

The ARIMA-GARCH model was widely applied in forecasting realizations of different time series. Mohammadi and Su (2010) analysed the performance of several variants of ARIMA-GARCH models to forecast crude oil prices. Four volatility models were tested – GARCH, EGARCH, APARCH, and FIGARCH. The results were not conclusive, but the APARCH model tended to perform better than the others in most cases. Yaziz et al. (2013) utilized the model to forecast gold prices on a 40-day window and compared it to other forecasting approaches. The result showed that ARIMA$(1,1,1)$-GARCH$(0,2)$ forecasts outperformed other compared methods (including standard ARIMA) in terms of MAE (mean absolute error) and MSE (mean squared error).

When it comes to forecasting stock prices, several pieces of research were conducted. For example, Mustapa and Ismail (2019) used the ARIMA-GARCH model to forecast the monthly prices of S&P500 index stocks in the period from January 2001 to December 2018. The authors did not fit the model simultaneously but rather fitted the ARIMA model and then applied GARCH on the mean model residuals. The results of the study showed that ARIMA$(2,1,2)$-GARCH$(1,1)$ model performed the best in forecasting the S&P500 index stock prices. What is interesting, most other ARIMA orders tested by the authors turned out to be insignificant at the 5% level. Vo and Ślepaczuk (2022) constructed an algorithmic investment strategy applying the ARIMA-(S)GARCH on S&P500 stocks. The authors found that adding the GARCH component to the ARIMA model increases the precision of the forecasts and causes underlying investment strategies to outperform benchmarks. Moreover, their findings seemed not to be sensitive to a variety of parameters, including error distribution or GARCH model type (SGARCH, EGARCH). Based on the literature, we can conclude that the addition of GARCH components can potentially increase the precision of the forecasts.

1.3. XGBoost

XGBoost is a modern machine-learning model ensembling many decision trees using the technique called gradient boosting. It is a versatile approach that has many applications (Chen and Guestrin, 2016). The model is trained on a set of observations including features and response variable and then it can be used on out-of-sample data to predict values of response variable based on the values of feature variables. The response variable can be either numerical or categorical. Many studies were conducted utilizing the model in financial applications, the most recent ones are going to be described in the following section.

Nobre and Neves (2019) applied the XGBoost binary classifier to predict the direction of stock price movement to make trading decisions. The authors used processed financial data and technical indicators as features. The data processing included normalization, PCA (Principal Component Analysis), and wavelet transform. Then the XGBoost model hyperparameters were optimized using multi-objective optimization, and a trading signal was received from the optimized model. The authors tested the strategy on different financial markets including future contracts, stocks, and the SP500 index. The results showed that the proposed strategy was able to outperform buy and hold in a tested period (2014-2017) in three out of five tested markets.

In their study, Nabipour et al. (2020) compared machine-learning and deep-learning models' performance when predicting the Iranian stock market groups' performance. The authors investigated both tree-based and artificial neural network (ANN) models. As the model features, the authors used different technical indicators, including moving averages and oscillators. The predictions were performed up to 30 days in advance. The results showed that XGBoost performed the best out of all tree-based models for one-day ahead predictions in diversified stock groups but still worse than most ANN models. As the forecast period extended, the performance of XGBoost was no longer the best among tree-based models. On the other hand, the average performance of XGBoost was still competitive among tree-based models concerning stocks divided in groups by industry.

Jabeur et al. (2021) attempted to predict gold prices using other financial time series as features. They include other metals' prices (silver, iron ore), currency exchange rates (USD/EUR, USD/CHY), oil prices, S&P500 index and inflation rate in the US. The authors benchmarked XGBoost against other models, including classical linear regression and machine-learning models like neural networks, random forests, light gradient boosting (LightGBM) or the CatBoost algorithm. Moreover, the authors used SHAP (Shapley additive explanations) to help interpret models' predictions. The results of the empirical study showed that XGBoost performs the best out of tested models (in terms of RMSE, MSE and MAE) and, used with SHAP, can be a successful tool to predict gold prices. Jabeur et al. concludes that policymakers could benefit from using the proposed approach.

To assess the importance of feature selection, Yun et al. (2021) proposed a novel approach to stock performance forecasting with XGBoost using three stage process. The first stage consisted of generating technical indicators of historical stock prices and volume and normalizing the data. Next, the authors utilized GA (genetic algorithm) to reduce the dimensionality and prevent overfitting by choosing features with the highest chance of positively contributing to model performance. Finally, they performed forecasts to predict up or down stock price movement, thus reducing the prediction problem to binary classification. Yun et al. assessed the model using accuracy metric, F1 or AUROC (area under the ROC) among others. The empirical study based on the Korea composite stock price index 200 (KOSPI) showed that the proposed feature engineering approach influences the results greatly. The authors also showed that it is possible to reduce dimensionality without affecting performance.

Recent literature findings show that XGBoost is a powerful tool in financial time series performance prediction. Studies put much emphasis on the feature selection stage of the forecasting project and show that despite its flexibility and adaptability XGBoost is not a panacea for forecasting problems.

II. METHODOLOGY AND DATA

A total of 152 strategies were tested in the base case scenario and the sensitivity analysis. The strategies may be divided into three groups:

- MV – mean-variance portfolio based on historical returns,
- FMV-AG – mean-variance portfolio based on returns forecasted using the ARIMA-GARCH model,
- FMV-XGB – mean-variance portfolio based on returns forecasted using the XGBoost model.

Each strategy can be described by a set of parameters that are presented in Table 1 (values for the base case scenario are marked in **bold**). All of the strategies groups are going to be described later in this chapter.

As benchmarks, we decided to use:

- EW – the equally weighted portfolio,
- DJIA – the ”buy and hold” the Dow Jones Industrial Average index.

Table 1. Strategies' parameters.

Parameter	Values	Applies to
Estimation window (in months)	$t_{ew} \in \{6;\ \mathbf{12};\ 24\}$	MV, FMV-AG, FMV-XGB
Portfolio rebalancing period (in months)	$f \in \{0{,}5;\ \mathbf{1};\ 3\}$	MV, FMV-AG, FMV-XGB
Number of stocks	$n \in \{5; 10;\ \mathbf{30}\}$	MV, FMV-AG, FMV-XGB
Transaction cost level (in %)	$c \in \{0{,}05; \mathbf{0{,}1};\ 0{,}2;\ 0{,}5\}$	MV, FMV-AG, FMV-XGB
ARIMA order	$(p, d, q) \in \{(1, 0, 0);\ (0, 0, 1); (1, 0, 1);$ $\mathbf{(2, 0, 1)}; (1, 0, 2);\ (2, 0, 2)\}$	FMV-AG
GARCH error distribution	**normal**, skew-normal, student-t, skew-student, GED, skew-GED	FMV-AG
XGBoost feature lag	$l \in \{6;\ \mathbf{11};\ 21\}$	FMV-XGB
XGBoost hyperparameters		
Numer of rounds	$\texttt{nrounds} \in \{50;\ \mathbf{100};\ 200\}$	FMV-XGB
Learning rate	$\texttt{eta} \in \{0{,}05;\ 0{,}1;\ \mathbf{0.3};\ 0{,}4\}$	FMV-XGB
Lambda	$\texttt{lambda} \in \{0{,}5;\ \mathbf{1};\ 5;\ 10\}$	FMV-XGB
Max depth	$\texttt{max\_depth} \in \{3;\ \mathbf{6};\ 9;\ 12\}$	FMV-XGB
Subsample	$\texttt{subsample} \in \{0{,}5;\ 0{,}75;\ \mathbf{1}\}$	FMV-XGB

Note: Values for the base case scenario are marked in **bold**, other parameters are tested in the sensitivity analysis. MV – mean-variance portfolio based on historical returns, FMV-AG – mean-variance portfolio based on returns forecasted using the ARIMA-GARCH model, FMV-XGB – mean-variance portfolio based on returns forecasted using the XGBoost model.

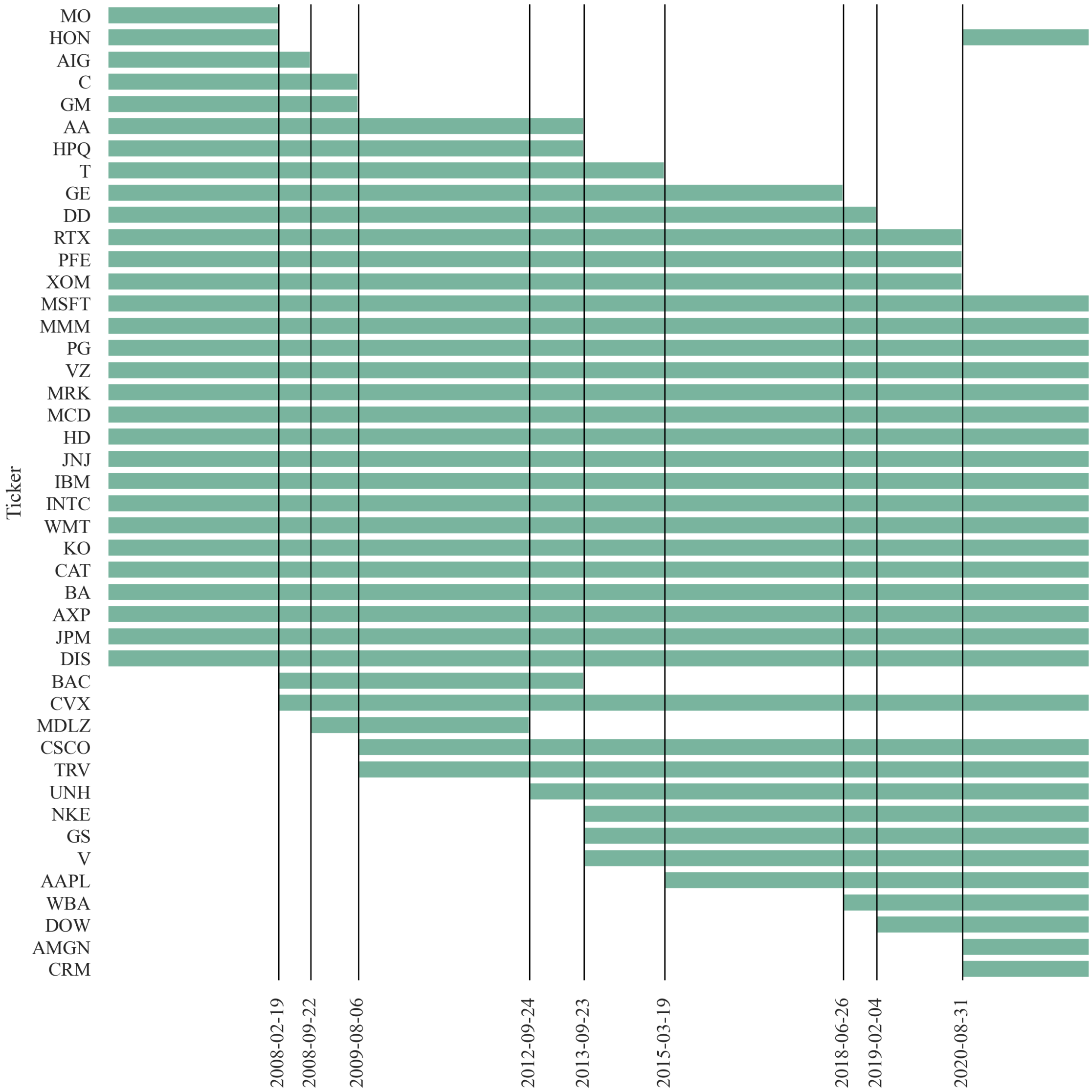
2.1. Data description

The strategies were tested on the Dow Jones Industrial Average stocks. We have decided to use DJIA stocks because their number is relatively small, allowing us to track changes in the index constituents over our testing period. Moreover, data about their historical returns is easily available. This equity index can also be assumed to be representative of the market. In the base case scenario and most of the sensitivity analysis, we used all DJIA components. It is worth noting that because of the financial problems of the "old" General Motors in 2009, a new company – "new" General Motors was established and emitted stock under the same ticker. Because of that, we found it impossible to obtain stock price data preceding the company reorganization. That is why we have decided to omit GM stock in the analysis, and we are aware of the fact that it may lead to a slight survivorship bias before June 2009.

There were some changes to the Dow Jones Industrial Average composition during the research period. All of them are presented in Figure 1. For each rebalancing period we chose the stocks that had been the components of the DJIA on the day before the first day of the rebalancing period. If any stock ceased to be a part of DJIA during the rebalancing period, it would not be excluded until the beginning of the new rebalancing period.

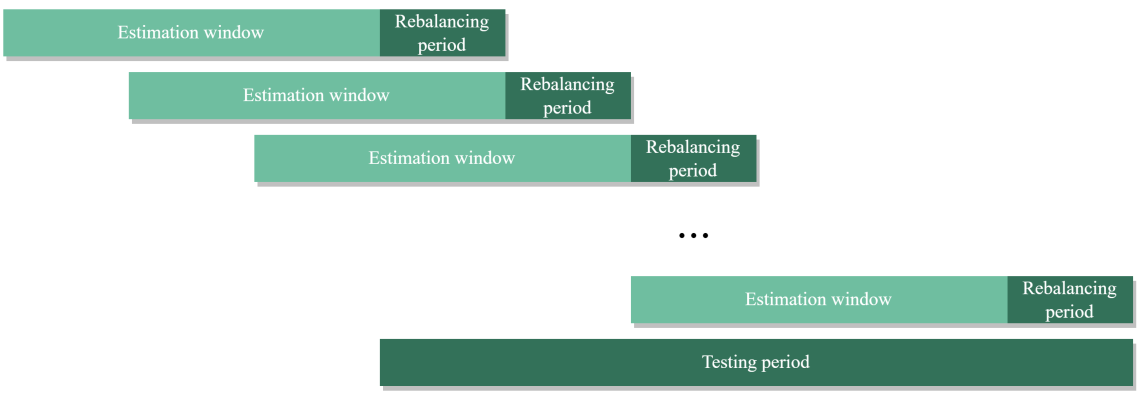
The testing period begins 2007-01-01 and ends 2022-12-31, which gives 16 years. The earliest data used in the study comes from the 2005-01-01. It is used in sensitivity analysis for a 2-year estimation window. All of the financial data used in the study comes from Yahoo! Finance. To better understand how the estimation window, rebalancing period, and testing period are related, it is beneficial to look at their graphical representation on Figure 2.

Figure 1. Components of Dow Jones Industrial Average from 2005-01-01 to 2022-12-31.



Note: Own preparation based on the list of changes in the DJIA from https://en.wikipedia.org/

Figure 2. Graphic representation of estimation window, rebalancing period and testing period.



Note: For the base case scenario first estimation windows started on 01/01/2006 and ended on 31/12/2006 (estimation window $T_{ew} = 12$ months), and the first rebalancing period started on 01/01/2007 and ended on 31/02/2007 (rebalancing window $f = 1$). Then both periods were recursively shifted by one month ($f = 1$). Summing all estimation windows gives us the testing period starting on 2007-01-01 and ending on 2022-12-31.

2.2. Mean-variance portfolio framework

Let $\boldsymbol{r}^*$ be a matrix of T historical daily log returns of n stocks:

$$\boldsymbol{r} = \begin{bmatrix} r_{t-1,1} & \cdots & r_{t-1,n} \\ \vdots & \ddots & \vdots \\ r_{t-T,1} & \cdots & r_{t-T,n} \end{bmatrix} \tag{3}$$

where:
$T \approx 21 \cdot t_{ew}$,
t_{ew} is the length of estimation window in months,
n is the number of stocks,
$r_{t-i,n}$ is the historical log return for n-th stock for day $t-i$ for $i \in \{1,\ldots,T\}$.

Then $\boldsymbol{\mu}^*$ will be a vector of means of historical log returns:

$$\hat{\boldsymbol{\mu}} = \begin{bmatrix} \hat{\mu_1} \\ \vdots \\ \hat{\mu_n} \end{bmatrix} = \begin{bmatrix} \frac{1}{T}\sum_{i=1}^{T} r_{t-i,1} \\ \vdots \\ \frac{1}{T}\sum_{i=1}^{T} r_{t-i,n} \end{bmatrix} \tag{4}$$

and $\hat{\boldsymbol{\Sigma}}$ a variance-covariance matrix of historical log returns:

$$\hat{\boldsymbol{\Sigma}} = \begin{bmatrix} \hat{\sigma}_1^2 & \cdots & \hat{\sigma}_{1,n} \\ \vdots & \ddots & \vdots \\ \hat{\sigma}_{1,n} & \cdots & \hat{\sigma}_n^2 \end{bmatrix} \tag{5}$$

where $\hat{\sigma}_i^2$ is an empirical variance of historical log returns of i-th stock and $\hat{\sigma}_{i,j}$ is an empirical covariance between historical log returns of i-th and j-th stock.

The goal is to find a vector of weights $\boldsymbol{w}$ which optimizes the objective function. Two portfolio optimization problems are considered, a global maximum information ratio portfolio and a global minimum variance portfolio.

Global maximum information ratio portfolio (GMIR)

The information ratio is a measure of trade-off between returns and risk. A GMIR portfolio is equivalent to the global maximum Sharpe ratio portfolio with the risk-free rate equal to 0. The optimization problem is:

$$\max_{\boldsymbol{w}} \frac{\boldsymbol{w}'\hat{\boldsymbol{\mu}}}{\boldsymbol{w}'\hat{\boldsymbol{\Sigma}}\boldsymbol{w}} \qquad s.t. \sum_{i=1}^{30} w_i = 1 \quad \wedge \quad \forall i : w_i \in [0,1] \tag{6}$$

Global minimum variance portfolio (GMV)

A GMV portfolio should minimize risk. The objective function does not take the returns into account. The optimization problem is:

$$\min_{\boldsymbol{w}} \boldsymbol{w}'\hat{\boldsymbol{\Sigma}}\boldsymbol{w} \qquad s.t. \sum_{i=1}^{30} w_i = 1 \quad \wedge \quad \forall i : w_i \in [0,1] \tag{7}$$

Please note that we do not allow for short selling as all weights must be positive. In this case (as well as in the mean-variance framework with forecasted stock returns) in period t resulting portfolios have gross returns of:

$$r_t^{gross} = \boldsymbol{w}'\boldsymbol{r}^{gross} = \begin{bmatrix} w_{t,1} & \cdots & w_{t,n} \end{bmatrix} \cdot \begin{bmatrix} r_{t,1}^{gross} \\ \vdots \\ r_{t,n}^{gross} \end{bmatrix} = \sum_{i=1}^{n} w_{t,n} \cdot r_{t,n}^{gross} \tag{8}$$

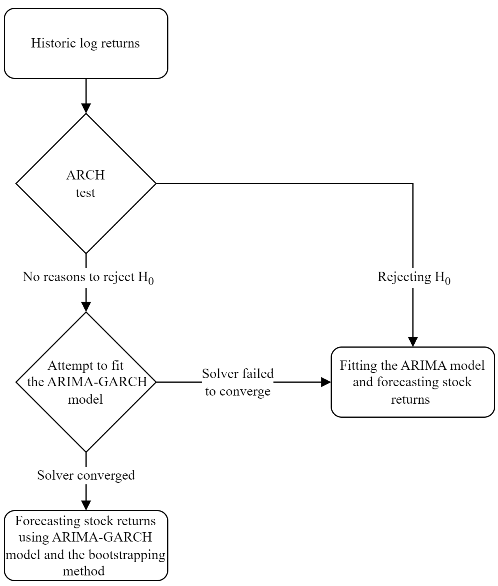
where $r_{t,n}^{gross}$ is the gross return on n-th asset on day t.

2.3. Stock return forecasting methods

ARIMA-GARCH

In the base case scenario, the ARIMA-GARCH model had a fixed order. This decision was made to decrease the computation time. ARIMA part had an order of $(p, d, q) = (2, 0, 1)$ which was determined based on the results from financial literature. On the other hand, the GARCH order of $(a, b) = (1, 1)$ was chosen because Engle (2001) claimed that it performed the best. In sensitivity analysis, other ARIMA orders were tested. We have not decided to difference the log returns (thus $d = 0$ for each case), as they are usually already stationary (contrary to stock prices). The algorithm for forecasting single returns for a single stock is presented in Figure 3.

Figure 3. Algorithm for forecasting stock returns.



Note: The algorithm is applied to all stocks in all periods.

For each stock in a given rebalancing period, the model was fitted on historical returns from 1 up to $T = 21 \cdot t_{ew}$ days before the first day of the rebalancing period (where t_{ew} is the length of estimation window in months). The forecasts were made for $T_f = \max(21 \cdot f, n)$ days ahead, where f is the length of the rebalancing period in months and n is the number of stocks. Both the mean and the variance equation were fitted simultaneously.

Forecasts are made using the bootstrapping method described by Pascual et al. (2006). It allows to fully benefit from GARCH component when forecasting next series realizations. The method involves drawing (with replacement) values from the empirical distribution of residuals and applying them to forecasted values. Then, the set of B bootstrap replications $(y^{*(1)}_{T+k}, \ldots, y^{*(B)}_{T+k})$, for the predicted value y_{T+k} is obtained, giving approximation of y_{T+k} density function. A similar approach is used for volatilities. The empirical density functions enable the construction of prediction intervals. Their means can be used as forecasts for future series realizations. It is worth noting that the series forecasts quite quickly approach the mean.

XGBoost

Since XGBoost is a non-parametric method for time series forecasting it is crucial to engineer features that will maximize the chances for accurate forecasts. In this case, the only features were the lagged returns. The number of lags l was equal to 11 (about half the number of trading days in a month) in the base case scenario. So if label were r_t then features were $r_{t-1}, \ldots, r_{t-11}$. In the sensitivity analysis, other numbers of lags (6 and 21) were tested. Various types of feature combinations were tried. What is interesting, including basic technical indicators or variables describing days of week or days of month actually made the performance of the strategy worse. That is why we have decided to use only lagged returns.

Forecasts for T_f days ahead were performed recursively. That means that the forecast for r_{t+1} was based only on historical values but for r_{t+2} on one forecasted return r^*_{t+1}, for r_{t+3} on two forecasted returns, and so on. A consequence of that is the fact that the forecast of r_{t+12} and further forecasts were performed exclusively using priorly forecasted data.

Similarly to ARIMA-GARCH, XGBoost hyperparameters were fixed. The best approach would be to grid search or random search optimal hyperparameter combination, but due to the big number of fitted models (5730 only in FMV-XGB base case scenario), significant computing power would be needed in order to finish tuning in a reasonable time.

2.4. Mean-variance portfolio framework with forecasted stock prices

Let $\boldsymbol{r}^*$ be a matrix of n forecasted stock log returns for T_f days ahead which is a result of forecasts described above:

$$\boldsymbol{r}^* = \begin{bmatrix} r^*_{t+1,1} & \cdots & r^*_{t+1,n} \\ \vdots & \ddots & \vdots \\ r^*_{t+T_f,1} & \cdots & r^*_{t+T_f,n} \end{bmatrix} \tag{9}$$

where:
$T_f = \max(21 \cdot f, n)$,
f is the length of rebalancing period in months,
n is the number of stocks,
$r^*_{t+i,n}$ is the forecasted log return for n-th stock for day $t+i$ for $i \in \{1, \ldots, T_f\}$.

Then $\boldsymbol{\mu}^*$ will be a vector of means of forecasted log returns:

$$\boldsymbol{\mu}^* = \begin{bmatrix} \mu_1^* \\ \vdots \\ \mu_n^* \end{bmatrix} = \begin{bmatrix} \frac{1}{T_f} \sum_{i=1}^{T_f} r_{t+i,1}^* \\ \vdots \\ \frac{1}{T_f} \sum_{i=1}^{T_f} r_{t+i,n}^* \end{bmatrix} \tag{10}$$

and $\boldsymbol{\Sigma}^*$ a variance-covariance matrix of forecasted log returns:

$$\boldsymbol{\Sigma}^* = \begin{bmatrix} \sigma_1^{2*} & \cdots & \sigma_{1,n}^* \\ \vdots & \ddots & \vdots \\ \sigma_{1,n}^* & \cdots & \sigma_n^{2*} \end{bmatrix} \tag{11}$$

where σ_i^{2*} is the variance of forecasted log returns of i-th stock and $\sigma_{i,j}^*$ is the covariance between log returns of i-th and j-th stock.

Similarly as in basic mean-variance portfolio framework based on historical log returns, the goal is to find a vector $\boldsymbol{w}$ to optimize the objective function.

Global maximum information ratio portfolio

The optimization problem is:

$$\max_{\boldsymbol{w}} \frac{\boldsymbol{w}'\boldsymbol{\mu}^*}{\boldsymbol{w}'\boldsymbol{\Sigma}^*\boldsymbol{w}} \qquad s.t. \sum_{i=1}^{30} w_i = 1 \quad \wedge \quad \forall i : w_i \in [0,1] \tag{12}$$

Global minimum variance portfolio

The optimization problem is:

$$\min_{\boldsymbol{w}} \boldsymbol{w}'\boldsymbol{\Sigma}^*\boldsymbol{w} \qquad s.t. \sum_{i=1}^{30} w_i = 1 \quad \wedge \quad \forall i : w_i \in [0,1] \tag{13}$$

2.5. Performance and diversification metrics

All metrics were calculated using either simple net returns (after transaction costs) or the equity line calculated using the given formula:

$$X(t) = K \cdot \prod_{t=1}^{T} (1 + r_t^{gross} - \sum_{n=1}^{N} |w_{t,n} - w_{t-1,n}| \cdot c) = K \cdot \prod_{t=1}^{T} (1 + r_t) \tag{14}$$

where:
$X(t)$ – the portfolio value at moment t,
K – the invested capital (here $K = 1$),
T – the number of trading days in the testing period,
r_t^{gross} – the portfolio's gross returns on day t,
N – the number of stock in the optimized portfolio,
$w_{t,n}$ – the weight of n-th asset on day t ($\forall\ n : w_{0,n} = 0$),
c – the percentage level of transaction costs,

r_t – the portfolio's net returns on period t.

Please note that we used log returns for forecasting, estimation and optimization due to their better statistical properties, but to measure performance, we chose regular returns for proper interpretation of real changes in the value of our portfolio. Performance and diversification metrics were based on Tsay (2005), Vo and Ślepaczuk (2022) and are described below.

Absolute rate of return

$$ARR\% = (\prod_{t=1}^{T}(1 + r_t) - 1) \cdot 100\% \tag{15}$$

where:
T – the number of trading days in the testing period,
r_t – the portfolio's net returns on day t.

Annualized rate of return

$$ARC\% = ((\prod_{t=1}^{T}(1 + r_t))^{\frac{1}{Y}} - 1) \cdot 100\% \tag{16}$$

where:
T – the number of trading days in the testing period,
r_t – the portfolio's net returns on day t,
Y – the length of testing period in years.

Annualized standard deviation

$$ASD\% = \sqrt{\frac{252}{T}\sum_{t=1}^{T}(r_t - \bar{r})^2 \cdot 100\%} \tag{17}$$

where:
$\bar{r} = \frac{1}{T} \cdot \sum_{t=1}^{T} r_t$,
T – the number of trading days in the testing period,
r_t – the portfolio's net returns on day t.

Maximum drawdown

$$MDD\% = \max_{\tau \in (0,T)} [\max_{t \in (\tau, T)} \frac{X(\tau) - X(t)}{X(\tau)}] \cdot 100\% \tag{18}$$

where:
$X(t)$ – portfolio value on day t,
T – the number of trading days in the testing period.

Information ratio*

A simplified version of the Sharpe Ratio not including a risk-free rate. The measure of pay-off between returns and volatility.

$$IR^* = \frac{ARC\%}{ASD\%} \tag{19}$$

where $ARC\%$ and $ASD\%$ defined as above.

Information ratio**

A modified version of information ratio* including the maximum drawdown.

$$IR^{**} = \frac{ARC\% \cdot |ARC\%|}{ASD\% \cdot MDD\%} \tag{20}$$

where $ARC\%$, $ASD\%$, and $MDD\%$ defined as above.

While we used information ratio* as an auxiliary metric, information ratio** was the main performance metric in this study used for results interpretation, as it synthesises information obtained from all other previously mentioned metrics.

Mean number of stocks constituting at least 75% of the portfolio

A measure of the diversification of the portfolio. $MN^{75\%} \in [1; 23]$. Lower values suggest that the portfolio is concentrated only on a few assets, while greater values imply more diversification. For the equally weighted portfolio with 30 assets, its value is equal to 23.

Let $w_{t,1}, ..., w_{t,30}$ be portfolio weights of stocks on period t so that $w_{t,1} \geq ... \geq w_{t,30}$. A number of stocks constituting at least 75% of the portfolio on period t ($N_t^{75\%}$) can be calculated using following algorithm:

if $w_{t,1} \geq 0.75$ **then** $N_t^{75\%} = 1$
else if $w_{t,1} + w_{t,2} \geq 0.75$ **then** $N_t^{75\%} = 2$
else if $w_{t,1} + w_{t,2} + w_{t,3} \geq 0.75$ **then** $N_t^{75\%} = 3$
and so on ...

Then, the mean number of stocks constituting at least 75% of the portfolio is defined as:

$$MN^{75\%} = \frac{1}{T_p} \cdot \sum_{i=1}^{T_p} N_t^{75\%} \tag{21}$$

where:
T_p – the number of portfolio rebalancing periods.

Mean number of stocks constituting at least 90% of the portfolio

$MN^{90\%}$ is defined analogically to $MN^{75\%}$, but with the 90% threshold. $MN^{90\%} \in [1; 27]$. For the equally weighted portfolio with 30 assets, its value is equal to 27.

III. RESULTS

3.1. Global maximum information ratio portfolio

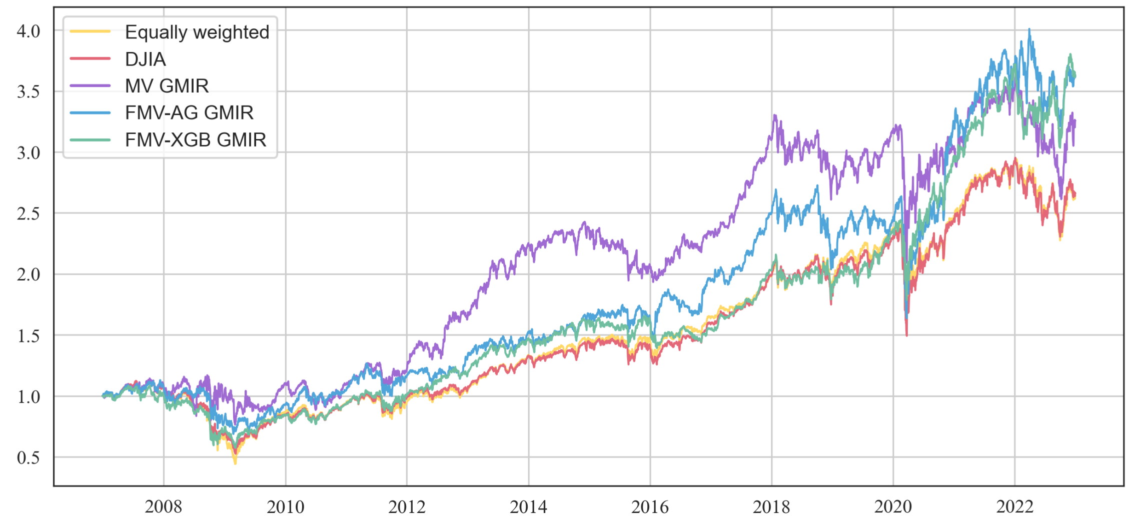
Analysing Table 3 and Figure 4, we can see that for GMIR portfolio both FMV strategies performed better than MV and benchmarks in terms of information ratio**. They had similar returns, but FMV-XGB performed better in terms of standard deviation, while FMV-AG had a lower maximum drawdown. MV also outperformed benchmarks in terms of information ratio**, but less than strategies with forecasted returns. We can see that all strategies promoted portfolios with a relatively low level of diversification, mostly focusing on less than 6 stocks.

Table 3. Performance statistics of the base GMIR strategies.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV GMIR	225,75%	7,66%	24,96%	**36,72%**	0,307	0,064	3,95	5,19
FMV-AG GMIR	261,62%	8,37%	22,56%	40,02%	0,371	0,077	3,66	6,35
FMV-XGB GMIR	**262,63%**	**8,38%**	**19,02%**	46,97%	**0,441**	**0,079**	5,97	8,68

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – …90% of the portfolio. MV GMIR – mean-variance global maximum information ratio portfolio based on historical returns, FMV-AG GMIR – mean-variance global maximum information ratio portfolio based on returns forecasted using the ARIMA-GARCH model, FMV-XGB GMIR – mean-variance global maximum information ratio portfolio based on returns forecasted using the XGBoost model.

Figure 4. Equity lines of the base GMIR strategies.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. MV GMIR – mean-variance global maximum information ratio portfolio based on historical returns, FMV-AG GMIR – mean-variance global maximum information ratio portfolio based on returns forecasted using the ARIMA-GARCH model, FMV-XGB GMIR – mean-variance global maximum information ratio portfolio based on returns forecasted using the XGBoost model.

Although FMV-AG, FMV-XGB and MV strategies outperformed benchmarks regarding information ratio**, we cannot infer that these strategies' returns distribution expected values are higher than the benchmarks' expected values. That is why we decided to test it using statistical inference. We performed two types of tests. The first test is t-test for paired samples (Devore et al., 2012) with the following hypotheses:
$H_0 : \mu_D = \mu_{strategy} - \mu_{benchmark} = 0$
$H_1 : \mu_D > 0$
Where μ_D is the difference between the expected values of strategy returns and benchmark returns. The p-values of the paired t-tests are presented in Table 4.

Table 4. P-values for the paired t-test for the base GMIR strategies.

		$\mu_{benchmark}$		
		EW	DJIA	MV GMIR
$\mu_{strategy}$	FMV-AG GMIR	0,220	0,179	0,491
	FMV-XGB GMIR	0,222	0,160	0,563
	MV GMIR	0,255	0,238	-

Note: P-values lower than 0.1 are marked in **bold**.

We can notice that the differences between strategies' expected values were not significant at the 0.1 significance level for any strategy pair. We also performed simple linear regression in the form of:

$$r_{strategy,t} = \alpha + \beta \cdot r_{benchmark,t} + \varepsilon_t \tag{22}$$

and performed a right-sided t-test for the significance of the intercept (Wooldridge, 2015) with the following hypotheses:
$H_0 : \alpha = 0$
$H_1 : \alpha > 0$
The p-values of the t-tests are presented in Table 5.

Table 5. P-values for the t-test for the intercept significance for the base GMIR strategies.

		$r_{benchmark}$		
		EW	DJIA	MV GMIR
$r_{strategy}$	FMV-AG GMIR	0,157	0,183	0,169
	FMV-XGB GMIR	**0,061**	**0,071**	0,111
	MV GMIR	0,259	0,289	-

Note: P-values lower than 0.1 are marked in **bold**.

The results of the t-test showed that FMV-XGB strategy returns had intercepts significantly higher than 0 when regressed on returns from equally weighted and buy-and-hold portfolios at a 0.1 significance level. For other pairs, there were no reasons to reject the null hypothesis.

3.2. Global minimum variance portfolio

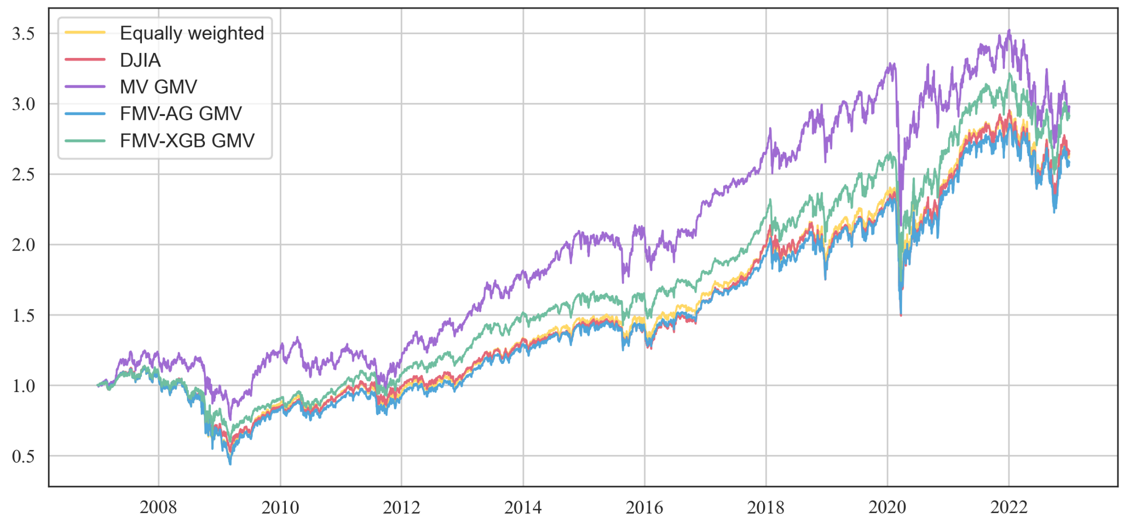
In the case of GMV strategies (which results are shown in Table 6 and Figure 5), we can see that the differences between them and benchmarks were much smaller. The FMV-XGB strategy and the MV strategy performed better than both the benchmarks and the FMV-AG. The MV strategy was the best one in terms of information ratio**. The results of the FMV-AG may be relatively poor because this portfolio was very similar to the equally weighted portfolio. We can see that when we compare $MN^{75\%}$ and $MN^{90\%}$ metrics. In this scenario, it is questionable if it is beneficial to forecast stock returns.

Table 6. Performance statistics of the base GMV strategies.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^*	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV GMV	**197,28%**	**7,05%**	20,12%	**39,71%**	0,350	**0,062**	5,79	8,19
FMV-AG GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
FMV-XGB GMV	190,98%	6,90%	**18,76%**	47,50%	**0,368**	0,053	12,81	17,44

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^* – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ... 90% of the portfolio

Figure 5. Equity lines of the base GMV strategies.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. MV GMV – mean-variance global minimum variance portfolio based on historical returns, FMV-AG GMV – mean-variance global minimum variance portfolio based on returns forecasted using the ARIMA-GARCH model, FMV-XGB GMV – mean-variance global minimum variance portfolio based on returns forecasted using the XGBoost model.

Statistical tests analogical to the GMIR case were performed for the GMV portfolios. Its results are presented in Tables 7 and 8.

Table 7. P-values for the paired t-test for the base GMV strategies.

		$\mu_{benchmark}$		
		EW	DJIA	MV GMV
$\mu_{strategy}$	FMV-AG GMV	0,592	0,458	0,609
	FMV-XGB GMV	0,442	0,358	0,568
	MV GMV	0,468	0,454	-

Note: P-values lower than 0.1 are marked in **bold**.

Also in the GMV case, the differences between strategies' expected values were not significant at the 0.1 significance level for any strategy pair.

Table 8. P-values for the t-test for the intercept significance for the base GMV strategies.

		$r_{benchmark}$		
		EW	DJIA	MV GMV
$r_{strategy}$	FMV-AG GMV	0,628	0,619	0,398
	FMV-XGB GMV	0,137	0,191	0,301
	MV GMV	0,223	0,262	-

Note: P-values lower than 0.1 are marked in **bold**.

In the GMV case, no intercepts were significantly higher than 0 at the 0.1 significance level. Please note that in both tests we assumed that both series values were independent of their previous values. This assumption may not hold for the financial time series, which is why the test results are inconclusive.

IV. SENSITIVITY ANALYSIS

All of the strategies depend on many parameters and even a slight change in them can have substantial consequences. In this chapter, we inspect strategies' performance with different parameter combinations. An extensive sensitivity analysis is crucial to answering the research questions and testing the robustness of the base case strategies.

4.1. Estimation window length

In the base case, we estimated both models' parameters on a 12-month estimation window. In this section, we show how the portfolio performance was affected by both the increase and decrease of estimation window length.

Global maximum information ratio portfolio

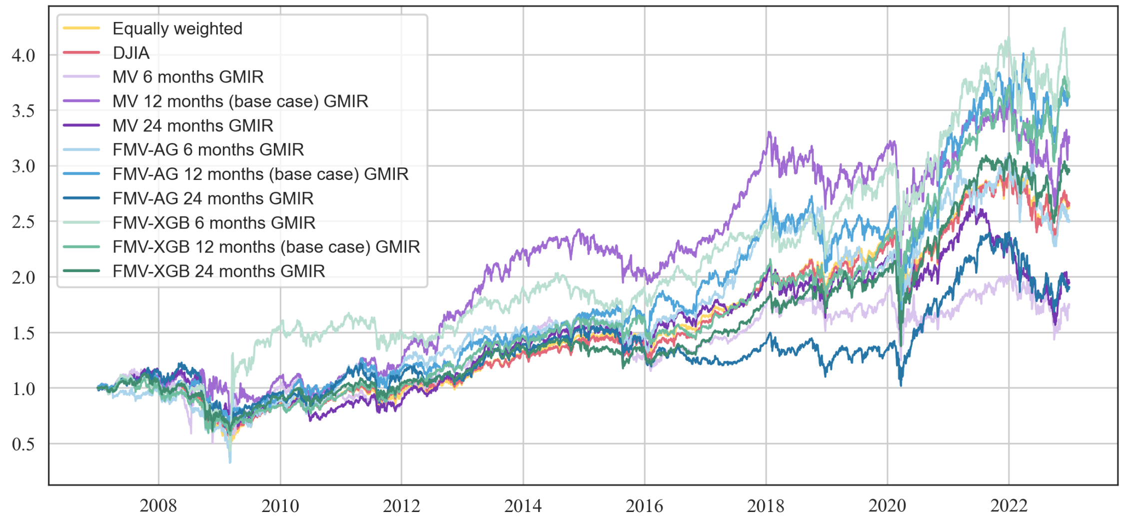
Table 9 and Figure 6 show that for GMIR portfolios, the best estimation window length was 12 months. For every strategy, it overperformed other tested estimation window lengths in terms of information ratios. That means all the strategies were sensitive to estimation window length. The FMV-XGB strategy with a 6-month estimation window performed the best in terms of returns, but due to its relatively high standard deviation and maximum drawdown, this strategy was worse than base case strategy (FMV-XGB 12 m GMIR).

Table 9. Performance statistics of the base GMIR strategies with different lengths of estimation window.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 6 months GMIR	74,96%	3,56%	28,54%	57,19%	0,125	0,008	3,42	4,53
MV 12 months (b.c.) GMIR	225,75%	7,66%	24,96%	**36,72%**	0,307	0,064	3,95	5,19
MV 24 months GMIR	94,62%	4,25%	21,72%	50,14%	0,196	0,017	5,32	6,79
FMV-AG 6 m GMIR	151,01%	5,92%	28,43%	68,34%	0,208	0,018	3,38	5,63
FMV-AG 12 m (b.c.) GMIR	261,62%	8,37%	22,56%	40,02%	0,371	0,077	3,66	6,35
FMV-AG 24 m GMIR	90,63%	4,11%	23,10%	51,82%	0,178	0,014	4,14	7,08
FMV-XGB 6 m GMIR	**274,83%**	**8,61%**	26,10%	61,94%	0,330	0,046	4,06	6,02
FMV-XGB 12 m (b.c.) GMIR	262,63%	8,38%	**19,02%**	46,97%	**0,441**	**0,079**	5,97	8,68
FMV-XGB 24 m GMIR	195,36%	7,00%	19,25%	45,81%	0,364	0,056	7,16	10,28

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 3, 12, and 24 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, b.c. – the base case

Figure 6. Equity lines of the base GMIR strategies with different lengths of estimation window.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 3, 12, and 24 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Global minimum variance portfolio

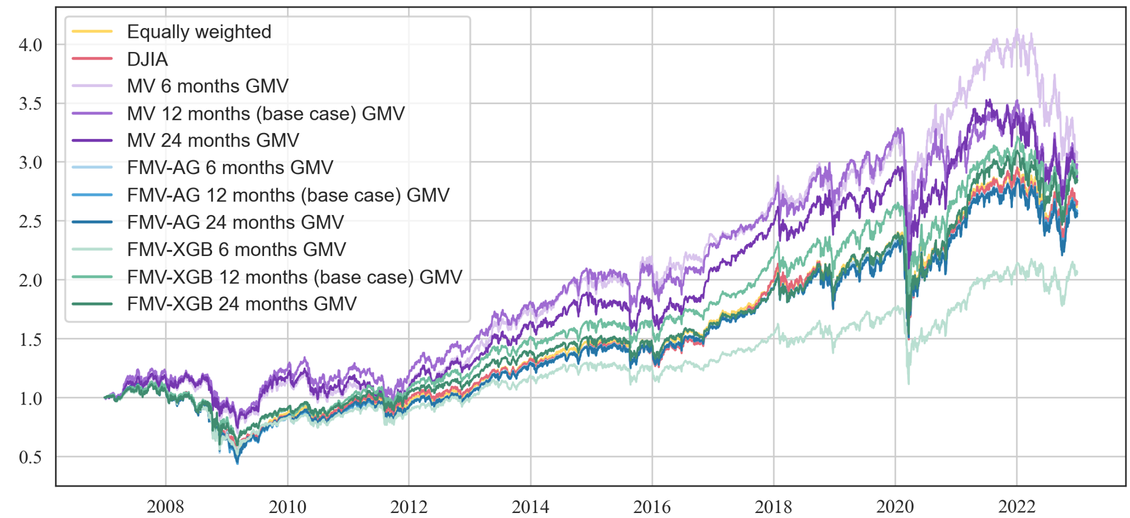
The situation is quite different when it comes to GMV portfolios. Looking at Table 10 and Figure 7, we can see that results were less sensitive to the estimation window length. The MV and FMV-AG strategies seemed to be robust to this parameter. The FMV-AG portfolios were strikingly similar to the equally weighted portfolio, independent of estimation window length. The FMV-XGB strategy was less robust. Similarly to GMIR portfolios, the FMV-XGB portfolios performed best for the 12-month estimation window (although the difference between 12 and 24-month portfolios was very small).

Table 10. Performance statistics of the base GMV strategies with different lengths of estimation window.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 6 months GMV	**207,46%**	**7,27%**	20,52%	42,91%	0,354	0,060	5,17	7,41
MV 12 months (b.c.) GMV	197,28%	7,05%	20,12%	39,71%	0,350	**0,062**	5,79	8,19
MV 24 months GMV	190,53%	6,89%	20,22%	**38,63%**	0,341	0,061	7,10	9,69
FMV-AG 6 m GMV	156,76%	6,07%	21,11%	61,03%	0,288	0,029	21,90	25,92
FMV-AG 12 m (b.c.) GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
FMV-AG 24 m GMV	156,16%	6,06%	20,80%	59,84%	0,291	0,029	22,74	26,74
FMV-XGB 6 m GMV	106,35%	4,63%	19,06%	54,22%	0,243	0,021	11,18	15,48
FMV-XGB 12 m (b.c.) GMV	190,98%	6,90%	**18,76%**	47,50%	**0,368**	0,053	12,81	17,44
FMV-XGB 24 m GMV	184,16%	6,75%	19,40%	46,72%	0,348	0,050	13,95	18,74

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 3, 12, and 24 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, b.c. – the base case

Figure 7. Equity lines of the GMV strategies with different lengths of estimation window.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 3, 12, and 24 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

4.2. Rebalancing period length

Rebalancing period length is a very important parameter. It may be tricky to find the best value, as a too short rebalancing period may result in high transactional costs, and a too long rebalancing period may make it harder to react to changes in the market. In the base case, portfolio weights are changed every month. Below, we show how the portfolios with a 2-week and a 3-month estimation window performed.

Global maximum information ratio portfolio

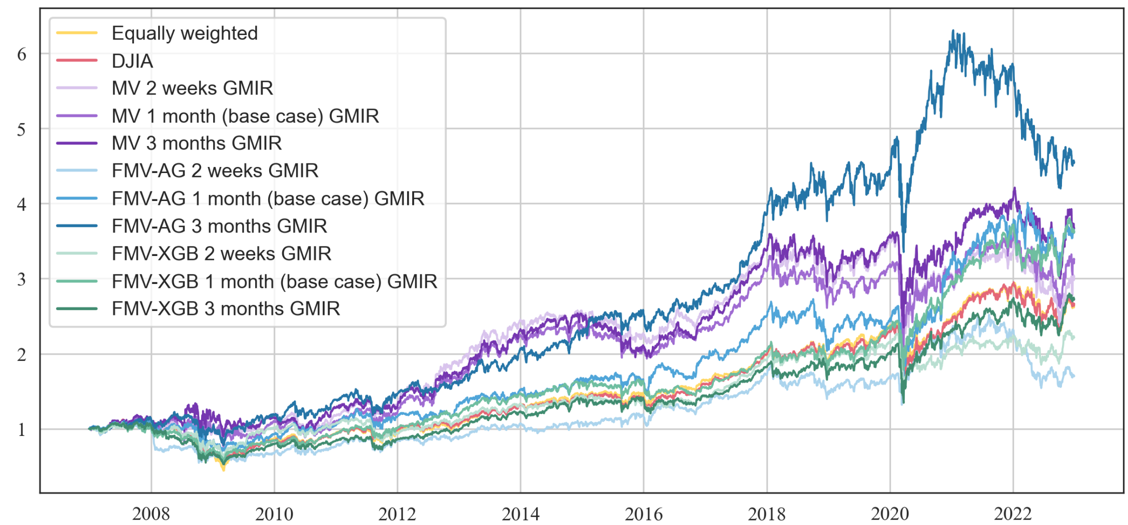
Table 11 and Figure 8 show that in the GMIR case, the MV and the FMV-AG strategies performed better when the rebalancing period was longer. FMV-AG strategy with a 3-month rebalancing period had an exceptionally good performance. On the other hand, for the FMV-XGB strategy, the best rebalancing period was a month. Both longer and shorter rebalancing periods performed worse in the case of FMV-XGB.

Table 11. Performance statistics of the base GMIR strategies with different lengths of rebalancing period.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 2 weeks GMIR	202,39%	7,16%	26,07%	37,78%	0,275	0,052	4,05	5,28
MV 1 month (b.c.) GMIR	225,75%	7,66%	24,96%	36,72%	0,307	0,064	3,95	5,19
MV 3 months GMIR	266,75%	8,46%	24,14%	**35,13%**	0,351	0,084	3,88	5,11
FMV-AG 2 w GMIR	70,18%	3,38%	21,37%	50,77%	0,158	0,011	3,91	6,53
FMV-AG 1 m (b.c.) GMIR	261,62%	8,37%	22,56%	40,02%	0,371	0,077	3,66	6,35
FMV-AG 3 m GMIR	**354,89%**	**9,93%**	21,81%	36,28%	**0,455**	**0,125**	3,23	5,55
FMV-XGB 2 w GMIR	121,71%	5,10%	19,87%	42,35%	0,257	0,031	6,64	9,26
FMV-XGB 1 m (b.c.) GMIR	262,63%	8,38%	**19,02%**	46,97%	0,441	0,079	5,97	8,68
FMV-XGB 3 m GMIR	172,48%	6,47%	19,98%	51,73%	0,324	0,040	5,88	8,77

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 2 weeks, 1 month, 3 months, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ...90% of the portfolio, b.c. – the base case

Figure 8. Equity lines of the GMIR strategies with different lengths of rebalancing period.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 2 weeks, 1 month, 3 months, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Global minimum variance portfolio

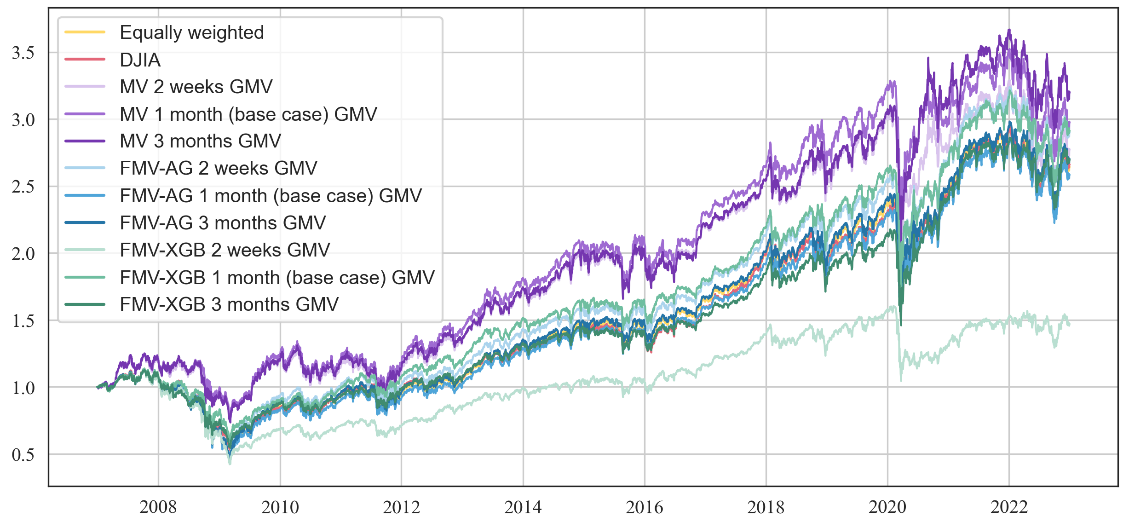
Table 12 and Figure 9 show that in the GMV case, the MV strategy performed better when the rebalancing period was longer. For the FMV-XGB strategy, the best rebalancing period was again a month (although the 3-month rebalancing period performed not much worse). What is interesting about the FMV-AG strategy, the 1-month estimation window performed the worst. Still, all the FMV-AG strategies were very similar to an equally weighted portfolio.

Table 12. Performance statistics of the base GMV strategies with different lengths of rebalancing period.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 2 weeks GMV	177,01%	6,58%	20,06%	**39,62%**	0,328	0,054	5,86	8,19
MV 1 month (b.c.) GMV	197,28%	7,05%	20,12%	39,71%	0,350	0,062	5,79	8,19
MV 3 months GMV	**219,59%**	**7,53%**	20,17%	41,30%	**0,373**	**0,068**	5,63	8,05
FMV-AG 2 w GMV	189,88%	6,88%	20,48%	58,71%	0,336	0,039	22,48	26,51
FMV-AG 1 m (b.c.) GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
FMV-AG 3 m GMV	170,10%	6,41%	20,66%	60,23%	0,310	0,033	22,70	26,75
FMV-XGB 2 w GMV	46,97%	2,44%	**18,49%**	62,44%	0,132	0,005	12,07	16,33
FMV-XGB 1 m (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
FMV-XGB 3 m GMV	169,41%	6,39%	18,41%	51,69%	0,347	0,043	14,59	19,50

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 2 weeks, 1 month, 3 months, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ... 90% of the portfolio, b.c. – the base case

Figure 9. Equity lines of the GMV strategies with different lengths of rebalancing period.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 2 weeks, 1 month, 3 months, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

4.3. Number of stocks

4.3.1. 10 stocks (in 3 groups)

Portfolios consisting of all DJIA stocks were optimized in the base case scenario. In this section, we divided stocks into groups by their estimated market capitalization. On the day before the first day of every rebalancing period, we ordered stocks by their decreasing estimated market capitalization and assigned them to the "1-10" (ten highest), "11-20" (ten middle), and "21-30" (ten lowest[1]) groups in the case of portfolios consisting of 10 stocks and analogically "1-5", "6-10", "11-15", "16-20", "21-25" and "26-30" [2] groups in the case of 5 stock portfolios. Due to the fact that the data for market capitalization were not available for free, we had to assess it by the average value of the price of given stock multiplied by its traded volume over the period of last year. The exact formula used when estimating market capitalization is shown below;

$$\hat{Cap} = \frac{1}{252} \sum_{i=1}^{252} P_{t-i} \cdot Vol_{t-i} \tag{23}$$

where:
$\hat{Cap}$ – the estimated market capitalization of a company,
P_t – the closing price of a stock on day t,
Vol_t – the volume of a stock on day t.

Global maximum information ratio portfolio

In Table 13 and Figure 10, we can see that there is one group of stock that tended to outperform the others - the group of lowest capitalization stocks. This group performed especially well when paired with the MV strategy. Other strategies still performed slightly better than benchmarks in this stock group. What is interesting, the other two groups performed worse than the benchmarks, no matter which strategy was utilized. We can see from the comparison of equally weighted portfolios that the "21-30" group performs better than "1-10" and "11-20" groups, and the difference was the biggest when using the MV strategy.

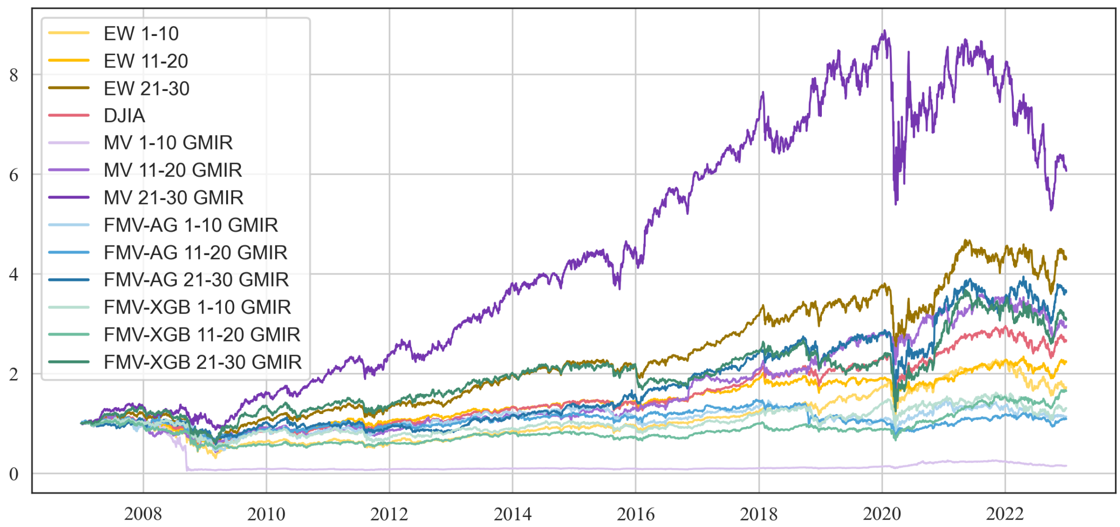
[1] nine lowest in the period when GM was a part of the DJIA
[2] this group consisted of four stocks in the period when GM was a part of the DJIA

Table 13. Performance statistics of the GMIR strategies in 3 groups of stocks by estimated market capitalization.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^*	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
EW 1-10	72,37%	3,46%	25,84%	72,00%	0,134	0,006	8,00	9,00
EW 11-20	123,29%	5,15%	18,86%	55,81%	0,273	0,025	8,00	9,00
EW 21-30	330,38%	9,55%	20,94%	54,73%	0,456	0,080	7,84	9,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 1-10 GMIR	-42,26%	-3,37%	28,77%	80,80%	-0,117	-0,005	2,98	4,23
MV 11-20 GMIR	97,18%	4,33%	21,98%	66,31%	0,197	0,013	3,54	5,02
MV 21-30 GMIR	**480,00%**	**11,61%**	22,57%	**42,25%**	**0,515**	**0,141**	3,38	4,64
FMV-AG 1-10 GMIR	57,84%	2,89%	25,48%	70,84%	0,114	0,005	7,82	8,88
FMV-AG 11-20 GMIR	113,85%	4,87%	18,65%	54,02%	0,261	0,023	7,85	8,88
FMV-AG 21-30 GMIR	338,59%	9,68%	20,79%	53,28%	0,466	0,085	7,74	8,86
FMV-XGB 1-10 GMIR	47,57%	2,46%	22,22%	62,19%	0,111	0,004	4,65	6,27
FMV-XGB 11-20 GMIR	82,92%	3,85%	**17,35%**	50,08%	0,222	0,017	4,95	6,64
FMV-XGB 21-30 GMIR	304,53%	9,13%	19,78%	48,57%	0,462	0,087	4,67	6,31

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^* – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 10. Equity lines of the GMIR strategies in 3 groups of stocks by estimated market capitalization.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Global minimum variance portfolio

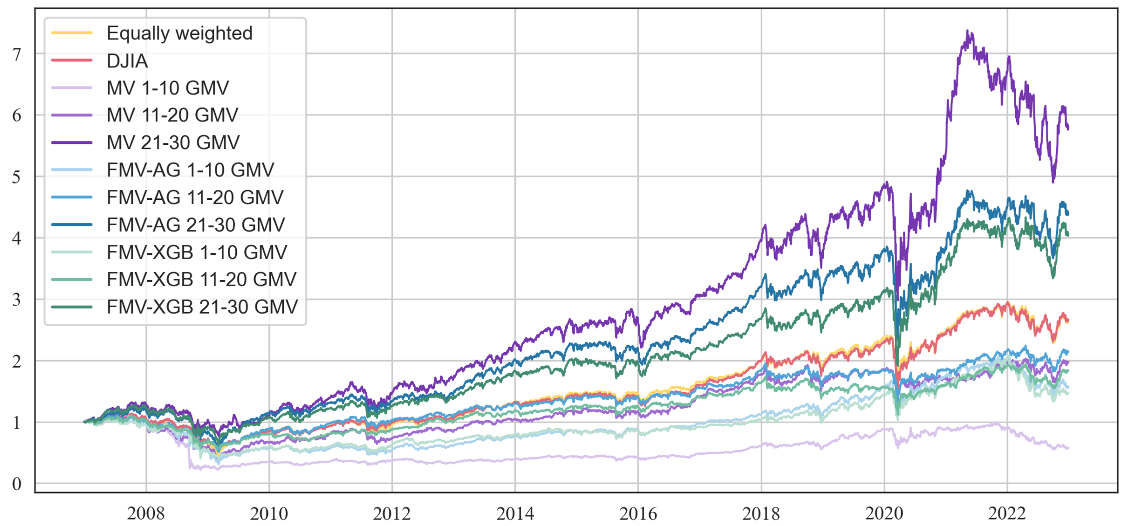
Table 14 and Figure 11 show that difference between MV for the lowest capitalization stocks and other portfolios was even bigger in the case of GMV portfolios. This strategy outperformed others significantly. In the case of the GMV portfolio, FMV strategies in the "21-30" group no longer outperformed the benchmarks. An interesting observation is that FMV-AG portfolios did not tend to resemble equally weighted portfolios, as was the case with all 30 stocks.

Table 14. Performance statistics of the GMIR strategies in 3 groups of stocks by estimated market capitalization.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
EW 1-10	72,37%	3,46%	25,84%	72,00%	0,134	0,006	8,00	9,00
EW 11-20	123,29%	5,15%	**18,86%**	55,81%	0,273	0,025	8,00	9,00
EW 21-30	330,38%	9,55%	20,94%	54,73%	0,456	0,080	7,84	9,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 1-10 GMV	-85,13%	-11,23%	40,19%	95,56%	-0,279	-0,033	2,03	2,58
MV 11-20 GMV	195,49%	7,01%	25,13%	64,30%	0,279	0,030	2,07	2,62
MV 21-30 GMV	**505,97%**	**11,92%**	23,36%	**40,70%**	**0,510**	**0,149**	2,40	2,99
FMV-AG 1-10 GMV	13,49%	0,79%	26,56%	58,34%	0,030	0,000	2,11	2,95
FMV-AG 11-20 GMV	8,54%	0,51%	19,52%	52,71%	0,026	0,000	2,07	2,93
FMV-AG 21-30 GMV	263,76%	8,41%	23,69%	49,35%	0,355	0,060	2,02	2,91
FMV-XGB 1-10 GMV	30,36%	1,67%	23,90%	54,63%	0,070	0,002	2,62	3,56
FMV-XGB 11-20 GMV	64,67%	3,17%	20,24%	58,94%	0,156	0,008	2,72	3,66
FMV-XGB 21-30 GMV	209,00%	7,31%	23,90%	54,26%	0,306	0,041	2,60	3,46

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ...90% of the portfolio, EW - equally weighted

Figure 11. Equity lines of the GMV strategies in 3 groups of stocks by estimated market capitalization.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

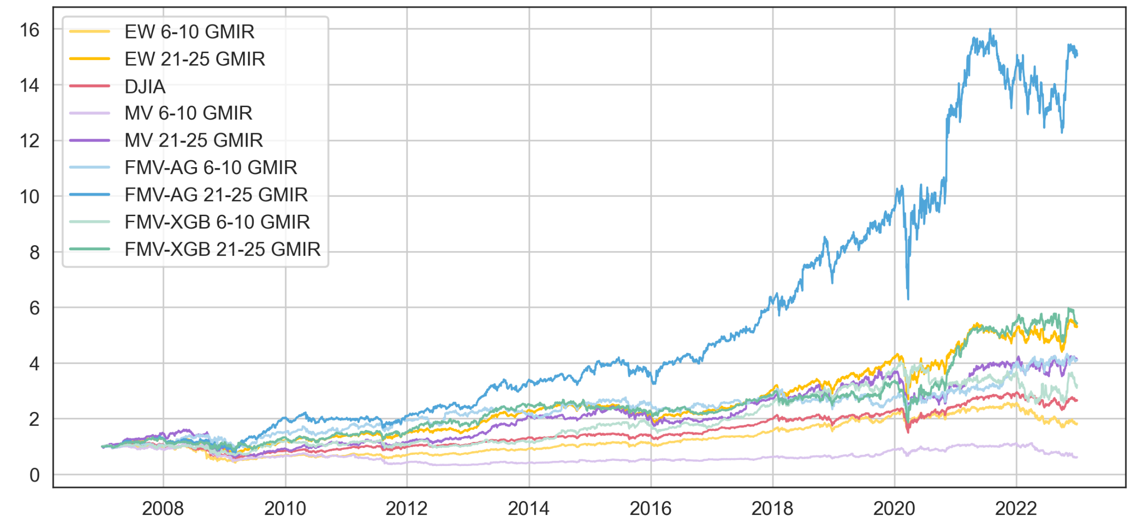
Tables 13 and 14 show that the relation between information ratios and market capitalisation was strictly monotonic in both cases. Because of that, we decided to split them into 5 stock groups and check whether the monotonic relation still holds.

4.3.2. 5 stocks (in 6 groups)

Global maximum information ratio portfolio

Figure 12 and Table 15 show that the relation between information ratios** and market capitalization was no longer monotonous. The ”21-25” group outperformed other groups not only in the case of FMV and MV portfolios but also in equally weighted portfolios. It suggests that the ”21-25” group contained stocks that performed exceptionally well in the tested period. The FMV-AG strategy in this group performed particularly well, yielding relatively high returns and information ratio**. This strategy utilized the well-performing stocks the most. We can also see that the performance of some stock groups differs between strategies which suggests that some strategies may perform better for different sets of stocks. For example, the ”1-5” group generated losses in the case of MV and FMV-AG strategies but generated profits in the case of EW and FMV-XGB portfolios. It is worth noting that for every FMV and MV strategy, there were stock groups that generated losses.

Figure 12. Examples of equity lines of the GMIR strategies in 6 groups of stocks by estimated market capitalization.



Note: For transparency, only the ”6-10” and ”21-25” (best performing for FMV strategies) groups' equity lines are shown. Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Table 15. Performance statistics of the GMV strategies in 6 groups of stocks by estimated market capitalization.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
EW 1-5	42,31%	2,23%	30,64%	81,65%	0,073	0,002	5,00	6,00
EW 6-10	82,70%	3,84%	23,74%	60,63%	0,162	0,010	5,00	6,00
EW 11-15	114,26%	4,88%	**18,28%**	43,79%	0,267	0,030	5,00	6,00
EW 16-20	118,21%	5,00%	21,40%	66,61%	0,234	0,018	5,00	6,00
EW 21-25	429,95%	10,99%	21,10%	50,36%	0,521	0,114	5,00	6,00
EW 26-30	220,19%	7,54%	22,49%	60,41%	0,335	0,042	5,00	6,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 1-5 GMIR	-45,66%	-3,74%	37,42%	90,24%	-0,100	-0,004	1,60	1,88
MV 6-10 GMIR	-37,96%	-2,94%	25,76%	70,96%	-0,114	-0,005	1,61	1,97
MV 11-15 GMIR	-18,11%	-1,24%	22,92%	58,75%	-0,054	-0,001	1,59	1,91
MV 16-20 GMIR	113,35%	4,85%	26,37%	78,49%	0,184	0,011	1,58	1,95
MV 21-25 GMIR	310,81%	9,23%	24,17%	61,94%	0,382	0,057	1,64	1,99
MV 26-30 GMIR	108,77%	4,71%	25,57%	46,83%	0,184	0,019	1,75	2,13
FMV-AG 1-5 GMIR	-58,87%	-5,40%	38,06%	88,85%	-0,142	-0,009	1,63	2,06
FMV-AG 6-10 GMIR	311,43%	9,24%	23,71%	38,27%	0,390	0,094	1,57	2,05
FMV-AG 11-15 GMIR	102,44%	4,51%	19,79%	44,25%	0,228	0,023	1,63	2,14
FMV-AG 16-20 GMIR	-32,91%	-2,46%	23,27%	69,73%	-0,106	-0,004	1,65	2,15
FMV-AG 21-25 GMIR	**1403,24%**	**18,46%**	24,64%	44,74%	**0,749**	**0,309**	1,59	2,07
FMV-AG 26-30 GMIR	56,92%	2,86%	25,23%	57,97%	0,113	0,006	1,54	1,93
FMV-XGB 1-5 GMIR	141,31%	5,66%	28,96%	78,08%	0,195	0,014	1,64	2,03
FMV-XGB 6-10 GMIR	220,76%	7,56%	24,08%	**38,06%**	0,314	0,062	1,67	2,16
FMV-XGB 11-15 GMIR	140,42%	5,64%	20,95%	43,69%	0,269	0,035	1,63	2,05
FMV-XGB 16-20 GMIR	108,29%	4,69%	22,72%	60,67%	0,207	0,016	1,75	2,15
FMV-XGB 21-25 GMIR	442,74%	11,15%	23,36%	44,83%	0,477	0,119	1,69	2,05
FMV-XGB 26-30 GMIR	-12,69%	-0,84%	26,43%	63,92%	-0,032	0,000	1,55	1,86

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 5. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ...90% of the portfolio, EW - equally weighted

Global minimum variance portfolio

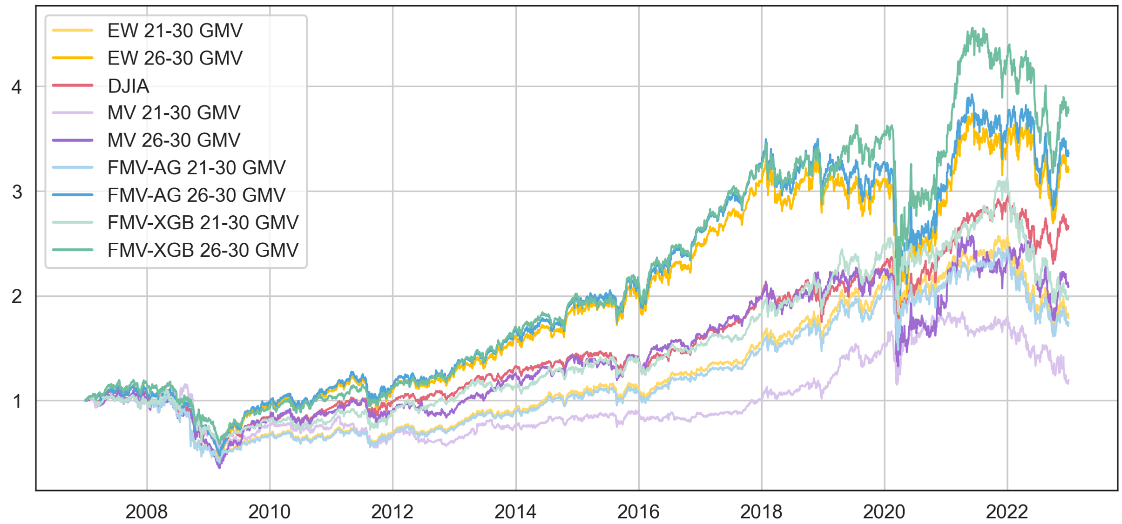
Table 16 and Figure 13 show that the relation between information ratios** and market capitalization was also no longer monotonous in the GMV case. The "21-25" group was still the best in the MV and FMV-AG strategies, but for the FMV-XGB strategy, the "26-30" group performed the best. It is worth noting that in the GMV case, there were no portfolios with losses. FMV-AG portfolios also did not tend to resemble equally weighted portfolios, as was the case with all 30 stocks.

Table 16. Performance statistics of the GMV strategies in 6 groups of stocks by estimated market capitalization.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^*	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
EW 1-5	42,31%	2,23%	30,64%	81,65%	0,073	0,002	5,00	6,00
EW 6-10	82,70%	3,84%	23,74%	60,63%	0,162	0,010	5,00	6,00
EW 11-15	114,26%	4,88%	18,28%	43,79%	0,267	0,030	5,00	6,00
EW 16-20	118,21%	5,00%	21,40%	66,61%	0,234	0,018	5,00	6,00
EW 21-25	429,95%	10,99%	21,10%	50,36%	0,521	**0,114**	5,00	6,00
EW 26-30	220,19%	7,54%	22,49%	60,41%	0,335	0,042	5,00	6,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 1-5 GMV	25,48%	1,43%	28,88%	76,82%	0,049	0,001	2,33	3,14
MV 6-10 GMV	19,57%	1,12%	23,23%	54,85%	0,048	0,001	2,32	2,97
MV 11-15 GMV	207,42%	7,27%	20,36%	48,29%	0,357	0,054	2,55	3,36
MV 16-20 GMV	46,14%	2,40%	23,91%	69,97%	0,100	0,003	2,51	3,37
MV 21-25 GMV	**488,12%**	**11,71%**	22,36%	53,64%	**0,524**	**0,114**	2,43	3,38
MV 26-30 GMV	108,54%	4,70%	24,25%	67,73%	0,194	0,013	2,39	3,16
FMV-AG 1-5 GMV	30,93%	1,70%	30,12%	80,60%	0,056	0,001	3,90	4,87
FMV-AG 6-10 GMV	74,85%	3,55%	23,92%	61,37%	0,149	0,009	3,98	4,96
FMV-AG 11-15 GMV	128,21%	5,29%	**18,02%**	40,65%	0,294	0,038	3,97	4,97
FMV-AG 16-20 GMV	79,00%	3,71%	21,37%	65,54%	0,173	0,010	3,96	4,95
FMV-AG 21-25 GMV	410,93%	10,73%	21,08%	49,83%	0,509	0,110	3,98	4,98
FMV-AG 26-30 GMV	235,45%	7,86%	22,33%	58,71%	0,352	0,047	3,82	4,81
FMV-XGB 1-5 GMV	28,87%	1,60%	25,65%	71,15%	0,062	0,001	2,91	3,83
FMV-XGB 6-10 GMV	99,48%	4,41%	20,60%	47,30%	0,214	0,020	2,96	3,86
FMV-XGB 11-15 GMV	152,99%	5,97%	17,07%	**40,21%**	0,350	0,052	3,04	3,88
FMV-XGB 16-20 GMV	154,59%	6,01%	18,70%	53,62%	0,322	0,036	2,95	3,88
FMV-XGB 21-25 GMV	268,55%	8,49%	20,41%	48,54%	0,416	0,073	3,02	3,94
FMV-XGB 26-30 GMV	276,86%	8,65%	21,18%	51,16%	0,408	0,069	2,90	3,78

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 5. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^* – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 13. Examples of equity lines of the GMIR strategies in 6 groups of stocks by estimated market capitalization.



Note: For transparency, only the "21-25" and "26-30" (best performing for FMV strategies) groups' equity lines are shown. Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 10. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

4.4. Transaction costs level

In the base case, we assumed that transaction costs are at the level of 0.1% of the transaction value. We made this assumption after analysing transaction fees at the biggest brokers in the United States (Charles Schwab, Fidelity Investments, TD Ameritrade)[3]. Transaction costs may vary over time due to macroeconomic circumstances, which is why in this section, we check how the results change when transaction costs fall to 0.05% or rise to 0.2% and 0.5%.

Global maximum information ratio portfolio

By analysing Table 17 and Figure 14, we can see that higher transaction costs influenced the performance of all strategies in a significant way. It is not surprising that higher costs decreased the information ratios of all portfolios. The effect was the smallest for the FMV-XGB strategy. With 0.2% transaction costs, this strategy still performed better than the benchmarks. For the MV and FMV-AG strategies with the same transaction costs, the information ratio** was also higher than the benchmarks. When costs rose to 0.5%, no strategy performed better than the benchmarks. When costs fell to 0.05%, all strategies performed better in terms of information ratio**. The increase of IR** was higher for FMV strategies.

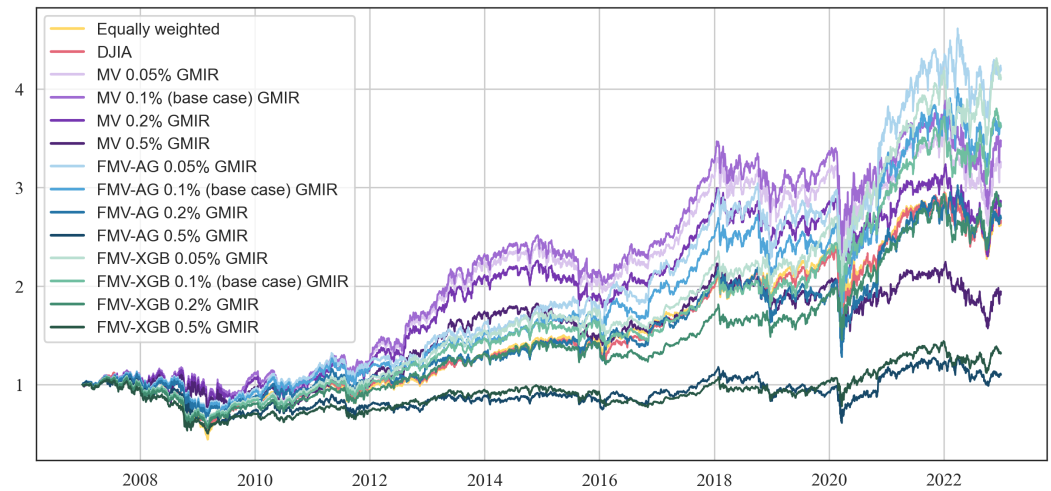
[3]https://www.investopedia.com/articles/professionals/110415/biggest-stock-brokerage-firms-us.asp

Table 17. Performance statistics of the GMIR strategies with different levels of transaction costs.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 0.05% GMIR	247,43%	8,09%	24,95%	**36,36%**	0,324	0,072	3,95	5,19
MV 0.1% (b.c.) GMIR	225,75%	7,66%	24,96%	36,72%	0,307	0,064	3,95	5,19
MV 0.2% GMIR	186,33%	6,80%	24,97%	37,43%	0,272	0,049	3,95	5,19
MV 0.5% GMIR	94,32%	4,24%	25,01%	39,51%	0,170	0,018	3,95	5,19
FMV-AG 0.05% GMIR	319,38%	9,37%	22,56%	39,24%	0,415	0,099	3,66	6,35
FMV-AG 0.1% (b.c.) GMIR	261,62%	8,37%	22,56%	40,02%	0,371	0,077	3,66	6,35
FMV-AG 0.2% GMIR	168,77%	6,37%	22,57%	41,65%	0,282	0,043	3,66	6,35
FMV-AG 0.5% GMIR	10,04%	0,60%	22,67%	48,29%	0,026	0,000	3,66	6,35
FMV-XGB 0.05% GMIR	**311,35%**	**9,24%**	**19,02%**	46,40%	**0,486**	**0,097**	5,97	8,68
FMV-XGB 0.1% (b.c.) GMIR	262,63%	8,38%	**19,02%**	46,97%	0,441	0,079	5,97	8,68
FMV-XGB 0.2% GMIR	181,76%	6,69%	19,03%	48,36%	0,352	0,049	5,97	8,68
FMV-XGB 0.5% GMIR	31,89%	1,74%	19,11%	52,34%	0,091	0,003	5,97	8,68

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio.

Figure 14. Equity lines of the GMIR strategies with different levels of transaction costs.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Global minimum variance portfolio

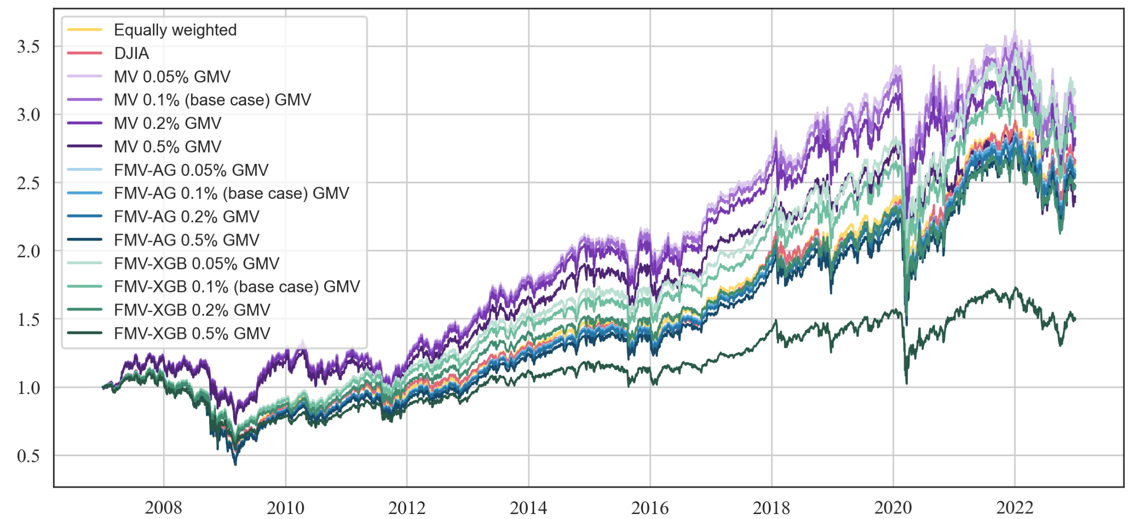
Looking at Table 18 and Figure 15, it is easy to notice that the FMV-AG strategy is very insensitive to change in transaction costs. The explanation is straightforward. As before, the FMV-AG GMV portfolio was similar to the equally weighted portfolio. The FMV-XGB strategy with higher transaction costs than in the base case scenario did not perform better than benchmarks. The FMV-XGB strategy with 0.2% transaction costs outperformed the benchmarks in terms of both information ratios. When the costs rose to 0.5%, the advantage disappeared. When the costs fell to 0.05% both MV and FMV-XGB strategies had higher information ratios**, but the change was smaller than in the GMIR case.

Table 18. Performance statistics of the GMV strategies with different levels of transaction costs.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
MV 0.05% GMV	205,44%	7,23%	20,11%	**39,55%**	0,359	**0,066**	5,79	8,19
MV 0.1% (b.c.) GMV	197,28%	7,05%	20,12%	39,71%	0,350	0,062	5,79	8,19
MV 0.2% GMV	181,60%	6,68%	20,12%	40,03%	0,332	0,055	5,79	8,19
MV 0.5% GMV	139,30%	5,60%	20,14%	40,98%	0,278	0,038	5,79	8,19
FMV-AG 0.05% GMV	159,66%	6,15%	21,02%	61,24%	0,292	0,029	22,40	26,44
FMV-AG 0.1% (b.c.) GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
FMV-AG 0.2% GMV	155,62%	6,04%	21,02%	61,47%	0,287	0,028	22,40	26,44
FMV-AG 0.5% GMV	147,70%	5,83%	21,03%	61,92%	0,277	0,026	22,40	26,44
FMV-XGB 0.05% GMV	**216,15%**	**7,46%**	**18,76%**	47,02%	**0,398**	0,063	12,81	17,44
FMV-XGB 0.1% (b.c.) GMV	190,98%	6,90%	**18,76%**	47,50%	0,368	0,053	12,81	17,44
FMV-XGB 0.2% GMV	146,47%	5,80%	**18,76%**	48,45%	0,309	0,037	12,81	17,44
FMV-XGB 0.5% GMV	49,64%	2,55%	18,80%	51,19%	0,136	0,007	12,81	17,44

Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 15. Equity lines of the GMV strategies with different levels of transaction costs.



Note: Testing period: 2007-01-01 - 2022-12-31. All strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30. FMV-AG parameters: ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal. FMV-XGB parameters: number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

4.5. ARIMA order

It is not easy, if not impossible, to find one ARIMA order that will suit all the data. Only for the base case ARIMA-GARCH strategy, 5730 forecasts needed to be performed. As we stated before, the best approach would be to perform a grid search for order that minimises chosen information criterion, but it would require significant computation power. Because of that, in the base case scenario, we selected an order of $(2, 0, 1)$. In this section, other orders are checked.

Global maximum information ratio portfolio

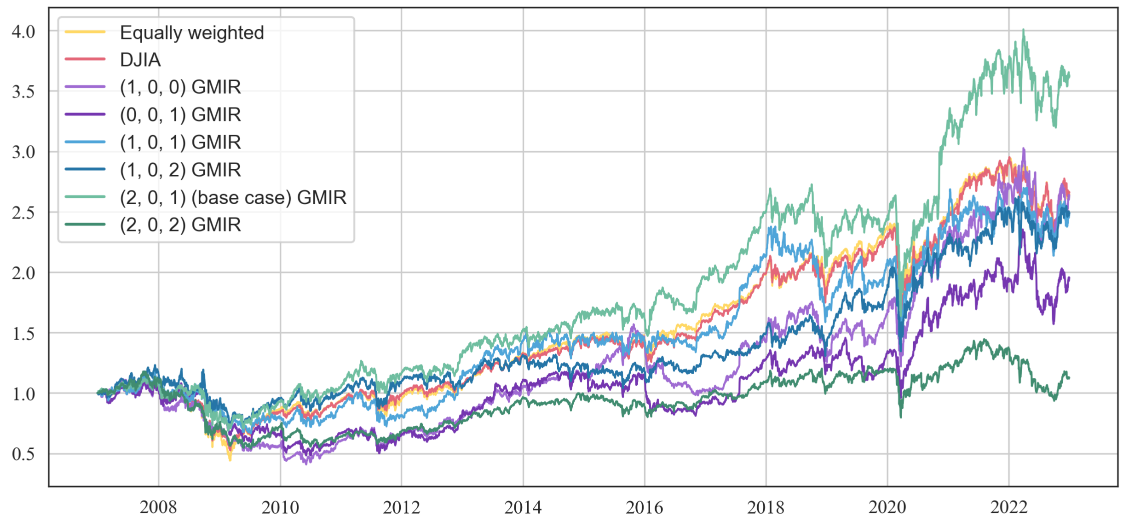
Table 19 and Figure 16 show that the FMV-AG strategy was sensitive to the ARIMA order. We can see that as the order increased the strategy performed better except for order $(2, 0, 2)$. All orders other than the base case order of $(2, 0, 1)$ performed worse than the DJIA buy and hold portfolio regarding the information ratio**. $(1, 0, 1)$ and $(1, 0, 2)$ orders performed better than equally weighted portfolio. In further research, checking how higher orders perform would be interesting.

Table 19. Performance statistics of the FMV-AG GMIR strategies with different ARIMA orders.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	**19,55%**	53,78%	0,323	0,038	-	-
(1, 0, 0) GMIR	163,71%	6,25%	24,86%	63,67%	0,251	0,025	4,34	5,76
(0, 0, 1) GMIR	95,59%	4,28%	24,89%	55,98%	0,172	0,013	6,74	8,83
(1, 0, 1) GMIR	149,18%	5,87%	22,99%	43,24%	0,255	0,035	2,61	4,41
(1, 0, 2) GMIR	148,52%	5,85%	24,14%	41,55%	0,243	0,034	3,38	5,71
(2, 0, 1) (b.c.) GMIR	**261,62%**	**8,37%**	22,56%	**40,02%**	**0,371**	**0,077**	3,66	6,35
(2, 0, 2) GMIR	12,80%	0,76%	21,78%	54,24%	0,035	0,000	3,08	5,38

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(1, 0, 0)$, $(0, 0, 1)$, $(1, 0, 1)$, $(2, 0, 1)$, $(1, 0, 2)$, $(2, 0, 1)$, $(2, 0, 2)$, GARCH error distribution: normal. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 16. Equity lines of the FMV-AG GMIR strategies with different ARIMA orders.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(1, 0, 0)$, $(0, 0, 1)$, $(1, 0, 1)$, $(2, 0, 1)$, $(1, 0, 2)$, $(2, 0, 1)$, $(2, 0, 2)$, GARCH error distribution: normal.

Global minimum variance portfolio

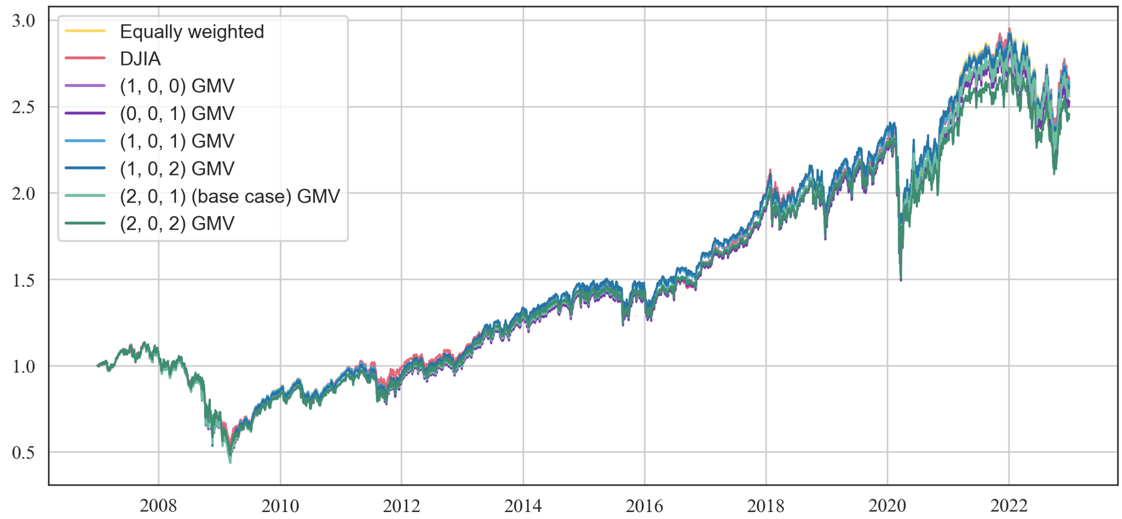
Table 20 and Figure 17 show that changing the ARIMA order did not prevent portfolios from resembling the equally weighted portfolio what can be noticed by looking at the $MN^{75\%}$ and $MN^{90\%}$ metrics. It suggests that the variance-covariance matrix estimated based on forecasted returns promoted diversification independent of ARIMA order. No FMV-AG GMV portfolio outperformed the benchmarks and all of them were characterized by very similar results.

Table 20. Performance statistics of the FMV-AG GMV strategies with different ARIMA orders.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	**165,96%**	**6,30%**	**19,55%**	**53,78%**	**0,323**	**0,038**	-	-
(1, 0, 0) GMV	152,49%	5,96%	20,71%	60,52%	0,288	0,028	22,80	26,81
(0, 0, 1) GMV	152,66%	5,96%	20,74%	60,57%	0,288	0,028	22,79	26,80
(1, 0, 1) GMV	163,79%	6,25%	20,59%	59,13%	0,304	0,032	22,55	26,59
(1, 0, 2) GMV	161,59%	6,19%	21,19%	60,21%	0,292	0,030	22,19	26,29
(2, 0, 1) (b.c.) GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
(2, 0, 2) GMV	145,32%	5,77%	20,26%	57,60%	0,285	0,029	21,37	25,52

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(1,0,0)$, $(0,0,1)$, $(1,0,1)$, $(2,0,1)$, $(1,0,2)$, $(2,0,1)$, $(2,0,2)$, GARCH error distribution: normal. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ...90% of the portfolio, EW - equally weighted.

Figure 17. Equity lines of the FMV-AG GMV strategies with different ARIMA orders.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(1,0,0)$, $(0,0,1)$, $(1,0,1)$, $(2,0,1)$, $(1,0,2)$, $(2,0,1)$, $(2,0,2)$, GARCH error distribution: normal.

4.6. GARCH error distribution

In the base case scenario, we assumed that errors follow a normal distribution. In fact, it does not always hold and other distributions may be applied to errors. In this section, we check some of them and their influence on the FMV-AG strategies results.

Global maximum information ratio portfolio

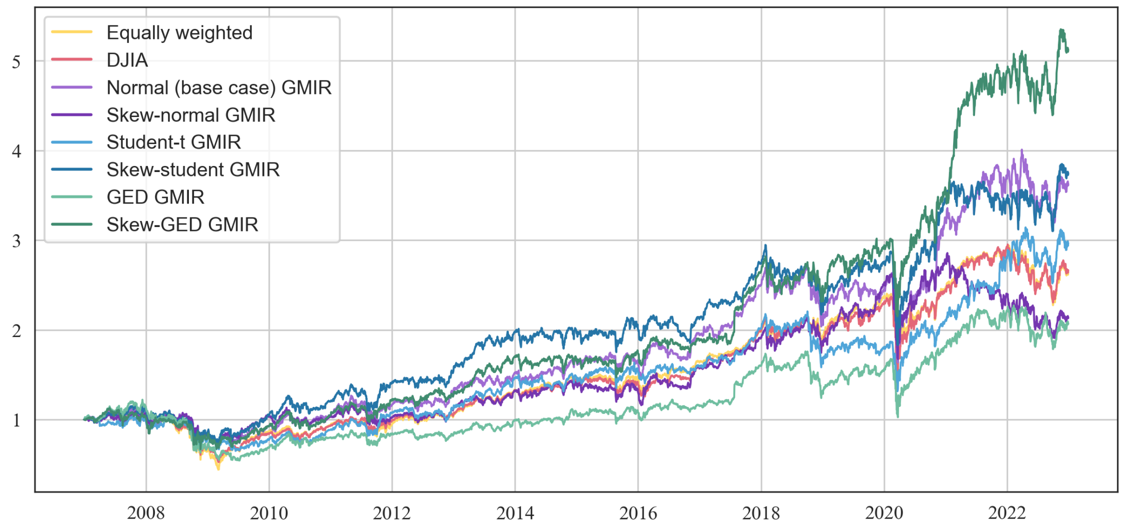
The results are very sensitive to GARCH error distribution. By looking at Table 21 and Figure 18, we can tell that assuming skewed Generalized Error Distribution significantly increased the performance of the portfolio and results in relatively high values of information ratios. The skew-GED distribution was then followed by skewed student-t and normal distribution, but their performance was much worse. These distributions outperformed the benchmarks. Student-t distribution was comparable to benchmarks. It had the information ratio* smaller than one of the benchmarks, but its information ratio** was higher than both. Other distributions (skewed normal and Generalized Error Distribution) performed worse than benchmarks.

Table 21. Performance statistics of the FMV-AG GMIR strategies with different GARCH error distributions.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	**19,55%**	53,78%	0,323	0,038	-	-
Normal (b.c.) GMIR	261,62%	8,37%	22,56%	40,02%	0,371	0,077	3,66	6,35
Skew-normal GMIR	113,28%	4,85%	21,67%	**37,65%**	0,224	0,029	3,63	6,14
Student-t GMIR	196,65%	7,03%	22,06%	40,41%	0,319	0,055	3,89	6,59
Skew-student GMIR	273,91%	8,59%	21,91%	37,84%	0,392	0,089	3,74	6,43
GED GMIR	106,99%	4,65%	22,11%	55,52%	0,210	0,018	3,68	6,21
Skew-GED GMIR	**411,69%**	**10,74%**	22,31%	40,78%	**0,482**	**0,127**	3,82	6,27

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal, skew-normal, student-t, skew-student, GED, skew-GED. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – ...90% of the portfolio.

Figure 18. Equity lines of the FMV-AG GMIR strategies with different GARCH error distributions.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal, skew-normal, student-t, skew-student, GED, skew-GED.

Global minimum variance portfolio

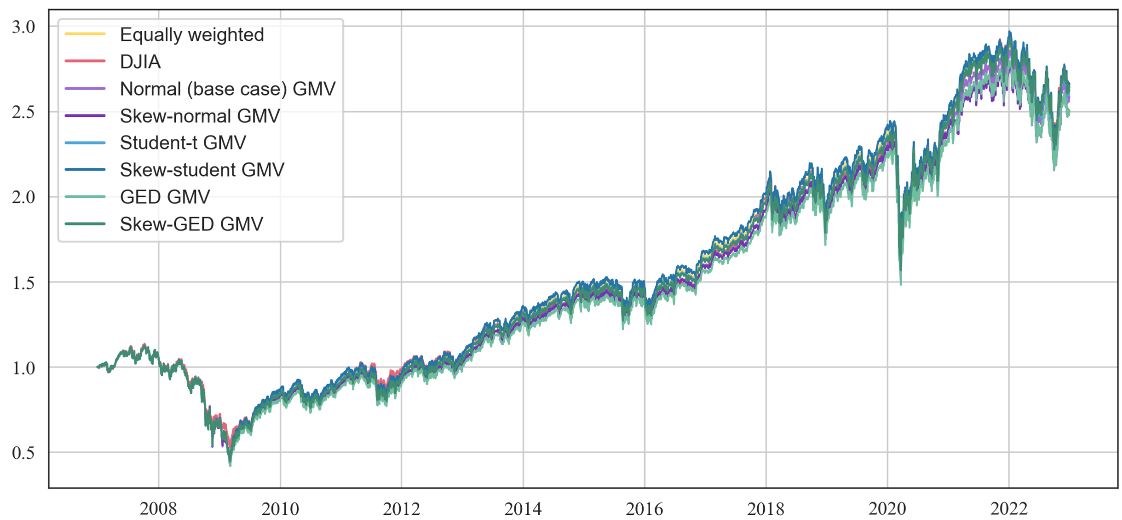
Table 22 and Figure 19 suggest that regardless of the GARCH error distribution all FMV-AG strategies converged to equally weighted portfolio. All of them performed worse than DJIA buy and hold and only skew-student distribution performed better than equally weighted portfolio (although the difference in information ratios** is very small).

Table 22. Performance statistics of the FMV-AG GMV strategies with different GARCH error distributions.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	264,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	**265,96%**	**6,30%**	**19,55%**	**53,78%**	**0,323**	**0,038**	-	-
Normal (b.c.) GMV	158,31%	6,11%	21,02%	61,32%	0,291	0,029	22,40	26,44
Skew-normal GMV	150,17%	5,90%	20,99%	61,00%	0,281	0,027	22,35	26,41
Student-t GMV	162,19%	6,21%	20,89%	60,54%	0,297	0,030	22,44	26,48
Skew-student GMV	165,96%	6,30%	20,87%	60,05%	0,302	0,032	22,40	26,45
GED GMV	150,30%	5,90%	20,88%	62,83%	0,283	0,027	22,49	26,53
Skew-GED GMV	164,09%	6,26%	20,69%	60,43%	0,302	0,031	22,45	26,47

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal, skew-normal, student-t, skew-student, GED, skew-GED. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 19. Equity lines of the FMV-AG GMV strategies with different GARCH error distributions.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, ARIMA order: $(2, 0, 1)$, GARCH error distribution: normal, skew-normal, student-t, skew-student, GED, skew-GED.

4.7. XGBoost feature lag

As we stated in the methodology section, the only features for the XGBoost model were the lagged returns from the last 11 days. Below, we show how the performance of the strategy changes when we change the number of lags to 6 and 21.

Global maximum information ratio portfolio

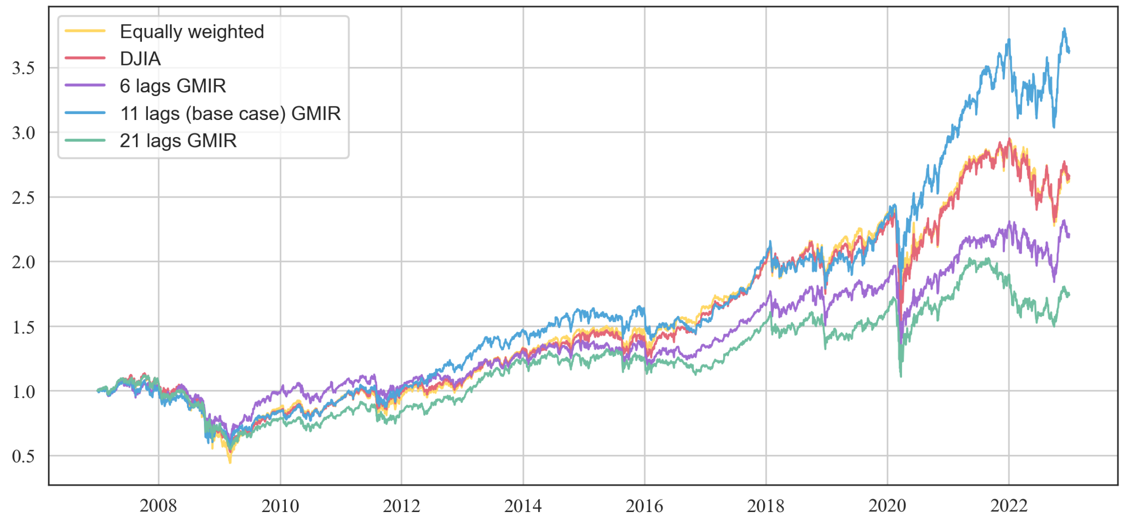
Table 23 and Figure 20 show that number of feature lags used to train the XGBoost model affected the results of underlying strategies. The best value was 11 lags. This strategy outperformed both of the benchmarks. For other numbers of lags (6 and 21), results were worse than benchmarks.

Table 23. Performance statistics of the FMV-XGB GMIR strategies with different feature lags.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
6 lags GMIR	121,05%	5,08%	19,26%	**45,66%**	0,264	0,029	6,80	9,84
11 lags (b.c.) GMIR	**262,63%**	**8,38%**	**19,02%**	46,97%	**0,441**	**0,079**	5,97	8,68
21 lags GMIR	74,28%	3,53%	19,89%	51,22%	0,178	0,012	5,63	8,20

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 6, 11, 21, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 20. Equity lines of the GMIR strategies with different feature lags.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 6, 11, 21, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Global minimum variance portfolio

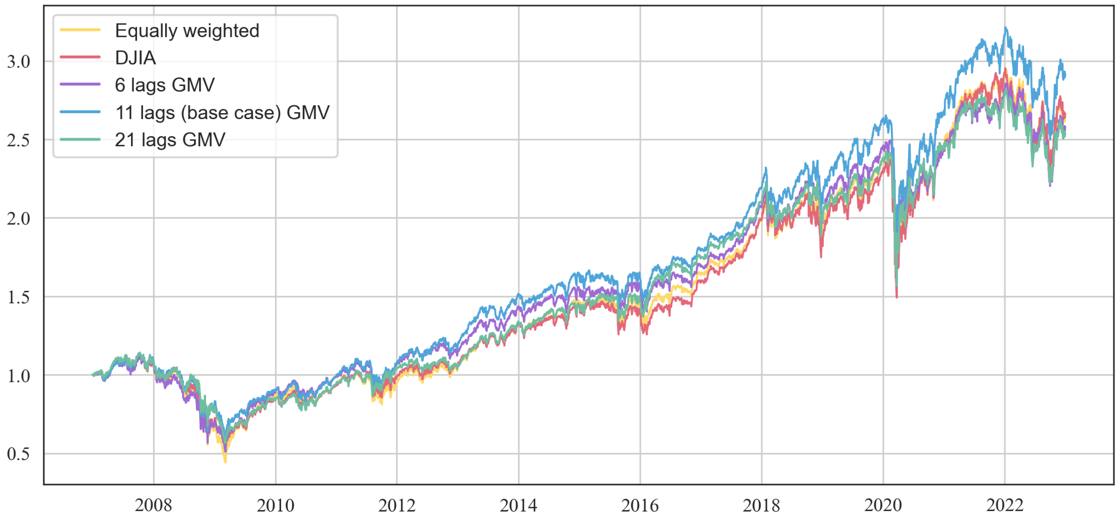
A similar pattern as in GMIR portfolios applies to the GMV portfolio, which can be seen by analysing Table 23 and Figure 20. The best value was still 11 lags, and such a strategy still outperformed both benchmarks (but by less than in the case of the GMIR portfolio). For 6 and 21 lags, information ratios were at a similar level to benchmarks.

Table 24. Performance statistics of the FMV-XGB GMV strategies with different feature lags.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
6 lags GMV	157,57%	6,09%	18,97%	54,47%	0,321	0,036	12,88	17,55
11 lags (b.c.) GMV	**190,98%**	**6,90%**	18,76%	**47,50%**	**0,368**	**0,053**	12,81	17,44
21 lags GMV	153,78%	5,99%	**18,34%**	50,61%	0,327	0,039	12,20	16,70

Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 6, 11, 21, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio.

Figure 21. Equity lines of the GMV strategies with different feature lags.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 6, 11, 21, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

4.8. XGBoost hiperparameters

XGBoost is a model that can be tuned with many parameters. In this section, we show how changes to the different parameters affected strategy performance.

Global maximum information ratio portfolio

Analysing Table 25 and Figures 22-26, we can conclude that the selection of XGBoost hyperparameters did indeed significantly affect the results. In this section, we will only compare information ratios**. When it comes to `nrounds` hyperparameter, we can see that results did not differ that much for different values. All strategies outperformed the benchmarks. The model probably converged around the 50th round, and additional iterations were not needed (in fact, they slightly decreased the information ratio**). For the learning rate (`eta`) hyperparameter, the information ratio approached its maximum for the 0,3 value. Lower values (0,05 and 0,1) minimally outperformed the benchmarks, and the higher value (0,4) was comparable to them. When it comes to the `lambda` hyperparameter, the results were best for the value of

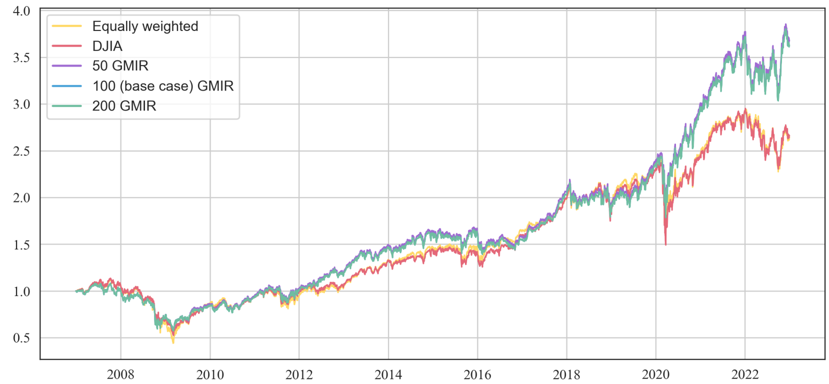
1. The strategies with `lambda` of 0,5 and 10 had a lower information ratio** than the benchmarks, and the strategy with `lambda` 5 had a comparable information ratio to benchmarks. The best `max-depth` hyperparameters in our dataset were 6 and 12. The values of 3 and 9 underperformed. For the `subsample` hyperparameter, the higher its value was, the better the strategies performed.

Table 25. Performance statistics of the FMV-XGB GMV strategies with different values of XGBoost parameters.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
nrounds								
50 GMIR	**267,96%**	**8,48%**	19,02%	46,87%	**0,446**	**0,081**	5,99	8,69
100 (b.c.) GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68
200 GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68
eta								
0,05 GMIR	134,08%	5,46%	**18,09%**	**38,84%**	0,302	0,042	6,00	9,44
0,1 GMIR	165,58%	6,29%	19,51%	50,35%	0,323	0,040	5,83	8,63
0,3 (b.c.) GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68
0,4 GMIR	155,86%	6,05%	19,57%	55,71%	0,309	0,034	6,24	8,96
lambda								
0,5 GMIR	125,52%	5,21%	19,13%	55,19%	0,273	0,026	6,01	8,80
1 (b.c.) GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68
5 GMIR	128,07%	5,29%	19,44%	43,07%	0,272	0,033	5,82	8,31
10 GMIR	66,70%	3,25%	19,89%	42,16%	0,163	0,013	5,63	8,10
max-depth								
3 GMIR	115,50%	4,92%	18,86%	43,84%	0,261	0,029	6,71	9,64
6 (b.c.) GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68
9 GMIR	42,55%	2,24%	19,88%	50,57%	0,113	0,005	4,64	6,64
12 GMIR	253,62%	8,21%	22,30%	46,06%	0,368	0,066	3,53	5,01
subsample								
0,5 GMIR	80,58%	3,76%	19,82%	51,83%	0,190	0,014	5,73	8,18
0,75 GMIR	170,84%	6,43%	19,66%	55,11%	0,327	0,038	5,91	8,59
1 (b.c.) GMIR	262,63%	8,38%	19,02%	46,97%	0,441	0,079	5,97	8,68

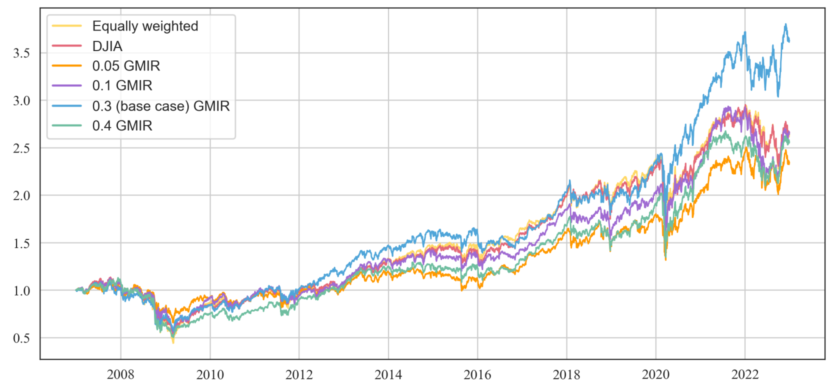
Note: If one hyperparameter was the subject of the sensitivity analysis, then other parameters were at the default level (in bold). Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 50, **100**, 200, `eta` = 0,05; 0,1; **0.3**; 0,4, `lambda` = 0,5; **1**; 5; 10, `max-depth` = 3, **6**, 9, 12, `subsample` = 0,5; 0,75; **1**. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 22. Equity lines of the FMV-XGB GMIR strategies with different values of `nrounds` parameter.



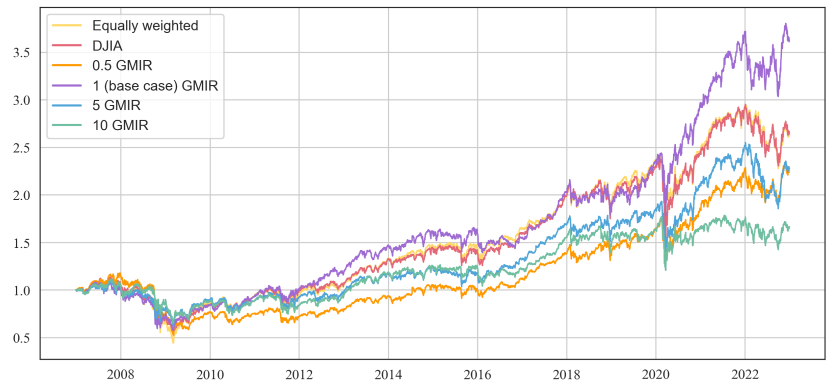
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 50, 100, 200, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Figure 23. Equity lines of the FMV-XGB GMIR strategies with different values of `eta` parameter.



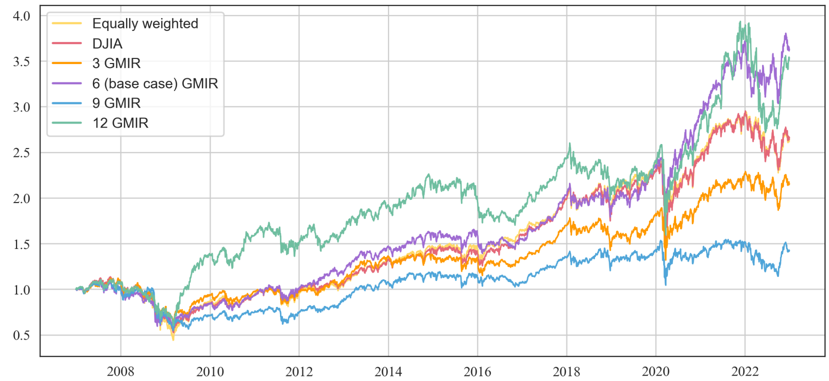
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,05; 0,1; 0.3; 0,4, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Figure 24. Equity lines of the FMV-XGB GMIR strategies with different values of `lambda` parameter.



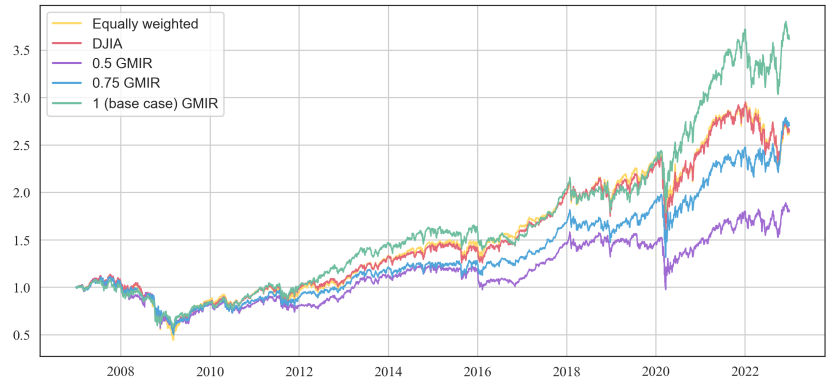
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 0,5; 1; 5; 10, `max-depth` = 6, `subsample` = 1.

Figure 25. Equity lines of the FMV-XGB GMIR strategies with different values of `max-depth` parameter.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,3, `lambda` = 1, `max-depth` = 3, 6, 9, 12, `subsample` = 1.

Figure 26. Equity lines of the FMV-XGB GMIR strategies with different values of `subsample` parameter.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,3, `lambda` = 1, `max-depth` = 6, `subsample` = 0,5; 0,75; 1.

Global minimum variance portfolio

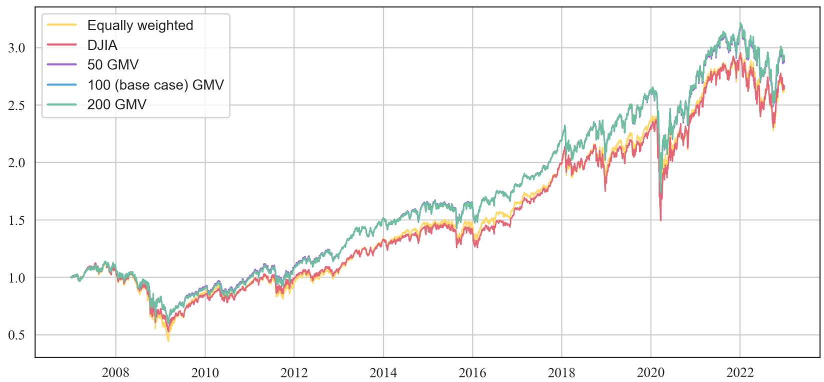
Analysing Table 26 and Figures 27-31, we can see that our assumption of the significance of the hyperparameter choice still holds for GMV portfolios. In this section, the same as previously, we will only compare information ratios**. For different values of `nrounds` hyperparameter, we can see that results still did not differ that much. All strategies outperformed the benchmarks. The model likely converged around the 100th round, and additional iterations did not improve the performance. For the learning rate (`eta`) hyperparameter, the information ratio approached its maximum for the lowest value (0,05). The values of 0,1 and 0,3 outperformed the benchmarks, and the higher value (0,4) was comparable to them. Concerning the `lambda` hyperparameter, the results were best for the values of 1 and 5. The strategy with lambda of 10 was similar to the better one of the benchmarks and the lambda of 0,5 had the lowest information ratio** (much lower than benchmarks). The best `max-depth` hyperparameter in our dataset was 6. The values of 3 and 9 performed comparably to benchmarks. The best `subsample` hyperparameter value in our case was 1. The second best value was 0,5 (similar information ratio** to benchmarks). For the value of 0,75, the strategy performed the worst.

Table 26. Performance statistics of the FMV-XGB GMV strategies with different values of XG-Boost parameters.

Strategy	$ARR\%$	$ARC\%$	$ASD\%$	$MDD\%$	IR^{*}	IR^{**}	$MN^{75\%}$	$MN^{90\%}$
Equally weighted	164,21%	6,26%	20,84%	60,82%	0,300	0,031	22,84	27,00
DJIA	165,96%	6,30%	19,55%	53,78%	0,323	0,038	-	-
nrounds								
50 GMV	188,46%	6,85%	18,79%	47,64%	0,364	0,052	12,74	17,34
100 (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
200 GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
eta								
0,05 GMV	**207,13%**	**7,26%**	18,92%	51,68%	**0,384**	**0,054**	15,96	20,41
0,1 GMV	175,03%	6,53%	18,53%	53,08%	0,352	0,043	13,93	18,65
0,3 (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
0,4 GMV	155,86%	6,05%	19,57%	55,71%	0,309	0,034	6,24	8,96
lambda								
0,5 GMV	91,71%	4,15%	18,58%	56,12%	0,223	0,017	13,11	17,86
1 (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
5 GMV	180,67%	6,66%	18,45%	**45,56%**	0,361	0,053	12,14	16,65
10 GMV	149,43%	5,88%	18,46%	49,22%	0,319	0,038	11,94	16,43
max-depth								
3 GMV	162,67%	6,22%	18,41%	54,31%	0,338	0,039	12,78	17,44
6 (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44
9 GMV	141,26%	5,66%	**18,32%**	50,35%	0,309	0,035	11,77	16,29
12 GMV	116,94%	4,96%	19,11%	56,40%	0,260	0,023	10,84	15,11
subsample								
0,5 GMV	143,93%	5,73%	18,53%	53,43%	0,309	0,033	10,80	15,22
0,75 GMV	120,60%	5,07%	18,55%	54,28%	0,273	0,026	12,07	16,66
1 (b.c.) GMV	190,98%	6,90%	18,76%	47,50%	0,368	0,053	12,81	17,44

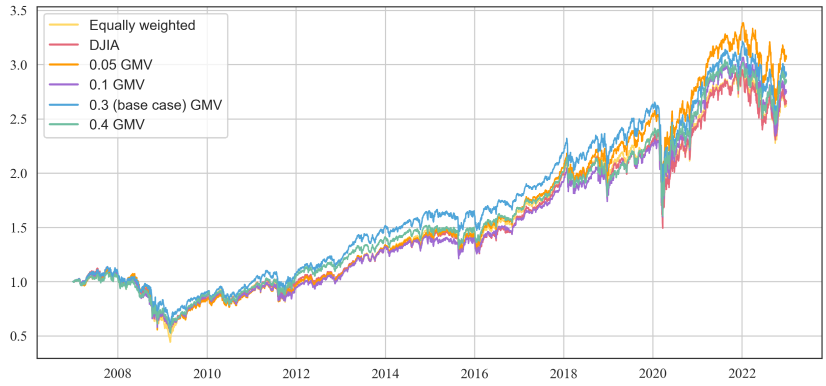
Note: If one hyperparameter was the subject of the sensitivity analysis, then other parameters were at the default level (in bold). Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 50, **100**, 200, `eta` = 0,05; 0,1; **0.3**; 0,4, `lambda` = 0,5; **1**; 5; 10, `max-depth` = 3, **6**, 9, 12, `subsample` = 0,5; 0,75; **1**. Performance and diversification metrics: $ARR\%$ – absolute rate of return, $ARC\%$ – annualized rate of return, $ASD\%$ – annualized standard deviation, $MDD\%$ – maximum drawdown, IR^{*} – information ratio*, IR^{**} – information ratio**, $MN^{75\%}$ – mean number of stocks constituting at least 75% of the portfolio, $MN^{90\%}$ – . . . 90% of the portfolio, EW - equally weighted

Figure 27. Equity lines of the FMV-XGB GMV strategies with different values of `nrounds` parameter.



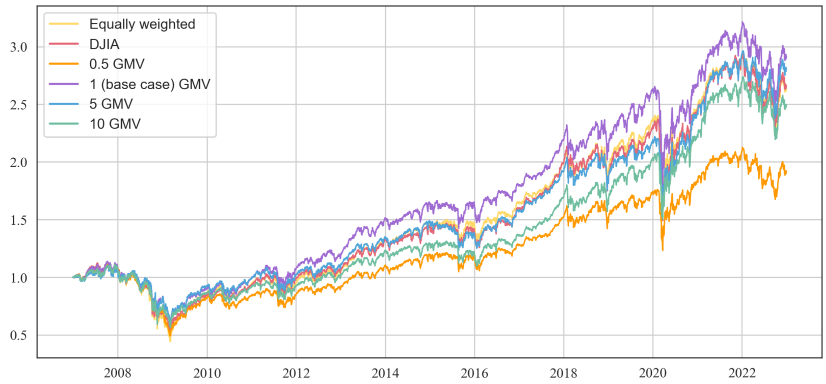
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 50, 100, 200, `eta` = 0.3, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Figure 28. Equity lines of the FMV-XGB GMV strategies with different values of `eta` parameter.



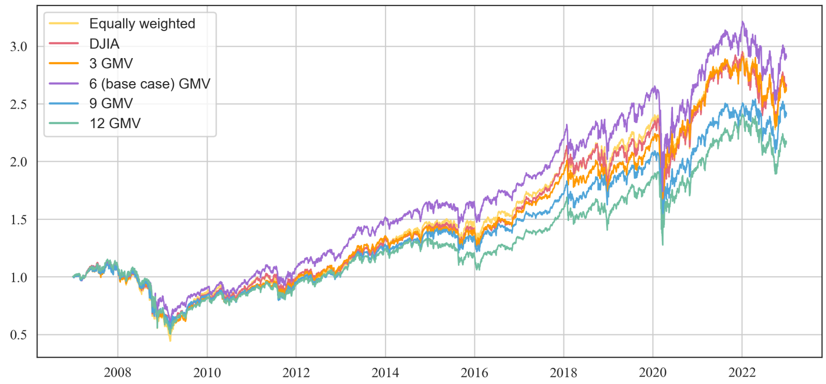
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,05; 0,1; 0.3; 0,4, `lambda` = 1, `max-depth` = 6, `subsample` = 1.

Figure 29. Equity lines of the FMV-XGB GMV strategies with different values of `lambda` parameter.



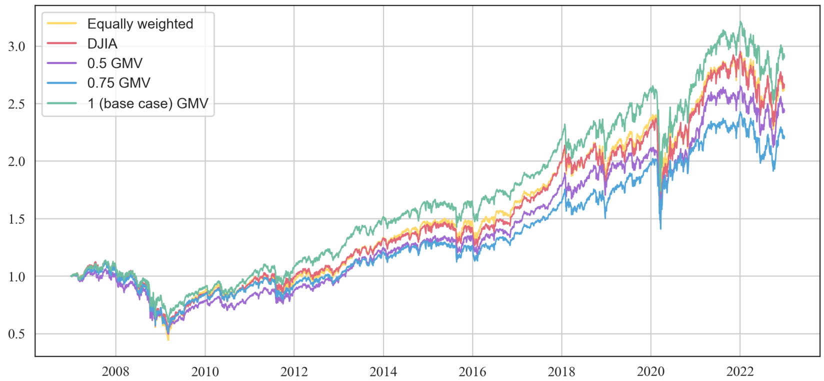
Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0.3, `lambda` = 0,5; 1; 5; 10, `max-depth` = 6, `subsample` = 1.

Figure 30. Equity lines of the FMV-XGB GMV strategies with different values of `max-depth` parameter.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,3, `lambda` = 1, `max-depth` = 3, 6, 9, 12, `subsample` = 1.

Figure 31. Equity lines of the FMV-XGB GMV strategies with different values of `subsample` parameter.



Note: Testing period: 2007-01-01 - 2022-12-31. Strategies' parameters: estimation window: 12 months, rebalancing period: 1 month, number of stocks: 30, number of lags: 11, hyperparameters: `nrounds` = 100, `eta` = 0,3, `lambda` = 1, `max-depth` = 6, `subsample` = 0,5; 0,75; 1.

The sensitivity analysis showed that, in most cases, the results were not robust to the change of parameters. Some of the parameters had a lower impact on the strategies information ratios** (e.g. the `nrounds` parameter in the case of the FMV-XGB portfolios), and some did significantly alter the results (e.g. the estimation window length for the MV GMIR and FMV-AG GMIR strategies). The only portfolios robust to parameter changes were the FMV-AG GMV ones because they resembled equally weighted portfolios. It is difficult to find the optimal parameters for the FMV and MV strategies especially that it seems that we cannot choose them independently. In our study, due to computation power constraints, we could change only one parameter at a time relative to the base case. The better approach would be to test many combinations of different parameters. Then, it is possible that on other data sets, the parameters that yielded the best results in our study could have different performance, but it will be the subject of another study.

CONCLUSIONS

The main objective of this thesis was to check if it is beneficial to forecast stock returns before using them in the mean-variance portfolio optimization strategy. Two hypotheses were formulated:
H1: *The strategies based on forecasted stock returns outperform (having higher values of information ratio**) the strategies based on historical stock returns.*
H2: *The strategies based on forecasted stock returns outperform (having higher values of information ratio**) an equally weighted portfolio and buy and hold on the equity index.*
Moreover, three research questions have been asked:
RQ1: *Which portfolio optimization method will perform better in terms of information ratio**?*
RQ2: *Which forecasting model performs better in the framework in terms of information ratio**?*
RQ3: *Are the results sensitive to the number of assets, the estimation and the rebalancing windows' lengths, the transaction costs, and the forecasting models' parameters?*

To test the stated hypotheses and answer the research questions, we conducted empirical research. We created 152 strategies divided into three groups. These groups were: strategies using "raw" historical returns, strategies using the ARIMA-GARCH forecasts, and strategies using the XGBoost forecasts. We tested these strategies on Dow Jones Industrial Average stocks between 1 January 2007 and 31 December 2022, covering two recent financial crises. We compared these strategies to two benchmarks – an equally weighted portfolio and buy and hold on the DJIA index.

In the GMIR base case scenario, the proposed strategies outperformed both the benchmarks and the MV strategy in terms of information ratio**. In the GMV base case, the FMV-XGB strategy outperformed the benchmarks but did not outperform the MV strategy. We also performed two statistical tests to test whether the expected values of FMV strategies return distributions were significantly higher than benchmarks. At the 0.1 significance level, one test showed that FMV-XGB GMIR returns were higher than equally weighted and DJIA buy-and-hold portfolios. Other tests performed on other pairs showed a lack of statistical significance. Moreover, analysing the sensitivity analysis results, it is fair to conclude that the choice of parameters highly impacted the results of each strategy.

We reject the first hypothesis that the FMV strategies outperform the MV strategies. There was clear evidence of FMV strategies yielding better outcomes in certain situations, but it is also easy to find contradictive examples. The sensitivity analysis raised further questions. There were undoubtedly circumstances where forecasting stock returns might have resulted in better portfolio performance, but it was not a universal solution. The second hypothesis, stating that the FMV portfolios outperform the benchmarks, should also be rejected. On average, strategies involving forecasting returns had similar results to the benchmarks. Again, we found situations in which the FMV strategies outperformed benchmarks by far, but there were also circumstances where FMV strategies generated losses.

Answering the first research question, we cannot say whether the Global Maximum Information Ratio or the Global Minimum Variance portfolios perform better. Under different circumstances, one or the other may perform better. A clear pattern emerges when we look at GMV Portfolios of the FMV-AG strategy. In most cases, they converged to the equally weighted portfolio, no matter which parameters were chosen. This observation suggests that the variance-covariance matrix derived from the ARIMA-GARCH forecasts promotes diversification. Also, no forecasting model was clearly better than the other. Similarly, both models had evidence of strong and weak performance, which answers the second research question.

Concerning our third research question, the sensitivity analysis showed that each strategy is highly sensitive to the change in model parameters. Reaction to even a slight adjustment of certain parameters can result in a significant change in strategy outcomes. The only notable exceptions were the FMV-AG GMV portfolios mentioned above. We were able to find parameter combinations that had relatively satisfactory results, but it would not be possible to choose them beforehand. There is also no guarantee that such combinations would perform well on the other data.

In further research, it would be beneficial to check how grid-searching optimal parameters for each strategy would affect their performance. It would be especially important for the forecasting model parameters (eg. XGBoost hyperparameters or ARIMA-GARCH order). Minimizing the forecasting error can result in better-optimized portfolios that yield better returns. As a result of restricted computing capacity and an enormous number of fitted models, we were unable to undertake any form of tuning. Additionally, it is worth noting that while ARIMA-GARCH and XGBoost are popular examples, there is a broad range of both traditional econometric and modern machine-learning models available. More could be tested, to check whether there is one that performs the best in this specific application.

REFERENCES

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Box, G. E., & Jenkins, G. (1976). *Time Series Analysis, Forecasting and Control*.

Chen, C., & Zhou, Y.-s. (2018). Robust multiobjective portfolio with higher moments. *Expert Systems with Applications*, *100*, 165–181.

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). New York, NY, USA: ACM.

Chen, W., Zhang, H., Mehlawat, M. K., & Jia, L. (2021). Mean–variance portfolio optimization using machine learning-based stock price prediction. *Applied Soft Computing*, *100*, 106943.

Devore, J. L., Berk, K. N., & Carlton, M. A. (2012). *Modern mathematical statistics with applications* (Vol. 285). Springer.

Engle, R. (2001). GARCH 101: The use of ARCH/GARCH models in applied econometrics. *Journal of economic perspectives*, *15*(4), 157–168.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica: Journal of the econometric society*, 987–1007.

Fabozzi, F. J., Gupta, F., & Markowitz, H. M. (2002). The legacy of modern portfolio theory. *The journal of investing*, *11*(3), 7–22.

Ghalanos, A. (2022). rugarch: Univariate GARCH models. [Computer software manual]. (R package version 1.4-9.)

Harvey, C. R., Liechty, J. C., Liechty, M. W., & Müller, P. (2010). Portfolio selection with higher moments. *Quantitative Finance*, *10*(5), 469–485.

Jabeur, S. B., Mefteh-Wali, S., & Viviani, J.-L. (2021). Forecasting gold price with the XGBoost algorithm and SHAP interaction values. *Annals of Operations Research*, 1–21.

Ma, Y., Han, R., & Wang, W. (2021). Portfolio optimization with return prediction using deep learning and machine learning. *Expert Systems with Applications*, *165*, 113973.

Markowitz, H. (1952). The utility of wealth. *Journal of political Economy*, *60*(2), 151–158.

Mohammadi, H., & Su, L. (2010). International evidence on crude oil price dynamics: Applications of ARIMA-GARCH models. *Energy Economics*, *32*(5), 1001–1008.

Mustapa, F. H., & Ismail, M. T. (2019). Modelling and forecasting S&P 500 stock prices using hybrid Arima-Garch Model. In *Journal of physics: Conference series* (Vol. 1366, p. 012130).

Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., & Salwana, E. (2020). Deep learning for stock market prediction. *Entropy*, *22*(8), 840.

Nobre, J., & Neves, R. F. (2019). Combining principal component analysis, discrete wavelet transform and XGBoost to trade in the financial markets. *Expert Systems with Applications*, *125*, 181–194.

Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics & Data Analysis*, *50*(9), 2293–2312.

Tsay, R. S. (2005). *Analysis of financial time series*. John Wiley & Sons.

Vo, N., & Ślepaczuk, R. (2022). Applying hybrid ARIMA-SGARCH in algorithmic investment strategies on S&P500 Index. *Entropy*, *24*(2), 158.

Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Cengage learning.

Yaziz, S., Azizan, N., Zakaria, R., & Ahmad, M. (2013). The performance of hybrid ARIMA-GARCH modeling in forecasting gold price. In *20th international congress on modelling and simulation, adelaide* (pp. 1–6).

Yun, K. K., Yoon, S. W., & Won, D. (2021). Prediction of stock price direction using a hybrid GA-XGBoost algorithm with a three-stage feature engineering process. *Expert Systems with Applications*, *186*, 115716.

University of Warsaw
Faculty of Economic Sciences
44/50 Długa St.
00-241 Warsaw
www.wne.uw.edu.pl