Handelshøyskolen BI

GRA 19703 Master Thesis

Thesis Master of Science 100% - W

Predefinert informasjon

Startdato:	09-01-2023 09:00 CET	**Termin:**	202310
Sluttdato:	03-07-2023 12:00 CEST	**Vurderingsform:**	Norsk 6-trinns skala (A-F)
Eksamensform:	T		
Flowkode:	202310||11184||IN00||W||T		
Intern sensor:	(Anonymisert)		

Deltaker

Navn:	Rikke Nordvoll og Marte Nystad Rise

Informasjon fra deltaker

Tittel *:	Exploring the drivers and prediction of fund flow in the Norwegian Bond Fund Market: Deploying classic statistics and machine learning methods
Navn på veileder *:	Paolo Giordani

Inneholder besvarelsen konfidensielt materiale?:	Nei	**Kan besvarelsen offentliggjøres?:**	Ja

Gruppe

Gruppenavn:	(Anonymisert)
Gruppenummer:	313
Andre medlemmer i gruppen:	

BI Norwegian Business School
Oslo, Spring 2023

Exploring the drivers and prediction of fund flow in the Norwegian Bond Fund Market: Deploying classic statistics and machine learning methods

Master Thesis

Marte Rise and Rikke Nordvoll
Supervisor: Paolo Giordani

Master of Science in Business Analytics
BI NORWEGIAN BUSINESS SCHOOL

Oslo, July 3, 2023

Abstract

This paper aims to analyze the drivers and predict flow of Norwegian bond funds one month in advance using machine learning models and macro variables. The study investigates a sample of 79 bond funds and a subset of medium credit risk funds. We utilize XGBoost feature importance to investigate the factors that influence net flow of Norwegian bond funds. Additionally, through the use of OLS regression, we address our hypothesis concerning the significant impact of performance and macro variables on net flow of the Norwegian bond fund market. Our findings reveal only a significant negative linear relationship between change EUR/NOK and Norwegian bond funds. To harness the capabilities of machine learning models, partial dependency plots are also examined in search for non-linear relationships. XGBoost reveals non-linear relationship among predicted net flow and changes in VIX, the models (XGBoost and MLP) show varying impacts of change EUR/NOK, and contradiction patterns in lagged net flow. Due to poor accuracy scores across prediction models, we are unable to achieve effective models for predicting bond fund flow one month ahead using top selected features.

Acknowledgements

First of all, we would like to thank Paolo Giordani, our supervisor, for his advice and helpful feedback on our thesis. Secondly, we want to thank Norwegian Fund and Asset Management Association (VFF) for providing us with the data on Norwegian bond funds. Finally, we would like to express our gratitude to Thomas Eitzen, Chief Analyst of Fixed Income at SEB, for his suggestion to explore capital flow of Norwegian bond funds.

Content

Introduction

In recent years there has been growing attention to the topic of fund flow as it offers valuable insights into investor behaviour, market dynamics and performance evaluation. Financial institutions, including the largest corporate bank in Norway, SEB, have shown an interest in understanding the flow of Norwegian bond funds. In Norway, fund data is published with a one-month lag by Norwegian Fund and Asset Management Association (VFF), making it relevant to predict fund flows before the official release date. This could be interesting for financial institutions and investors in general. By utilizing this information in conjunction with other models and their expertise, they can gain understanding of the markets they operate in, while other investors can make more informed investment decisions. Machine learning models are known to be effective for prediction and is fairly unexplored in work contexts. This serves as the motivation behind our question of whether we can predict bond fund flow in an efficient manner using machine learning methods.

Existing international literature on fund flow focuses on performance and its impact on capital flow, primarily with the perspective of equity funds. Studies that specifically examine bond funds tend to concentrate on performance measures such as raw returns and risk returns. Some papers have found significant relationships using bond indices, changes in VIX and interest rates. Findings of these papers, along with the investor sentiment theories of risk tolerance and the tendency of investors to chase returns, motivates our inclusion of factors reflecting conditions in other financial markets as explanatory variables. Thus, these formal observations have led us to the question whether these affects also apply to the capital flow of Norwegian bond market. Accordingly, this is forming our hypothesis that cash flow into and out of bond fund market is significantly affected by performance and macro variables.

The objective of this thesis is to analyze the drivers and predict Norwegian bond fund flow one month ahead. With limited research on variables effect on Norwegian bond fund flow, we utilize a comprehensive set of 17 measurements that reflect the conditions of various financial markets. Through the application of XGBoost, we identify the top five features with the strongest relationship with net flow of Norwegian bond funds. Furthermore, we investigate the quantifiable relationship between these selected features and net flow employing OLS regression. Considering that OLS assumes linear relationships, we explore whether machine learning models can improve prediction accuracy by addressing non-linearities. We will answer this by comparing accuracy scores across OLS, XGBoost and MLP. Additionally, we utilize partial dependence plots to gain deeper insights into the drivers of the machine learning models predictions and understand the reasons for their potential differences. Our analysis is conducted on two samples; one includes 79 bond funds with different credit risk levels, while the other comprises 29 bond funds classified as medium credit risk.

In our results, we observe a discrepancy in top features between bond fund and bond fund with medium credit risks indicates the presence of distinct of explanatory variables depending on the credit risk profile. However, we uncover little evidence that there is significant relationship between bond fund flow and measurements reflecting the conditions in other financial markets. We only find a significant linear negative relationship with flow of capital for bond funds and change in EUR/NOK. For funds with medium credit risk, no clear linear relationships are identified. Regarding the prediction using machine learning methods, our findings suggest that we are unable to achieve an effective model for predicting bond fund flows one month ahead using the top selected features. Further insights from the partial dependency plots illustrate that our models use different signals in their predictions. Particularly noteworthy, XGBoost highlights non-linear relationship with change VIX, the models show varying impact of change in EUR/NOK and contrasting patterns in lagged net flow.

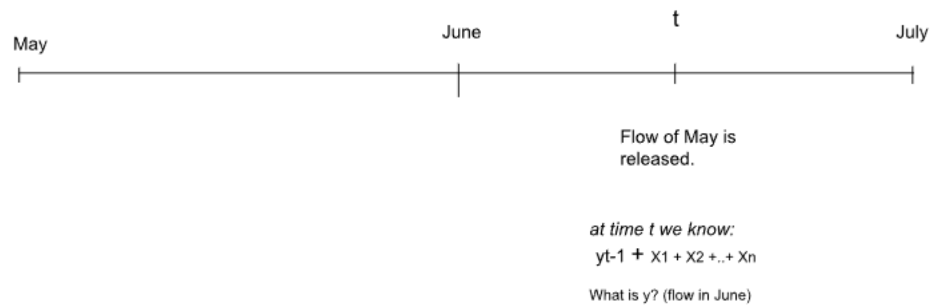
This thesis is organized by the following sections. First, we will describe our motivation for the chosen topic of interest. Secondly, we present a literature review. Third, we go through data collection and sample description. In the fourth section, we provide the used methodology and theoretical content we found suitable for this thesis. In the fifth part, we present the results from the research and complement it with a discussion related to limitations of our study. The last, and final section contains a conclusion with suggestive alternatives for future research related to our topic.

Motivation

In discussions with Thomas Eitzen, Chief Analyst of Fixed income at SEB Markets, Norway's largest corporate bank, there is an increasing interest in understanding the capital flow of Norwegian bond funds. Recent research has indicated a relationship between fund flow and fund performance as well as macro variables, studying other geographical areas. Hence, it is of interest to see if some of this also holds for the Norwegian bond fund market.

For SEB, having a predictive model capable of forecasting Norwegian bond funds one month in advance would provide practical benefits. In Norway, fund flow data is reported on a monthly basis, although one month lagged. Figure 1 illustrates the release structure of Norwegian bond fund reporting.

Figure 1 – Structure of Norwegian bond fund reporting



By developing a predictive model that considers the conditions of other financial markets, SEB would gain valuable market insights. Business vise this would enhance their ability to provide informed advice on this particular topic. Machine learning models is relatively unfamiliar field among professionals working in the banking industry, but they are becoming increasingly relevant. This motivates our question of whether we can predict bond fund flow in an efficient manner using machine learning methods.

Challenges

There are several challenges and possible drawbacks to doing this. Firstly, it is difficult to find trustworthy factors that may be used for precise forecasts because of the limitation of study in this area. Furthermore, there is a limited amount of data available on Norwegian bond funds. In addition, there might be some biases in the fund flow data available that are difficult to address.

To summarize, there is no doubt that money flows into bond funds is an interesting topic for both financial institutions and investors in general. International authors and theory on investor sentiment also support our chosen measures under investigation for our models, but not studied much on a Norwegian bond market. Particularly the link between macro variables and the Norwegian bond funds is little studied. In addition, as modern machine learning methods have become more popular due to real time usage this is also becoming a relevant method for financial time series data. Together, this builds a convincing case for our thesis topic.

Literature Review

To further supplement the previous section, this section provides a more detailed overview of previous theories in the field.

Few have looked into the drivers and prediction of Norwegian bond fund flow. However, there is extensive research examining mutual fund flow with the perspective of equity funds. (Warther, 1995) finding a positive relationship between flows and aggregated subsequent returns. (Shrider, 2009) look at redemptions and purchases instead of net flow and find evidence that raw returns and risk adjusted returns are important for flows into mutual funds. (Ippolito, 1992) finds a positive relationship between fund growth and recent performance. (Sirri & Tufano, 1998), find a convex-performance flow relationship in equity funds, i.e as performance of a fund increases, the flow of capital gained from each additional return also increases, while poor performance does not have an outflow effect. (James & Karceski, 2006) look into drivers of flow of retail and institutional funds and find that institutional fund flow is less sensitive to past raw return compared to retail fund flows, but more sensitive to risk adjusted performance. (Gruber, 1996) show that return and excess return predict cash flow, defining cash flow as change in total net assets. (W. Chen, n.d.) examine equity sentiment, using cash flow and its effect on bond returns, and find a negative relationship.

(Y. Chen & Qin, 2017; Edwards & Zhang, 1998; Grose, n.d.; Zhao, 2005) and (Kopsch et al., 2015) are among the few that have studied flow of bond funds. (Y. Chen & Qin, 2017) study US corporate bond funds with data from 1991-2014.

Similarly, as with equity funds, they show that corporate bond fund flows chase recent returns, but not in a convex manner. Rather, they demonstrate that poor performance leads to the same amount of capital outflow as good performance leads to capital inflow. Their results argue that funds experiencing subsequent inflow outperform funds with outflow, presumably because of performance persistence. In their study they also look at significant relationships between corporate fund flow and macro variables. Particularly, they find that flow to corporate bonds is positively related to recent VIX, return bond index, default spread and a negative relationship with T-bill rate. In contrast to (Y. Chen & Qin, 2017) and (Edwards & Zhang, 1998), who do not find that bond net sales have had an impact on bond returns. (Zhao, 2005) study the determinants of retail bond fund and bond funds with different investment objectives. The study focuses on measurements such as sharp ratio, fund size and maturity, finding that investors chase risk adjusted performance leaders instead of raw return leaders.

(De Lange et al., 2018) investigates the supply and demand dynamics of the Norwegian bond market. They focus on determinants for credit spreads and find that equity volatility VIX only matters when spread movements are above their average and the yield curve is not significant. In addition, they explain how change in EUR/NOK could be related to the Norwegian bond fund market. Even though they look at credit spread rather than flow, their results are of interest in understanding the Norwegian bond fund market and hence also to our analysis.

Some papers have investigated bond fund flow in smaller markets. (Grose, n.d.) investigate determinants of capital flow in and out of Greek bond funds and find that risk-weighted returns and not high mean returns are important drivers for capital flows. (Kopsch et al., 2015) study the determinants of aggregate fund flow in equity and hybrid funds in the Swedish economy and find some evidence for change in VIX having a negative relationship with fund flow, and fund flow to have a positive relationship with exchange rate USD/SEK.

A popular paper (Da et al., 2013) argue that investor sentiment can be directly measured through the Internet search behavior of households. Using internet search queries with the keywords, “recession, unemployment and bankruptcy” as a proxy for fear, they find fear predicts fund flows out of equity funds and into bond funds. This is an interesting finding supporting investor sentiment theory about risk appetite and diversification. As risk increases investors might seek to reduce their risk exposure. Presumably by going from equity to bond investments. This is highly relevant for our thesis as it supports investor sentiment theory that risk drives the choice of an investor's allocation across financial markets.

The findings of existing research are a formal foundation for connecting bond fund flow to performance and macro variables.

In terms of forecasting the author (Fama & French, 1989) uses Fama French regression both in and out of sample for forecasting bond returns. (Yue et al., 2021) proposes a fund flow prediction model using ARIMA model based on the historical purchase and redemption of individual equity funds. We build on this by investigating if market conditions at time t can be used in forecasting Norwegian bond funds using machine learning models.

To our knowledge we build on the existing literature by investigating the Norwegian bond fund flow to variables reflecting conditions in other financial markets and macro variables, as well as utilizing machine learning techniques. Based on related theory and the motivation behind the thesis topic, there is a convincing case for our thesis.

Data and sample description

In this section we will present our dependent variable, our choice of explanatory variables, as well as a description of our sample of data. Our primary data sources are The Norwegian Fund Asset Management Association (VFF), Bloomberg, Infront and Yahoo Finance.

The dependent variable (y) for our study will be log percentage flow at time t. For our independent variables we first turn to relevant theory and then use a feature selection method to obtain the most relevant features to reduce noise and create a simple to use and accurate model. The method for feature selection is described in more detail during the next section. Previous research and well established theories on investor sentiment such as risk appetite, sector rotation and asset class preferences are the foundation for our choice of variables. The overall intuition is that investors have the option to allocate their money into different financial markets and that this decision is determined on the performance and conditions of the markets at the allocation time. Thus, our focus will be on creating a model of flow of capital going into and out of Norwegian bond fund market using measurements that reflect conditions in other financial markets. In our main dataset, there are 232 observations of data for each variable, which cover the period from May/June 2003 to August 2022 in a monthly frequency.

Dependent variable (y)

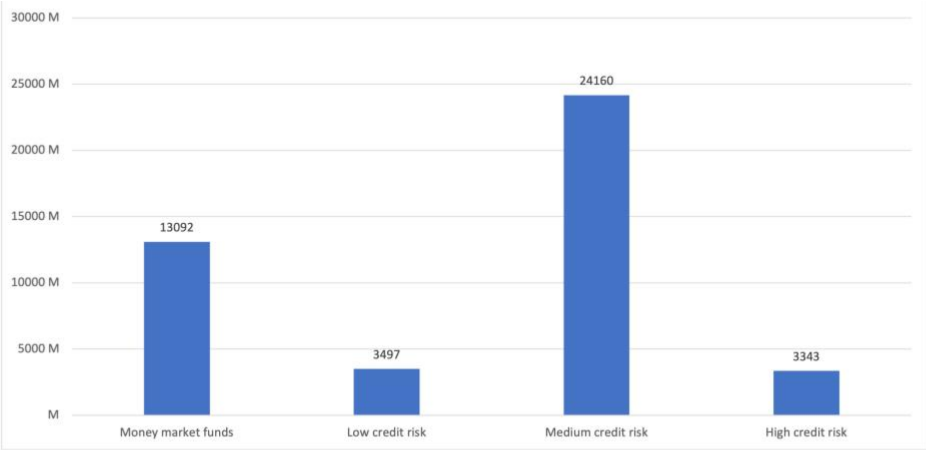
Average log percentage flow of Norwegian bond funds

Data on Norwegian fixed income funds flows were provided by The Norwegian Fund Asset Management Association (VFF). We were provided with historical data in 3 separate excel sheets, 2003-2010,2011-2018,2018-2022, respectively. This included specifications on individual funds subscription, redemptions, net flow and total assets under management (AuM) each month. These were merged and structured for a tabular data analysis. We then selected those funds that had absolute observations in net flow and aum for the time period 2003-2022. We are left with a dataset on Norwegian fund flow consisting of 81 active fixed income funds with different time series lengths.

To account for potential name changes of funds during the sample period, we utilize the ISIN number as the identifier in our research. Additionally, it is worth noting that the raw dataset includes instances where funds have a net flow approximately equal to the AuM at the start of the time series. This can result in net flow values which represent cumulative flow at establishment rather than monthly flow, contradicting the purpose of our analysis. We thus remove these rows. Further 2 funds, Skagen Avkastning (ISIN:NO0008000452) and Eika OMF (ISIN:NO0010479066), had incidents of unnaturally large net flows, in a large extent during the time series and thus were excluded from the analysis.

After finding the most active funds Thomas Eitzen classified them based on their credit risk level (low, medium, high) and whether they were money market funds. This classification is relevant for analyzing if there are different drivers influencing fund flow depending on the risk profiles of the funds. In our analysis we limit our analysis to the medium credit risk bond fund. Figure 2 shows our sample divided into their total AuM for the entire time period, reflecting the market size of the respective groups.

Figure 2 - Distribution of total AuM for each bond fund subgroup



To construct our dependent variable of interest at aggregate level, we construct the market weighted percent flow from the individual funds flow to be consistent with other papers. $flow_t \ = \frac{Net\ Flow_{i,t}}{AuM_{i,t-1}} * \frac{AuM_{i,t}}{\sum_{i=0}^{i} AuMi,t}$

In our case market weighted aggregation is important as our data set includes both small, medium and large size funds. In comparison to an outflow of 10% from a large fund, a 10% outflow from a small fund does not have the same impact on the entire market. In order to ensure that our data is as precise as possible, it is crucial that this be taken into account in the model. After this process we are left with market weighted fund flow using our 79 funds as an example, to represent the flow of capital for the Norwegian bond funds market. After the aggregation we make it into logarithmic form. Figure 3 and 4 shows the time series of our dependent variable both as growth and cumulative growth.

Figure 3 – Market weight net flow Norwegian bond funds

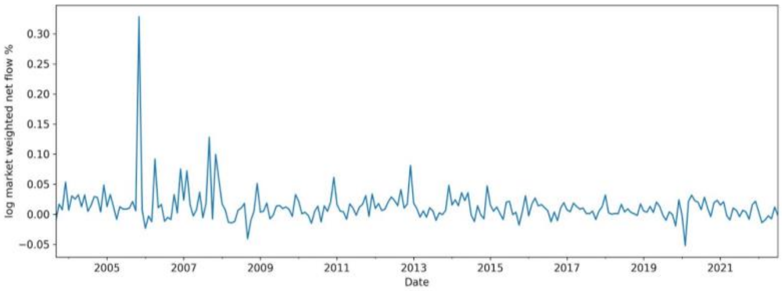
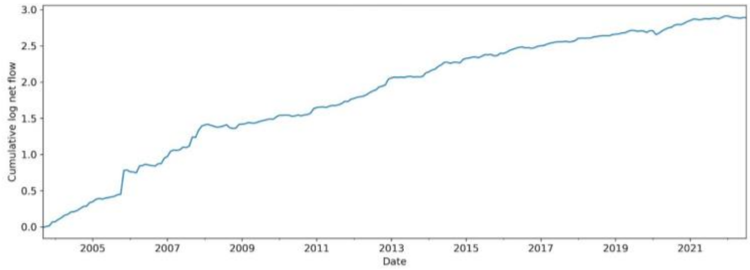


Figure 4 – Market weight cumulative net flow Norwegian bond funds



Independent variables (X)

According to (Alexakis et al., 2005) stocks, bonds and cash equivalents make up the three main asset classes. There are several researchers that have researched mechanisms between financial markets. As a result, we have incorporated a variety of macro variables that represent various financial markets. All the independent variables are one month lagged as our purpose is to predict money flow of Norwegian bond funds one month later. All our variables are in logarithmic form. The formula for the log return X variables as followed $\ln\left(\frac{X_t}{X_{t-1}}\right)$and for the log change X variables the formula is $\ln\left((X_t - X_{t-1}) + 1\right)$ To deal with outliers, we use a winsorize approach using scipy package in python, where outliers become nearest the largest value at selected threshold. To be consistent with other researchers, we winsorize all outliers beyond 3 standard deviations away from the mean. Our full selection of macro variables are as follows:

Log return Norwegian Stock Market - OSEBX

As mentioned, researchers have found evidence that investors chase returns. Looking at the Norwegian bond fund market, it is interesting to include the log returns of the Oslo stock exchange index as an independent variable. This represents the performance in the Norwegian equity market.

Log return Nasdaq and S&P500

US stock market is one of the largest financial markets in the world and thus is expected to influence the Norwegian market. Therefore, we have included log return on the Nasdaq Market, where Nasdaq Composite is used as the measure. Nasdaq Composite is a stock index that includes all stock and associated securities listed on Nasdaq, Additionally, we include S&P500, which is a stock index representing the 500 largest American corporations. Both indexes can be a good measure to represent the US Stock market.

Log return real estate

We further chose to include real estate as a variable because it has historically been a desirable asset class for investors both nationally and internationally. From Zillow, we have used the monthly average for single-family homes. After that, we calculated the log return of the monthly price averages.

Log percentage change of EUR/NOK and USD/NOK

Furthermore, we have included the percentage log change in exchange rate EUR/NOK and USD/NOK. In accordance with existing literature, the exchange rate significantly affects fund flows. (Kopsch et al., 2015)

US Yield Curve and Norwegian Yield Curve

The yield curve gives an idea of future interest changes and economic activity. Including yield curves for Norway and the United States in our prediction model could provide valuable information about interest rates at different maturities. These curves influence investor expectations and risk tolerance, which, in turn, affect capital flows in the Norwegian bond market. The relationship between yield curves and capital flow can be positive or negative depending on various factors such as market conditions, investor expectations, and economic circumstances. Generally, a flatter yield curve implies lower risk and increases the inflow of money into the bond market, while a steeper curve attracts investors seeking higher returns and may lead to money flowing out of the market.

Therefore, we have decided to include this as a variable both for the US and Norwegian market. The US yield curve is obtained from FRED as US 10-year

Treasury constant maturity minus 3-month treasury, constant maturity, at monthly frequency in percent, not seasonally adjusted. The 10-year Norwegian treasury and NIBOR is obtained from the central bank of Norway at monthly frequencies. We take the 10-year Norwegian treasury rate minus NIBOR. In the data we obtained, there were missing values from September 2010 to April 2011 for the Norwegian 10-year bond yield. We filled them with the average of the previous six months to prevent missing value.

Log change NIBOR

NIBOR is a rate based on what a bank charges another banks for an unsecured loan in Norwegian kroner. Generally, we believe that when change in NIBOR is positive it is more attractive for investors to store money in bank deposits and hence lowers the money flow into bond funds.

Log change VIX

VIX measures the US stock market expectations on volatility based on the S&P 500 index. As mentioned, previous authors have found some relationships between the Norwegian bond market and VIX. Thus, we have included log change VIX as a variable.

Log return VBTLX

Several authors have stated that investors seem to buy equity fund that have performed well in the past (Sirri & Tufano, 1998). Similar result did also (Edwards & Zhang, 1998) refer to measuring the effect of flow of capital on past bond return. Thus, we would like to see if past return could also be an explanatory variable for Norwegian bond funds. As we were only able to find return data on the funds for our model back to 2009, we decided to use Vanguard total bond market index as a measure on the performance of the Norwegian bond fund market. Vanguard total bond market index measures performance of a wide broad of range of public, investment-grad, taxable fixed income securities in the US, including government corporate and foreign dollar-dominated bonds, as well as mortgage- and asset-backed securities, all of which have maturities of more than a year (*Vanguard Total Bond Market Index Fund (VBTLX) Stock Price, News, Quote & History - Yahoo Finance*, n.d.). This index metric is frequently employed by Norwegian businesses, like Storebrand, as a benchmark to determine their own performance on funds and thus seems reasonable approximation for our aggregate level study (Storebrand, n.d). Semi-annual and yearly returns are calculated summing log monthly returns.

Log return medium credit risk bond fund

To represent return on medium credit risk bond fund, we have used Bloomberg global aggregate corporate index, LGCPTRUU. This index is a measure of global investment grade, fixed-rate corporate debt. The index covers bonds from developed and emerging markets issuers within the industrial, utility and financial sectors. (Source: Bloomberg).

Lagged log percentage flow

In a similar manner as (Grose, n.d.) we include log percentage flow lagged one month, two months and 3 months. The intuition is that a positive/negative net flow previous month could be followed by positive/negative net flow upcoming month.

To end the data section, a comprehensive overview of the descriptive statistics for all variables are found in Table 1 on next page.

Table 1

Descriptive statistics of all variables

Variable	*Count*	*Mean*	*STD*	*Min*	*25%*	*50%*	*75%*	*Max*	*Skewness*	*Kurtosis*
<u>Full sample</u>										
Net Flow	232	0,01	0,03	-0,05	0,00	0,01	0,02	0,33	6,02	59,53
Return OSEBX	232	0,01	0,05	-0,20	-0,01	0,02	0,04	0,13	-0,87	2,53
Return Nasdaq	232	0,01	0,05	-0,14	-0,02	0,02	0,05	0,12	-0,53	0,41
Return SP500	232	0,01	0,04	-0,12	-0,02	0,01	0,03	0,10	-0,63	0,89
Return Real Estate	232	0,00	0,01	-0,01	0,00	0,00	0,01	0,02	-0,10	0,08
Change EUR/NOK	232	0,00	0,02	-0,04	-0,01	0,00	0,01	0,06	0,59	1,09
Change USD/NOK	232	0,00	0,03	-0,05	-0,01	0,00	0,02	0,07	0,28	0,18
US yield curve	232	0,02	0,01	0,00	0,01	0,02	0,02	0,04	-0,14	-0,85
Change NIBOR	232	0,00	0,00	-0,01	0,00	0,00	0,00	0,00	-1,69	5,46
Change VIX	232	0,00	0,22	-0,46	-0,14	-0,02	0,10	0,65	0,41	0,36
Return Bond Funds	232	0,00	0,01	-0,03	0,00	0,00	0,01	0,03	-0,47	0,83
Norwegian yield curve	232	0,02	0,01	0,00	0,01	0,02	0,03	0,06	1,39	1,60
Yearly Return Bond Funds	221	0,04	0,04	-0,12	0,01	0,04	0,06	0,12	-0,72	1,40
Lag Net Flow	231	0,01	0,03	-0,05	0,00	0,01	0,02	0,33	6,01	59,34
Return3 Bond Funds	230	0,01	0,02	-0,07	0,00	0,01	0,02	0,05	-0,84	1,38
Return6 Bond Funds	227	0,02	0,03	-0,10	0,00	0,02	0,04	0,08	-0,92	2,12
Lag2 Net Flow	230	0,01	0,03	-0,05	0,00	0,01	0,02	0,33	6,00	59,08
Lag3 Net Flow	229	0,01	0,03	-0,05	0,00	0,01	0,02	0,33	6,00	59,00
<u>Medium Credit Risk</u>										
Net Flow	232	0,01	0,03	-0,08	0,00	0,01	0,02	0,23	2,77	17,03
Return Medium Credit Risk bond fund	232	0,00	0,02	-0,09	-0,01	0,00	0,01	0,06	-0,86	3,96

Methodology

To answer the question of what is driving the prediction of net flow of capital in or out of the Norwegian bond market, we will first apply feature selection using XGBoost. Then we will conduct an OLS analysis with the top 5 features to obtain some quantifiable linear relationships to net flow using the whole sample. We will use this as a base for comparing differences between the models. Particularly, we will compare accuracy scores to investigate if machine learning methods can do a better job at prediction than classical OLS. To better understand the drivers of the machine learning models we also will apply partial dependency plots. Overall, our model should be accurate and not overly complex. Thus, this method part will discuss the concepts just mentioned in the following order: First, we will describe the machine learning algorithms and OLS. Then we will discuss the train test split and parameters for the machine learning algorithms and present our tuning procedure. We will also present the intuition and math behind XGBoost feature selection. For our analysis, we use Sklearn and XGBoost API in python that has a large library of packages for classic statistics and machine learning, and we will refer to the package's documentation when describing these concepts. (*User Guide*, n.d.; *XGBoost Parameters — Xgboost 1.7.6 Documentation*, n.d.)

Describing OLS and machine learning algorithms

For our models we have decided to use OLS linear regression model and two machine learning models, XGBoost and MLPRegressor. OLS falls naturally due to its widespread application in research and its ability to extrapolate. XGBoost is a machine learning model built on the principles of decision tree theory. It is particularly useful in handling missing values and can effectively uncover nonlinear relationships between variables. Its gradient boosting framework enables improved predictive performance. MLPRegressor, on the other hand, is based on neural network architecture. It offers the advantage of identifying complex nonlinear patterns in the data. MLPRegressor's multi-layer perceptron structure allows it to approximate continuous functions given sufficient data and training time. By incorporating these three models, we aim to leverage their respective strengths. OLS provides a straightforward and interpretable benchmark, while XGBoost and MLPRegressor offer more advanced techniques to handle nonlinearities in the data. We will further give a short description of each method.

Ordinary least squares (OLS)

One of the most used linear regression techniques, ordinary least squares (OLS) regression, is used to estimate a model's unknown parameters. The OLS approach is based on minimizing the sum of squared residuals between the observed values and the model's predicted values. The difference between the actual value and the predicted value is represented by the residual. Additionally, this sum is known as the residual sum of squares (RSS). By identifying coefficient values that produce the least RSS, the OLS approach minimizes the RSS. The resulting line, which is referred to as the regression line, represents the data's best fit.

It is also crucial to recall that the OLS approach is valid only under certain conditions. First, the independent and dependent variables need to be linearly related. Furthermore, the observations need to be independent of one another. On top of that, the variance of the residuals ought to be constant for all values of the independent variables. Additionally, the residuals must follow a normal distribution. Finally, there should be no multicollinearity between the independent variables which implies that they should not have a strong correlation with one another. We will in our analysis assume that researchers would meet these conditions when finding the best prediction model, hence we apply a dataset adjusted for outliers, seasonality, and trend in the OLS model. The assumed relationship in an OLS is thus the following (Kumar, 2023):

$$\mathrm{y_i} = \beta_0 + \beta_1 \mathrm{X}_1 + \beta_2 \mathrm{X}_2 + \ldots + \beta_\mathrm{n} \mathrm{X}_\mathrm{n} + \mathcal{E}_\mathrm{i}$$

y: Dependent variable

β_0: Intercept

X: Independent variables

n: Number of independent variables

β_n: Slope coefficient

$\mathcal{E}$: Error term

Extreme Gradient Boosting (XGBoost)

XGBoost is a python API based on a gradient boosting algorithm.

XGBoost is an advanced machine learning algorithm that performs well at handling complex non-linear relationships and captures high-order interactions between variables. The machine learning model is based on a technique that combines multiple weak decision trees in order to create a strong and more precise ensemble model. Decision trees are built sequentially in the XGBoost model. Assigning weights to all of the independent variables, which are then fed into the decision tree that forecasts outcomes, is a crucial part of the model. All variables whose values were incorrectly predicted by the tree have their weights boosted and are then fed into the second decision tree. These different predictions are then combined to produce a robust and accurate model ('XGBoost', 2021).

The model can be mathematically written as follow (*Introduction to Boosted Trees — Xgboost 1.7.6 Documentation*, n.d.):

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i), f_k \in Ƒ$$

K: the number of trees,
Ƒ: the set of possible CARTs
f: the functional space of Ƒ

$$obj(\theta) = \sum_{i}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \omega(f_k)$$

$\omega(f_k)$: the complexity of tree f_k

XGBoost and feature selection

According to feature importance, a gain level is assigned to each x-variable, indicating the percentage contribution of each feature to the mode depending on the overall gain of this feature split. A higher score indicates a stronger predictive feature. To put it another way, it establishes the level to which a certain variable is beneficial for the current model and forecast. This is also known as gain. We use XGBoost for the feature importance as there are some missing values in our data, which XGBoost is able to handle effectively. (*Xgb.Importance*, n.d.)

The mathematically model behind gain is as followed:

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda}\right] - \Upsilon$$

Multi-Layer Perceptron Neural Network (MLP)

Multi-Layer Perceptron Regressor, also known as MLPRegressor is a type of artificial neural network model. It consists of multiple layers of interconnected nodes (neurons) and is capable of learning complex relationships between variables. There are input layers, hidden layers, and output layers in neural networks. The MLP model is trained using input that comes from one layer, passes through the hidden layer, and then creates an output layer. The neurons between these layers represent weights (the degree of coupling between each neuron) and biases (the threshold at which a neuron is activated). By modifying their weights and biases over time, neural network models learn to reduce errors in their outputs when compared to predicted results. MLPRegressor is particularly useful in scenarios where there might be non-linearities, as it can capture such complexities through its hidden layers and activation functions, which is the reason for why we decided to include this model in our study. (*Sklearn Neural Network Example - MLPRegressor - Data Analytics*, n.d.)

A simplified explanation on the mathematics behind the MLP Logistic regression model goes as follow:

First, a linear model that represent the weighted sum of inputs at a neuron

$$z(x) = w_1X_1 + w_2X_2 + \cdots + w_nX_n + b = w^Tx + b$$

Here w_i represents the weights, X_i represents feature inputs and b is a bias term. Then the linear inputs are passed through the activation function to form our predictor. Since we have adopted the logistic function, it will be as follows:

$$g(z) = \frac{1}{1 + e^{-z}}$$

Final output is then functional composition: g(z(x)) and the complexity depends on number of hidden layers. Non-linearities are spotted by creating many linear models passed through activation functions. (*1.17. Neural Network Models (Supervised)*, n.d.; *Neural Networks*, 2017; *Sklearn.Neural_network.MLPRegressor*, n.d.)

Machine Learning Considerations

Train test split

Train-test split is a common technique used in machine learning to evaluate the performance of a model. It involves dividing the available dataset into two separate sets: the training set and the test set. The training set is used to train the machine learning model. It is the portion of the data on which the model learns the underlying patterns and relationships between the input features and the corresponding target variable. The model adjusts its parameters based on the training set to optimize its performance. The test set, on the other hand, is used to evaluate the model's performance and assess its ability to generalize to unseen data. The test set is kept separate from the training set and is not used during the training phase. By evaluating the model on unseen data, we can get an estimate of its performance in real-world scenarios. Thus, the train test split procedure is an important factor to succeed in creating a good prediction model using machine learning.

The train-test split is typically done randomly, ensuring that both sets represent the overall dataset's characteristics and maintain the same distribution of data. Commonly, around 70-80% of the data is allocated for training, while the remaining 20-30% is used for testing. However, for our purposes it is crucial to divide the dataset based on time, to prevent data leakage and ensure accurate predictions. Data leakage in our case refers to using future information, leading to overoptimistic models. To avoid this, we employ expanding window validation, conducting 5 consecutive train-test splits. Starting with 50% training and 10% testing, we gradually increase the training window by 10% until reaching 90% training. This method strikes a balance between generating sufficient training-test pairs and incorporating new data. As we have limited data per period, this approach avoids overemphasizing outdated patterns that may change (Filho, 2022). One weakness of machine learning models is that the accuracy scores are highly dependent on the train test split. With this approach we can evaluate the model on each test set and average them to get a more robust measure of our model performance.

Accuracy metrics

The accuracy metrics used for our prediction models are mean squared error (MSE) and R-squared. The models will be trained on training sets, where they learn the underlying patterns and relationships, and then deployed on test sets to evaluate their performance. Using MSE and R-squared as accuracy metrics helps us gauge the quality of our models' predictions and understand the level of fit between the predicted and true values. Lower MSE indicates smaller prediction errors, while a higher R-squared value signifies a better fit between the model and the data.

$$MSE(y, \hat{y}) = \frac{1}{n}\sum_{i=0}^{n-1}((y_i - \hat{y}_i)^2$$

$$R^2(y, \hat{y}) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

By employing these accuracy metrics and evaluating the models' performance on test sets, we can assess their effectiveness in making accurate predictions and choose the most suitable model for prediction.

Hypertuning procedure

Hyperparameter tuning is a critical step in machine learning model development that aims to optimize the performance and generalization of the model. In machine learning, hyperparameters are parameters that are not learned from the data but are set prior to the training process. They influence the behavior and performance of the model. It can be a challenging task to obtain the optimal hyperparameters, however there are various techniques developed for automating this process. The process of hyperparameter tuning involves systematically searching for the best combination of hyperparameter values that yields the optimal model performance. We used sklearn gridsearchCV packages for hypertuning with a time series split cross validation strategy. This function loops through a range of predefined hyperparameters and fits several models on training set. Since we have adopted and time series split strategy with 5 number of splits for cross validation the gridsearchCV function will loop through a range of predefined parameters for each of the 5 training splits and use the parameters that generate the smallest measure of mean squared error. The biggest weakness with this automated function is the bias that could occur in the predetermination parameters to be tested. XGBoost and MLP (Multi-Layer Perceptron) have different parameters to tune. We have focused on those that are most frequently used in other examples and resources.

For XGBoost, the parameters we use are presented below (*XGBoost Parameters — Xgboost 1.7.6 Documentation*, n.d.):

- Learning rate: This parameter controls the shrinkage of variables weights. A larger value makes the boosting process more conservative.
- Max depth: Setting this parameter to 0 means there is no limit on the debt of tree. Increasing this measure makes the model more complex.
- Gamma: It represents the minimum loss reduction before making a further partition on a leaf node, default value is 0. The larger the more conservative the model. Thus, we leave it at default since we get worse fit when increasing it.
- Subsample: This parameter determines the fraction of training instances to be randomly sampled before growing trees. This is often used to prevent overfitting. Setting this to 0.5 means XGBoost will randomly sample half of the training data. The family parameter of the subsample is colsample by tree which decides the fraction of columns to be used when constructing each tree.

In MLP, commonly used parameter include:
Number of hidden layers: These layers perform nonlinear transformation of the inputs. For each layer an activation function is specified. In our case, we use the logistic sigmoid function, which returns:

$$g(z) = \frac{1}{1 + e^{-z}}$$

In addition, there needs to be specified a solver for weight optimization. Options include stochastic gradient descent, Adam (a stochastic gradient- base optimizer), and lbfgs (based on quasi-Newton methods finding zeros or local maxima and minima of functions). We use lbfgs as it tends to work well on smaller datasets, as stated in the documentation. Alpha is a measure of ridge regression, which adds a penalty to the loss function. It includes the square magnitude of the coefficient and helps reduce the impact of large weights. MLP is sensitive to feature scaling. We thus use standard scaler function in Sklearn to address this in our approach (*Sklearn.Preprocessing.StandardScaler*, n.d.).

Our hypertuning procedure gives us the predetermined parameters presented in Table 2.

Table 2 Hyperparameter space and other parameters used for machine learning algorithms training.

Algorithm	**Approach**	**Hyperparameters and parameters names**	**Hypertuning measures**
Classic statistics	OLS		
Gradient Boosting	XGBoost	Learning rate (eta)	0.07
		Max_debth	2
		Min child weight	6
		Subsample	0.5
		colsample_bytree	0.7
		n_estimators	100
		gamma	0
Multilayer Perceptron neural network	Sklearn MLPRegressor	Solver	lbfgs
		Hidden layer sizes	500
		Max_iter	200
		Activation	Logistic
		Alpha	0.05
		Feature scaling	True

Results and main analysis

This section provides our results and main analysis as described in methodology chapter. Starting with XGBoost feature selection to maintain most relevant features. Then we use OLS to examine linear relationships and linear model fit. Subsequently, we compare the accuracy scores of our machine learning models to assess their performance in predicting net flow using top selected features. Additionally, partial dependency plots will be discussed for both XGBoost and MLP with the perspective of obtaining a better understanding of what are driving their prediction and spot potential non-linearities.

XGBoost Feature Selection

The results of our feature selection on the entire dataset using XGBoost with the objective to minimize squared error is shown in figure 5 below. We observe that Change VIX, Net flow in a time lag of 2 months, change EUR/NOK, Net flow in a time lag of 3 months, and the semi-annual return of bond funds is most important when predicting Norwegian bond fund flow. However, when we narrow our focus to the sample of bond funds with medium credit risk, the feature important results differ. In this subset, Return Nasdaq, lag net flow, Norwegian yield curve and change NIBOR have the most impact on the models' predictions (as illustrated in figure 6 below). Interestingly, only Change EUR/NOK appears consistently important for bond funds and the subsample of bond funds with medium credit risk. These findings highlight the varying importance of different features in predicting Norwegian Bond fund flow, depending on the risk profile under consideration.

Figure 5 - Feature Importance of all independent variables using XGBoost for bond funds

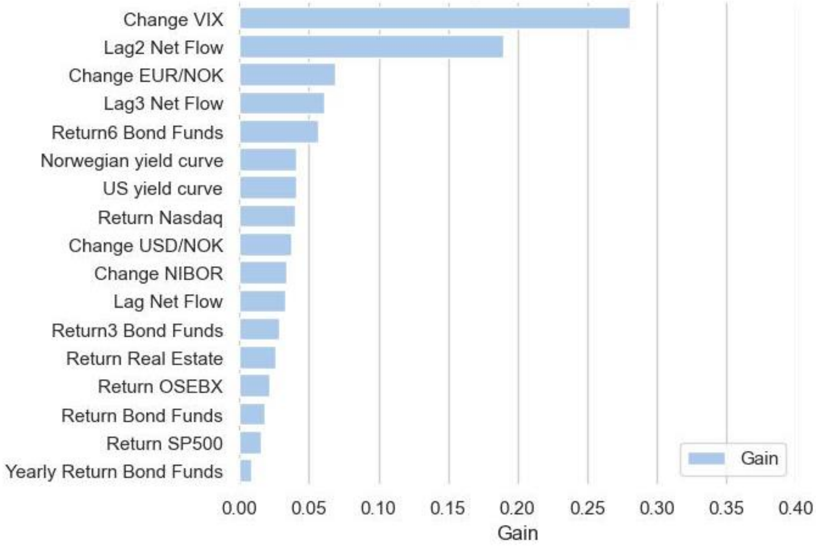
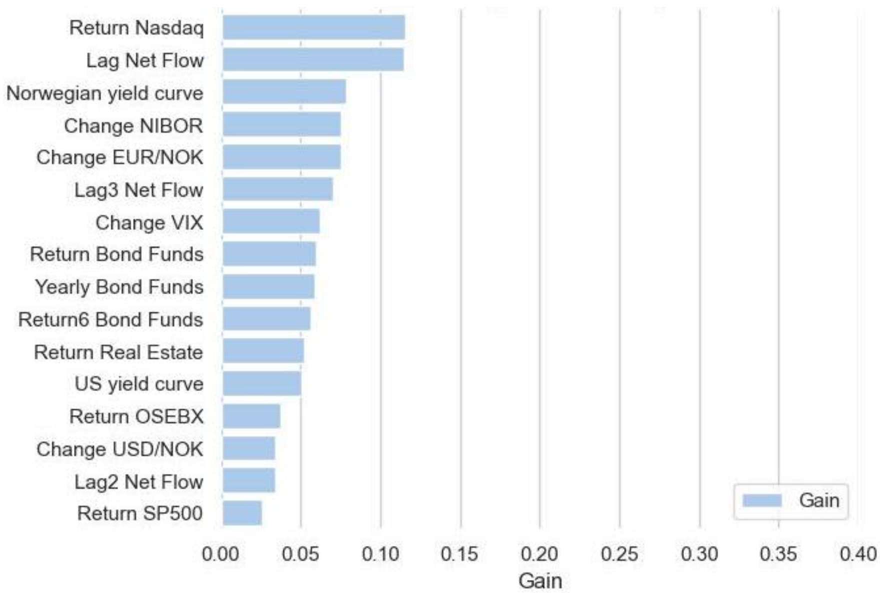


Figure 6 - Feature Importance of all independent variables using XGBoost for medium risk bond funds



<u>OLS result bond funds</u>

As mentioned, there is little theory on determinants of Norwegian bond funds. Thus, we apply classical OLS to understand better the selected features' sensitivity to net flow. For the full sample of bond funds using the top important features above our OLS model is as follows:

$$Net\ Flow_{t+1} = \beta_0 + \beta_1 Change\ VIX_t + \beta_2 Change\ EUR/NOK_t + \beta_3 Lag2\ Net\ Flow_t\ + \beta_4 Lag3\ Net\ Flow_t + \beta_5 Return6\ Bond\ Funds_t$$

Since both our dependent and independent variables are log transformed our regression captures the movement on net flow in percent to a one percentage change in independent variables. For the purpose of predicting net flow one month ahead the independent variables are all lagged one month. We also have corrected for seasonality and trend where necessary. After performing necessary adjustments and model diagnostics (refer to the appendix for details), we find the OLS results presented in Table 3 to be reliable for interpretation.

Table 3: OLS factors determining cash flow into bond mutual funds seasonal adjusted etc.
The table shows the findings of an OLS regression analysis for calculating cash flow into bond mutual funds for the full sample. The r-squared level and whether the variables are seasonally adjusted or not are reported in the table.

	Full Sample	
Dependent variable	Log Percent Net Flow	Seasonally adjusted
Change VIX	-0.0051	
Change EUR/NOK	-0.1470**	
Lag2 Net Flow	-0.0518	
Lag3 Net Flow	-0.1090	
Return6 Bond Funds	0.1040	Seasonally adjusted
Intercept	0.0008	
Overall	0.053	Observations 215

Notes: Significant at 90% level **

Change VIX is negative related to net flow, but not significant. Many researchers have found a negative relationship between VIX when significant, but this is particularly for flow of stock funds.

Interestingly, change EUR/NOK is significant and negatively related to net flow in our regression. Such that a 1% increase in EUR/NOK in the previous month is associated with a decrease in net flow of 0.14%. There is limited research that could explain such a correlation into meaningful interpretation. However, the authors (De Lange et al., 2018) suspect the after the financial crises the Norwegian bond market become dependent on euro-liquidity.

Lagged variables of net flow are not significant. (Grose, n.d.) find that lagged flows are significant and have a positive relationship to net flow of Greek funds. A possible explanation why our results differ might be that he looks at subgroups of bond funds while our data of net flow includes all funds and thus is not able to pick up a significant relationship.

Semi-annual bond funds are not significant, this is different from the (Y. Chen & Qin, 2017) paper that finds a 1% return on bond index previous month is associated with net money flow of approximately 0.22%. They use the Barcleys aggreagate bond index while we use Vanguard, which can be a natural explanation of the different results.

OLS result bond funds with medium credit risk

For the sample of bond funds with medium credit risk using the selected features from XGBoost our OLS model is as follows:

$$Net\ Flow_{t+1} = \beta_0 + \beta_1 Return\ Nasdaq_t + \beta_2 Lag\ Net\ Flow_t + \beta_3 Norwegian\ Yield\ Curve_t + \beta_4 Change\ NIBOR_t + \beta_5 Change\ EUR/NOK_t$$

Similarly for bonds with medium credit risk, our dependent and independent variables are log transformed. Our regression captures the movement on net flow in percent to a one percentage change in independent variables. Also here, the independent variables are all lagged one month and corrected for seasonality and trend where necessary. After performing necessary adjustments and model diagnostics (refer to the appendix for details), we find the OLS results presented in Table 4 to be reliable for interpretation.

Table 4: OLS factors determining cash flow into bond mutual funds seasonal adjusted etc.
The table shows the findings of an OLS regression analysis for calculating cash flow into bond mutual funds for the medium risk bond fund group.

	Medium Credit Risk Funds	
Dependent variable	Log Percent Net Flow	Seasonally adjusted
Return Nasdaq	-0.0026	
Lag Net Flow	-0.0880	Seasonally adjusted
Norwegian Yield Curve	-0.1213	Seasonally adjusted
Change NIBOR	-0.6067	
Change EUR/NOK	-0.1033	
Intercept	0.0000	
Overall	0.017	Observations 219
Notes: Significant at 90% level **		

Our analysis reveals that none of the variables exhibit significant relationships with net flow one month ahead of Norwegian bond fund with medium credit risk. Additionally, the model's goodness of fit, as indicated by the r-squared value is only 1.7%. This result is lower compared to our bond funds model and also falls short of the 16% r-squared obtained by (Y. Chen & Qin, 2017). Implying that the predictive power of our model is relatively weak in comparison. However, it is important to note Chen and Qin conducted their analysis on US corporate bond funds using different explanatory variables. Therefore, direct comparisons between their results and our findings should be made with caution but show that for our topic there might be potential of finding a better fitting model under OLS using other variables. However, since this study focuses on the top selected features, we will refrain from developing deeper into that here.

Generally, we find little evidence that net flow of Norwegian bond fund is affected by factors other than change EUR/NOK. The explanatory power of our whole sample model (r-squared of 0,053) is less than that found in similar research. Further, we find no statistically significant correlations between the selected explanatory factors and net flow in our analysis of medium credit risk bond funds. This model also obtains a poor goodness of fit (r-squared of 0,017).

Exploring addition variables could enhance explanatory power for both full sample model and medium risk bond fund model.

Net flow prediction with machine learning

Since OLS assumes linear relationships, it is of interest to see if machine learning models can do a better job in predicting fund flow one month ahead by addressing non-linearities. More particularly it is of interest to see if R-squared increases using XGBoost and MLP. To assess this, we follow the procedure outlined in the methodology section and compare the performance of these machine learning models to that of OLS. In the case of OLS, we adhere to the assumption of linearity in the measurements, while dataset introduced for the machine learning models are kept log transformed only. The tables below present model accuracy for both bond fund and bond funds with medium credit risk. The accuracy scores are presented by the mean values from each train-test split, accompanied by the standard deviation to measure the variability in the accuracy scores. Large standard deviation suggests that the accuracy scores differ in each split, indicating potential instability in the model's performance when introduced to sequential future time periods. We have included standard deviation in the tables but will not comment on them further, as our attention is on the mean accuracy scores.

Table 5: Model Accuracy on bond funds with different credit risk

	OLS		XGBoost		MLP	
	Mean	Sd	Mean	Sd	Mean	Sd
$R^2\ Test$	-0.0527	0.1428	-2.0823	1.4901	-0.6707	0.3701
$R^2\ Train$	0.0548	0.0046	0.4102	0.0165	0.0252	0.0049
$MSE\ Test$	0.0001	-	0.0006	0.0003	0.0003	0.0001
MSE $Train$	0.0003	-	0.0007	0.0001	0.0011	0.0002

Notes: OLS is conducted on dataset transformed to meet the OLS assumptions of normality, endogeneity, stationary and non-seasonality. Sample is shorter due to rolling averages of trend and seasonality transformations.

Table 6: Model Accuracy on medium credit risk mutual bond funds						
	OLS		XGBoost		MLP	
	Mean	Sd	Mean	Sd	Mean	Sd
R^2 $Test$	-0.0664	0.0510	-0.4817	0.1466	-0.3258	0.1044
R^2 $Train$	0.0169	0.0024	0.5073	0.0212	0.0229	0.0106
MSE $Test$	0.0002	-	0.0003	0.0001	0.003	0.0001
MSE $Train$	0.0005	0.0001	0.0005	0.0000	0.0010	0.0001

Notes: OLS is conducted on dataset transformed to meet the OLS assumptions of normality, endogeneity, stationary and non-seasonality. Sample is shorter due to rolling averages of trend and seasonality transformations.

Table 5 presents the model accuracy for bond funds with different credit risks. XGBoost model increase predictability, achieving approximately 40% r-squared in the training sample. On the other hand, MLP model exhibits lower predictability compared to OLS, decreasing from around 5% to 2%. When these models are applied to the unseen test set, all of them result in negative R-squared values, indicating a poor fit. To evaluate model performance, we use the mean squared error (MSE), which measures the average magnitude of the residuals. In our case, the models exhibit relatively small MSE values on both train and test sets, but negative r-squared values.

Transitioning to Table 6, which focuses on bond funds with medium credit risk, we observe similar results. The XGBoost model demonstrates increased predictability, achieving approximately 50% r-squared in the training sample. In contrast to the bond fund analysis, where OLS had a r-square of ca 5% it dropped to 1.7% for the bond funds with medium credit risk. Thus, MLP model performs better in this case while maintaining a similar predictability level of around 2%. Notably, all the models demonstrate minimal disparity in MSE between the train and test sets, as evident from the small values displayed in the table, but also here we get negative r-squared values.

In short, all the models show relatively little difference in MSE between the train and test sets. However, despite the small MSE values, the model's regression line explains little of the variation (negative R-squared), indicating a poor overall model fit. According to online resources (Wei, 2022), R-squared can be negative when we use test data to evaluate models built on train data and where working with non-linear models. Both apply in our case, hence seems reliable.

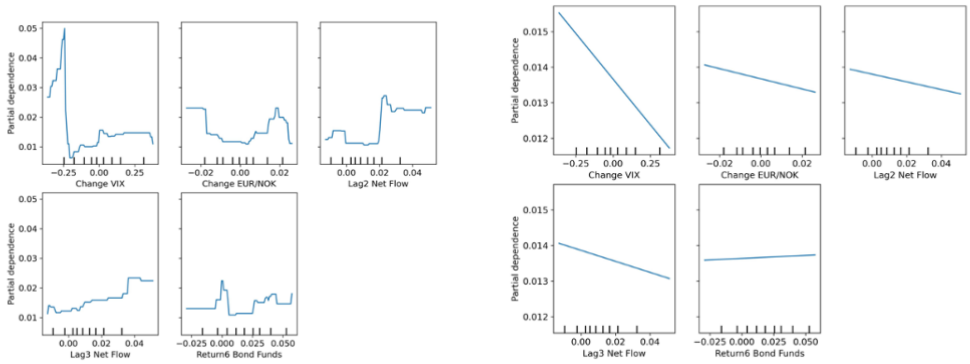
One potential challenge we encounter is the limited availability of consistent signals over time in the variables, which can pose difficulties for our models when applied to test set representing future time periods. As we have mentioned, we observe low r-sqaured for OLS and MPL already in training set indication that the explanatory variables have little prediction power on net flow. Consequently, causing poor results on test set. As for XGBoost (larger r-squared values on train but substantially higher negative r-square on test) it seems like the algorithm is picking up additional signals and do not generalize enough. This indicates that there could be some presence of overfitting despite our efforts to address this through the application of hypertuning process. We conducted tests using alternative parameters manually, only to observe poorer results. Taking into account our aim to maintain consistency across models, we have made the decision to retain the results obtained from the hypertuning using gridsearchCV with time series split as the chosen cross-validaton strategy.

To gain further insights into the drivers of the different model predictions and their results, we will further examine their partial dependency plots. Moreover, we will compare these observations with the relationships obtained in OLS and consider the findings of other authors in our analysis. Overall, this will help our aim in understanding the performance variations among models in predicting Norwegian bond fund flow.

Understanding the drivers of prediction results in our Machine learning models

One critique of machine learning models compared to classical statistical methods is the degree of interpretability of the prediction results. Partial dependence plots are often used by researchers in understanding the models in more detail. Optimally, this should be used when the models have good accuracy on the test set, however for our purposes in understanding what is driving the predictions of our models we find it sufficient. More particularly, we will use partial dependency plots to understand each feature's impact on the model's prediction of net flow of Norwegian Bond Funds one month ahead and see if it picks up some non-linear behaviors.

Figure 7 - Feature dependency plot for bond funds using XGBoost (left) and MLP (right)



The figure above shows the partial dependency plots for XGBoost and MLP for both bond funds, revealing interesting observations that will be further discussed.

Starting on the top left we have the partial dependence plots for bond funds, under the XGBoost model. We observe that net flow increases when the change in VIX is within the interval +- 0.25 but predicts negative net flow at extreme drops in VIX below -0.25, inferring a non-linear relationship. A possible explanation could be that at extreme events such as under the financial crisis, the uncertainty is so high in the overall economy such that investment activity drops, hence also a decrease in net flow for bond funds. In comparison, MLP model shows a steep downward sloping relationship. There is no indication of non-linear relationships here, and the straight-line show there are smoother predictions. OLS show an insignificant negative relationship.

Change EUR/NOK has a negative relationship to net flow in both machine learning models, similarly as the OLS model. Particularly, at a 90% confidence level a 1% increase in EUR/NOK relates to a decrease in net flow of 0.14%, holding all other variables constant. As mentioned, a surprising relationship with foreign exchange rate might be explained by the Norwegian bond market becoming more dependent on the euro-liquidity after the Financial Crisis.

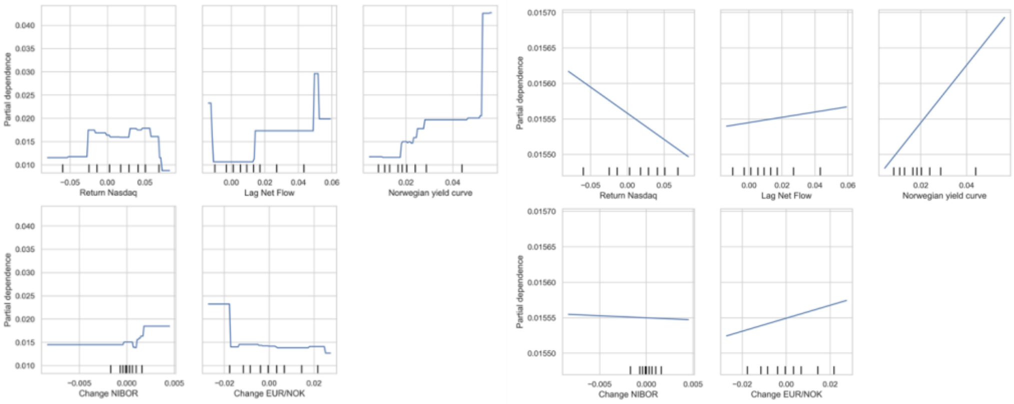
Another, possible explanation is that Norwegian bond funds hold bonds in EUR so when the NOK strengthens (positive change EUR/NOK), the values of these bond holdings in NOK terms will decline. Consequently, investors see a drop in the funds NAV which might affect their capital allocation. The XGBoost model captures additional dynamics. As NOK weakens relative to EUR (as positive change in EUR/NOK increase), the model associated this with a capital inflow to Norwegian bond funds. Conversely, As NOK strengthens relative to EUR (as negative change in EUR/NOK decreases), the model also captures an increase in net flow to Norwegian bond funds.

XGBoost and MLP exhibit contrasting patterns when it comes to capturing signals from lagged variables. XGBoost identifies a significant threshold of 0.02 in net flow with 2 months lag, which is associated with a substantial increase in growth inflow.

Furthermore, XGBoost shows a positive relationship with net flow lag 3 months. This differs from the MLP model, that indicates negative relationships. Further, semi-Annual returns of the bond index have a slight decreasing relationship in MLP and no clear direction in XGBoost. Our OLS shows a positive and insignificant result.

We will proceed to investigate the partial dependency plots pertaining to the Norwegian medium credit risk bond funds, limiting to the most dissimilarities between the models.

Figure 8 - Feature dependency plot for medium risk bond funds XGBoost (left) and MLP (right)



We observe a negative relationship between the return of Nasdaq and the predicted net flow in MLP. In the XGBoost model there is a slight negative relationship where most of the observations lie (as indicated by the vertical lines), and a sharp inflow of flow when experiencing Nasdaq returns drop 0.02%. OLS was negative and insignificant.

Further, an interesting observation is the steep positive relationship with Norwegian yield curve. As yield curve increases, the machine learning models associate this with an inflow of capital to bond funds. While the OLS show a negative relationship, however insignificant.

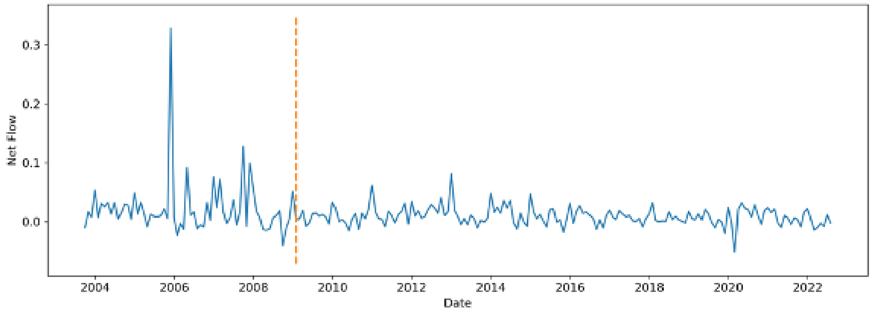
For the medium credit risk change bond funds, XGBoost with a slight decreasing change in EUR/NOK, while MLP demonstrated a slight positive relationship. OLS shows a negative relationship to net flow, but insignificant.

Partial dependency plots illustrate that our models use different signals in their predictions. The accuracy of these models on the test set is limited, preventing us from establishing relationships as an absolute fact and instead treating it as an observational finding. Particularly noteworthy, XGBoost highlights non-linear relationship with change VIX, the models show varying impact of change in EUR/NOK and contrasting patterns in lagged net flow that can explain their differing accuracy measures.

Discussion

In light of the prediction models accuracy on the test set, it is important to acknowledge the limitations and challenges associated with forecasting financial time series. Our models have shown weak results, which can be attributed to several factors discussed in relevant literature. Both (Boot & Pick, 2020) and (Clements & Hendry, 1995) highlights the presence of structural changes and non-stationarity effects in financial data can lead to models capturing time specific signals that are not to be precent in the future, even in the case of machine learning models. Considering our data sample including the period of financial crisis in 2008, where capital easing was incorporated causing demand for bonds to drop, there is natural to believe that we have a structural change at this threshold. In addition, the COVID -19 pandemic had tremendous impact on asset prices, and large redemptions added stress in the corporate bond and treasury market (Liang, n.d.). Figure 9 shows the time series of net flow of bond funds. Here we see that before 2009 there is large volatility compared to the time after. There is also a dip in 2020 to then stabilize again.

Figure 9 - Net Flow 2003 - 2022



In order to examine weak results further, we conduct an experiment focusing solely on the data from 2009 to 2022. Although this approach resulted in a reduction in the number of available observations, it was undertaken with the expectation of potentially obtaining more promising results. However, our models still yielded negative r-squared values.

Extending on the study's limitations, one notable aspect is the absence of a model specifically designed to address net flow in stressed periods. Even though, in our PDP analysis, it comes out that our XGBoost appeared to capture certain signals associated with stressed periods, in relation to large fluctuations in VIX, there are other measures that are not included in our modelling. Specifically, researchers have investigated the bond market during stressed using factors like trade prices, bond yields, and transaction cost (Friewald et al., 2012). Consequently, it is evident that a separate study focusing on these variables effects on net flow both in stressed times and across different types of bond funds could be beneficial in obtaining better results.

It is important to acknowledge the limitations imposed by the available dataset, compromising 227 observations. While this number is sufficient for traditional OLS models, machine learning models typically benefit from larger sample sizes. Additionally, due to data unavailability, we relied on using indexes as proxy for performance the entire timeseries in our sample. Precious research has found reasonable findings linking fund return and net flow, thus we are confident that building on our analysis using raw returns of the bond funds will enhance prediction power of our models.

Conclusion

In this thesis we aimed to understand the drivers of net flow in Norwegian bond funds and compare the predictive performance of different machine learning models to traditional OLS. With little research on this particular topic, we used XGBoost in selecting the most important features among our pool of 17 variables measuring the condition in other financial markets. The difference in top features indicate that there are differing explanatory variables for bonds depending on its credit risk profile. Our findings, indicate that the factors influencing net flow in Norwegian bond funds, as identified by the OLS analysis, are limited with change in EUR/NOK showing a significant negative relationship with bond funds. Thus, little supports our hypothesis that bond fund flow is sensitive to performance and macro variables. The OLS models showed relatively weak explanatory power, both in the bond funds and when analyzing bonds with medium credit risk only. Regarding the prediction result, measured by comparing accuracy scores across models, all the models show relatively little difference in MSE between the train and test sets. However, despite the small MSE values, the model's regression line explains little of the variation (negative R-squared), indicating inadequate model fit. Further insights from the partial dependency plots illustrate that our models use different signals in their predictions. Particularly noteworthy, XGBoost highlights non-linear relationship with change VIX, the models show varying impact of change in EUR/NOK and contrasting patterns in lagged net flow. The non-linear relationship identified by the XGBoost and differing relationships can explain the differing results in our models. Ultimately, our findings suggest that we are unable to achieve an effective model for predicting bond fund flows one month ahead using the top selected features. The analysis and experiment indicate that this could be due to the impact of variables on bonds varying over time and is contingent on their risk profiles, making it challenging for the performance of our models. However, despite this outcome, our study provides valuable insights into the field and provides a solid foundation for further investigation.

Contribution for further research

Building upon the results and limitations of this study, future research can expand on the following to enhance the understanding of drivers and prediction of capital flow in Norwegian bond funds. Although our study yielded weak predictive power results, our feature selection results provided findings that there are different variables affecting bond funds depending on their risk profile. Thus, research looking into explanatory variables of different Norwegian bond funds groups, such as corporate, investment grade (highly rated bond funds), high yield and government bond funds would be interesting for our understanding in this topic. Additionally, further exploration of the significant negative relationship between change EUR/NOK and capital flow, particularly related to financial crisis due to the euro-liquidity imposed to the Norwegian bond market, would contribute to a better understanding of their interplay. Further, other variables than what are included in our models are most likely to affect net flow during stressed markets. Combining models that incorporate this might yield better prediction results. It is important to acknowledge the limitations imposed by the available dataset. In retrospect, we believe that one might explain to a greater extent and obtain better prediction results by including fund returns rather than return from bond market indexes. Earlier research emphasizes a significant relationship here, we are therefore confident that this could increase performance of models. While our dataset obtains sufficient observations for traditional OLS models, machine learning models typically benefit from larger sample sizes. Furthermore, conducting the same analysis using a dataset of individual bond fund net flow would provide a larger number of observations that could be advantageous for machine learning models. To enhance prediction accuracy, one could combine a classification model yielding a probability of inflow or outflow. In this way the models consider both the direction and magnitude of net bond fund flows. Overall, these observations present interesting avenues for future research and have the potential to deepen our understanding of the capital flow of Norwegian bond funds.

Bibliography

1.17. Neural network models (supervised). (n.d.). Scikit-Learn. Retrieved 29 June 2023, from https://scikit-learn/stable/modules/neural_networks_supervised.html

Alexakis, C., Niarchos, N., Patra, T., & Poshakwale, S. (2005). The dynamics between stock returns and mutual fund flows: Empirical evidence from the Greek market. *International Review of Financial Analysis*, *14*(5), 559–569. https://doi.org/10.1016/j.irfa.2004.10.019

Boot, T., & Pick, A. (2020). Does modeling a structural break improve forecast accuracy? *Journal of Econometrics*, *215*(1), 35–59. https://doi.org/10.1016/j.jeconom.2019.07.007

Chen, W. (n.d.). *Equity Investor Sentiment and Bond Market Reaction: Test of Overinvestment and Capital Flow Hypotheses*.

Chen, Y., & Qin, N. (2017). The Behavior of Investor Flows in Corporate Bond Mutual Funds. *Management Science*, *63*(5), 1365–1381. https://doi.org/10.1287/mnsc.2015.2372

Clements, M. P., & Hendry, D. F. (1995). Macro-Economic Forecasting and Modelling. *The Economic Journal*, *105*(431), 1001–1013. https://doi.org/10.2307/2235166

Da, Z., Engelberg, J., & Gao, P. (2013). *The Sum of All FEARS: Investor Sentiment and Asset Prices* (SSRN Scholarly Paper No. 1509162). https://doi.org/10.2139/ssrn.1509162

De Lange, P. E., Aamo, P. E., & Westgaard, S. (2018). The Norwegian financial bond market: A comprehensive market review with an empirical analysisof the main drivers of spreads. *Beta*, *32*(1), 59–84. https://doi.org/10.18261/issn.1504-3134-2018-01-04

Edwards, F. R., & Zhang, X. (1998). Mutual Funds and Stock and Bond Market Stability. *Journal of Financial Services Research*, *13*(3), 257–282. https://doi.org/10.1023/A:1008084311260

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics*, *25*(1), 23–49. https://doi.org/10.1016/0304-405X(89)90095-0

Filho, M. (2022, September 21). *How To Do Time Series Cross-Validation In Python*. https://forecastegy.com/posts/time-series-cross-validation-python/

Friewald, N., Jankowitsch, R., & Subrahmanyam, M. G. (2012). Illiquidity or credit deterioration: A study of liquidity in the US corporate bond market during financial crises. *Journal of Financial Economics*, *105*(1), 18–36. https://doi.org/10.1016/j.jfineco.2012.02.001

Grose, C. (n.d.). *The Determinants of Cash Flows in Greek Bond Mutual Funds*.

Gruber, M. J. (1996). Another Puzzle: The Growth in Actively Managed Mutual Funds. *The Journal of Finance*, *51*(3), 783–810. https://doi.org/10.2307/2329222

Introduction to Boosted Trees—Xgboost 1.7.6 documentation. (n.d.). Retrieved 1 July 2023, from https://xgboost.readthedocs.io/en/stable/tutorials/model.html?fbclid=IwAR02nLgRRYxhXO8TM4gSy5tvlMC9yTRnPHSUjmPy0g0ivgcr1T2gperqdQk

James, C., & Karceski, J. (2006). *Investor monitoring and differences in mutual fund performance q*.

Kopsch, F., Song, H.-S., & Wilhelmsson, M. (2015). Determinants of mutual fund flows. *Managerial Finance*, *41*(1), 10–25. https://doi.org/10.1108/MF-06-2013-0161

Kumar, A. (2023, April 14). Ordinary Least Squares Method: Concepts & Examples. *Data Analytics*. https://vitalflux.com/ordinary-least-squares-method-concepts-examples/

Liang, J. N. (n.d.). *Corporate Bond Market Dysfunction During COVID-19 and Lessons from the Fed's Response*.

Neural networks: Representation. (2017, June 29). Jeremy Jordan. https://www.jeremyjordan.me/intro-to-neural-networks/

Sirri, E. R., & Tufano, P. (1998). Costly Search and Mutual Fund Flows. *The Journal of Finance*, *53*(5), 1589–1622.

Sklearn Neural Network Example—MLPRegressor—Data Analytics. (n.d.). Retrieved 29 June 2023, from https://vitalflux.com/sklearn-neural-network-regression-example-mlpregressor/

Sklearn.neural_network.MLPRegressor. (n.d.). Scikit-Learn. Retrieved 29 June 2023, from https://scikit-learn/stable/modules/generated/sklearn.neural_network.MLPRegressor.html

Sklearn.preprocessing.StandardScaler. (n.d.). Scikit-Learn. Retrieved 1 July 2023, from https://scikit-learn/stable/modules/generated/sklearn.preprocessing.StandardScaler.html

User guide: Contents. (n.d.). Scikit-Learn. Retrieved 21 June 2023, from https://scikit-learn/stable/user_guide.html

Vanguard Total Bond Market Index Fund (VBTLX) Stock Price, News, Quote & History—Yahoo Finance. (n.d.). Retrieved 2 July 2023, from https://finance.yahoo.com/quote/VBTLX/

Wei, T. N. (2022, June 9). *Explaining negative R-squared*. Medium. https://towardsdatascience.com/explaining-negative-r-squared-17894ca26321

xgb.importance: Importance of features in a model. in xgboost: Extreme Gradient Boosting. (n.d.). Retrieved 29 June 2023, from https://rdrr.io/cran/xgboost/man/xgb.importance.html

XGBoost. (2021, September 18). *GeeksforGeeks*. https://www.geeksforgeeks.org/xgboost/

XGBoost Parameters—Xgboost 1.7.6 documentation. (n.d.). Retrieved 21 June 2023, from https://xgboost.readthedocs.io/en/stable/parameter.html

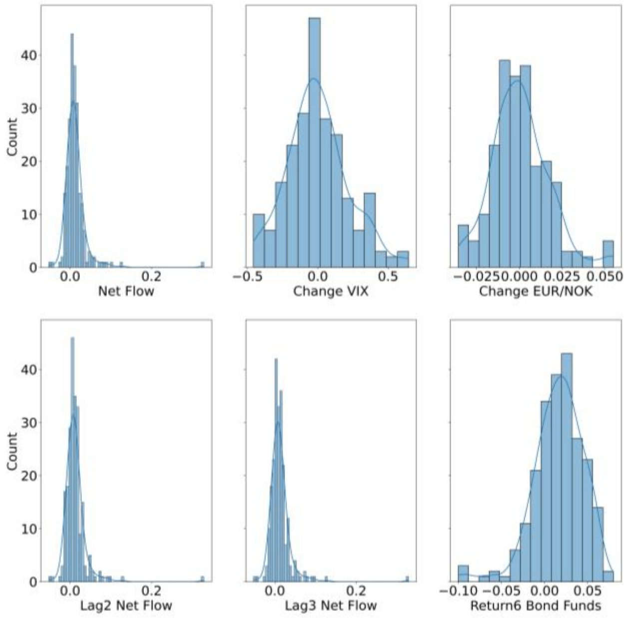
Yue, Q., Hu, Z., & Li, D. (2021). Research on Intelligent Analysis and Prediction Model of Funds Flow Based on Non-stationary Time Series. *Journal of Physics: Conference Series*, *1927*(1). https://doi.org/10.1088/1742-6596/1927/1/012027

Zhao, X. (2005). Determinants of Flows into Retail Bond Funds. *Financial Analysts Journal*, *61*(4), 47–59.

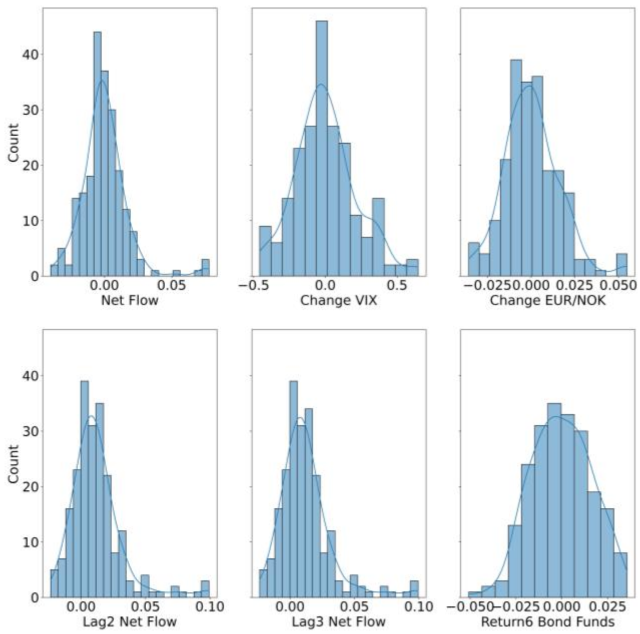
Appendix

Appendix 1: Histogram

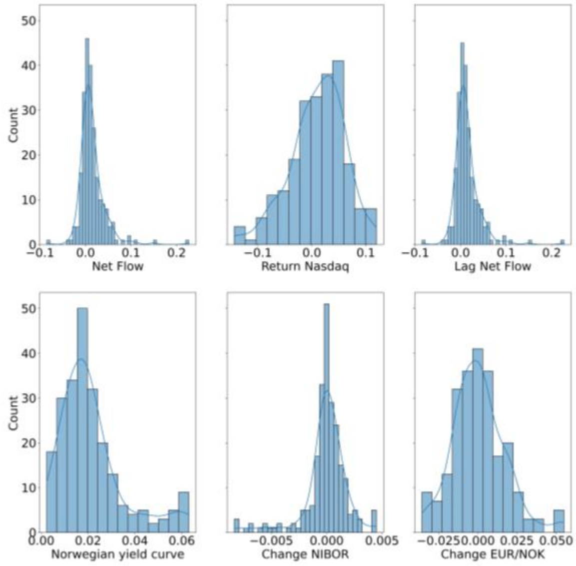
Histograms of bond funds showing distributions of variables used in Machine Learning



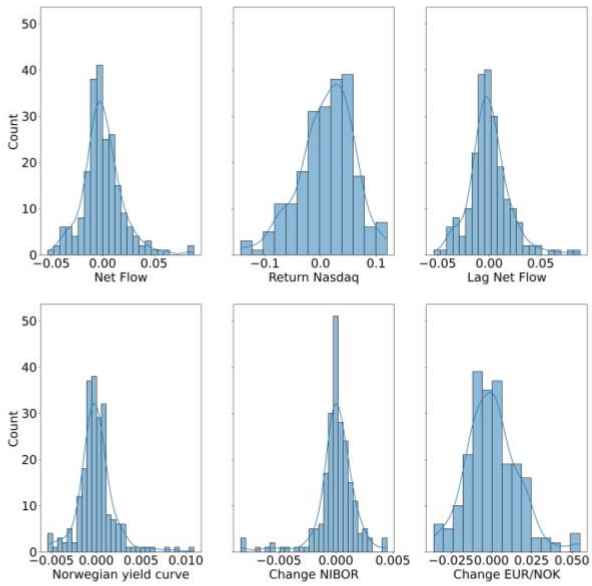
Histograms of bond funds showing distributions of variables used in OLS



Histograms of bond funds with medium credit risk showing distributions of variables used in Machine Learning



Histograms of bond funds with medium credit risk showing distributions of variables used in OLS



Appendix 2: Correlation matrix

Correlation matrix for bond funds

	Net Flow	Change VIX	Change EUR/NOK	Lag2 Net Flow	Lag3 Net Flow	Return6 Bond Funds
Net Flow						
Change VIX	-0.12					
Change EUR/NOK	-0.10	0.18				
Lag2 Net Flow	0.04	0.04	0.01			
Lag3 Net Flow	0.01	-0.05	-0.09	0.02		
Return6 Bond Funds	0.10	-0.01	-0.09	0.06	0.06	

Correlation matrix for bond funds with medium credit risk

	Net Flow	Return Nasdaq	Lag Net Flow	Norwegian yield curve	Change NIBOR	Change EUR/NOK
Net Flow						
Return Nasdaq	0.00					
Lag Net Flow	0.03	-0.06				
Norwegian yield curve	0.21	-0.16	0.19			
Change NIBOR	-0.01	-0.09	-0.02	0.04		
Change EUR/NOK	-0.06	-0.14	-0.08	0.12	-0.01	

Appendix 3: OLS model diagnostics

OLS model diagnostics for bond funds

```
Variable: Net Flow
1. ADF :  -8.119987087105656
2. P-Value :  1.1630604625031148e-12
3. Num Of Lags :  11
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 203
5. Critical Values :
         1% :  -3.462980134086401
         5% :  -2.875885461947131
         10% :  -2.5744164898444515

Variable: Change EUR/NOK
1. ADF :  -8.60289454075841
2. P-Value :  6.793306799922988e-14
3. Num Of Lags :  3
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 211
5. Critical Values :
         1% :  -3.46172743446274
         5% :  -2.8753374677799957
         10% :  -2.574124089081557

Variable:Change VIX
1. ADF :  -8.52205418139192
2. P-Value :  1.0939464371078895e-13
3. Num Of Lags :  5
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 209
5. Critical Values :
         1% :  -3.4620315036789666
         5% :  -2.8754705024827127
         10% :  -2.5741950726860647

Variable: Lag2 Net Flow
1. ADF :  -13.378318936523375
2. P-Value :  5.036222363173076e-25
3. Num Of Lags :  0
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 214
5. Critical Values :
         1% :  -3.4612821203214907
         5% :  -2.875142613826617
         10% :  -2.574020122281422

Variable: Lag3 Net Flow
1. ADF :  -13.51066083966481
2. P-Value :  2.857603472283617e-25
3. Num Of Lags :  0
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 214
5. Critical Values :
         1% :  -3.4612821203214907
         5% :  -2.875142613826617
         10% :  -2.574020122281422

Variable: Return6 Bond Funds
1. ADF :  -5.746897917675518
2. P-Value :  6.091887685925368e-07
3. Num Of Lags :  15
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 199
5. Critical Values :
         1% :  -3.4636447617687436
         5% :  -2.8761761179270766
         10% :  -2.57457158581854
```

Table 7: Breach Pegan test for heteroscedasticity bond funds

Ho: Homoscedasticity is present

Ha: Homoscedasticity is not present

Lagrange multiplier statistics	**F-value**
7.6045	1.5327
p-value	**p-value**
0.1794	0.1810

Table 8: Breach Pegan test for heteroscedasticity bond funds	
Ho: No autocorrelation at any order less than or equal to p	
Ha: There exists autocorrelation at some order less than or equal to p	
1 test stats	**2 test stats**
5.6550	1.8549
p-value	**p-value**
0.1297	0.1383

OLS model diagnostics for bond funds with medium credit risk

```
Variable: Net Flow
1. ADF :  -7.709935300392158
2. P-Value :  1.2717978268180782e-11
3. Num Of Lags :  13
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 205
5. Critical Values :
        1% :  -3.4626576734812318
        5% :  -2.8757444215841326
        10% :  -2.5743412314098753

Variable: Return Nasdaq
1. ADF :  -14.065466505028095
2. P-Value :  3.0041911444428707e-26
3. Num Of Lags :  0
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 218
5. Critical Values :
        1% :  -3.460707667106296
        5% :  -2.874891213486339
        10% :  -2.573885987711472

Variable: Lag Net Flow
1. ADF :  -7.551251600537643
2. P-Value :  3.1844784210889974e-11
3. Num Of Lags :  12
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 206
5. Critical Values :
        1% :  -3.4624988216864776
        5% :  -2.8756749365852587
        10% :  -2.5743041549627677

Variable: Norwegian yield Curve
1. ADF :  -6.595487878933046
2. P-Value :  6.942619668677972e-09
3. Num Of Lags :  1
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 217
5. Critical Values :
        1% :  -3.460849270544952
        5% :  -2.87495318813585
        10% :  -2.5739190539191745

Variable: Change NIBOR
1. ADF :  -4.726092586799117
2. P-Value :  7.515094569159297e-05
3. Num Of Lags :  7
4. Num Of Observations Used For ADF Regression and Critical Values Calculation : 211
5. Critical Values :
        1% :  -3.46172743446274
        5% :  -2.8753374677799957
        10% :  -2.574124089081557
```

Table 9: Breach-Godfrey test for order lags 3 subsample medium credit risk bond funds	
Ho: No autocorrelation at any order less than or equal to p	
Ha: There exists autocorrelation at some order less than or equal to p	
Lagrange multiplier statistics	**F-value**
6.2694	1.2555
p-value	**p-value**
0.2809	0.2845

Table 10: Breach-Godfrey test for order lags 3 subsample of medium credit risk bond funds	
Ho: No autocorrelation at any order less than or equal to p **Ha: There exists autocorrelation at some order less than or equal to p**	
1 test stats	**2 test stats**
5.8829	1.9323
p-value	**p-value**
0.1174	0.1254

Appendix 5: Structural changes experiment

Table 11: Model Accuracy after 2009 for bond funds

	OLS		XGBoost		MLP	
	Mean	Sd	Mean	Sd	Mean	Sd
$R^2\ Test$	-0.2046	0.3063	-0.6268	0.5751	-0.4132	0.5489
$R^2\ Train$	0.0613	0.0069	0.3247	0.0266	0.0502	0.0220
$MSE\ Test$	0.0003	0.0001	0.0002	0.0001	0.0002	0.0001
MSE $Train$	0.0006	0.0001	0.0002	-	0.0002	-

Notes: OLS is conducted on dataset transformed to meet the OLS assumptions of normality, endogeneity, stationary and non-seasonality.

Table 12: Model Accuracy on medium credit risk mutual bond funds after 2009

	OLS		XGBoost		MLP	
	Mean	Sd	Mean	Sd	Mean	Sd
$R^2\ Test$	-0.1397	0.0859	-0.5099	0.6112	-0.0944	0.0762
$R^2\ Train$	0.0353	0.0094	0.5134	0.0426	0.0629	0.0191
$MSE\ Test$	0.0002	0.0001	0.0002	0.0001	0.003	0.0001
MSE $Train$	0.0005	-	0.0001	0.0000	0.0003	-

Notes: OLS is conducted on dataset transformed to meet the OLS assumptions of normality, endogeneity, stationary and non-seasonality