Applied Finance Project

Multi-Asset Portfolio Construction and Return Prediction with Dynamic Features GitHub

1 Abstract

Our endeavor is twofold: to develop a deeper understanding of factors that drive multi-asset returns and volatility, and to use this understanding to construct dynamic and optimal portfolios. We have developed a holistic multi-asset factor model to predict the returns of 5 asset classes: Commodities, Equities, Treasuries, FX, and REITS. We employ features built on Macro Fundamentals, Technical Analysis, Pattern Recognition, Cross Sectional Returns and Regime classifiers based on Hidden Markov Model. As a part of this paper we have explored the feature importance across the asset classes and how the addition or removal of a category of factors impacts the R^2 of returns and Sharpe of the portfolio. We have explored the scope of ensemble methods in variance prediction and how it improves the OOS R^2 for realized volatility. Our investigation has exemplary Out of Sample prediction from 2017 to 2023, with ensemble of all factors giving an OOS average R^2 of 0.17. Patterns particularly look promising in improving predictive accuracy for currencies, equities and commodities. We have implemented utility maximization functions including Sharpe and MVO for portfolio construction. This methodology marks a significant advancement, establishing a comprehensive framework for analyzing multiple asset classes by combining feature creation, pattern and regime identification, and the application of non-linear analysis techniques. We demonstrate superior portfolio performance with our non-linear returns and volatility predictions.

Contents

2 Introduction

In the rapidly evolving financial markets, the quest for diversified multi-asset portfolios has become increasingly paramount among investors. This raises two important questions: How can we construct profitable portfolios, and what factors drive their performance? This study represents our endeavor to develop a deeper understanding of factors potentially impacting multi-asset return profiles and to build dynamically evolving portfolios. Our three-fold contribution includes advancing understanding of return and volatility drivers beyond current literature, incorporating regime importance for return/ vol predictions, and constructing dynamic portfolios with these drivers.

First, we delve into the relative importance of factors for return and vol predictions. For returns predictions, four broad categories of variables include macroeconomic indicators, technical analysis metrics, cross-sectional return analyses, and pattern recognition. We demonstrate the application of linear models, machine learning, and deep learning models for return/ volatility prediction and to study factor importance.

Second, we attempt to study the relative importance of regimes for returns and volatility predictions. For detecting regimes, we use clustering and HMM-based approaches. They are especially important as they capture the intuition of business cycles and macro factors interplay of the markets.

Third, leveraging our enhanced understanding, we construct portfolios optimized for a utility function of choice. We construct multi-asset portfolios and study the importance of how each category of factor can broadly impact the utility. These portfolios can differ based on holding periods, methods used for predictions, etc. For example, risk-adjusted returns such as the Sharpe ratio.

We also attempt to build this as a flexible framework that can be a toolkit for any potentially interested portfolio manager to just plug and play for multi-asset portfolio construction and assessment.

3 Literature Review

3.1 Returns

Literature has two fundamental ways to predict returns. First approach focuses on employing expected returns difference as forecasted by stock characteristics. This approach was popularized by Famma French in their paper [1] Second way focuses on forecasting a broad cross-section of returns using a set of macro-variables. This was studied in depth by Kojen et.al[2] and Rapach et.al[3]. Koijen's paper also attempts to understand other asset classes beyond For currency predictions, interest rates, US industrial production and exchange rates have been shown to be important currency return predictors[4] and that these returns are countercyclical. Bansal et. al[5] show that the predictability of bond returns and currencies based on bond risk premia, expected growth, and expected inflation. Bond risk premia is important earlier in [6]

3.2 Volatility Prediction

Volatility predictions differ from returns predictions on account of persistence. Engle (1982) propose ARCH and Bollerslev (1986) built upon Engle to propose GARCH. Success of GARCH on energy commodities has been discussed by Humprey [7]. Shiku Kui[8] intermixes XGBoost with GARCH to show better results with mixture models. We used this to run variance with GARCH as proposed by Bollerslev, and intermixed ML with GARCH to get better estimates.

3.3 Regimes

Regimes have gained a lot of traction in last few years and there have been efforts to include regimes in portfolio selection decisions. Uchida et al. [9] demonstrated improved portfolio selection performance when using regime-switching means and volatilities for construction.

Matthew et al, in general discusses Hidden Markov Model as a regime feature[10] where they identify different market regimes in the US stock market to propose factor investment models depending on the detected regime.

3.4 Portfolio Allocation and Optimization

Portfolio allocation has typically two components which are risk premia and volatility prediction. There are broad studies focusing on intra-asset class optimizations or multi-asset class allocation[11]. We focus on the latter and work to create optimal multi-asset portfolios. Cochrane[12] earlier work focused on the multifactor efficient frontier. Brandt et. al.[13] in their paper discussed a useful framework for intertemporal utility maximization for investor. Brandt's paper built on another paper from Brandt which was more fundamental and introduced a framework to optimize asset weights in portfolio based on asset 's characteristics[14]. Moreover, Bayesian approach Black Litterman has been contrasted with mean variance optimization and have shown promising improvements for out-of-sample multi-asset portfolio optimizations[15]

4 Data

We have developed 5 sets of features, which we have employed for prediction. These factors are price/ cross-sectional returns, Technical Features, Patterns, Regime, and Macro Factor Data. Our data, consists of primarily two parts: i) Gathering asset-specific prices and ii) gathering macro data. Our daily prices and macro data are gathered from 1st January 2001 through December 2023. We mention data sources and definitions below:

4.1 Price and Cross-Sectional Returns

We have considered assets from commodities, equities, FX, and fixed income to get broad exposure. We get the price data from Bloomberg. The list of assets considered in this case study are:

REITS	EUR/USD	USD/JPY
GBP/USD	10YR TREASURY	EQUITY INDEX
COMMODITY INDEX	OIL	SOY
CORN	SILVER	GOLD

4.2 Macro

The features have been developed using the following:

ID	Macro Variable	Description	Source
1	Inflation	The percentage Consumer Price Index (CPI) numbers released by the Federal Reserve monthly.	Bloomberg, FRED
2	Industrial Production MoM Changes	This feature captures the percentage change in industrial production from one month to the next.	Bloomberg
3	Unemployment Numbers	Monthly unemployment figures reflecting the percentage of the labor force that is unemployed and actively seeking employment.	Bloomberg
4	Uncertainty	A measure of economic uncertainty, potentially represented by indices such as the Economic Policy Uncertainty Index.	WRDS/CRSP
5	AAA Rated Bonds	Performance or yields of bonds rated AAA by credit rating agencies.	Bloomberg
6	BAA Rated Bonds	Performance or yields of bonds rated BAA by credit rating agencies.	Bloomberg
7	CAY	Consumption–wealth ratio as developed by Martin et al.	Martin Lettau Website
8	Fama French 5 Factors	Factors including market risk, size, value, profitability, and investment, commonly used in asset pricing models.	Fama-French Website
9	Risk-Free Rate	Interest rates on low-risk, short-term government securities such as Treasury bills, often used as a proxy for the risk-free rate in financial models.	Bloomberg
10	Money1 and Money2 Supply changes	The percentage change in Broad and Narrow Money supply	FRED

Table 1: List of Macro Variables and Their Descriptions

4.3 Patterns, Technicals, and Regimes

These are constructed from price data and macro data of the assets. No separate data requirement for these categories.

5 Methodology

With a focus on studying the importance of parameters for predictability and superior performance, we divide our study into four parts as listed below:

- Feature Construction
- Returns Model Construction
- Variance Construction
- Portfolio Construction

5.1 Feature Construction

We list all the features and how they were constructed in the table below for perusal. We also write how many features from each category were created.

Table 2: Feature Construction

Major Feature	**Constructed Sub Features**	**Total Features**
Macro	Raw values and three months lagged values	75
Pattern	Ta-LIB Patterns as mentioned in Appendix 1	61
Technical	15-day lagged self-returns, Denoised returns, skew/kurtosis/volatility, RSI, ACX, momentum, and reversals	28
Cross Section	Features derived from other assets including 15 lags	1000+
Regime	Hidden Markov Model and Clustering Method with 15 Lags	90

Denoising Returns: Our paper uses base returns in combination with denoised returns derived from a wavelet transformation at level 4. We applied this across all our 12 assets. This is a standard approach as mentioned in Zhifeng Dai, Huan Zhu, Jie Kang (2021) for the inclusion of technical indicators in factor models.

5.1.1 Patterns

Patterns reveal useful information about buying and selling behavior in the markets also representing dynamic buy-sell support and resistance levels. We use different patterns as provided by TA-Lib. Ta-LIB library is used to identify various candlestick patterns. These patterns include for example doji star, 3-black crows, etc. An exhaustive list of all the patterns used for the study is provided in the appendix. here.

5.1.2 Technical

For technical parameters, we have used the following standard technical parameters to capture momentum and reversals at different levels.

- Momentum of last 7 days, 15 days, 30 days, 45 days.
- Reversals of last 15 days and 30 days.
- RSI, ADX, Moving Average, Skewness, and Kurtosis.
- 15-day lags

5.1.3 Regime

Regimes represent an important idea and captures the intuition of changing global macro and business cycles. Two inter-related algorithms were used in this project to detect and cluster market regimes based on our set of features constructed. First, an unsupervised k-means clustering approach was taken, however our final algorithm was the Hidden Markov Model (HMM). HMM is a statistical model that assumes an underlying process to be a Markov process with unobserved (hidden) states. Compared to K-Means which creates clusters independently to the temporal order of the dataset (and therefore require several lagged features to do this), HMMs explicitly model the temporal dependencies in the data, making them suitable for time series analysis.

The hidden Markov model has finite sets of states, which are associated with probability distributions, and transitions among states are governed by a set of transition probabilities.

The Python library, HMMlearn, has various Markov Chain algorithms. For completeness, below we explain the most commonly used HMM algorithm, the Viterbi algorithm. This is a dynamic programming algorithm for obtaining the maximum posteriori probability estimate of the most likely sequence of hidden states called the Viterbi path that results in a sequence of observed events.

Given a sequence $O = (o_1, o_2, \ldots, o_T)$ and known transition and emission probabilities matrices A, B. For each hidden sequence S, the joint probability that both O and S happen is

$$\begin{aligned} P(O, S \mid A, B) &= P(O \mid S, A, B) \cdot P(S \mid A, B) \\ &= \prod_{i=1}^{T} P(o_t \mid s_t) \cdot \prod_{i=1}^{T} P(s_t \mid s_{t-1}) \\ &= \prod_{i=1}^{T} B(s_t, o_t) \cdot \prod_{i=1}^{T} A(s_{t-1}, s_t) \end{aligned}$$

We want to find the sequence S that maximizes the quantity above, that is

$$S^* = \arg\max_S P(O, S \mid A, B)$$

While this algorithm is implementable, it requires a strong awareness of the 3 model parameters pi, A and omega. pi being the transition probabilities, omega being the emission probabilities.

To avoid having to make uninformed prior estimates, instead we approached the HMM using a Gaussian mixture emissions model from the provided HMM library. The Gaussian emissions model assumes that the values in X are generated from multivariate Gaussian distributions (i.e. N-dimensional Gaussians), one for each hidden state. Each multivariate

Gaussian distribution is defined by a multivariate mean and covariance matrix. One can place restrictions on the covariance matrix, but in our case we used the 'full' criteria. The advantage of this model is that it fits a prior distribution to the training data itself.

For this project, the approach used was to run both K-means clustering and a Gaussian HMM model for a range of 3 to 7 market states. The models were trained from 2006 to 2016 with the entire universe of features and returns as clustering points to fit to. Then the market states/regimes were predicted for 2017-2023. We chose the criterion BIC to select the best model and states for each of the methods. BIC is defined as the Bayesian information criterion. It is a model selection tool given by the following formula:

$$\mathcal{BIC} = -2 \cdot \log(L) + p \cdot \log(T) \tag{1}$$

where log L is the log-likelihood of the fitted model, and k denotes the number of features of the model.

It was found that the 3 state and 5 state models were the ones with the lowest AIC/BIC combinations.

Below, is a cross-section for the Equity Index against a 3-state Hidden Markov clustering. It looks to capture the high volatility regime states in 2008 and 2020. Interestingly, this corresponds with high unemployment, low inflation, macro variables suggesting that there are bull and bear cycles captured by just these factors considered.

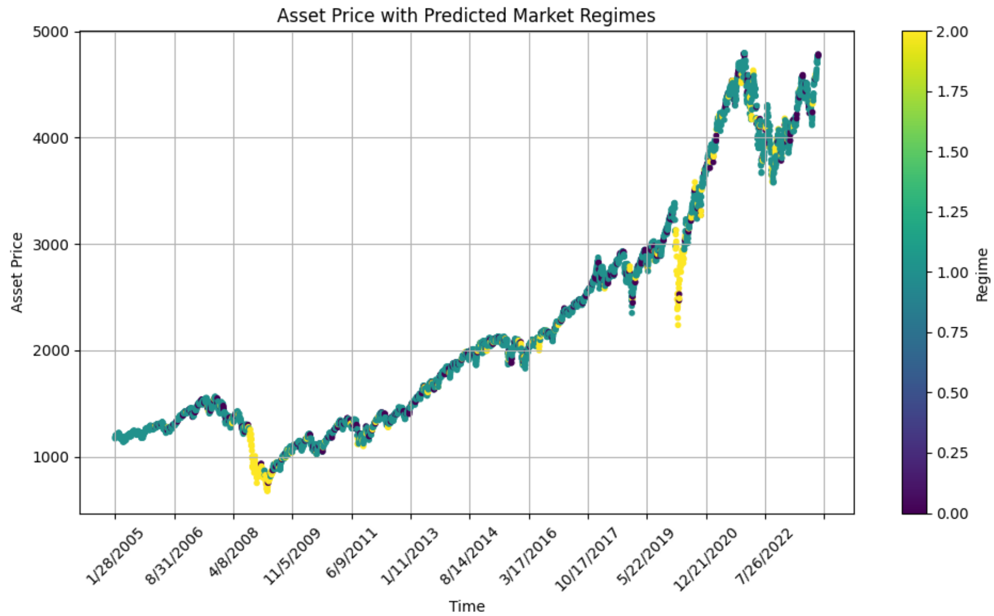

Figure 1: 3-state HMM with SPX index overlay

5.1.4 Cross-Sectional Features

We have two ways of predicting the returns of an asset class. Either we can use the features and factors of the asset class or we can add the features and factors generated from other asset classes as well. When we are using the technical and pattern-derived features of other asset classes we are terming it as using "Cross-Sectional Features".

- Use Technical and Pattern of Other Assets

5.2 Return Model Construction

The comprehensive model proposed by us takes 5 set of input features, employs dimension reduction and performs a non-linear prediction of returns.

Shown in Figure 2 the input contains all the features derived from the 5 factors. These are sent into a Lasso Feature Selection Model, which reduces the input predictors to a manageable number. The features are reduced to the tune of 90% from 1200 to < 20. Once we have the selected features they are given as an input to XGBoost Predictor. Prediction is

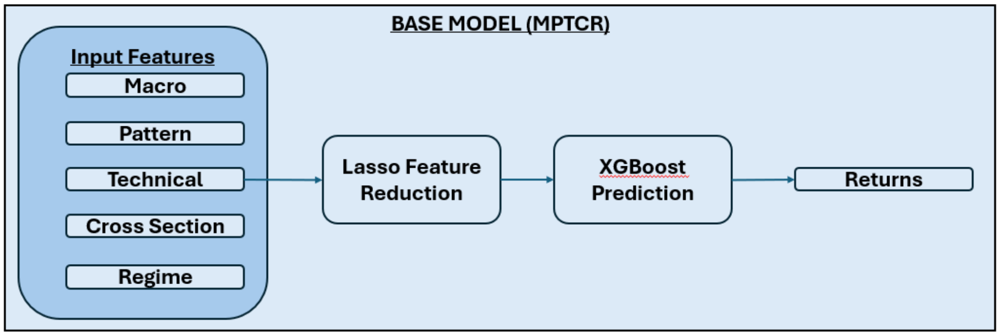

Figure 2: Comprehensive Model with all features

done for one holding period using a training set of the last 600 periods. For the prediction of the next period, the training period is rolled over by one holding period. For example, in the results which we will discuss below, our holding period is weekly. We take a training period of 600 weeks and roll it forward to predict out-of-sample returns for one-week returns.

We run four other models where we remove the Input Features one by one. This gives us in total five models which are summarized in the table Table 3:

Model Short Name	Input Features
MTCR	Macro, Technical, Cross Section, Regime
MPCR	Macro, Pattern, Cross Section, Regime
MPTC	Macro, Pattern, Technical, Cross Section
PTCR	Pattern, Technical, Cross Section, Regime
MPTCR	Macro, Pattern, Technical, Cross Section, Regime

Table 3: Return Models Used

5.3 Variance Model Construction

The comprehensive model for estimating variance uses Returns, Regime, and Macro as inputs. Returns are used directly for estimating conditional volatility by GARCH(1,1). GARCH(1,1) can be used as a standard estimator or can be fed along with a series of other predictors for prediction. Regime being a state variable is converted to dummy variables, and Macro variables are considered as they are.

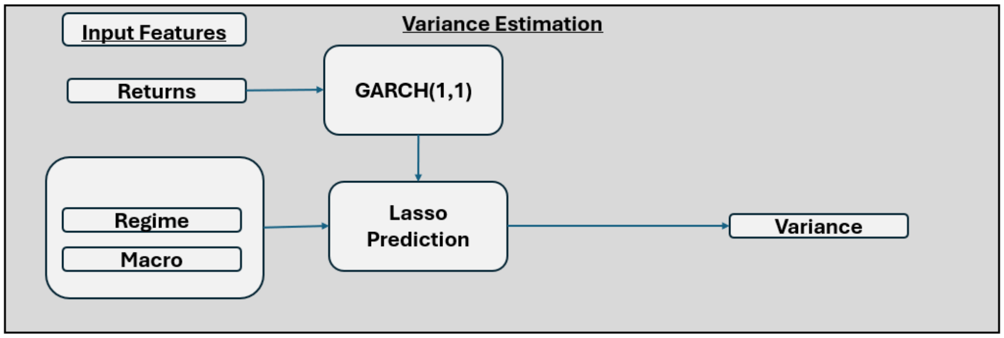

Figure 3: Variance Model

We have tried three other methods for variance estimation, and the detailed output is discussed in the results. A summary of all methods is provided in Table 4:

Table 4: Volatility Models

Model Name	Model Details 2
GARCH(1,1)	Pure GARCH Values
Linear Regression	Linear Regression on GARCH
Linear Regression	GARCH, Regime
Lasso Model	GARCH, Macro, Regime

For variance estimation, we have used studied two training periods, with a training period of 200 lags, and a training period of 22 lags. We estimate expected variances for a holding period of one week. The reason for using two horizons is, to conclude whether short-term variances or long-term variances are a better predictor of short-horizon (in our case) volatility.

5.4 Portfolio Construction

In this section, we consider portfolios based on the expected returns and expected variances calculated from previous models. We consider a utility function to maximize the utility under given constraints. This gives us the optimum weights. The correlation matrix is calculated using lagged returns of the last 26 holding periods. Since our holding period is one week, this would correspond to a half year.

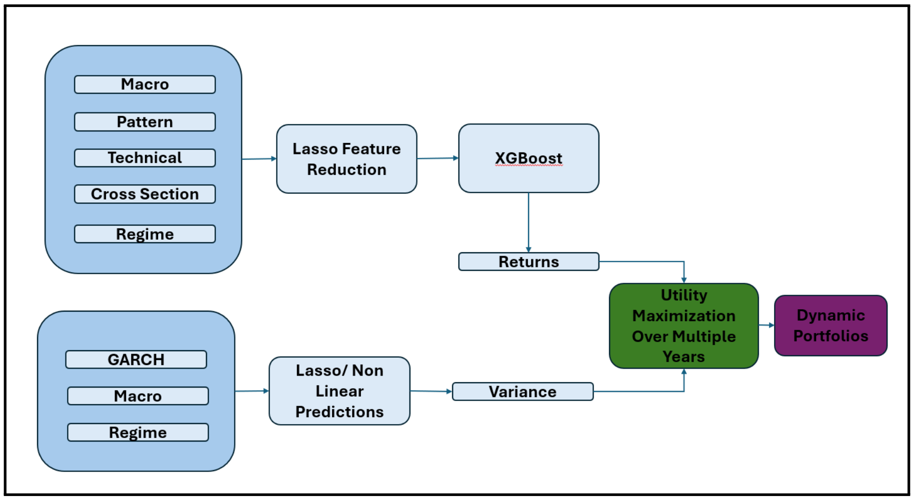

Figure 4: Portfolio Construction

Table 5: Optimization Models

Name	**Description**
Sharpe Maximization	$\max_w \left\{ \frac{w^T\mu - r_f}{\sqrt{w^T\Sigma w}} \right\}$
Mean Variance Maximization	$\max_w \left\{ w^T\mu - \frac{1}{2}\lambda w^T\Sigma w \right\}$

6 Results

6.1 Returns Results

We study out-of-sample performance as shown by R^2 criteria. We study the gain in R^2 due to each factor. We use a kitchen sink approach and start with the base model that contains all the factors. We remove each factor one by one to identify the loss in R^2. We see some fairly interesting results across the board. We show our predictions for a holding period of one week and rolling trained across 600 weeks. Following is a table for gain in R square.

Gain in R Square					
Gain	**Technical**	**Pattern**	**Regime**	**Cross-Asset**	**Macro**
REITS	-0.0353	-0.9580	0.0064	-0.1533	-0.1428
EURUSD	-0.0846	3.6467	-0.0007	0.0677	-0.0142
USDJPY	0.0284	46.2517	0.0039	0.0470	-0.1735
GBPUSD	-0.1984	7.8638	-0.0001	0.1324	0.0657
10Y Treasury	-0.6949	-0.6164	0.0036	-0.5070	-0.4181
Equity Index	0.3502	56.7345	0.0000	0.0641	0.2911
Commodity Index	-0.0938	2.6278	-0.0004	0.2356	0.0329
Oil	1.6104	-0.1257	-0.0114	0.4536	-0.1353
Soy	-0.0861	-0.2750	-0.0019	0.0621	0.0045
Silver	-0.2161	-0.1200	-0.0351	0.0156	0.0746
Gold	-0.0268	-0.2732	-0.0002	0.0198	0.0551
Corn	-0.4167	-0.5045	-0.0029	-0.0450	-0.2854

Figure 5: R Square Gain for Returns wrt factors

- At a broad level, cross-sectional returns and patterns add the highest average gains. Macro has the least impact overall.
- The regime variable that we are studying is a positive gain for Real Estate and 10-year Treasury but overall its impact we note is muted.
- Patterns are particularly important for FX markets, Equity index, and Commodity Index. These markets hence respect patterns and can form the basis for pattern-based strategies.

- The impact of Technical was minimal and was the highest for Oil and Real Estate.
- Macro Factors also caused small gains in R^2 and were most prominent for Silver and Gold.

6.2 Variance Results

For our case of volatility prediction, we employed two methods, which are a 22-day training period and a 200-day training period. Most of the assets perform similarly on both scales.

22 days training: R Square				
Asset	GARCH	Linear Regression on Garch	Linear Regression on GARCH and Regime	Linear Regression with feature selection
REITS	0.171	0.491	0.434	0.330
EURUSD	-5.698	0.261	0.103	0.268
USDJPY	-2.054	-0.141	-0.253	0.080
GBPUSD	-2.150	-3.536	-5.108	0.073
10Y Treasury	0.027	-1.171	-1.571	-0.006
Equity Index	0.201	0.341	-9.37E+23	0.081
Commodity Index	-0.362	0.166	0.038	0.221
Oil	0.000	-0.122	-0.136	-0.049
Soy	0.120	0.033	-4.30E+26	0.077
Silver	0.090	-0.151	-1.92E+24	0.056
Gold	-0.301	-0.397	-1.395	0.137
Corn	0.048	-0.028	-8.19E+24	-0.010

200 day training: R Square				
	GARCH	Linear Regression on Garch	Linear Regression on GARCH and Regime	Linear Regression with feature selection
REITS	0.171	-0.010	-0.047	-0.212
EURUSD	-5.698	0.128	0.111	-0.162
USDJPY	-2.054	0.120	0.103	-0.031
GBPUSD	-2.150	0.033	0.019	-0.005
10Y Treasury	0.027	0.026	0.028	-0.006
Equity Index	0.201	0.192	0.200	-0.032
Commodity Index	-0.362	0.215	0.214	-0.137
Oil	0.000	-0.032	-0.040	-0.003
Soy	0.120	0.159	0.155	-0.112
Silver	0.090	0.113	0.101	-0.107
Gold	-0.301	0.091	0.059	-0.327
Corn	0.048	0.056	0.017	-0.062

Figure 6: R Square Gain for Volatilities

- At a broad level, regression with macro-features added provided the best insights for volatility for short-term training for REITS and currencies, gold and commodities for a shorter time frame training.
- The regime variable that we are studying showed interesting contrast in shorter period vs longer training periods. This we can attribute to a higher importance of training regimes over a longer time horizon.
- GARCH with linear regression is quiet robust for a longer training period used for predictions.
- The impact of Technical was minimal and was the highest for Oil and Real Estate.

- Macro Factors also caused small gains in R^2 and these were most prominent for Silver and Gold.

6.3 Sharpe Maximization

We show in Figure 7 the results with different features as input. For this exercise, we tested returns/ volatility based on both linear(lasso) and nonlinear models. We present results for xgboost for this exercise. We compare the Sharpe generated with two XGBoost portfolios. In one (shown in Cyan below) there is no constraint on weights and we can have weights between -0.2 to +0.2. On the contrary, XGBoost Long only (shown in Orange Colour) has constraints on weights between -0.2 and 0.2. We conclude that Patterns have the highest impact on Sharpe. MTCR (without Patterns) has a very small insignificant Sharpe for both the methods and the adding of patterns leads to a significant increase. It is also interesting to note that while PTCR has higher returns, MPTCR has lower returns despite using xgboost to capture non-linearities. This suggests that non-linearities are not getting captured as intended.

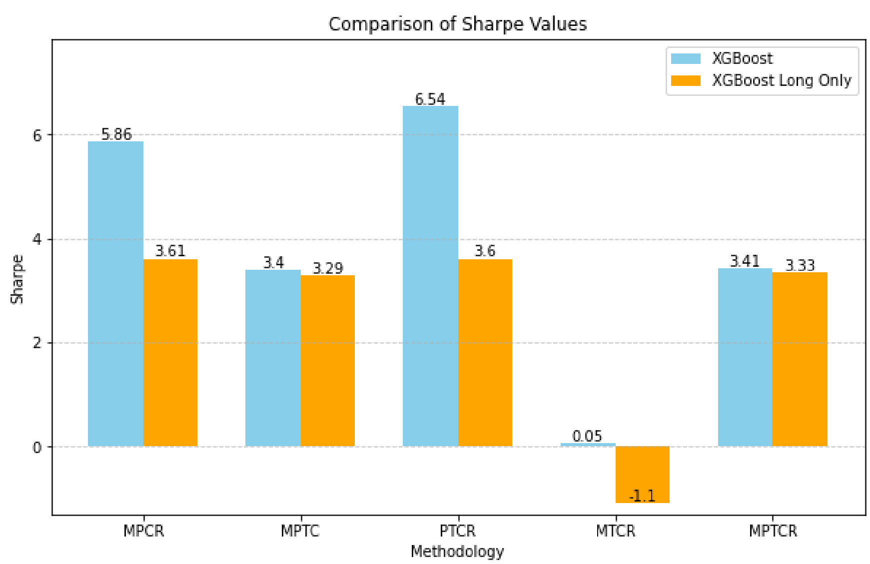

Figure 7: XGB Sharpe Comparison

We have further shown in the below figure, how returns vary for different strategies. For the best performing Sharpe in Figure 8 the returns are positive in all the years. For Figure 9 we see that the returns are still positive but they are capped.

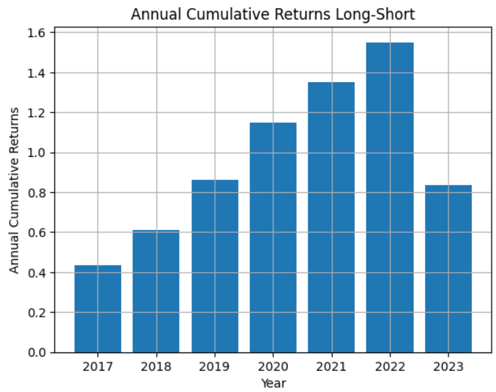

Figure 8: PTCR Long Short Portfolio

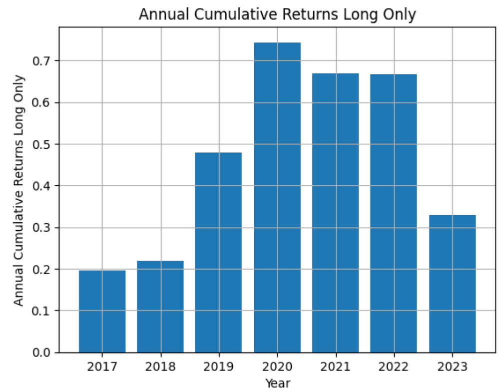

Figure 9: PTCR Long Only Portfolio

We have further shown Figure 10, how returns vary for the worst Sharpe portfolio: MTCR. Again the performance of uncapped weights is overall better than Long only portfolios.

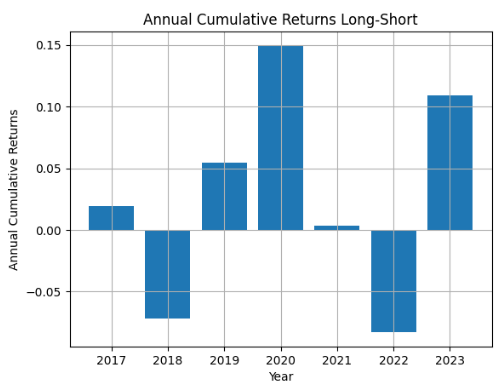

Figure 10: MTCR Long Short Portfolio

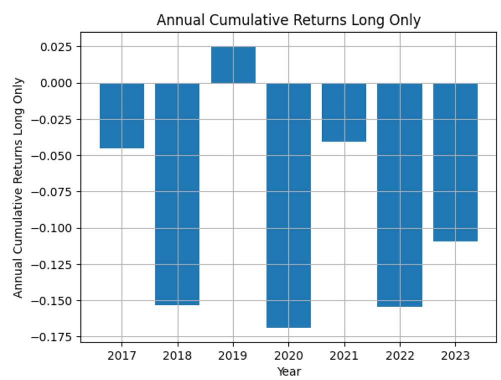

Figure 11: MTCR Long Only Portfolio

Portfolio Sharpe is significantly impacted by pattern inclusion. Given below are the annualized returns of the average case, corresponding to the portfolio formed using MPTC.

The capped version of weights on the positive side as seen in Figure 13 has lower annualized returns than the long-short portfolio shown in figure Figure 12.

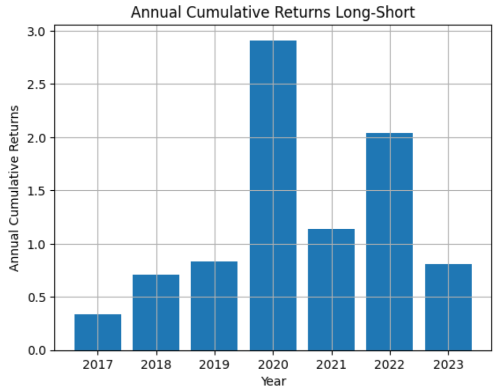

Figure 12: MPTC Long Short Portfolio

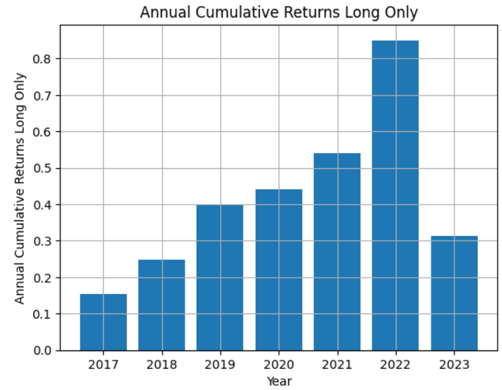

Figure 13: MPTC Long Only Portfolio

In conclusion, the five factors: Macro Fundamentals, Technical Analysis, Pattern Recognition, Regime Classifier and Cross Section Returns are robust predictors of returns. We experimented with different Machine Learning algorithms and the results were equally robust. However, there is no single factor that can alone explain all the asset classes returns; but in combination they do an outstanding work.

GARCH has been shown to have very good R^2 for realized volatility prediction. For some of the asset classes Machine Learning ensemble methods when clubbed with GARCH and Regime shows improved results.

Given the good estimates generated by the above two methods we can successfully create any Utility function. We have shown the results for Sharpe Maximization, but similar results can also be drawn for Mean Variance Maximization.

7 Appendix

7.1 Patterns

This section enlists all the pattern combinations that were considered for the study.

Pattern	Description	Pattern	Description
CDL2CROWS	Two Crows	CDL3BLACKCROWS	Three Black Crows
CDL3INSIDE	Three Inside Up/Down	CDL3LINESTRIKE	Three-Line Strike
CDL3OUTSIDE	Three Outside Up/Down	CDL3STARSINSOUTH	Three Stars In The South
CDL3WHITESOLDIERS	Three Advancing White Soldiers	CDLABANDONEDBABY	Abandoned Baby
CDLADVANCEBLOCK	Advance Block	CDLBELTHOLD	Belt-hold
CDLBREAKAWAY	Breakaway	CDLCLOSINGMARUBOZU	Closing Marubozu
CDLCONCEALBABYSWALL	Concealing Baby Swallow	CDLCOUNTERATTACK	Counterattack
CDLDARKCLOUDCOVER	Dark Cloud Cover	CDLDOJI	Doji
CDLDOJISTAR	Doji Star	CDLDRAGONFLYDOJI	Dragonfly Doji
CDLENGULFING	Engulfing Pattern	CDLEVENINGDOJISTAR	Evening Doji Star
CDLEVENINGSTAR	Evening Star	CDLGAPSIDESIDEWHITE	Up/Down-gap side-by-side white lines
CDLGRAVESTONEDOJI	Gravestone Doji	CDLHAMMER	Hammer
CDLHANGINGMAN	Hanging Man	CDLHARAMI	Harami Pattern
CDLHARAMICROSS	Harami Cross Pattern	CDLHIGHWAVE	High-Wave Candle
CDLHIKKAKE	Hikkake Pattern	CDLHIKKAKEMOD	Modified Hikkake Pattern
CDLHOMINGPIGEON	Homing Pigeon	CDLIDENTICAL3CROWS	Identical Three Crows
CDLINNECK	In-Neck Pattern	CDLINVERTEDHAMMER	Inverted Hammer
CDLKICKING	Kicking	CDLKICKINGBYLENGTH	Kicking - bull/bear determined by the longer marubozu
CDLLADDERBOTTOM	Ladder Bottom	CDLLONGLEGGEDDOJI	Long Legged Doji
CDLLONGLINE	Long Line Candle	CDLMARUBOZU	Marubozu
CDLMATCHINGLOW	Matching Low	CDLMATHOLD	Mat Hold
CDLMORNINGDOJISTAR	Morning Doji Star	CDLMORNINGSTAR	Morning Star
CDLONNECK	On-Neck Pattern	CDLPIERCING	Piercing Pattern
CDLRICKSHAWMAN	Rickshaw Man	CDLRISEFALL3METHODS	Rising/Falling Three Methods
CDLSEPARATINGLINES	Separating Lines	CDLSHOOTINGSTAR	Shooting Star
CDLSHORTLINE	Short Line Candle	CDLSPINNINGTOP	Spinning Top
CDLSTALLEDPATTERN	Stalled Pattern	CDLSTICKSANDWICH	Stick Sandwich
CDLTAKURI	Takuri (Dragonfly Doji with very long lower shadow)	CDLTASUKIGAP	Tasuki Gap
CDLTHRUSTING	Thrusting Pattern	CDLTRISTAR	Tristar Pattern
CDLUNIQUE3RIVER	Unique 3 River	CDLUPSIDEGAP2CROWS	Upside Gap Two Crows
CDLXSIDEGAP3METHODS	Upside/Downside Gap Three Methods		

Table 6: Technical Indicator Patters used

References

[1] Eugene F. Fama and Kenneth R. French. "Dissecting Anomalies". In: *The Journal of Finance* 63.4 (2008), pp. 1653–1678. ISSN: 00221082, 15406261. URL: http://www.jstor.org/stable/25094486 (visited on 02/29/2024).

[2] Ralph S.J. Koijen and Stijn Van Nieuwerburgh. "Predictability of Returns and Cash Flows". In: *Annual Review of Financial Economics* 3.1 (2011), pp. 467–491. DOI: 10.1146/annurev-financial-102710-144905. eprint: https://doi.org/10.1146/annurev-financial-102710-144905. URL: https://doi.org/10.1146/annurev-financial-102710-144905.

[3] David Rapach and Guofu Zhou. "Chapter 6 - Forecasting Stock Returns". In: Handbook of Economic Forecasting 2 (2013). Ed. by Graham Elliott and Allan Timmermann, pp. 328–383. ISSN: 1574-0706. DOI: https://doi.org/10.1016/B978-0-444-53683-9.00006-2. URL: https://www.sciencedirect.com/science/article/pii/B9780444536839000062.

[4] Hanno Lustig, Nikolai Roussanov, and Adrien Verdelhan. "Countercyclical currency risk premia". In: *Journal of Financial Economics* 111.3 (2014), pp. 527–553. ISSN: 0304-405X. DOI: https://doi.org/10.1016/j.jfineco.2013.12.005. URL: https://www.sciencedirect.com/science/article/pii/S0304405X1300305X.

[5] Ravi Bansal and Ivan Shaliastovich. "A Long-Run Risks Explanation of Predictability Puzzles in Bond and Currency Markets". In: *The Review of Financial Studies* 26.1 (2013), pp. 1–33. ISSN: 08939454, 14657368. URL: http://www.jstor.org/stable/23355402 (visited on 03/01/2024).

[6] John H. Cochrane and Monika Piazzesi. "Bond Risk Premia". In: *American Economic Review* 95.1 (Mar. 2005), pp. 138–160. DOI: 10.1257/0002828053828581. URL: https://www.aeaweb.org/articles?id=10.1257/0002828053828581.

[7] H. Brett Humphreys. “Applications of GARCH models to energy commodities”. In: *Dissertation Abstracts International* 58-07 (1997).

[8] Cui1 Peiyang Zhao Shikun Cui1. “Comparison of Stock price prediticon based on XGBoost and GARCH”. In: *BCP Business Management* 36 (1997).

[9] Hiroshi Ishijima and Masaki Uchida. “The Regime Switching Portfolios”. In: *Asia-Pacific Financial Markets* 18 (May 2011), pp. 167–189. DOI: 10.1007/s10690-010-9129-x.

[10] Mikhelson Matthew Wang Yi-Hong L. “Regime-Switching Factor Investing with Hidden Markov Models”. In: *Risk and Financial Management* 36 (2020). DOI: https://doi.org/10.3390/jrfm13120311.

[11] Jessica A. Wachter. “Asset Allocation”. In: *Annual Review of Financial Economics* 2.1 (2010), pp. 175–206. DOI: 10.1146/annurev-financial-073009-104026. eprint: https://doi.org/10.1146/annurev-financial-073009-104026. URL: https://doi.org/10.1146/annurev-financial-073009-104026.

[12] John Cochrane. “Portfolio Advice for a Multifactor World”. In: *SSRN Electronic Journal* (Jan. 2000). DOI: 10.2139/ssrn.218871.

[13] Michael Brandt. “Portfolio Choice Problems”. In: *Handbook of Financial Econometrics, Vol 1* 1 (Dec. 2010). DOI: 10.1016/B978-0-444-50897-3.50008-0.

[14] Michael W. Brandt, Pedro Santa-Clara, and Rossen Valkanov. “Parametric Portfolio Policies: Exploiting Characteristics in the Cross-Section of Equity Returns”. In: *The Review of Financial Studies* 22.9 (2009), pp. 3411–3447. ISSN: 08939454, 14657368. URL: http://www.jstor.org/stable/40247666 (visited on 03/01/2024).

[15] Wolfgang Bessler, Heiko Opfer, and Dominik Wolff. “Multi-Asset Portfolio Optimization and Out-of-Sample Performance: An Evaluation of Black-Litterman, Mean Vari-

ance and Naïve Diversification Approaches". In: *The European Journal of Finance* 23 (June 2012). DOI: 10.2139/ssrn.2081636.