Forecasting Treasury Inflation-Protected Securities Yield

Artem Tsalkovich, artermts@stanford.edu

Project Overview

The primary objective of this project is to compare the effectiveness of different machine learning models in predicting time series of TIPS, utilizing market data and macroeconomic variables as underlying features.

Figure 1. Historical time series for 5, 7 and 10y TIPS yields (%)

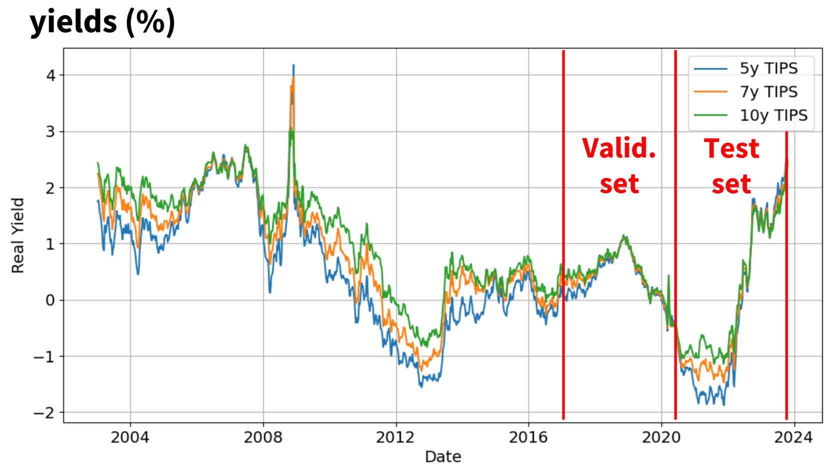


Figure 2. Historical time series for 10y TIPS yield differences (%, non-normalized)

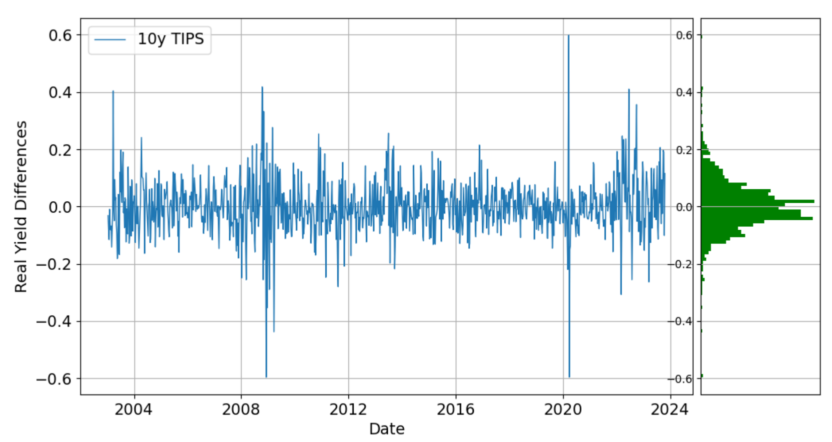


Table 1. Statistics for 5, 7 and 10 year TIPS (stationary, non-normalized)

Data	std	skew	kurtosis	min	25%	75%	max
5y TIPS	0.13	-5.69	119.1	-2.46	-0.06	0.06	0.80
7y TIPS	0.12	-4.60	91.1	-2.04	-0.05	0.05	0.71
10y TIPS	0.09	0.07	6.3	-0.60	-0.05	0.05	0.60

- Treasury Inflation-Protected Securities ("**TIPS**") are debt instruments issued by the US Treasury whose coupons and principal payments are indexed to the non-seasonally adjusted consumption price index ("CPI").
- These securities are among the few instruments available to investors that can provide **a hedge against unexpected surges in headline inflation**, similar to what occurred during 2021-2023.
- Similar to other fixed income instruments TIPS have a fixed lifespan in the end of which they are repaid back. The cash flows can be presented as

$$\sum_{i=1}^{T}(\text{coupon\%}\cdot\text{DCF}_i + 1\{i=T\})\cdot\text{CPI}_i\cdot\text{Notional}$$

- The dataset includes **1081 time points (weeks)**, covering the period from **Mar 2003 to Sep 2023**. The training, validation, and test sets are split in a **70:15:15 ratio**. The test set spans the period from Aug 2020 to Sep 2023, a time marked by heightened market volatility and a sudden surge in inflation.

Methods

Distinctive features of financial time series for market data:

- **Nonstationarity of joint distribution** – In addition to heteroskedasticity (time-dependent variance), the dependencies between different time series (correlations, causality, etc.) evolve over time.
- **Fat tails** – Most time series are not normally distributed and exhibit high values of kurtosis, i.e. relatively frequent occurrences of extreme values.
- **Low signal-to-noise ratios**, especially noticeable for time series with higher frequencies (daily or intraday).
- Models tend to demonstrate **both high bias and variance**; finding a delicate balance between them is a never-ending challenge for both researchers and market practitioners

In our study we considered **multiple machine learning models** split along three classes

Linear Models

- PCA regression. We considered 8 first principal components explaining 80% of variance.
- Regularized linear models (Ridge, Lasso, Elastic Net).

Ensemble Models

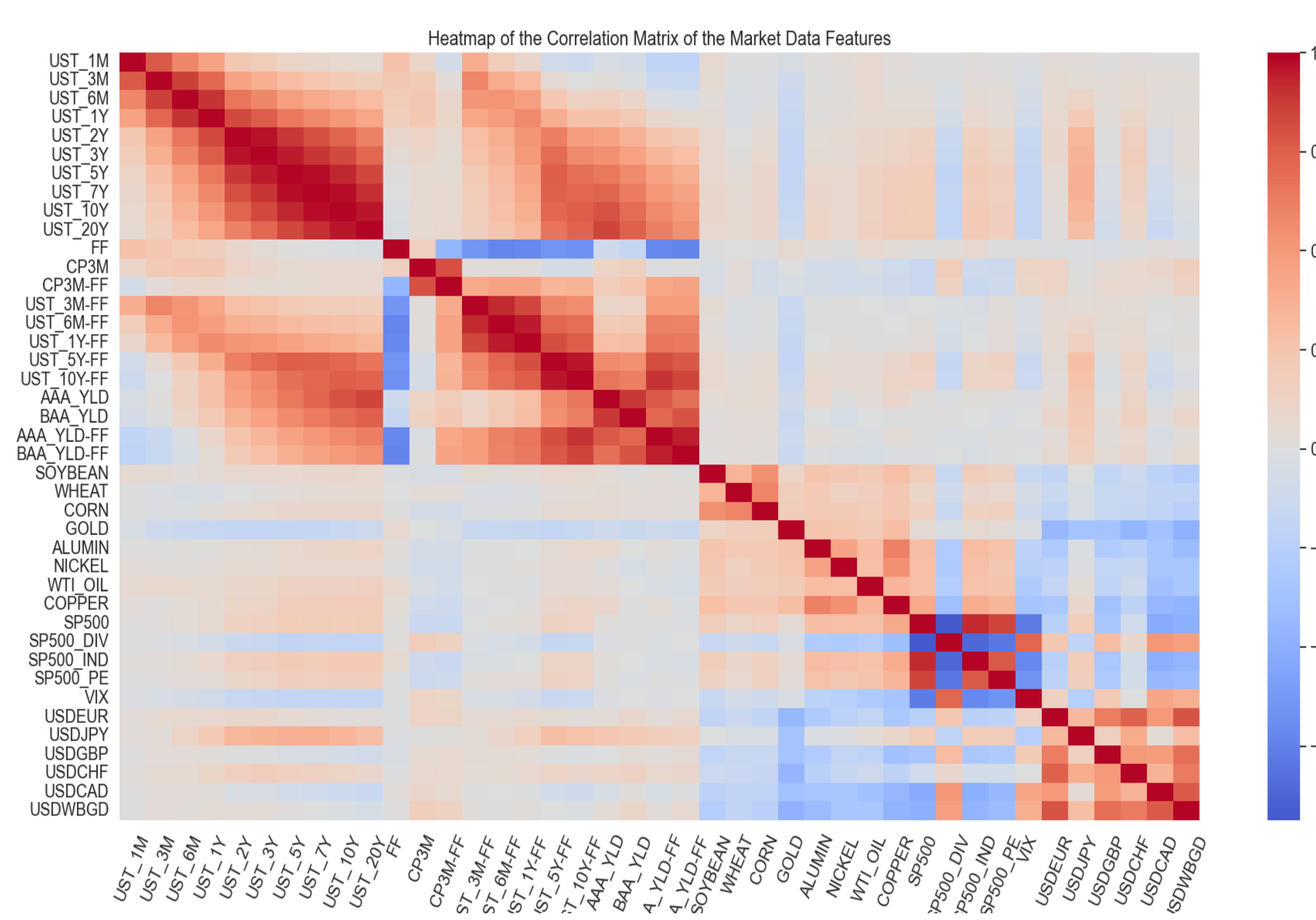
- ADABoost. Used as a baseline model for ensemble class of models, predictors – trees with max depth 3.
- Gradient Boost. Best configurations with up to 100 trees and max depth up to 6.
- XGBoost. Best configurations with up to 100 trees and max depth up to 8. Regularization parameter is dependent on TIPS maturity (higher for shorter tenors).
- Random Forest ("RF"). Best configurations with up to 100 trees and max depth up to 10.
- Extra-Trees ("ET"). Similar configurations to RF, trees are marginally deeper.

Neural Networks

- 1, 2 and 3 layer NN. We considered dense and shallow NN structures w/o dropout.

Dataset & Feature Selection

Figure 3. Heatmap of the correlation matrix of the market data features



Target variables – par real yields for TIPS with 5, 7 and 10 year maturities (3 time series)

- Time series for par yields are extracted from the actual daily trading data on prices as published by US Treasury. All time series are subject to stationarity adjustments and normalization.

Feature subset 1 – Market Data (41 time series). Source: US Treasury, Bloomberg

- **US Rates and Credit (22 features).** The historical data on the Treasury par yield curve rates for 1, 3, 6 month, 1, 2, 3, 5, 7, 10 and 20 year maturities. 3 month AA Financial CP Rate, FedFunds Effective Rate ("FF"), Moody's Aaa and Baa Corporate Bond Yields and 8 rate spreads between nominal yields and FF.
- **US Equity (5 features).** Time series for S&P 500 index, its industrial sub-index, actual PE ratios, dividend yields and VIX.
- **Commodities (8 features).** Gold, soft (corn, soybean and wheat) and industrial (aluminum, copper, nickel) commodities and WTI oil.
- **FX (6 features).** USDEUR, USDJPY, USDGBP, USDCHF, USDCAD and Trade Weighted U.S. Dollar.

Feature subset 2 – Macroeconomic Data from FRED-MD database (99 time series). Source: Fed

- Panel data of major U.S macroeconomic variables, maintained by the St. Louis Federal Reserve Bank. The database includes monthly reading of multiple features, which can be divided in the following six groups: (1) Output and income, (2) Labor market, (3) Housing, (4) Consumption, orders, and inventories, (5) Money and credit and (6) Prices.

Results & Discussion

Table 2. R-squared results for different models for the TIPS with 5,7 and 10 year maturities using only market data

Model	Train dataset			Test dataset		
	5y TIPS	7y TIPS	10y TIPS	5y TIPS	7y TIPS	10y TIPS
PCA (8 components)	37.0%	47.4%	64.9%	47.4%	53.2%	63.5%
OLS	49.9%	55.3%	71.4%	38.5%	48.1%	65.8%
Ridge	37.5%	45.0%	67.8%	52.6%	58.0%	66.3%
Lasso	47.4%	51.8%	68.5%	49.8%	57.3%	66.6%
Elastic Net	44.6%	45.9%	68.3%	50.2%	57.9%	66.4%
AdaBoost	78.1%	77.5%	73.6%	48.7%	55.1%	60.6%
GBoost	82.2%	96.5%	89.0%	54.0%	56.8%	63.6%
XGBoost	48.5%	63.2%	86.3%	54.2%	60.8%	65.7%
Random Forest	72.7%	85.3%	84.1%	60.2%	64.9%	65.2%
Extra-Trees	79.8%	85.1%	88.7%	58.3%	62.0%	63.5%
NN 1 layer	24.8%	47.6%	67.5%	16.7%	63.4%	55.6%
NN 2 layers	92.1%	93.1%	72.5%	46.9%	39.0%	63.2%
NN 3 layers	63.1%	64.8%	68.2%	44.1%	54.1%	63.9%

Table 3. R-squared results for different models for the TIPS with 5,7 and 10 year maturities using market data together with macroeconomic data from FRED-MD database

Model	Train dataset			Test dataset		
	5y TIPS	7y TIPS	10y TIPS	5y TIPS	7y TIPS	10y TIPS
PCA (8 components)	36.6%	45.3%	61.3%	53.1%	58.9%	61.4%
OLS	59.3%	64.5%	75.3%	1.7%	17.2%	58.8%
Ridge	25.0%	50.5%	71.2%	31.8%	58.5%	66.4%
Lasso	40.1%	47.1%	70.2%	57.2%	63.1%	67.8%
Elastic Net	37.9%	45.0%	71.0%	54.2%	59.7%	66.7%
AdaBoost	71.1%	75.3%	77.7%	40.4%	42.9%	59.1%
GBoost	12.3%	16.4%	80.3%	1.6%	4.0%	63.7%
XGBoost	48.9%	72.8%	87.0%	49.6%	46.5%	62.7%
Random Forest	83.5%	76.4%	93.0%	42.9%	42.4%	63.8%
Extra-Trees	82.7%	87.9%	88.9%	55.4%	60.3%	65.2%
NN 1 layer	68.8%	60.3%	81.1%	21.4%	8.3%	63.7%
NN 2 layers	35.5%	50.8%	88.8%	25.6%	45.6%	63.6%
NN 3 layers	56.0%	97.4%	86.1%	49.1%	33.9%	63.7%

- Ensemble models, such as **RF and ET regressors**, are better suited to fit the data for time series with higher patterns of nonstationarity. We have observed **statistically significant improvement in R^2_{OOS} versus baseline regularized linear models**.
- Time series with relatively close tenors (5, 7, and 10 year in our case) exhibit very distinctive behaviors, requiring separate analyses.
- The larger feature set (Table 3, market data + macro variables) resulted in noticeable overfitting. Such results can be potentially explained by significant spurious correlations between time series within each category of macro data, negatively affecting the weights for features with higher signal-to-noise ratios. Feature selection is crucial for financial time series, as even economically justified features can lead to increased variance.

Future Research

- Perform more scholastic analysis of feature set while focusing on FT and ET models.
- Consider larger datasets, encompassing daily or even intraday time series, which would likely yield more reliable models with reduced bias.
- Explore various NN architectures beyond dense and shallow models.