RWTH Aachen University

Bachelor's Thesis

The present work was submitted to the
Chair of Computational Economics

Determinants of Swap Curve Dynamics: Insights from Feature Importance Analysis

Author:
Mathis Makarski, B.Sc.

Examiners
Prof. Dr. Thomas S. Lontzek
Prof. Dr. Almut Balleer

Supervisor
Arthur Enders, M.Sc.

October 31, 2023

1 Introduction

Interest rate swaps are widely acknowledged as practical tools for hedging interest rate risk and trading on directional interest rate movements. Given the significant expansion and magnitude of the market for these derivatives, there is an essential need to understand the factors that determine swap rates. An important issue in this context is how the steepness of the term structure for interest rate swaps changes over time, which is connected to the evolving dynamics of interest rates and the broader economic landscape. Understanding changes in the steepness of the swap curve is critical for market participants because it can inform interest rate risk management methods and investment decisions. As a result, this research aims to shed light on the links between macroeconomic indicators and the steepness of the euro swap curve, adding to a more comprehensive understanding of interest rate movements in the European market.

1.1 Literature Review

The literature review contains an overview of existing research on the relationship between macroeconomic variables and the swap curve while highlighting the methodologies commonly used in the current literature. While there are studies that examine the forecasting of the term structure with variables derived from the term structure itself, often using techniques such as the dynamic Nelson-Siegel model (Luo et al. 2021) or stochastic models like Heath-Jarrow-Morton (HJM) (Shreve 2004), there are also studies that examine the explanation of macroeconomic variables with the term structure as an explanatory variable (Estrella 2005) or the explanation of the dynamics of the term structure with macroeconomic variables (Füss and Nikitina 2011). The not entirely unexpected aspect is that there is no consensus in the literature on the correct direction of causality, and it is entirely plausible that there is no single direction, as real-world effects can be complex and self-reinforcing. Nevertheless, there remains a clear need to investigate whether certain macroeconomic variables contain information relevant to changes in the term structure, regardless of the direction of causality.

This thesis focuses on the issue of explaining term structure movements with macroeconomic and financial variables. In the context of interest rate swaps, previous research has focused mainly on explaining the spread between swaps and corresponding government bonds, while research explaining swap curve movements is relatively scarce. Given the similarity between interest rate swaps and bonds, this review also addresses the literature that attempts to explain treasury yield curve movements to review the macro indicators and models used.

Various methodologies have been used to analyze the relationship between macroeconomic variables and the yield curve. There exists extensive literature that uses vector autoregression (VAR) models, for example, a study using a VAR model to explain Treasury yields with inflation and real activity (Bikbov and Chernov 2010). VAR models using other variables have also been studied in the literature (Chang et al. 2011), (Evans and Marshall 2007). Moreover, the Nelson-Siegel approach can be extended to a macroeconomic dynamic factor model to forecast the yield curve (Koopman and van der Wel 2013), (Härdle and Majer 2014). Especially regarding swaps, researchers have applied dynamic factor models to estimate daily variations in swap rates based on macroeconomic fundamentals (Lu and Wu 2009). Furthermore, nonlinear econometric models have been developed to jointly analyze fluctuations in macroeconomic variables and the yield curve (Doh 2011). There also exist approaches using neural networks to forecast the yield curve, as detailed in the work by (Nunes et al. 2018).

1.2 Problem Statement

The prevalent models to explain yield curve movements in current literature, especially VAR and dynamic factor models, are trying to forecast yields at specific points in time by a linear combination of weighted explanatory variables, assuming that these variables contain valuable information for forecasting. Given that macro data can be noisy and complex, it becomes difficult to assess the true importance of each explanatory variable to the problem at hand. The more fundamental question of whether certain macroeconomic variables contain at least some valuable information for explaining yield curve phenomena may be better assessed by a model that does not attempt to forecast exact yields at exact points in time. To address further methodological challenges associated with researching the relationship between macroeconomic variables and the change of the curve steepness, it is essential to consider the characteristics of different modeling techniques. Traditional approaches in financial research often rely on p-values to determine statistical significance, but these have severe caveats, as elaborated in chapter 6 of (de Prado 2020). Furthermore, the standard logistic regression model is only valid when working with a data set where the classes are linearly separable. However, the interdependent and nonlinear nature of financial markets raises concerns about the suitability of such methods, specifically for exploratory research. As an alternative, supervised learning by decision trees offers a different approach, allowing flexible modeling without requiring predetermined functional forms. Furthermore, we use the triple barrier labeling method and the permutation feature importance of (de Prado 2018) and (de Prado 2020) to address whether specific macroeconomic indicators contain information to explain swap curve movements. To the best of the author's knowledge, no literature currently focuses on exploring the use of feature importance to determine whether macroeconomic variables contain information to explain swap curve movements.

The remainder of this thesis is organized as follows: Chapter 2 will explain the labeling technique and data engineering process to obtain the feature space. Chapter 3 will provide an analysis of the feature space to facilitate feature preselection. It will also include an explanation of the learning model used and the methodology for assessing the importance of features. Furthermore, it will present the results obtained and provide an interpretation of these findings. Finally, chapter 4 concludes the thesis.

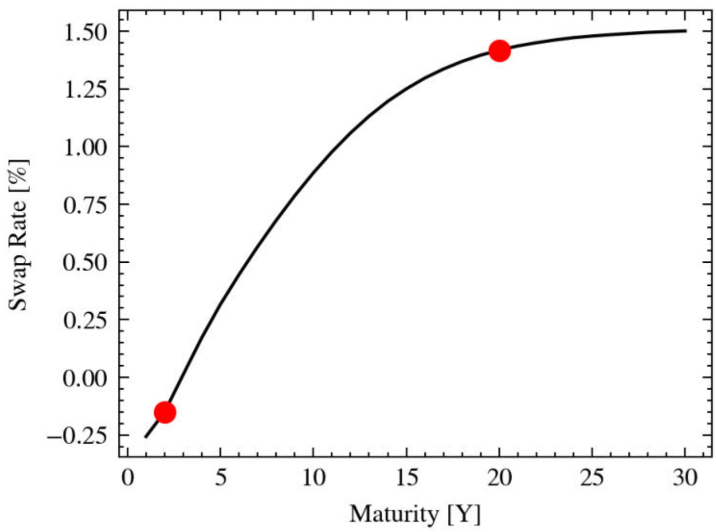
2 Data

This chapter provides an extensive analysis of the underlying data that supports this research thesis. We first uncover how the dynamics of the swap curve are modeled, followed by a review of the labeling method. Moreover, an examination of the feature engineering process is done. Throughout this chapter, we provide not only relevant statistics that reveal characteristics of the data but also use a special differentiation technique. We also reviewed the required data cleaning and modification procedures, highlighting their importance in guaranteeing the reliability of our study.

2.1 Swap Curve Dynamics

In this section, we will define the concept of swap curve dynamics, a fundamental component of our analysis. It is assumed that the reader has a basic understanding of interest rate swap agreements; however, comprehensive literature on the subject is available for reference for those who need a refresher or want to gain more insights (Sadr 2009). By plotting the swap rates of all quoted market instruments at one specific point in time to their corresponding maturity, as seen in Figure 1, one can obtain the term structure of interest rate swaps, also called swap curve.

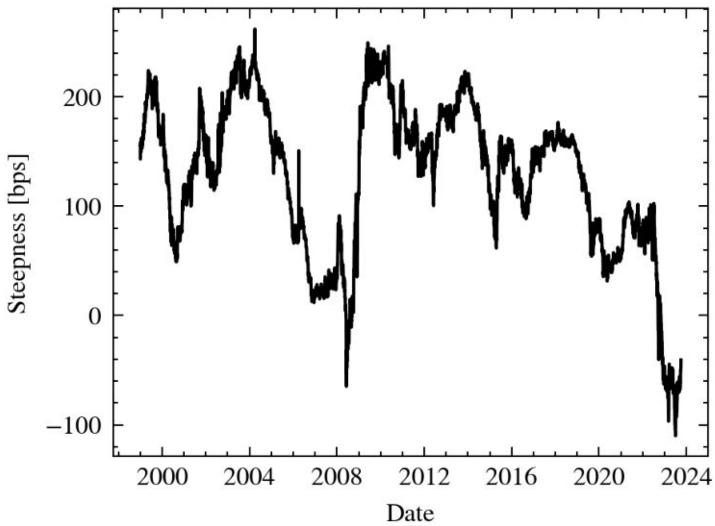
Figure 1: Swap Curve from 1st January 2018



This figure shows an example of a steep swap curve. The 2-year and 20-year maturity points are marked.

The steepness of the swap curve is derived by subtracting the interest rates at two specific points along the curve. This method serves as the basis for modeling curve dynamics in this thesis. In particular, we concentrate on the 2s20s Curve Steepness, defined by subtracting the interest rate at the 2-year maturity point from the interest rate at the 20-year maturity point.

Figure 2: 2s20s Curve Steepness



Higher values of the 2s20s Curve Steepness reflect steeper curves, which typically indicate expectations of future economic growth and occur in phases of economic expansion. In contrast, yield curves that are flatter or inverted, where the 2-year and 20-year rates are closer together, or where the 2-year rate is higher than the 20-year rate may signal economic uncertainty or pessimism, potentially foreshadowing economic contractions or recessions.

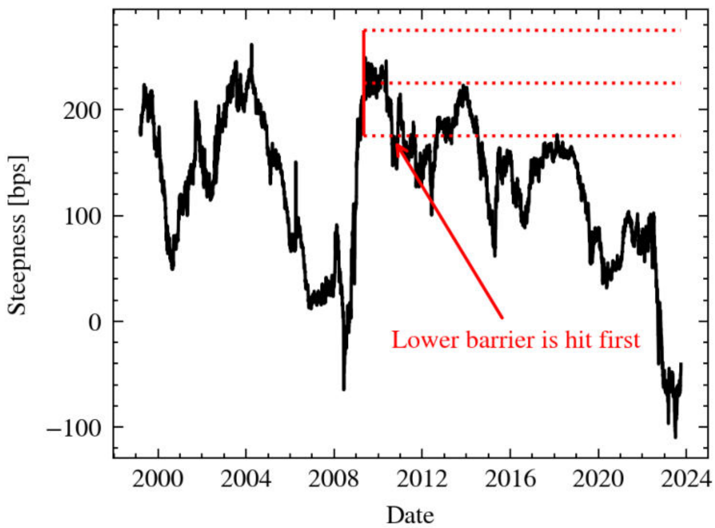
While we focus primarily on the 2s20s Curve Steepness, other popular combinations, such as the 2s10s and 2s30s, can also be explored in the context of swap curve dynamics. Additionally, the concept of a Fly, which shows the difference between three points on the swap curve, can provide further insight into market dynamics. However, for this thesis, we focus on the 2s20s Steepness as it provides a comprehensive lens to analyze and understand the dynamics of interest rate swaps.

2.2 Labeling

In this section, we address the aspect of labeling the 2s20s curve steepness by the triple barrier labeling technique, which was introduced in chapter 3 of (de Prado 2018). This procedure is central to defining the specific swap curve phenomenon since these labels are used to train the supervised learning algorithm. The triple barrier labeling technique is particularly relevant to our research because of its significant similarity to the perspective of an investment professional, which represents a fundamental difference from the typical forecasting problem in the prevalent literature.

The three barriers include two horizontal and one vertical barrier, each serving a specific purpose in the labeling process. The profit-taking and stop-loss targets dynamically define the horizontal barriers, and the vertical time barrier represents an expiration limit based on the number of bars that have passed since initiating the position. If the upper barrier is touched first, the observation is labeled "1". If the lower barrier is touched first, the observation is labeled "-1". If the vertical barrier is touched first, we have two options: Label the observation based on the sign of the outcome, reflecting realized profit or loss, or label the observation as "0," indicating that the position resulted in neither profit nor loss within the defined limits.

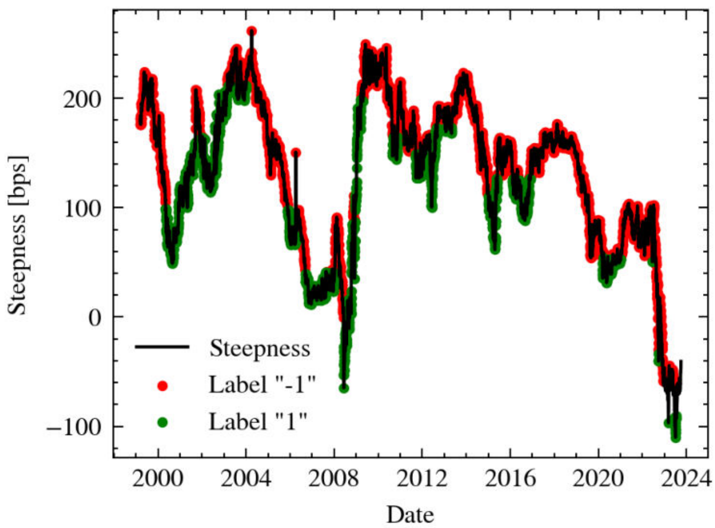
Figure 3: Barrier Labeling Method



This figure shows the barrier labeling technique of the 2s20s curve steepness at one example point with 50 bps targets and no time barrier. It is indicated that the lower barrier is hit first, so the observation is labeled with "-1".

In this study, we adapt the Triple Barrier Labeling Method with certain parameter choices to suit our specific research objectives. Our parameters include horizontal barriers with a target value of +/- 50 basis points (bps) and no time barrier. We choose specifically 50 bps as this produces a relatively balanced number of positive and negative labels. In addition, with such a parameter choice a barrier gets hit at least within a few years, but also significantly shorter when big price moves occur, thus making this the approximate time frame for which economic indicators are expected to contain valuable information for predicting that such a price move will take place. Furthermore, we omitted the time barrier because it would induce further complexity and make the results more challenging to understand. The labels generated from the barrier labeling method with certain parameter choices, as depicted in Figure 4, are essential to the exact problem statement because they serve as the basis for exploring whether the features in our dataset hold valuable information for predicting a 50 bps price movement.

Figure 4: Generated Labels



An important issue when using labels derived from future points in time is the need to prevent information leakage, which would lead to false discoveries in the case of irrelevant features. Thus, during the cross-validation process, all labels that intersect with the testing set were eliminated to ensure the model's integrity. This method is called purging and was introduced in chapter 7.4 of (de Prado 2018).

2.3 Feature Engineering

This chapter focuses on the feature engineering process, which consists of transforming and preparing the feature dataset to make it useable in the supervised learning algorithm. The feature space comprises a range of economic indicators that may be relevant to the problem statement defined through our labeling technique and can be retrieved from

Table 2 in the appendix. We have included the Core Consumer Price Index (CPI) and the Producer Price Index (PPI), standardized to 2010 prices. In addition, we have included the unemployment rate, gross domestic product (GDP), and three measures of money supply (M1, M2, and M3) denominated in euro. We also include the European Central Bank (ECB) refinancing rate. A detailed summary of the descriptive statistics of the dataset is presented in Table 1. Notably, our dataset includes price indices rather than percentage change indices. The only exceptions are the unemployment rate and the ECB refinancing rate, as these are already defined in percentage terms, which we choose to present in decimal format. The decision to use price indices rather than already differentiated series was made in preparation for the differentiation technique that we will use.

Table 1: Descriptive Statistics

Name	Count	Mean	Std	Min	Median	Max
CPI	333	98.6	11.2	79.9	99.4	124.0
PPI	512	88.0	22.3	50.8	82.8	172.4
Unemployment Rate	305	0.092	0.015	0.064	0.091	0.122
GDP	114	2.33e+12	5.41e+11	1.41e+11	2.39e+12	3.56e+12
M1 Money Supply	645	2.91e+12	3.10e+12	1.61e+11	1.43e+12	1.18e+13
M2 Money Supply	525	5.96e+12	4.07e+12	1.08e+12	4.57e+12	1.54e+13
M3 Money Supply	523	6.55e+12	4.27e+12	1.10e+12	5.38e+12	1.62e+13
ECB Refinancing Rate	264	0.0176	0.0151	0.0	0.0150	0.0475

Unemployment Rate and ECB Refinancing Rate are used in decimal format instead of percentage. The other variables are price indices representing an absolute level.

After compiling our feature set, our next steps involve data cleaning and manipulation. First, we use the forward fill technique to ensure that our feature dataset matches the daily resolution of the derived labels, which are based on the 2s20s steepness. Second, we restrict our dataset to the longest available time period in which each economic indicator and the 2s20s steepness have recorded data, resulting in a dataset that spans from January 10, 1999, to July 4, 2023. To expand our feature set, we introduce two additional variables, M1 and M2 money velocity, calculated as the ratio of GDP to M1 or M2 money supply, according to the methodology described in (Benati 2019). Third, we utilize fractional differentiation, a mathematical technique to balance the demands of stationarity in the features and preserve the maximum possible amount of memory. Finally, we apply standard scaling to each feature, which involves subtracting the mean and dividing by the standard deviation, which is necessary when dealing with features that are denoted in different scales, which is the case in our dataset.

Surprisingly, the fractional differentiation technique is rarely used in financial literature, so we cover the introduction to the mathematical background from chapter 5.4 (de Prado 2018) and provide a basic understanding of the fractional differentiation coefficient in this thesis. Consider the backshift operator B, that shifts a time series X_t by any integer $k \geq 0$ when applying it $B^k X_t = X_{t-k}$. The often used differentiation by $k = 1$ can be written as $(1 - B)X_t = X_t - X_{t-1}$ and represents the simple difference between timestep t and $t - 1$. To be able to choose fractional differentiation coefficients $d \in \mathbb{R}$ instead of integers the binomial series expansion is used

$$(1-B)^d = \sum_{k=0}^{\infty} \binom{d}{k} (-B)^k = 1 - dB + \frac{d(d-1)}{2!} B^2 - \frac{d(d-1)(d-2)}{3!} B^3 + ... \tag{1}$$

to define weights

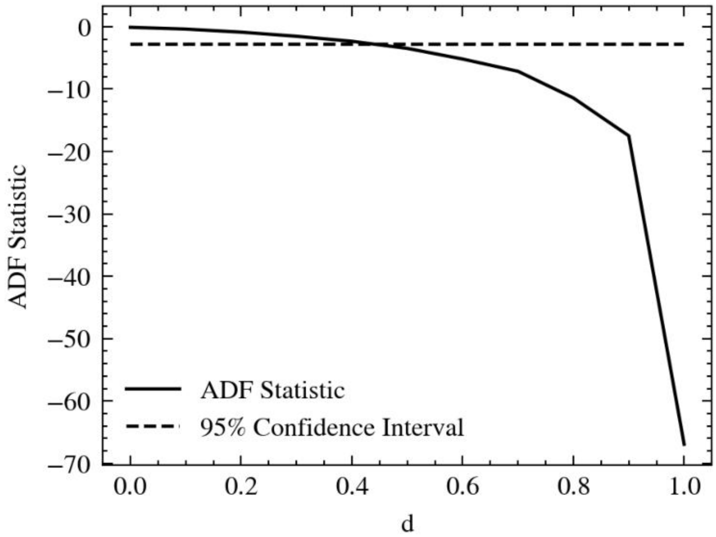
$$\omega = \left\{ 1, -d, \frac{d(d-1)}{2!}, -\frac{d(d-1)(d-2)}{3!}, ..., (-1)^k \prod_{i=0}^{k-1} \frac{d-i}{k!}, ... \right\} \tag{2}$$

that are used to finally obtain the fractional differentiated time series $\tilde{X}_t$.

$$\tilde{X}_t = \sum_{k=0}^{\infty} \omega_k X_{t-k} \tag{3}$$

For a more detailed derivation we refer to chapter 5 of (de Prado 2018). The determination of the fractional differentiation coefficient d for optimal balance between data stationarity and memory preservation is achieved through performing the Augmented Dickey-Fuller (ADF) test for different $d \in [0, 1]$ in 0.1-steps and computing d at the intersection of the linearly interpolated ADF test statistics with the 95% confidence interval as seen in Figure 5. The null hypothesis of the ADF test is that there is a unit root in a time series sample, while the alternative hypothesis is that the time series is stationary. As the ADF test statistic becomes more negative, the strength of the rejection of the null hypothesis increases, and the associated p-value decreases. Therefore, the more negative the ADF value, the more stationary the time series is considered to be.

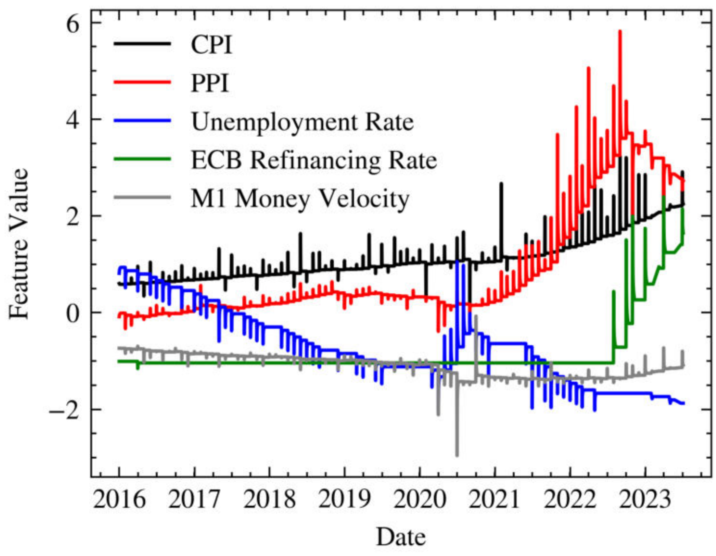
Figure 5: Finding the optimal differentiation coefficient



To achieve the optimal balance between stationarity and preservation of memory, our objective is to choose d to make the time series sufficiently stationary so that the test statistic reaches a 5% p-value. The 5% p-value was arbitrarily selected since it is a standard value for statistical tests. The resulting coefficients for the used data can be obtained from Table 3 in the appendix.

After completing all the data transformation steps, to provide an insight into the feature space, we include in Figure 6 a plot of selected features from 2016 to 2023, although the actual dataset ranges from 1999 to 2023 as previously mentioned.

Figure 6: Excerpt from the Feature Space



3 Results

In this section, we apply several methods to the previously derived feature space and present the results. First, we analyze the feature space by evaluating the correlation and the variation of information between each feature. We then make a feature preselection to prevent possible substitution effects in the feature importance analysis. Next, we apply the Permutation Feature Importance on the preselected feature subspace, which yields the main results in this thesis. Finally, the results of feature importance are interpreted in an economic context, followed by a critical discussion of the results and model limitations.

3.1 Feature Preselection

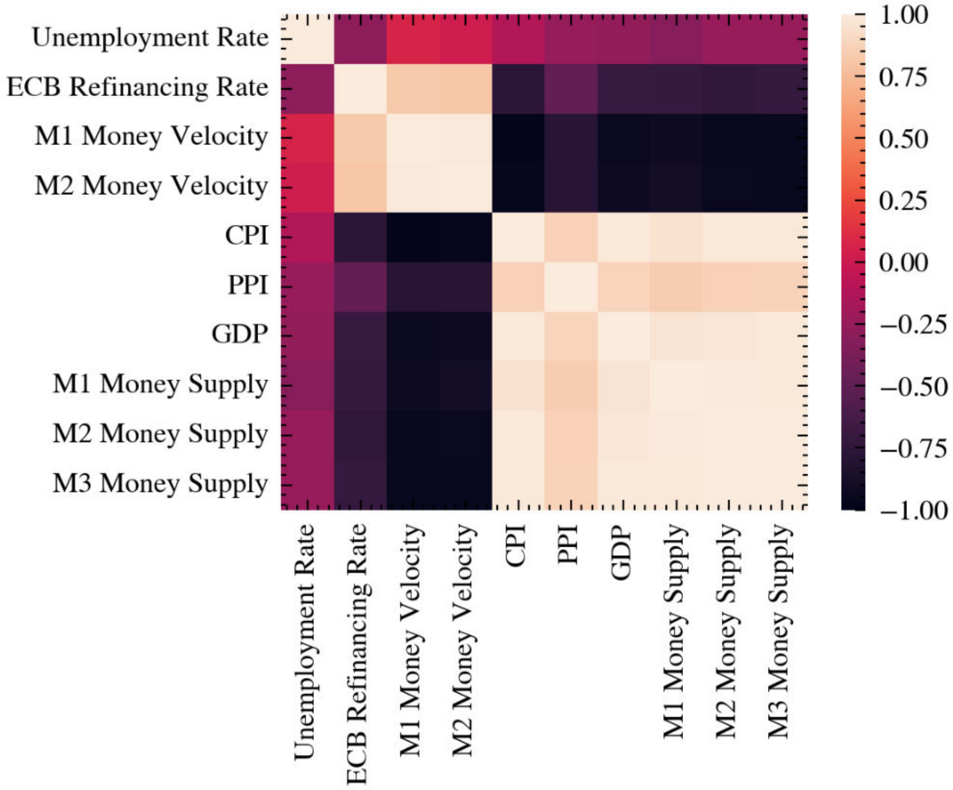
Before getting into the details of feature importance, it is essential to address the potential interplay between our features. When features share similar information, they can introduce a phenomenon known as substitution effects, which may affect the results of our feature importance methods. For example, in the context of Permutation Feature Importance, if two features are nearly identical, their importance will be halved due to their equal probability of selection, leading to an underestimation of importance even if they are critical (de Prado 2020). To mitigate these substitution effects and ensure the reliability of our feature importance analysis, we employ a careful feature preselection process by eliminating the simultaneous use of features that share a large amount of information. To evaluate the shared information between our features, we initiate a two-part evaluation. The first involves evaluating a correlation matrix, while the second focuses on the variation of the information metric. This helps us gain insight into the relationships between our features and identify potential sources of substitution effects.

The correlation matrix, shown in Figure 7, is calculated by the standard correlation coefficient[1] between two features for each combination and provides a visual representation of the relationships between the feature time series, which are sorted to reveal blocks of high correlation, allowing us to pinpoint critical findings. Three primary blocks emerge from this analysis. First, Unemployment stands alone in its block, as it has a low correlation with other features, suggesting its unique role in the dataset. The ECB Refinancing Rate and Money Velocity indicators form the second block. In particular, M1 and M2

[1] Also known as Pearson correlation coefficient.

Money Velocity show an almost perfect correlation. However, the ECB Refinancing Rate moderately correlates with the Money Velocity. In the third block, we find a cluster consisting of CPI, PPI, GDP, and all Money Supply indicators. In this block, PPI has a moderate correlation with the other characteristics, while the remaining indicators have a significantly high correlation.

Figure 7: Correlation Matrix



In scenarios where nonlinearity is prevalent, as argued by (de Prado 2020), the variation of information turns out to be a more appropriate distance metric than the correlation measure. This metric allows us to address questions about the unique information provided by a random variable, all without imposing specific functional assumptions. Since variation of information is rarely used in the financial literature, we provide the mathematical definition here for better understanding. Consider X a discrete random variable that takes a value x from the set S_X with probability $p[x]$. The entropy of X is defined as

$$H[X] = -\sum_{x \in S_X} p[x]log[p[x]] \tag{4}$$

The joint entropy of X and Y is

$$H[X,Y] = -\sum_{x,y \in S_X \times S_Y} p[x,y]log[p[x,y]] \tag{5}$$

The conditional entropy is defined as

$$H[X|Y] = H[X,Y] - H[Y] \tag{6}$$

The mutual information is defined as the decrease in uncertainty in X that results from knowing the value of Y:

$$I[X,Y] = H[X] - H[X|Y] = H[X] + H[Y] - H[X,Y] \tag{7}$$

Now we can define and transform the equation of the variation of information to

$$VI[X,Y] = H[X|Y] + H[Y|X] = H[X,Y] - I[X,Y] \tag{8}$$

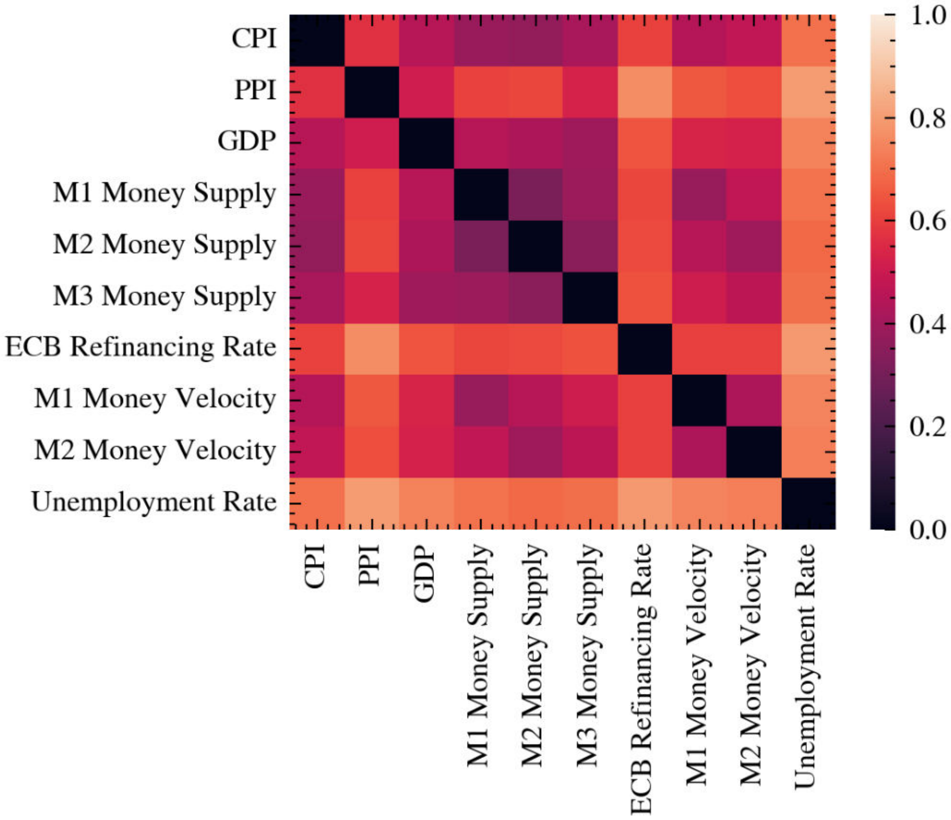
By standardizing the variation of information

$$\tilde{VI}[X,Y] = \frac{VI[X,Y]}{H[X,Y]} = 1 - \frac{I[X,Y]}{H[X,Y]} \tag{9}$$

we bound it between zero and one, thus making it compareable. Further, variation of information is a metric because it satisfies nonnegativity, symmetry $VI[X,Y] = VI[Y,X]$, and the triangle inequality. The variation of information measure can be interpreted as the uncertainty we expect in one variable when we know the value of the other. So, a lower value indicates that more information is shared between both variables.

As the variation of information matrix in Figure 8 shows, CPI and PPI appear as distinct entities in the matrix, indicating that they do not share a significant amount of common information. In contrast, GDP and all Money Supply Metrics form a visible block within the matrix, underscoring the significant common information among these economic indicators. Similarly, the Money Velocity indicators show a significant amount of shared information within each other and also, as expected with the Money Supply indicators, highlighting their interdependence. On the other hand, the ECB Refinancing Rate and the Unemployment Rate remain relatively isolated from the rest of the metrics, suggesting limited shared information with the broader set of variables.

Figure 8: Variation of Information Matrix



In light of the feature preselection analysis, we have determined that using Money Supply, GDP, and Money Velocity in combination is not advisable due to the significant amount of common information among these variables. As a result, we choose to retain Money Velocity as it is derived from both GDP and Money Supply and thus provides a more focused representation. Furthermore, the expected shared information between M1 and M2 Money Velocity leads us to retain M1 Money Velocity in line with the literature (Benati 2019). As a result, the features selected for the subsequent Permutation Feature Importance analysis include CPI, PPI, M1 Money Supply, ECB Refinancing Rate, and Unemployment Rate.

3.2 Baseline Model

The baseline model used in the feature importance analysis is constructed using a combination of the decision tree and the bagging classifier using the well-known sklearn library, as shown in the listing below. This approach is similar to Random Forests, although it has distinct characteristics and advantages. Further insights into their different strengths and properties are elaborated in chapter 6.4 of (de Prado 2018).

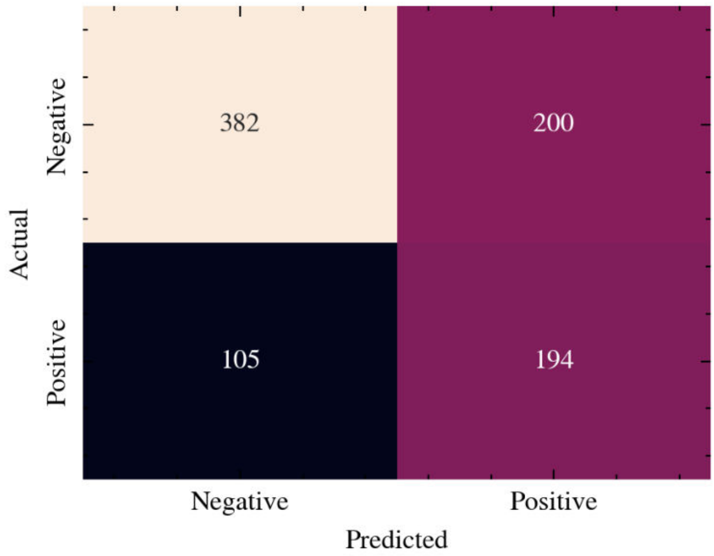
```
from sklearn.tree import DecisionTreeClassifier
from sklearn.ensemble import BaggingClassifier
clf = DecisionTreeClassifier(criterion='entropy', max_features=1,
    class_weight='balanced', min_weight_fraction_leaf=0)
clf = BaggingClassifier(estimator=clf, n_estimators=1000,max_features=1,
     max_samples=avgU, oob_score=False)
```

Listing 1: Implementation of the Bagging Classifier

In scenarios where a significant portion of the samples have non-identically and independently distributed (non-IID) characteristics, the problem of overfitting can persist. This is due to the process of sampling with replacement, which results in the construction of a significant number of essentially identical decision trees. Unfortunately, this overfitting phenomenon is a well-known weakness of decision trees, as each tree tends to capture noise and irregularities in the data. In (de Prado 2018), solutions for specific parameter settings are provided to counteract this problem. An essential technique for dealing with non-IID characteristics is to set the `max_samples` parameter of the bagging classifier to the average label uniqueness. This adjustment is critical because our label values are derived from different points in time, spanning different time intervals that overlap. A detailed description of this approach can be found in chapter 4 of (de Prado 2018), explaining how adjusting the `max_samples` parameter to label uniqueness effectively counteracts non-IID characteristics.

In the next step, we train the base classifier using the previously preselected subset of features, specifically M1 Money Velocity, ECB Refinancing Rate, Unemployment Rate, PPI, and CPI. To evaluate the performance of the base classifier, we present a confusion matrix in Figure 9 derived from a 10-fold purged cross-validation procedure. It provides an overview of the model's predictive capabilities, allowing us to evaluate its performance.

Figure 9: Confusion Matrix



Using the values from the confusion matrix, the accuracy of the base classifier is calculated to be 0.6538. While this accuracy level may need to be higher to present this model as a good classifier, it's important to note that this model serves a specific purpose in our analysis. Specifically, it is used for exploratory feature importance research, where our primary interest is in assessing the relative performance of the classifier.

3.3 Permutation Feature Importance

Permutation Feature Importance[2] (PFI) is a method for assessing the importance of individual features within a supervised learning model. The procedure begins by fitting the bagging classifier, which serves as a baseline model, to the data set. This classifier has already learned to make predictions. In the last section, the model's performance was evaluated with the accuracy metric because it is easy to interpret. However, for further purposes in the PFI analysis, we consider log loss[3] as the relevant performance metric. For a better understanding, we provide the definition of log loss as it was computed in the sklearn library. Let Y be a 1-of-K encoded binary indicator matrix, so $y_{i,k} = 1$ if sample i has taken label k from a set of K labels, and P be a matrix of probability estimates,

[2] Also called mean-decrease accuracy.
[3] Also called cross-entropy loss.

with $p_{i,k} = Pr(y_{i,k} = 1)$. Then, the log loss is

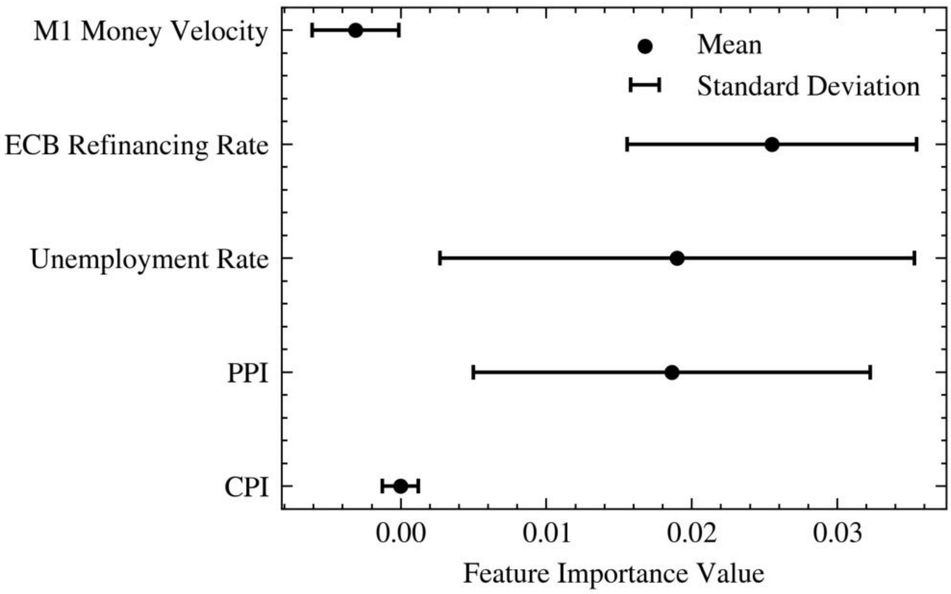
$$L[Y,P] = -logPr[Y|P] = -\frac{1}{N}\sum_{i=0}^{N-1}\sum_{k=0}^{K-1} y_{i,k} log p_{i,k} \tag{10}$$

The log loss considers not only how many labels are correctly classified, but also the probability with which they were chosen. The log loss profits from correct classifications with high confidence, while those with low confidence are weighted less. The reason why we use log loss and not accuracy as the scoring parameter is provided in chapter 9.4 in (de Prado 2018). However, in (de Prado 2020), it was suggested to use other scoring parameters such as probability-weighted accuracy, which we leave as a task for future research. The essential step in PFI is to isolate each feature individually and shuffle its values, effectively creating a randomized version of that particular feature while keeping all other features intact. With this modified data set, the bagging classifier is retrained to reflect its performance when the original relationship between the feature and the target variable is disrupted. By comparing the log loss of the new model L_1 to the baseline model L_0, we calculate the feature importance value F as the difference in log loss

$$F = L_0 - L_1 \tag{11}$$

which quantifies the impact of the shuffled feature on the model's predictive power. Thus, a positive feature importance value means a drop in model performance after shuffling, indicating that the shuffled feature contains important information. It is important to note that by using the 10-fold purged cross-validation procedure we obtain 10 feature importance values per feature from which we derive mean and standard deviation. The shuffling process is repeated for each feature in each cross-validation step. In Figure 10, we present the results of the feature importance analysis, which provides insight into the relevance of each macroeconomic feature in predicting a 50 bps change of the 2s20s steepness.

Figure 10: Feature Importance Results



Upon closer inspection, we see that M1 Money Velocity and CPI have feature importance means close to zero, accompanied by small standard deviations. These results strongly suggest that these features contain little substantial information critical for predicting the correct label, as their contributions to the model performance are negligible. The unemployment rate shows a positive mean in the feature importance analysis; however, it is essential to note the presence of a significant standard deviation. This means that the unemployment rate has a substantial but unreliable contribution to the model performance. On the other hand, both the ECB Refinancing Rate and the PPI have high feature importance values accompanied by lower standard deviations than the unemployment rate. Although the standard deviation of PPI may require a cautious interpretation, these results indicate that these features play a notable role in contributing to the correct classification of labels.

3.4 Economic Interpretation

The method presented, which is designed to examine economic indicators that may contain valuable information to explain changes in swap curve steepness, does not provide the ability to place the results in an economic context. Nevertheless, we provide an economic interpretation in this section, which is purely hypothetical and should be considered independent of the methodology used.

The objectives of the European Central Bank, which focus on achieving a balance between economic growth and job creation while maintaining price stability through monetary policy, support the idea that inflation-related features such as the CPI and the PPI, as well as the unemployment rate, could prove to be important indicators. These variables are likely relevant in monetary policy decision-making, as they are directly aligned with the ECB's objectives. The importance of PPI in the feature importance analysis may underscore the relevance of inflation-related indicators in explaining yield curve movements. The reason for the lower importance of the CPI on its own could be attributed to the fact that this indicator usually lags behind the PPI.

The well-established Phillips curve, which depicts the trade-off between inflation and unemployment, and the ECB objectives suggest that unemployment may be a valuable feature. However, the explanation for why unemployment does not emerge as a reliable feature, mainly due to its standard deviation, may be rooted in the well-known notion that unemployment tends to lag real economic activity. This lag is a notable characteristic of unemployment dynamics, as documented in (Levine 2013). While a viable economic indicator, the lagged response of unemployment to economic changes may explain its relatively high standard deviation and limited consistency in the feature importance analysis.

The research conducted by (Benati 2019) focuses on the characteristics of M1 Money Velocity, particularly its relationship with short-term interest rates. Their findings suggest that M1 Money Velocity is a permanent component of short-term interest rates. In particular, the paper finds that M1 Money Velocity has the characteristic of responding primarily to permanent shocks while largely ignoring transitory shocks in the short-term interest rate. This characteristic may be linked to our finding that M1 money velocity does not have a significant feature importance. The reason for this observation may be that by focusing on permanent shocks and ignoring transitory fluctuations, M1 money velocity may lack sufficient information to capture the dynamics of the swap curve effectively.

Insights from (Ejerskov et al. 2008) provide a supply-side analysis of the ECB monetary policy, revealing the relationship between the ECB's refinancing rate and liquidity management. According to their findings, the ECB smoothly manages liquidity imbalances by setting the refinancing rate at the level of the euro overnight index average (EONIA[4]), while trying to maintain a natural spread of 5 basis points. This spread serves as an indicator of market expectations regarding potential changes in the ECB's refinancing rate. A widening spread due to a change in the EONIA rate may indicate anticipation of a change in the ECB's monetary policy stance. At the same time, significant and unexpected changes in the ECB refinancing rate may also affect the EONIA rate. Considering the ECB refinancing rate as an indicator of supply-side liquidity, our findings of high feature importance with a relatively low standard deviation are supported by the current literature.

3.5 Discussion

The used baseline model may have a low predictive accuracy. Still, it is essential to note that the primary goal of this analysis is not prediction but rather the identification of features that contribute to the relative performance of the model. Furthermore, it is important to be cautious about the economic interpretation of our results. While our feature importance analysis provides valuable insights, false discoveries are always possible. In addition, any interpretations drawn from these results should be considered hypothetical and are subject to further investigation. Furthermore, it is essential to emphasize that the results from feature importance do not imply causality. Our model identifies the importance of economic variables in predicting changes in the swap curve steepness. However, it does not establish a causal relationship between the selected economic indicators and the behavior of the swap curve. Another critical aspect to consider is the substantial influence of the labeling process on the results. The choice of certain barrier thresholds in the labeling procedure defines the underlying question we are investigating, and variations in labeling may lead to different results. It is also worth noting that the results are subject to selection bias since the feature set used in our analysis was manually selected.

[4] The EONIA rate was replaced on 2nd October 2019 by ESTR (EURO Short Term Rate) as the standard reference rate

4 Conclusion

In this thesis, we have taken a distinctive approach to current financial research that differs from conventional methods that primarily focus on interest rate forecasting using predetermined functional models. Instead, we have adopted an innovative approach, as introduced in (de Prado 2018), that utilizes the power of machine learning without imposing a predefined functional form on the data. Our primary focus was to assess the feature importance of various macroeconomic indicators in understanding the dynamics of the swap curve, with the specific goal of explaining the factors contributing to a 50 basis point change in the 2s20s steepness of the curve. In pursuit of these objectives, we utilized ensemble learning techniques, explicitly a bagging classifier with decision trees. In addition, we incorporated advanced techniques rarely used in financial research, such as triple barrier labeling and fractional differentiation.

Our research shows that the ECB Refinancing Rate and the PPI exhibit significant feature importance, whereas the unemployment rate has a substantial but unreliable impact. In contrast, M1 Money Velocity and CPI do not exhibit notable feature importance values. This thesis pioneers a novel approach by applying the methods introduced in (de Prado 2018) and (de Prado 2020) to a specific financial problem related to interest rate swaps. In doing so, we have effectively set a new path in the field of research about yield curve determinants by demonstrating the applicability of these techniques in uncovering underlying determinants of swap curve dynamics.

On the one hand, future research can focus on refining the predictive capabilities of the classifier. This can be achieved through model selection, expanding the feature space, and a more systematic feature selection procedure. On the other hand, it is essential to emphasize that our methodology is limited in establishing causality. Future research can deepen the understanding of the important considered features and seek to uncover the economic mechanisms through which these macroeconomic indicators may influence swap curve dynamics.

A Appendix

Table 2: Data

Name	RIC	Start	End	Frequency
2Y Swap Rate	EURAB6E2Y=ICAP	1998-12-31	2023-07-04	Daily
20Y Swap Rate	EURAB6E20Y=ICAP	1998-12-31	2023-07-04	Daily
CPI	aXZCCORE/CA	1996-01-31	2023-09-30	Monthly
PPI	aXZCPPIE/CA	1981-01-31	2023-08-31	Monthly
Unemployment Rate	EUUNR=ECI	1998-04-30	2023-08-31	Monthly
GDP	aXZGDPV/CA	1995-03-31	2023-06-30	Quarterly
M1 Money Supply	aXZM1	1970-01-31	2023-09-30	Monthly
M2 Money Supply	aXZM2	1980-01-31	2023-09-30	Monthly
M3 Money Supply	EUMM3=ECI	1980-01-31	2023-08-31	Monthly
ECB Refinancing Rate	EUECBR=ECI	1998-12-31	2023-09-30	Monthly

Summary of the used data, including name, Refinitiv Instrument Code (RICs), start date, end date and data recording frequency.

Table 3: Differentiation Coefficients

Feature	d
CPI	0.84
PPI	0.58
Unemployment Rate	0.54
GDP	0.57
M1 Money Supply	0.83
M2 Money Supply	0.88
M3 Money Supply	0.86
ECB Refinancing Rate	0.45
M1 Money Velocity	0.66
M2 Money Velocity	0.64

B Code

The code used for the analyses can be found in the following GitHub repository. To gain access to this repository, reach out to the thesis author.

`https://github.com/4d6174686973/swap-curve-thesis`

References

L. Benati. Money velocity and the natural rate of interest. *Journal of Monetary Economics 116 (2020) 117–134*, 2019.

R. Bikbov and M. Chernov. No-arbitrage macroeconomic determinants of the yield curve. *Journal of Econometrics*, 2010.

C.-Y. Chang, C.-Y. Chen, and J.-H. Chou. Empirical evidence on the causality among yield curve factors and macroeconomic determinants. *International Review of Accounting, Banking and Finance Vol 3, No.3, Fall 2011 Page 48-69*, 2011.

M. L. de Prado. *Advances in Financial Machine Learning*. Wiley, 2018.

M. L. de Prado. *Machine Learning for Asset Managers*. Cambridge University Press, 2020.

T. Doh. Yield curve in an estimated nonlinear macro model. *Journal of Economic Dynamics Control 35 (2011) 1229–1244*, 2011.

S. Ejerskov, C. M. Moss, and L. Stracca. How does the ecb implement monetary policy? *Journal of International Money and Finance 27 (2008) 1199–1214*, 2008.

A. Estrella. Why does the yield curve predict output and inflation? *The Economic Journal, 115 (July), 722–744*, 2005.

C. L. Evans and D. A. Marshall. Economic determinants of the nominal treasury yield curve. *Journal of Monetary Economics 54 (2007) 1986–2003*, 2007.

R. Füss and O. Nikitina. Explaining yield curve dynamics. *The Journal of Fixed Income*, 2011.

W. K. Härdle and P. Majer. Yield curve modeling and forecasting using semiparametric factor dynamics. *European Journal of Finance, 12, 22, pages 1109-1129*, 2014.

S. J. Koopman and M. van der Wel. Forecasting the us term structure of interest rates using a macroeconomic smooth dynamic factor model. *International Journal of Forecasting 29 (2013) 676–694*, 2013.

L. Levine. Economic growth and the unemployment rate. *Congressional Research Service 7-5700*, 2013.

B. Lu and L. Wu. Macroeconomic releases and the interest rate term structure. *Journal of Monetary Economics 56 (2009) 872–884*, 2009.

D. Luo, T. Pang, and J. Xu. Forecasting u.s. yield curve using the dynamic nelson–siegel model with random level shift parameters. *Economic Modelling 94 (2021) 340–350*, 2021.

M. Nunes, E. Gerding, F. McGroarty, and M. Niranjan. A comparison of multitask and single task learning with artificial neural networks for yield curve forecasting. *Expert Systems With Applications 119 (2019) 362–375*, 2018.

A. Sadr. *Interest Rate Swaps and Their Derivatives: A Practitioner's Guide*. Wiley, 2009.

S. E. Shreve. *Stochastic Calculus for Finance II*. Springer, 2004.