Predicting Interest Rate Swap spreads behind the linear regression ECM and the Yield Curve, via Machine Learning

Luís Miguel Ribeiro Teixeira

Dissertation written under the supervision of Professor Dan Tran

Dissertation submitted in partial fulfilment of requirements for the
International MSc in Finance, at the Universidade Católica
Portuguesa, January 2024.

Abstract

Tittle: Predicting Interest Rate Swap spreads behind the linear regression ECM and the Yield Curve, via Machine Learning
Author: Luís Miguel Ribeiro Teixeira

This dissertation aims to forecast US 10-year Interest Rate Swap spreads out of sample using the last 20 years of data, which encompass significant events such as the 2008 financial crisis, the puzzle of negative spreads, liquidity shortages, and the COVID-19 pandemic.

This dissertation shifts from the traditional theory-driven approach to swap spreads, taking a statistical perspective aligned with investment banking practices that prioritize model performance and forecasting accuracy. Drawing on the work of Kobor et al. (2005) and Cortez (2003), it extends their linear regression Error Correction Model (ECM) to machine learning algorithms (Lasso, XGBoost, and Decision Tree Regressor), covering a wider time frame, and integrating new features to capture variations behind Treasury Supply-related features.

Main findings reveal a cointegration between U.S. dollar swap spreads and the supply of U.S. Treasury bonds, supporting prior evidence, while short-term deviations from the trend are associated with factors such as the AA spread, the repo rate, and the TED spread, but also to news data, sentiment, and uncertainty features. Another surprising key factor appears to be cointegrated with U.S. dollar swap spreads: the Google trend search for the term ‘Interest rate swap’.

Machine Learning models outperformed the linear regression ECM in predicting swap spreads, underscoring their potential in financial applications.

Keywords: Swap spreads, Lasso, Decision Tree Regressor, XGBoost, Machine Learning

Resumo

Título: Previsão de Interest Rate Swap spreads melhor do que o MCE de regressão linear e que a Yield Curve, via Machine Learning

Autor: Luís Miguel Ribeiro Teixeira

Esta dissertação tem como objetivo prever US 10-year Interest Rate Swap spreads fora da amostra, utilizando os últimos 20 anos de dados, que englobam eventos relevantes como a crise financeira de 2008, o puzzle dos spreads negativos, a escassez de liquidez e a pandemia COVID-19.

Esta dissertação afasta-se da tradicional abordagem teórica dos swap spreads, adoptando uma perspetiva estatística alinhada com as práticas da banca de investimento que dá prioridade ao desempenho dos modelos e à precisão das previsões. Inspirando-se no contributo de Kobor et al. (2005) e Cortez (2003), estende os seus Modelos de Correção do Erro (MCE) com regressão linear a algoritmos de machine learning, abrangendo um período de tempo mais alargado e integrando novas variáveis para captar as variações para além das variáveis relacionadas com a oferta de obrigações do Tesouro.

As principais descobertas revelam uma cointegração entre os USD swap spreads dos EUA e a oferta de obrigações do Tesouro dos EUA, corroborando evidências anteriores, enquanto os desvios de curto prazo da tendência estão associados a factores como o AA spread, a repo rate e o TED spread, mas também a dados sobre notícias, sentimento e incerteza relativamente ao mercado. Outro fator-chave surpreendente parece estar cointegrado com os USD swap spreads dos EUA: a tendência de pesquisas no Google do termo "Interest rate swap".

Os modelos de machine learning superaram o MCE de regressão linear na previsão dos swap spreads, sublinhando o seu potencial em aplicações financeiras.

Palavras-Chave: Swap spreads, Lasso, Decision Tree Regressor, XGBoost, Machine Learning

Acknowledgements

No fim, sempre escasseiam as palavras. As melhores que encontro, certamente limitadas, devem ser dedicadas na sua maioria à minha mãe, empreendedora, guerreira e justa. Cujo único limite que me impôs foi o limite do meu esforço. Umas outras à minha família abrangente, sem quem a maior vitória não teria sabor, não fossem deles os elogios e o reconhecimento. Sobram ainda umas palavras para aqueles que foram cruzando o meu caminho e com isso contribuindo para a minha formação e crescimento, de amigos a professores e outros quantos. E uma última palavra a mim. Ao pequeno eu. Que existia, já não existe, não da mesma forma, não com os mesmos horizontes de pensamento, mas que tinha em si a essência da luta, do sonho e da justiça, modestamente parte de uma herança genética.

Que este conhecimento adquirido me possibilite tornar o mundo num lugar melhor, desde logo com a mais pequena das ações do quotidiano.

Conhecimento é liberdade, e é reconhecer que pouco ou nada se sabe.

Porque os homens que mais admiro não são os homens perfeitos, mas aqueles que nunca se acabam.

Table of Contents

List of Tables

List of Figures

1. Introduction

An interest rate swap is a financial agreement between two parties to exchange interest rate payments without transferring the actual loan amounts. In a typical fixed/floating rate swap, one party commits to paying a fixed interest rate (known as the 'swap rate'), while the other party commits to paying a floating interest rate. These payments occur at regular intervals throughout the swap's duration and are calculated based on a predetermined notional amount. Typically, the floating rate is tied to the prevailing interbank interest rate for the relevant currency, such as the 6-month LIBOR for U.S. dollar swaps. At each settlement date, the floating rate is both paid and reset. The swap rate, representing the agreed-upon terms, is influenced by market conditions and essentially serves as the 'price' of the swap. Importantly, at each settlement date, the two parties don't exchange the full interest payments; instead, a single payment is made to account for the net difference between their agreed-upon payments.

Interest rate swaps have emerged as a fundamental financial innovation since the 1980s and continue to demonstrate remarkable growth on the global financial stage. Recent data from the Quarterly Report on Bank Trading and Derivatives Activities (June 2023) reveals that the notional value of interest rate swaps exceeded US$137 trillion by the first quarter of 2023, being the largest type of traded interest rate derivatives in the OTC[1] market with 86% of total notional value.

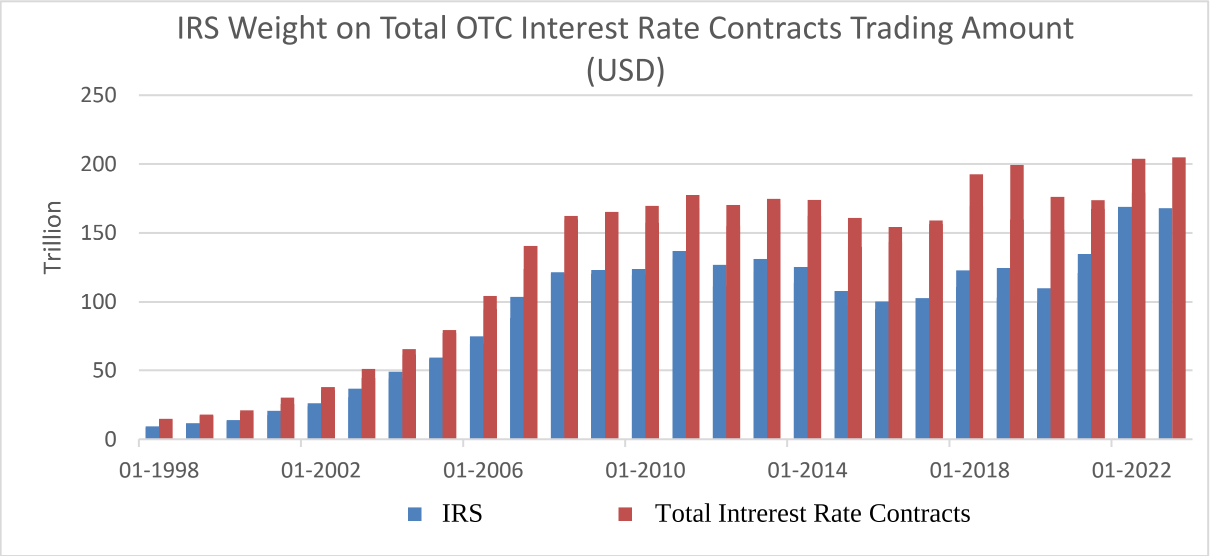


Figure 1 - IRS weight on Total OTC Interest Rate Contracts Trading Amount (USD)

---

[1] Over-the-counter means an asset that is not traded on an exchange but traded as a result of direct negotiation between buyers and sellers.

The swap spread plays a crucial role in determining the pricing of interest rate swaps, representing the difference between the swap rate and the par yield on a Treasury bond with the same maturity, i.e., the premium that investors require for holding a swap instead of a risk-free Treasury bond. For example, a market participant might agree to a market maker's bid, such as a swap spread of 40 basis points for a 5-year swap linked to 3-month LIBOR. By accepting this offer, the participant commits to paying the current LIBOR rate and, in return, receives a fixed interest rate of 5.40% (see In, Brown and Fang, 2003). These swap spreads are not static; they exhibit volatility and fluctuate over time. As the prevailing price of the swap, any changes in the swap spread can significantly impact the value of an existing swap position for all parties involved.

To make the concepts presented so far more visual, Figure 2 illustrates a plain vanilla[2] interest rate swap.

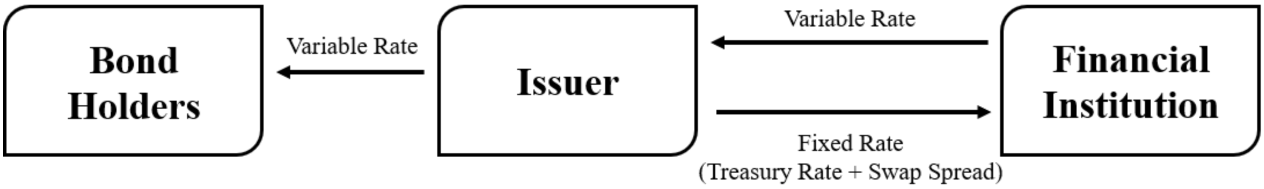


Figure 2 - Illustration of an Interest Rate Swap

Interest rate swap spreads are a key benchmark for pricing and hedging fixed-income securities. Consequently, financial institutions engaged in these markets require swap rate forecasts to fulfil various purposes (Moody's, 2017):

- Stress testing portfolios;
- Calculating expected long-term losses to meet accounting needs;
- Generating portfolio value projections for business-as-usual purposes;
- Forecasting value of bonds;
- Estimating market value of traded assets;
- Valuing cash flows contingent on future interest rates.

Swap spreads convey critical information about financial markets because they provide additional insights behind the swap rate, and their fluctuations have a decisive impact on some of the most essential financial operations. For example, an issuer of a bond constantly

---

[2] Another common type of swap is a currency swap involving the exchange of principal and interest payments in one currency for principal and interest payments in another currency.

exchanging its fixed payments by floating rate payments indexed to LIBOR, at issuance, will see its final borrowing cost as measured against the LIBOR curve (or swap curve), as a function of the relative evolution of the swap spread versus its own spread against Treasuries. A sovereign debt manager, on the other hand, will perform a comprehensive assessment of the medium-term evolution of swap spreads when determining whether to decrease the duration of public debt through a series of fixed-rate-receiver swaps (Kobor, A., Shi, L., & Zelenko, I., 2005). In any of the presented scenarios, the capacity to accurately forecast the spread level and its future direction holds considerable importance.

Investment bankers are also renowned for their practice of forecasting swap spreads as part of convergence trading strategies, contributing to the efficiency of the Interest Rate Swap (IRS) market. This strategy revolves around the concept that swap spreads will converge to a fundamental level. Traders who anticipate a decrease in swap spreads can potentially secure substantial profits with relatively low risk. They achieve this by taking a long position in an interest rate swap while simultaneously shorting a Treasury security, as discussed in the FRBNY Economic Policy Review (May 2006).

This combination of long and short positions serves to protect the trader from simultaneous and parallel shifts in swap and Treasury interest rates. The strategy relies on relative movements between these rates, and traders can realize profits if these movements align with their expectations. For example, if the swap rate decreases in comparison to the Treasury rate, the long swap position appreciates in value relative to the short Treasury position. Closing out the position results in a profit equivalent to the difference between these two positions.

Exhibited its relevance, this dissertation will delve into IRS spreads predictability, aiming to address the following research questions:

Q1: Can Interest Rate Swap spreads be reliably predicted out of sample, when controlling for look-ahead bias?

Q2: Are news data, sentiment and uncertainty features useful for predicting swap spreads behind the Yield Curve?

Q3: Can Machine Learning models surpass the predictive power of the linear regression Error-Correction Model (ECM)?

This research unfolds as follows. Chapter 2 consists of a Literature Review that synthesizes the most pertinent variables identified over the years. It encompasses an exploration of methodologies employed, shedding light on their strengths and limitations. Chapter 3 describes Data collection and description. The choice of predictors will be carefully justified, aligning with the variables deemed relevant by existing literature in explaining and forecasting swap spreads. Chapter 4, the Methodology section, details the application of the Error-Correction Model (ECM) and Machine Learning Algorithms. On Chapter 5, the obtained results will be analyzed and discussed. Key findings will be contextualized within the existing body of knowledge, offering a nuanced understanding of their implications. Finally, on Chapter 6, it will be presented the main conclusions from the analysis of the results as well as limitations and possible future research.

2. Literature Review

2.1. Theory Review

Early research on interest rate swaps aimed to establish a connection between swap spreads and their fundamental factors. The fair swap spread can be viewed as the present values of cash flows from both the fixed and floating legs and solving for the swap spread:

$$\sum_{t=1}^{T} \frac{Y_b + swap\_spread_t}{(1+r)^t} = \sum_{t=1}^{T} \frac{LIBOR_t}{(1+r)^t}$$

Where the swap_$spread_t$ is the difference between the swap rate and the Treasury bond yield for the same maturity; Y_b represents the yield of a Treasury bond with a given maturity; $LIBOR_t$ represents the floating interest rate for the same maturity as the Treasury bond at time t; r is the discount rate used to calculate the PV of cash flows; T is the number of payment periods until the swap matures.

This relationship tells us that the fair swap spread is influenced by the spread between the LIBOR and Treasury rates -reflecting a premium for the liquidity of the Treasury bond market and Banking Sector Default risk-, as well as the appropriate discount rate -which reflects counterparty default risk and the level of interest rates-.

2.1.1. Risk

Swap spreads function as a measure of relative pricing, prompting financial theory to scrutinize the structural disparities between swaps and government bonds, considering them as alternative financial assets. In fact, a swap can be thought of as a fixed rate bond funded by the short sell of a floating rate note, in a way that they exhibit the same type of market risk.

Regarding credit risk, however, the two assets differentiate from each other. Government bonds in general bear the risk of the country’s bankruptcy, which in the case of the US market can be said to be negligible. In contrast, two primary sources contribute to credit risk in swaps: Counterparty risk (embedded in the LIBOR index on the floating leg) and Banking Sector Credit risk.

Swap contracts entail counterpart default risk as they are traded OTC, lacking explicit backing from a clearinghouse or exchange. While the cash flows in an IRS contract are analogous to those of a par yield bond with the same maturity as the IRS, empirical evidence suggests that the counterparty default risk premium is typically higher in the bond market when compared to the IRS markets. Sun, Sundaresan, and Wang (1993) highlight critical distinctions between the two markets that support this claim:

1) In an IRS, there is no exchange of principal as it remains notional.
2) Immediate termination of the contract when a counterparty defaults.
3) Credit enhancement innovations over the years:
    a. Requirement of collateral by the parties involved in the swap.

       Litzenberger (1992) notes that weaker credit-rated counterparties are either simply rejected or required to collateralize the IRS contracts, rather than be quoted higher spreads. Litzenberger examined in detail credit-risk mitigation techniques associated with swaps and concluded that there was no reason for spreads to be sensitive to counterparties' credit quality.
    b. Agents may be demanded to marking-to-market.
    c. In contrast to a collateralized loan, where the lender is typically prohibited from liquidating the collateral once a bankruptcy petition is filed, the collateral supporting a swap can be liquidated and used by the solvent counterparty to offset a positive settlement amount (Johannes and Sundaresan, 2003).
    d. 10-year swaps usually contain credit triggers, meaning that if the credit rating of a counterparty falls below Investment Grade, the counterparty can ask for swap cash-settle.

Sorensen and Bollier (1994) and Duffie and Huang (1996) further argue that a swap between two-parties of similar credit quality should entail no default risk premium in either direction because of the symmetric nature of the contract.

Bomfim (2002) tests the robustness of the credit mitigation techniques at times of market stress. Over the period from 1993 to 2002, Bomfim (2002) compares swap dealers' quotes (midmarket) with "synthetic" swap rates resulting from the traditional identity made between swaps and portfolios of forward LIBORs. Bomfim (2002) estimates forward LIBORs from quoted futures contracts relying on different models to calculate the convexity adjustment. He finds no significant difference between swap-market rates and "synthetic" rates, a result hinting at the robustness of the credit mitigation techniques currently in use.

Regarding Banking Sector Default Risk, it arises from the difference between the LIBOR rate, which is the rate at which banks rated A to AA on average lend to each other in the short term, and the Treasury yield. This difference, which is influenced by a survivorship bias in the LIBOR indices, is expected to have a limited impact on swap spreads variation, except during times of economic turbulence when a generalized credit problem affects the entire economy.

Duffie, D., & Singleton, K. J. (1997) referred to this constant interbank credit rating that arises from the survivorship bias in the LIBOR as a refreshed LIBOR-homogeneous credit quality in swap rate. They further view swap rate as rates applicable to a bond issuer that would keep the same LIBOR-homogeneous credit quality over time. This paper served as a theoretical base of the LIBOR-swap curve, which is many times seen by companies as a better proxy for discount rates, rather than the US slope yield curve which is assumed to be credit-risk free.

One can therefore conclude that credit enhancement innovations seem to have mitigated counterparty default risk in IRS contracts over time, making it less relevant for pricing. Banking sector default risk, however, represented by the credit quality of the short-term interbank rate, contributes to swap spreads but has a minimal impact on their variability. During times of increased systemic risk, it may influence swap spreads, particularly in the long end of the curve (Sun, Sundaresan and Wang, 1993; Collin-Dufresne and Solnik, 2002; Feldhütter and Lando, 2008; He, Nagel and Song, 2022).

2.1.2. Liquidity Premium

Swap spreads also tend to widen as government bonds become significantly more liquid, which prompts alternative instruments like swaps to offer a liquidity premium over government bond yields. As explained by Grinblatt (2002), a convenience yield is earned due to the relative liquidity of Treasury notes and the ability to go "on special"[3] in the repo market, advantages not enjoyed by users of swaps. The author further argues that credit risk is irrelevant for the explanation of swap spreads, proposing that liquidity, through the convenience yield, is a more plausible determinant.

Liv, Longstaff, and Mandell (2002) provide further support for the idea that changes in swap spreads are primarily driven by liquidity risk. While default risk typically dominates, the liquidity aspect tends to be more volatile and, at times, can outweigh the default risk component in magnitude. Therefore, many fluctuations in swap spreads can be attributed to changes in the relative liquidity of swap instruments compared to Treasury Bonds.

Feldhütter and Lando (2008) also endorse the significance of the convenience yield as

---

[3] "On special" meaning that borrowers can get lower interest rates using Treasuries as collaterals.

a key factor in modeling swap spreads. They employ a six-factor model that encompasses Treasury bonds, corporate bonds, and swap rates to jointly price these components.

One can then conclude that the liquidity premium plays a significant role in determining swap spreads.

2.1.3. The Short Rate

As mentioned, swap spreads are also reliant on the risk-free short rate, which is not only used for discounting cash flows but also crucial in shaping corporate hedging and mortgage prepayment hedging decisions. Tuckman (2002) argues that the low interest rates in the turn of the millennium contributed to significant variations in swap spreads. Indeed, changes in Mortgage-Backed Security (MBS) durations, driven by lower interest rates and earlier debt repayment, made market participants to take long positions in the swap market (receiving fixed, paying floating) to increase durations. This, in turn, negatively affected the magnitude of the swap spread. Conversely, as interest rates increase, the opposite effect is also empirically observed.

2.1.4. Government Bond Issuance

During a cyclical economic slowdown, it's anticipated that tax revenues will decrease, leading to the government needing to borrow more money. This increased borrowing could cause government bond prices to drop due to the surplus supply, resulting in higher government bond interest rates and tighter swap spreads. Conversely, in periods of strong economic growth, governments typically reduce their debt issuance due to higher tax revenues, which could lead to wider swap spreads. Empirical evidence exhibits correlation between these expectations and swap spreads – when budget balance expectations are more positive, government bond issuance is expected to be smaller, causing swap spreads to widen. More recently, Du, Hébert and Li, 2023, explain how Treasury supply conducted to narrower spreads (even becoming negative).

2.1.5. Slope of the Yield Curve

When a recession is anticipated, economic agents seek stable sources of fixed income in turbulent periods, increasing the demand for long-term government bonds. To finance these purchases, they may sell short-term assets, leading to a flattening or even an inversion of the yield curve. In an inverted yield curve environment, swap spreads are likely to widen for two reasons. First, assuming that the term structure of swap rates remains constant, swap spreads for longer maturities tend to widen as long-term government bond yields decrease. Second, economic slowdowns typically bring increased risks to the stability of the financial system, raising expectations of future Libor-GC repo spreads and exerting pressure on swap spreads to widen.

2.2. Swaps Market Recent Milestones

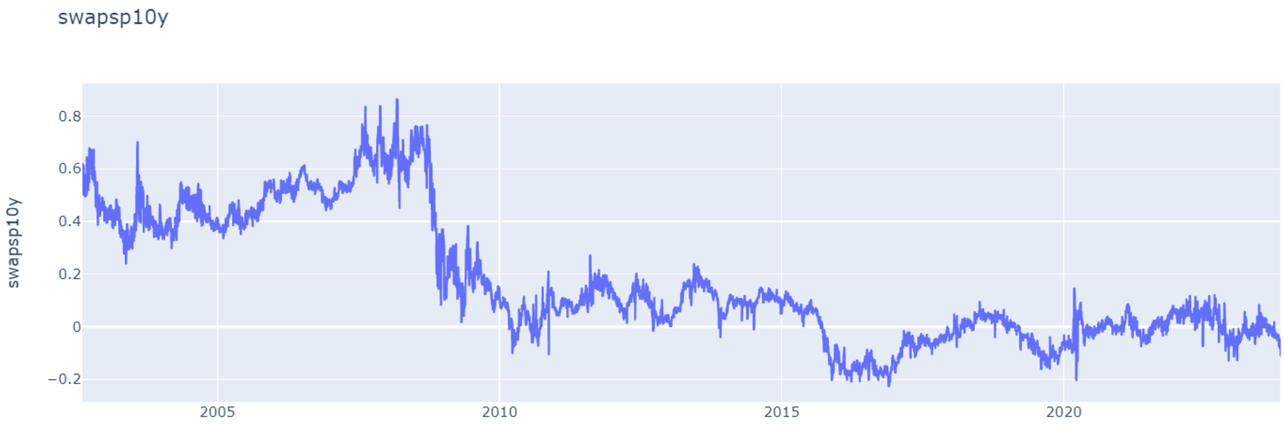


Figure 3 - Swapsp10y plotted over time

In the beginning of 21st century, the US Treasury made a significant announcement regarding debt buybacks, which had a profound impact on US dollar ten-year swap spreads. Market expectations were shaped by the belief that the US fiscal position would continue to improve. This optimistic outlook implied a reduction in the stock of US Treasury debt outstanding. The prospect of reduced Treasury supply had direct consequences, pushing down Treasury yields and widening swap spreads. These buybacks announcements highlight a possible influence of the changes in debt supply on the swap market.

With the rise of the 2008 crisis, the slope component of swap spread sharply increased and inverted, causing spreads to fall rapidly, even reaching negative territory. While Bank of England's response during the crisis explains part of this change, the persistent drop in spreads for terms beyond 10 years suggests that other factors, like financial market frictions, likely played a role.

The swap spread's level smoothly went back up during the stress periods of 2008 and 2012, which may indicate that the difference between swap rates and Treasury yields widened uniformly for a while. Similarly, the level went down during the post-crisis period of low-interest-rate policy, this may suggest that changes in yield levels and financial stress, like stock market volatility or the TED spread, can influence the overall swap spread.

Around 2020, US Federal Reserve announced the drop of its benchmark interest rate to zero, launching a new round of quantitative easing to respond to the global pandemic, which influenced swap spreads significantly. With interest rates at historic lows, investors sought higher yields, driving up demand for interest rate swaps. Simultaneously, the Federal Reserve's extensive asset purchases compressed yields on government securities, making swaps more appealing. These dynamics resulted in positive values for swap spreads, showcasing the profound impact of monetary policy on financial instruments.

2.3. Methodology Review

Academic papers generally provide evidence of a particular hypothesis, such as a theoretical component present in swap rates or spreads. Specifically, the credit risk or the liquidity premium (Kobor, A., Shi, L., & Zelenko, I., 2005). Investment banks, on the other hand, focus on top performer models. Therefore, their methodology can be distinguished.

2.3.1. Theory Driven Methodologies

Huang and Nefti (2002) employed a VAR model on daily data from January 1999 to March 2002. Their five-variable VAR model included the 10-year U.S. swap spread, the 10- to 2-year yield curve slope, six-month LIBOR, the market capitalization of the CSFB Liquid Corporate Bond Index, and the duration of this index. The study revealed that the slope, credit, and duration factors collectively explained a substantial portion (18% to 25%) of the variance in swap spreads. Notably, credit factor influence was more pronounced over longer horizons, and a widening of credit spreads tended to narrow swap spreads.

Francis, Brown, and Fang (2003), employ a multivariate EGARCH model to analyze volatility interactions across different maturities. The authors find that changes in swap spreads are positively related to interest rate volatility, shifts in default risk premiums in the corporate bond market, and alterations in the liquidity premium for government securities. Conversely, changes in swap spreads exhibit a negative association with shifts in interest rates and adjustments in the term structure slope.

Liu, Longstaff, and Mandell (2006) investigate the market pricing of default and liquidity risks embedded in interest rate swap spreads. Using a five-factor affine[4] framework, the study jointly models Treasury, repo, and swap term structures, estimating parameters through maximum likelihood. The credit spread is attributed to persistent liquidity and rapidly mean-reverting default intensity processes. The analysis assumes no counterparty credit risk and models the Libor rate's credit risk. Data from actively traded Treasury bonds and the general collateral government repo rate aid in identifying liquidity and default components. Results reveal that the credit spread in swaps comprises both liquidity and default components, with the latter being larger on average but the former exhibiting higher volatility. Liquidity risk is compensated with a significant premium, while the default risk premium ranges from zero to 30 basis points, with a flat term structure.

Feldhütter and Lando (2008) introduce a six-factor model to analyze Treasury bonds, corporate bonds, and swap rates. They break down swap spreads into three components: a convenience yield from holding Treasuries, a credit risk element from the underlying LIBOR rate, and a factor specific to the swap market. The convenience yield emerges as the largest contributor to spreads, followed by credit risk and a variable swap-specific factor linked to hedging in the mortgage-backed security market. The model provides insights into the relationship between hazard rates and LIBOR rates, shedding light on the riskless rate compared to swap and Treasury rates. The analysis employs an affine short-rate model for Treasuries and an intensity-based, affine framework for corporate bonds and swaps.

Azad, Batten, and Fang (2015) employ two regression models to explore the determinants of yen swap spreads. The main findings reveal that yen swap spreads are influenced not only by default and liquidity risks but also by business cycle risk, skewness risk, and correlation risk. The paper underscores the significance of the time-varying correlation between short-term market interest rates and longer-term government bond yields in shaping yen swap spreads. Additionally, it confirms the substantial role of liquidity risk, with default risk losing significance during the global financial crisis.

2.3.2. Statistical Driven Methodologies

Statistical driven models focus on developing top notch, highly accurate and robust econometric models, capable of making good predictions on swap rate or its spread.

Swap spreads can be particularly difficult to model given the nonstationary nature of the

---

[4] An affine model in finance is a mathematical framework describing interest rate changes over time. It assumes the expected return on a rate or bond yield is a linear function of the current rate and a set of underlying factors.

spread and its main recognised drivers. Nonstationary, i.e. the presence of trends, makes traditional econometric models like linear regression models inadequate in modelling series with this property.

2.3.2.1 Long-Term Determinants

Kocic and Quintos (2001) and Kurpiel (2003) explicitly point to the expectations on supply of Treasury bonds, based on U.S. fiscal deficit, as a key factor influencing swap spreads over the long term. Despite establishing a long-term correlation between the supply of government bonds and swap spreads, both studies fall short of confirming a cointegration relation due to the proxies, the data, and sample periods they use (1995–2001 and 1992–2002, respectively). When regressing U.S. dollar swap spreads on Treasury supply, Kocic and Quintos (2001) consistently find nonstationary residuals, interpreting them as signals of structural changes. To capture these changes, they opt for a nonlinear regression model that includes some feature engineering like multiplying the Treasury supply by other factors, like the slope of the yield curve, resulting in a long-term relation with stationary residuals.

Kurpiel (2003) studied both the European market and the US market. He reveals a cointegration relationship between EUR 10-year swap spreads and the German Treasury deficit, though he doesn't confirm the same connection in the U.S. dollar market. For U.S. dollar swap spreads, Kurpiel identifies evidence of a cointegration relation with U.S. Treasuries when introducing an additional long-term factor—the outperformance of bank stocks relative to the S&P index. However, despite this discovery, Kurpiel (2003) does not detail the implementation of an error correction model based on this long-term cointegration equation.

Cortes (2003) employs a Vector Error Correction Model (VECM) to analyze U.S. and sterling ten-year swap markets. Findings reveal that swap spreads, influenced by factors like the yield curve slope and equity-implied volatility, exhibit nuanced relationships. The Libor-GC repo spreads, on its turn, exhibits no correlation with swap spreads, something already documented in literature. In contrast, Mortgage-related factors play a key role in the U.S. market.

In Cortes (2006), the author employs a principal component analysis (PCA) before employing the Error Correction Model (ECM), to investigate the factors shaping the term structure of swap spreads. The key findings include evidence of international co-movement in swap spreads, indicating common factors such as a default term premium and expectations of government bond issuance. Additionally, idiosyncratic factors, including liquidity preferences and preferred habitat effects on both the government bond and swap curves, are identified. Variability in the term structure over time is attributed to factors like regulatory changes, demand and supply imbalances, and market inefficiencies.

Kobor, Shi and Zelenko (2005) employed an Error Correction Model to predict Interest Rate Swap spreads out of sample for all 2, 5, and 10 years, relying on the relevant features identified in the literature. While they argue for a good out-of-sample performance, their approach involved using moment t variations of predictors, which introduced a potential look-ahead bias.

Chowdhury and Wurm (2007) models and forecasts the term structure of interest rate swap spreads. The study builds 28 models for specific currencies, employing a two-step approach involving principal component analysis (PCA) and macroeconomic variables. The first step decomposes swap spreads into level and slope components, while the second step uses forecasts of macroeconomic factors to predict these components. Factors such as liquidity, default risk, yield curve shape, and demand for government debt are considered. The paper concludes by acknowledging the difficulty in predicting future swap rates due to unprecedented global interest rate trends.

2.4. Key Takeaways

Contemporary econometric models, developed post-2001, consistently integrate proxies for outstanding public debt as determinants of swap spreads. These models, predominantly based on linear regression frameworks, incorporate shared variables such as the yield curve slope (commonly regarded as a credit-risk proxy), a risk perception indicator (e.g., VIX), and a measure of the liquidity premium in Treasuries. However, the quality of fit in these models is not always transparent, and concerns about short observation periods, the use of levels instead of first differences, and look-ahead bias are noted.

In addition to the aforementioned factors, the expected signs of the relationships are presented in Table 1.

Relevant Factors	**Coefficient Sign**	**Initial Movement**	**Effect On Swap Spreads**
Liquidity	+	Increase	Widening
Yield Curve Slope	-	Steepening	Tightening
REPO	-	Increase	Tightening
Treasury Bond Supply	-	Increase	Tightening
Implied Volatility	+	Increase	Widening
Banking Risk	-	Increase	Widening
AA Credit Spreads	+	Increase	Widening

Table 1 - Relevant factors and their expected effect on swap spreads

3. Data

To conduct the research, data was initially retrieved from Refinitiv Eikon Datastream, the Federal Reserve Bank of St. Louis (FRED), and Google Trends (refer to Table 2). The goal was to gather data covering the widest range possible. However, data for all variables was only accessible for the period from 2004 to 2023, spanning 20 years.

To predict swap spread directions behind the yield curve, additional short-term variables were introduced to address volatility. These variables include market indexes (S&P 500), news data (aggregate Google search), sentiment features (market bearish sentiment from the American Association of Individual Investors (AAII) survey and VIX), and an uncertainty feature (Public Uncertainty Index).

Variable	**Description**	**Source**	**Frequency**
swapsp10y	USD 3M Middle Rate LIBOR IRS 10Y – 3-month LIBOR	Datastream	Daily
treassup	US Federal Budget Forecasts from Reuters, USD	Datastream	Quarterly
aasp10y	US Corporate AA 10Y Yield – 10Y Treasury Note Yield	Datastream	Daily
REPO	US 3-month Middle Rate Repo	Datastream	Daily
VIX	Implied 30-day volatility of the S&P 500 index options	CBOE	Daily
slope	10Y Treasury Note Yield – 3-month Treasury Bill Rate	Datastream	Daily
ted	3-month LIBOR – 3-month Treasury Bill	Datastream	Daily
sp500	S&P500 returns	Datastream	Daily
bearish	US Sentiment Survey by the American Association of Individual Investors (% Bearish)[5]	Datastream	Daily
PU	Public Uncertainty[6]	Datastream	Daily
trend_1	Search term count[7] in the US: “Swap spreads”	Google Trends	Monthly
trend_2	Search term count in the US: “Swap spread”	Google Trends	Monthly
trend_3	Search term count in the US: “Interest rate swap”	Google Trends	Monthly
trend_4	Search term count in the US: “Interest Rate swaps”	Google Trends	Monthly

---

[5] The AAII Sentiment Survey offers insight into the opinions of individual investors by weekly asking them their thoughts on where the market is heading in the next six months.

[6] The public policy-related economic uncertainty index is calculated using three components: counting how often 10 large US newspapers talk about economic policy uncertainty; tracking the annual dollar-weighted numbers of temporary federal tax code provisions that are set to expire over the next 10 years; and measuring the dispersion between experts' predictions on future levels of the Consumer Price Index, Federal Expenditures, and State and Local Expenditures.

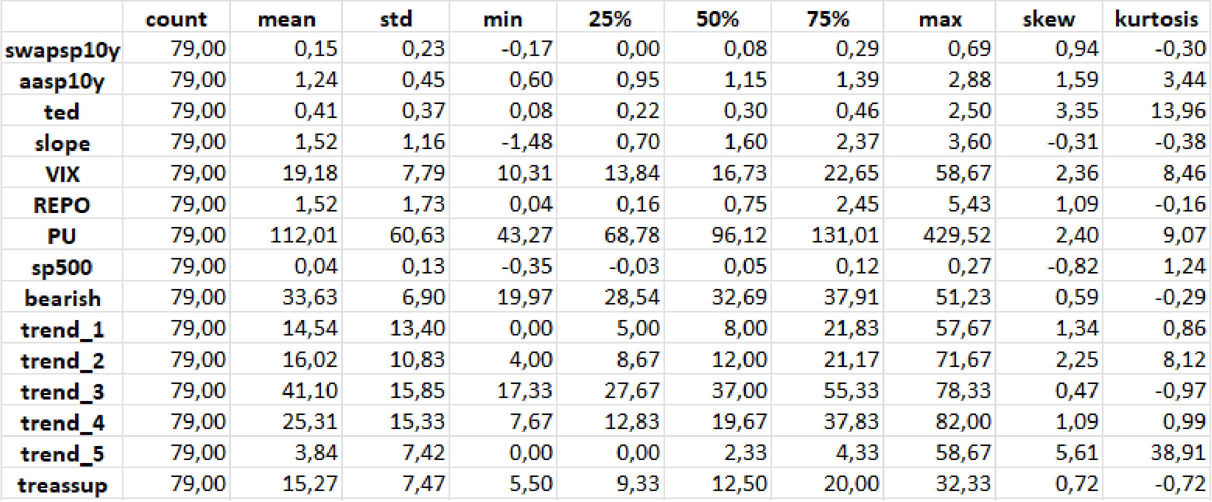
[7] Google Trends data is pulled from a random, unbiased sample of Google searches, meaning that Google doesn’t have exact numbers for any terms or topics. To give a value to terms, they index data from 1-100, where 100 is the maximum search interest for the time and location selected.

trend_5	Search term count in the US: “Interest rate swap spread”	Google Trends	Monthly

Table 2 - Variables description, source and frequency

Data frequencies were derived from daily data, with weekly data determined based on the last day of the working week (Friday). Monthly and quarterly data were then extrapolated from the weekly data. Quarterly Treasury supply data was interpolated to a monthly frequency. Following these transformations, the resulting dataset comprised a total of 1,108 observations at a weekly frequency, 235 observations on a monthly scale, and 79 observations at a quarterly frequency.

Descriptive statistics for all variables at a quarterly frequency from 2004 to 2023 are summarized in Table 3. Notably, the data displays a skewed distribution and is characterized by excess kurtosis, indicating a departure from a normal distribution. Visualizations of variable distributions (see Figure 5) underscore the 2008 regime switch, particularly evident in swapsp10y, treassup, trend_3, and trend_4, which display a bimodal distribution rather than adhering to a normal distribution.



	count	mean	std	min	25%	50%	75%	max	skew	kurtosis
swapsp10y	79,00	0,15	0,23	-0,17	0,00	0,08	0,29	0,69	0,94	-0,30
aasp10y	79,00	1,24	0,45	0,60	0,95	1,15	1,39	2,88	1,59	3,44
ted	79,00	0,41	0,37	0,08	0,22	0,30	0,46	2,50	3,35	13,96
slope	79,00	1,52	1,16	-1,48	0,70	1,60	2,37	3,60	-0,31	-0,38
VIX	79,00	19,18	7,79	10,31	13,84	16,73	22,65	58,67	2,36	8,46
REPO	79,00	1,52	1,73	0,04	0,16	0,75	2,45	5,43	1,09	-0,16
PU	79,00	112,01	60,63	43,27	68,78	96,12	131,01	429,52	2,40	9,07
sp500	79,00	0,04	0,13	-0,35	-0,03	0,05	0,12	0,27	-0,82	1,24
bearish	79,00	33,63	6,90	19,97	28,54	32,69	37,91	51,23	0,59	-0,29
trend_1	79,00	14,54	13,40	0,00	5,00	8,00	21,83	57,67	1,34	0,86
trend_2	79,00	16,02	10,83	4,00	8,67	12,00	21,17	71,67	2,25	8,12
trend_3	79,00	41,10	15,85	17,33	27,67	37,00	55,33	78,33	0,47	-0,97
trend_4	79,00	25,31	15,33	7,67	12,83	19,67	37,83	82,00	1,09	0,99
trend_5	79,00	3,84	7,42	0,00	0,00	2,33	4,33	58,67	5,61	38,91
treassup	79,00	15,27	7,47	5,50	9,33	12,50	20,00	32,33	0,72	-0,72

Table 3 - Quarterly Descriptive Statistics

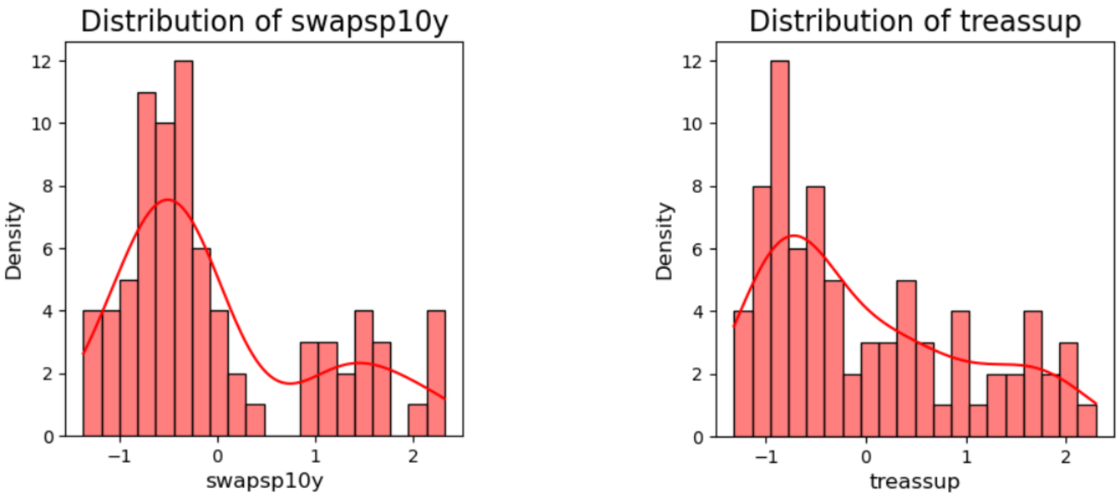


Figure 4 - Distribution of swapsp10y and treassup

Kobor, Shi, and Zelenko (2005) highlight that both the relative and absolute significance of the determinants of swap spreads can vary over time, with at least three distinct regimes identified historically. While regime-switching models can partially address this issue, they have limitations, primarily due to the exogenous nature of regime transitions. These models are effective in improving the in-sample fit of the data, offering a tailored approach to specific historical periods. However, their predictive power may diminish in out-of-sample scenarios. This limitation becomes particularly evident if unforeseen events trigger regime shifts in the future, potentially leading to degraded performance of the model in predicting new data. In this dissertation, the author aims to develop a model that more robustly explains swap spreads across various future cycles, focusing on out-of-sample applicability and prediction.

4. Methodology

The methodological framework overview is defined as follows. Initially, the data was divided into a training set and a test set. The rolling window technique was then applied for model validation within the training set. In each rolling window, the data was standardized before training the model, and the same standardization was used to transform the validation set. Following this, the long-term relationship was estimated using an Ordinary Least Squares (OLS) regression within each window.

The long-term equation is formulated as:

$$swapsp10y_t = \propto + \beta \; treassup_t + \gamma \; trend\_3_t + \varepsilon_t$$

The error-correcting term (ECT) is defined as the estimation error from this OLS equation:

$$ECT_t = swapsp10y_t - (\propto + \beta \; treassup_t + \gamma \; trend\_3_t)$$

The lagged values of the error-correcting term, along with the other lagged predictors, were used to estimate the short-term regression. The estimation for the short-term was achieved through an ensemble of models, including Lasso, Random Forest Regressor, and XGBoost. Subsequently, the fitted values from this ensemble model were used to predict the next quarter's Δ *swapsp10y*.

Once the optimal parameters were identified, an out-of-sample test with continuous update was conducted.

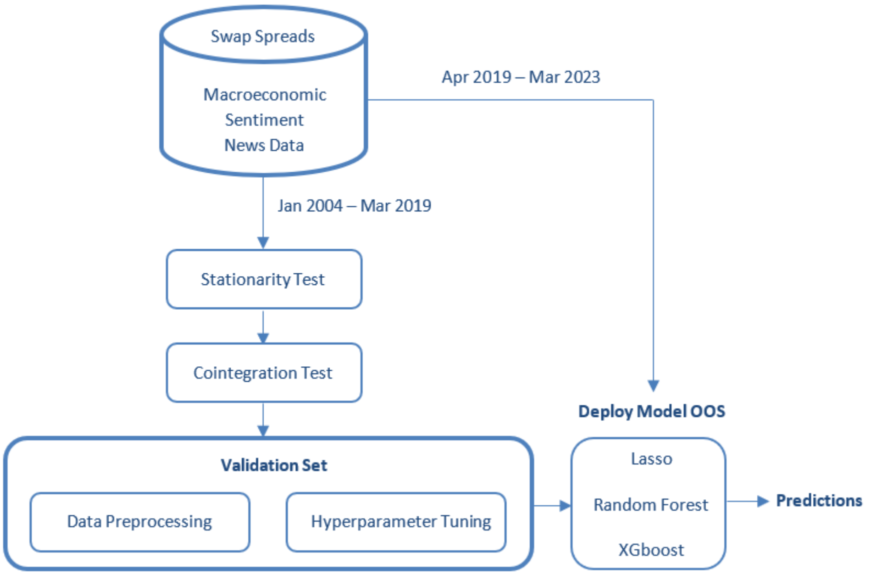


Figure 5 - Methodology Flowchart

4.1. Testing for Stationarity

Stationarity refers to a time series characteristic where statistical properties such as mean, variance, and autocorrelation are constant over time, essentially implying that the series does not exhibit trends or seasonal effects.

The identification of a regime switching effect and the data plotting strongly indicate that the variables are non-stationary, rendering typical linear regressions unsuitable for modelling purposes. Therefore, a rigorous statistical test for stationarity was imperative. To this end, the Augmented Dickey-Fuller test (ADF test) was conducted. The null hypothesis posits that the variable under consideration has a unit root, signifying nonstationarity. The alternative hypothesis, on the other hand, suggests that the series lacks a unit root, indicating stationarity. To assess stationarity at the level, a model incorporating a constant but no time trend was employed, considering the non-zero mean of the variables. For stationarity at first differences, a model without a constant or time trend was applied, given that the first differences of the variables exhibited a mean not significantly different from zero. The test p-values in Table 4 reveal that, for a 5% significance level, only aasp10y, ted, VIX, REPO, PU, sp500 and bearish can be deemed stationary at the level, while other variables exhibit stationarity at the first difference.

Variable	**P-Value for Level**	**P-Value for First Differences**
swapsp10y	0,456	0,003
aasp10y	0,005	0,000
ted	0,018	0,000
slope	0,213	0,000
VIX	0,005	0,000
REPO	0,031	0,004
PU	0,003	0,000
sp500	0,000	0,000
bearish	0,007	0,000
trend_1	0,696	0,030
trend_2	0,724	0,000
trend_3	0,818	0,000
trend_4	0,256	0,025
trend_5	0,071	0,000
treassup	0,488	0,000

Table 4 - ADF test p-values for level and first diferences

A nonstationary series is integrated of order 1, I(1), if the first difference of the series is stationary. This precondition is crucial for exploring a long-run relationship through cointegration methodology.

4.2. Testing for Cointegration

One approach to address the nonstationarity issue is to apply Ordinary Least Squares (OLS) on first differences, a method widely endorsed by the literature. The presence of a cointegration or long-term relationship would offer a more robust analytical method, therefore it is worth testing.

Johansen’s cointegration test was conducted, revealing that 10Y Swap Spreads are cointegrated with Treasury Supply and trend_3 at a quarterly level. While the long-term relationship with Treasury Supply aligns with expectations, the unexpected long-term relationship of swapsp10y with Google Trends search emerged. Subsequent tests for cointegration between the 2 predictors indicated that Treasury Supply and trend_3 are also cointegrated. Consequently, the long-run relationship among 10Y IRS Spreads, Treasury Supply, and trend_3 - representing the Google count of the search term 'Interest rate swap' - were leveraged to construct the Error Correction Model (ECM).

The ECM is defined as follows: first, the long-term relation of 10Y IRS Spreads with Treasury Supply and trend_3 is estimated using quarterly data. Subsequently, the lagged residuals of the long-term regression are incorporated into a second short-term linear regression. This second regression includes stationary variables at the level, and the first difference of non-stationary variables, capturing short-term fluctuations around the long-term trend.

4.3. The Selection Process for Variables and Lags

Certain variables exhibit some degree of autocorrelation in their first differences, a phenomenon well-documented in the literature, suggesting that swap spreads deviate from a random walk process. Consequently, it becomes imperative to scrutinize the optimal number of lags to retain in the short-term regression.

Given the absence of theoretical recommendations on the number of lags in autoregressive dependence (Kobor, 2005), researchers typically navigate a method of trial and error to identify the most suitable set of features and lags. Unlike the conventional approach of sequentially introducing variables and assessing their impact on model fit using the Akaike

Information Criterion (a metric for model's fit), the author has adopted Lasso Regression for feature selection and to address multicollinearity. For the initial analysis, variables were lagged up to a maximum of 4 periods.

4.4. Splitting The Data Before Further Steps

When handling time series data, it is imperative to prevent data leakage from future observations into the model training phase. To this end, it is crucial to set aside a portion of the dataset for final evaluation. In this case, the last sixteen quarterly observations constitute the test set, which remains untouched and isolated from any preprocessing or training activities.

4.5. Lasso Regression Model

Lasso Regression, short for Least Absolute Shrinkage and Selection Operator, is a refined version of linear regression that introduces a regularization technique known as L1 regularization, making it a powerful tool in statistical modelling, particularly when dealing with datasets characterized by multicollinearity. In conventional linear regression models, multicollinearity -where predictor variables are highly correlated with each other-, can lead to unstable and unreliable coefficient estimates. Lasso tackles this challenge by penalizing the absolute size of the regression coefficients. In Lasso regression, the sum of the absolute values of the model parameters is constrained, an effect controlled by a tuning parameter λ (lambda). At a lambda value of zero, Lasso regression is equivalent to an ordinary least squares regression model. However, as the value of lambda increases, so does the regularization effect, leading to the shrinking of the coefficients towards zero. This feature of Lasso regression is particularly beneficial when dealing with datasets with a large number of features, as it effectively zeroes out the coefficients of less important variables. This not only simplifies the model but also helps in pinpointing the most relevant features.

However, Lasso regression comes with its limitations. When predictors are highly correlated, Lasso often arbitrarily selects one variable from a group of correlated variables, ignoring others, which can result in the exclusion of potentially useful information.

$$\text{L1 Regularization: } \boldsymbol{Cost} = \boldsymbol{\lambda} \sum |\boldsymbol{w_j}|$$

4.6. Hyperparameter Tunning

The choice of the appropriate $\boldsymbol{\lambda}$ significantly influences the model's ability to balance between bias and variance. A too-small λ might lead to overfitting, where the model learns the training data too well, including its noise, but performs poorly on unseen data. Conversely, a too-large λ may oversimplify the model, resulting in underfitting, where the model fails to capture the underlying pattern of the data. Finding the right λ value is crucial for the model to generalize well on unseen data, and Optuna library facilitates this by efficiently searching through the hyperparameter space.

For hyperparameter tuning, the training data was split into several subsets, creating multiple training and test sets. By tuning the model across these different sets, we can more reliably determine the optimal λ value.

The evaluation metric used was the Mean Squared Error.

4.7. Cross Validation and TimeSeries

Cross-validation is an essential technique in model evaluation, used to assess the predictive performance of statistical models. It is particularly crucial in scenarios where we aim to understand how well a model will generalize to an independent dataset. One of the primary reasons for using cross-validation is to prevent overfitting, ensuring that the model doesn’t just echo the patterns in the training data, but also captures the underlying relationships effectively enough to perform well on unseen data. This is achieved by dividing the dataset into multiple smaller subsets or 'folds'. The model is trained on a combination of these folds and validated on the remaining fold, a process that is repeated several times. This methodology provides a more comprehensive evaluation, as it leverages multiple train-test splits, offering a broader insight into the model's performance across different subsets of data.

However, time series data present unique challenges that make standard cross-validation techniques unsuitable. Time series data are inherently sequential - the value at a given time point is often highly dependent on previous time points. This temporal structure means that random splitting of data, as done in traditional cross-validation, can lead to significant leakage of information from the future into the past, an unrealistic scenario in real-world applications. To address this, two variations of cross validation are typically used: Rolling Windows and Expanding Windows.

Both Rolling Windows and Expanding Windows maintain the temporal order of observations, dividing the dataset into training and testing sets in a way that respects the time series nature of the data. In each split, the training set consists of the initial segment of the sequence, while the test set comprises the subsequent segment, ensuring that the training set always precedes the test set temporally. These methods are particularly important in time series forecasting where the goal is to predict future values based on past observations.

In Rolling Windows, a specific subset of data, defined by a fixed window size, is used to train the model, and this window 'rolls' forward through the data for each iteration. Expanding windows, on the other hand, begins similarly to the rolling window approach but differs in its treatment of the window's size. In an expanding window, the initial window size starts small, at a point that is large enough to be representative of the dataset. However, unlike the rolling window, the size of the window increases with each iteration, continuously incorporating more data into the analysis or model training.

In this dissertation, the Rolling Window technique is used, setting each window to include 30 observations and testing on the next quarter.

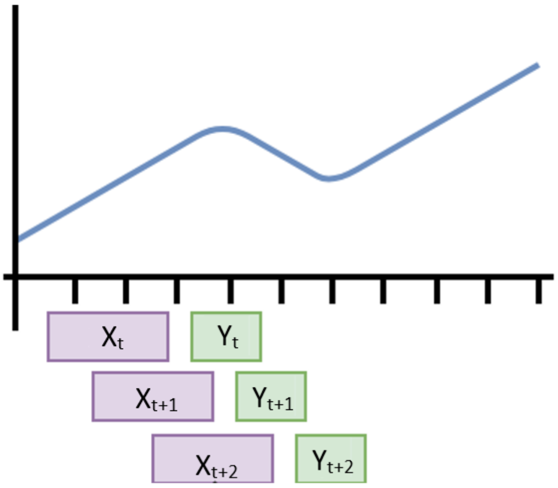


Figure 6 - Rolling Window Cross-Validation

4.8. Preprocessing

In Chapter 3, it was observed that features are on different scales. This can be problematic, as variables with larger scales may disproportionately influence the model's results. To address this, common techniques such as Min-Max scaling and standardization are employed to bring all data to a comparable scale.

Min-Max scaling adjusts the range of continuous data to fit between 0 and 1. While this alters the range of the feature, it preserves the original distribution's shape within the new scale.

In contrast, standardization, also known as Z-score normalization, transforms the data to have a mean of zero and a standard deviation of one. This process does not limit the range of the transformed data. Standardization can be particularly beneficial for models that rely on distance calculations, such as Lasso regression, as it ensures that each feature contributes equally to these calculations.

To empirically test the best method for the data, both Min-Max scaling and standardization methods were tested. Standardization proved to be more effective. This finding is consistent with the observation in Chapter 3 that the timeseries data were not normally distributed, making standardization logically more suitable in this context. For the standardization process, the StandardScaler from the sklearn library was utilized.

$$\textbf{Standardization}: x_{scaled} = \frac{x_i - \mu}{\sigma} \qquad \textbf{Min} - \textbf{Max Scaling}: x_{scaled} = \frac{x_i - x_{min}}{x_{max} - x_{min}}$$

To ensure the integrity of the model and avoid look-ahead bias, the preprocessing steps were fitted exclusively within the training set. Subsequently, these same transformations were used to transform the test data, thereby preventing any potential data leakage, such as using the overall mean from all observations to standardize the data, which could unfairly influence the training set.

4.9. Ensembling

Ensembling is a fundamental technique in machine learning that involves combining multiple models to improve predictive performance, often surpassing what any single model could achieve on its own. The primary strength of ensembling lies in its ability to leverage the diverse perspectives of various models, thereby reducing the risk of overfitting and enhancing the robustness of predictions. Each model in an ensemble may have its own strengths and weaknesses, and by combining them, ensembling effectively balances these out. This leads to more stable and accurate predictions, especially in complex scenarios where no single model can capture all the nuances of the data.

Moreover, ensembling can handle the variability present in different parts of the data by allowing individual models to specialize in certain areas. For instance, some models might perform better on specific segments of the data, while others might excel in different areas. Ensembling capitalizes on this by blending their predictions, ensuring a comprehensive coverage across the entire dataset. This is particularly beneficial in cases where the data is heterogeneous or contains intricate interactions that are challenging for a single model to capture.

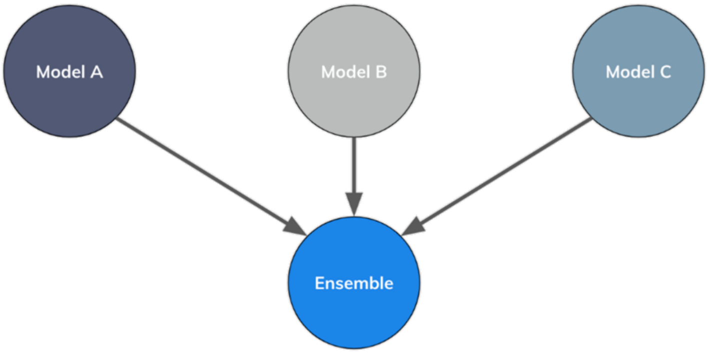


Figure 7 - Ensembling visual example

In this dissertation, the author ensembled three different models – Lasso, Random Forest, and XGBoost – to create a stronger predictive tool. This approach takes advantage of each model's strengths, aiming to achieve more accurate results than any single model could provide on its own.

Ensembling Method: $\hat{y} = \frac{1}{m}\sum_{i=1}^{m}\widehat{y_i}$, with m = 3

4.9.1. Random Forest

The Random Forest algorithm is a robust model that combines multiple decision trees to enhance prediction accuracy. It works by training these trees on various subsets of data, a process known as bootstrapping and aggregation. This method ensures each tree is unique, reducing the overall error of the algorithm. One of the key strengths of Random Forest is its ability to handle multicollinearity, where multiple features are correlated. Since it trains each tree on different feature subsets, it is not as affected by multicollinearity as some other models. This classifier generally outperforms many others in terms of accuracy and doesn't easily overfit, meaning it works well with new data. Additionally, it does not require feature scaling and is more robust against noisy data and the choice of training samples compared to a single decision tree. While it's harder to interpret, it's easier to fine-tune its settings for optimal performance.

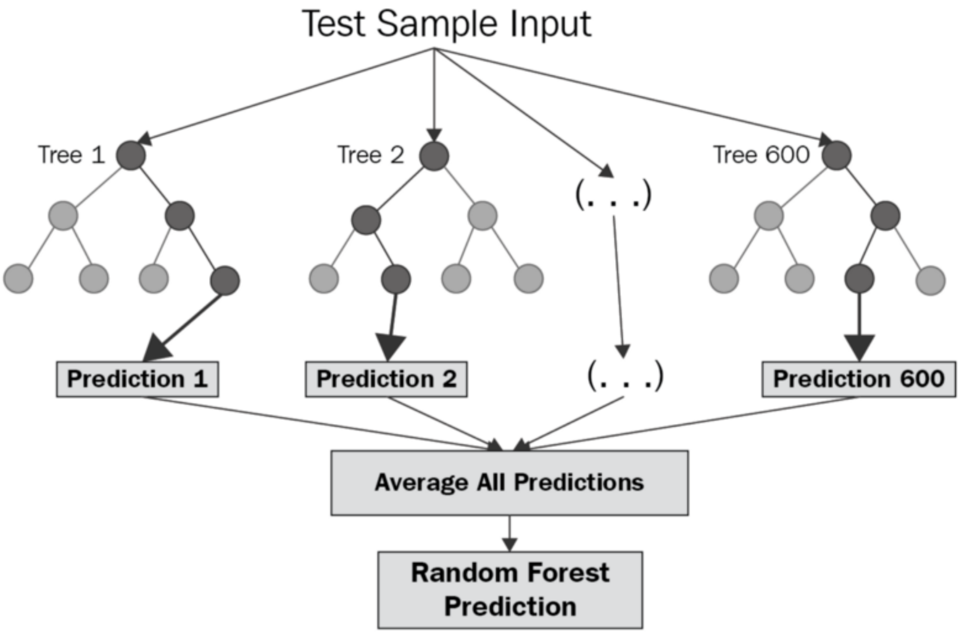


Figure 8 - Random Forest Regressor example, retrieved from Towards Data Science

4.9.2. XGBoost

XGBoost stands for eXtreme Gradient Boosting. It's a supervised algorithm designed to build decision trees sequentially, where each new tree aims to correct the errors of the previous ones. This approach, known as gradient boosting, contributes to the algorithm's high accuracy and efficiency in making predictions.

XGBoost effectively manages overfitting through regularization. It can work with various types of data and does not necessarily require feature scaling. The algorithm also handles multicollinearity well. Its learning method, which focuses on correcting previous errors, helps reduce the negative impact that multicollinearity might have on prediction accuracy.

However, despite its high accuracy, the complex ensemble structure of XGBoost may be more challenging to interpret compared to simpler models.

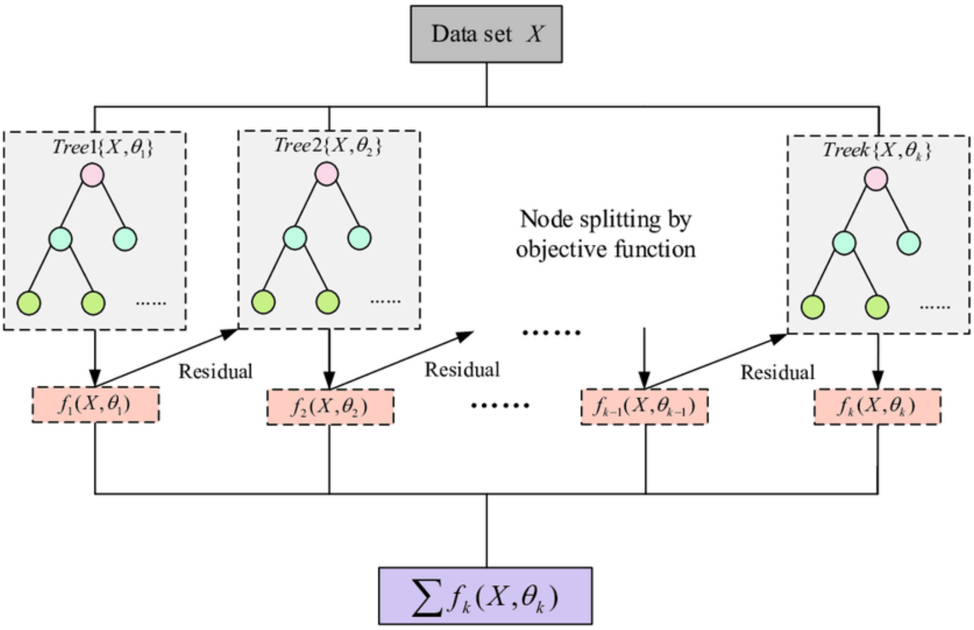


Figure 9 - XGBoost, retrieved from ResearchGate

5. Results and Discussion

The results section is structured as follows. Initially, the performance of the ensemble model is evaluated, highlighting the enhancements achieved as new elements were integrated into the process. Then, feature importance within each model comprising the ensemble will be discussed.

5.1. Model Performance

The initial model testing used Lasso regression without hyperparameter tuning and applied Min-Max scaling, yielding an out-of-sample R2 of 58%. Switching to Standard Scaler improved this performance to 66%. An ensemble of Lasso and Random Forest, however, resulted in a slight decrease in R2 to 63%. Combining Lasso with XGBoost maintained the performance at around 66%. However, when Lasso, Random Forest, and XGBoost were used together, the performance increased significantly to 72%. Additionally, experiments with LightGBM were conducted, but these did not result in performance enhancements.

Subsequently, the author conducted hyperparameter tuning using Optuna, leading to an increase in performance to 76%. This approaches the 80% R2 achieved by Kobor, A., Shi, L., & Zelenko, I. (2005) using a linear regression ECM, though their study showed strong evidence of lookahead bias, an essential factor to control for in such analyses. Additionally, the Mean Squared Error (MSE) of the model is 0.08.

The plot of the estimated values against the actual values show that the two lines move fairly together (Figure 8).

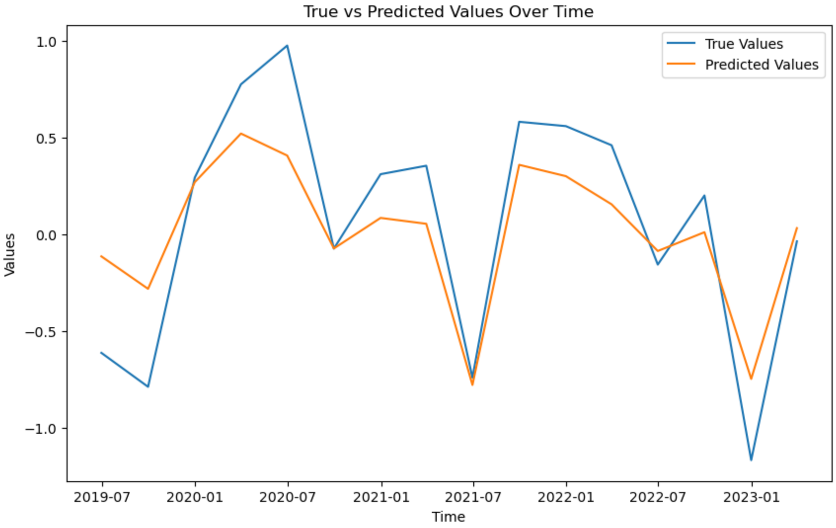


Figure 10 - 10-year Swap Spreads: Actual versus Predicted

It is worth noting that while the model may underestimate the magnitude of a positive or negative shock, it consistently predicts the direction of the shock accurately. One could argue that the direction of the shock is more important than its size, as it enables market participants, such as convergence traders, to take positions based on these directional cues.

The success of the model ensemble indicates its superiority compared to simple linear regression models, such as standalone LASSO.

5.2. Feature Importance

To gauge the significance of the selected features in forecasting swap spreads, it is essential to rank their relative importance. Clarifying the contribution of each variable to the results is crucial, especially since machine learning models are frequently labeled as black boxes. Fortunately, all three models employed facilitate feature ranking.

Top 10 selected features by each model are presented in Figures 9 to 11.

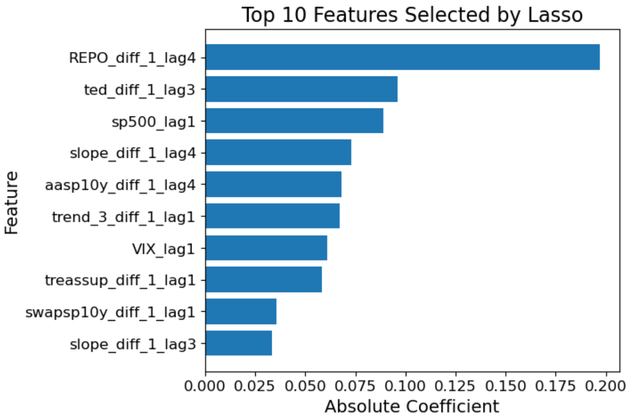


Figure 11 - Top 10 selected features by LASSO

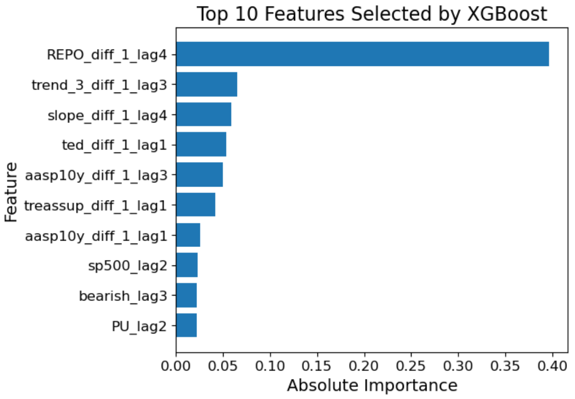


Figure 13 - Top 10 selected features by XGBoost

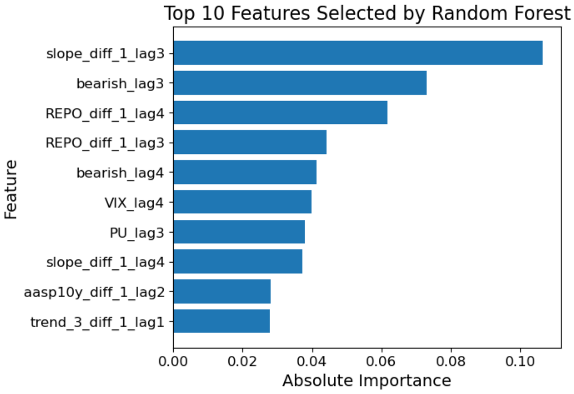


Figure 12 - Top 10 selected features by Random Forest

In Kobor, A., Shi, L., & Zelenko, I. (2005), the repo rate was excluded from the model because it consistently appeared to decrease model performance through the trial-and-error process. However, when feature selection is made through the models, it is discovered that the repo rate is consistently chosen across models. This suggests that the repo rate is indeed useful for prediction.

Lasso retained 17 out of 54 features, setting the coefficients of the remaining features to zero. Lagged residuals were among the variables kept not only in Lasso but also in other models, indicating that the error term significantly contributes to explaining variations in swap spreads. This is because the error-correction term measures the deviation of the current swap spread from its long-term equilibrium, towards which the swap spread is expected to converge.

Public uncertainty, bearish sentiment, VIX, and trend_3 were identified as relevant features across models, suggesting that the added variables for capturing short-term variation enhanced prediction beyond yield curve-related features.

6. Conclusion

This dissertation focused on the out-of-sample prediction of 10-year interest rate swap spreads. It explored three key research questions.

The first examined the feasibility of reliable out-of-sample prediction of swap spreads, refining the findings of Kobor, A., Shi, L., & Zelenko, I. (2005) and confirming that accurate predictions are possible, even when controlling for look-ahead bias.

The second question investigated the utility of market indexes, news data, sentiment, and uncertainty features in swap spread predictions, moving beyond traditional yield curve-focused features. The results highlighted the value of incorporating these additional features to enhance predictive accuracy.

Lastly, the study evaluated the performance of Machine Learning models, including ensemble methods, against a standard Error-Correction Model with linear regression. The results showed that Machine Learning models not only facilitate data-driven feature and lag selection but also significantly enhance prediction accuracy. This improvement is particularly notable when combining models, as opposed to using single models like Lasso, which still relies on linear regression.

This research, however, faced some limitations. Previous research had been conducted with access to private data, obliging the author to search for the best proxies. The data used may further raise concerns as much of it is sourced from Datastream Refinitiv Eikon, which does not always detail the nature of the data. Furthermore, even though this study has used a considerably larger timespan compared to Kobor, A., Shi, L., & Zelenko, I. (2005), nearly double the size, the limited number of observations at a quarterly level still prevent us from obtaining more robust results.

Further research could explore additional machine learning techniques like feature engineering to enhance model accuracy. Regime-switching models also present a promising area for investigation. A potential direction for future studies could involve implementing a Markov-Chain Regression model following feature selection with Lasso. Subsequently, other Machine Learning models, such as the Random Forest Classifier, could be employed to predict the state of the next regime.

References

Adrian, T., Crump, R. K., & Moench, E. (2013). Pricing the term structure with linear regressions. *Journal of Financial Economics*, *110*(1), 110–138. Scopus. https://doi.org/10.1016/j.jfineco.2013.04.009

Augustin, P., Chernov, M., Schmid, L., & Song, D. (2021). Benchmark interest rates when the government is risky. *Journal of Financial Economics*, *140*(1), 74–100. https://doi.org/10.1016/j.jfineco.2020.10.009

Azad, A. S. M. S., Batten, J. A., & Fang, V. (2015). What determines the yen swap spread? *International Review of Financial Analysis*, *40*, 1–13. Scopus. https://doi.org/10.1016/j.irfa.2015.04.001

Blanco, R., Brennan, S., & Marsh, I. W. (2005). An empirical analysis of the dynamic relation between investment-grade bonds and credit default swaps. *Journal of Finance*, *60*(5), 2255–2281. Scopus. https://doi.org/10.1111/j.1540-6261.2005.00798.x

Breach, T., D'Amico, S., & Orphanides, A. (2020). The term structure and inflation uncertainty. *Journal of Financial Economics*, *138*(2), 388–414. Scopus. https://doi.org/10.1016/j.jfineco.2020.04.013

Brown, K., Harlow, W., & Smith, D. (1994). An Empirical Analysis of Interest Rate Swap Spreads. *The Journal of Fixed Income*, *3*, 61–78. https://doi.org/10.3905/jfi.1994.408095

Chung, H.-L., & Chan, W.-S. (2010). Impact of credit spreads, monetary policy and convergence trading on swap spreads. *International Review of Financial Analysis*, *19*(2), 118–126. https://doi.org/10.1016/j.irfa.2010.01.004

Collin-Dufresn, P., Goldstein, R. S., & Martin, J. S. (2001). The Determinants of Credit Spread Changes. *The Journal of Finance*, *56*(6), 2177–2207. https://doi.org/10.1111/0022-1082.00402

Collin-Dufresne, P., & Goldstein, R. S. (2001). Do Credit Spreads Reflect Stationary Leverage Ratios? *The Journal of Finance*, *56*(5), 1929–1957. https://doi.org/10.1111/0022-1082.00395

Collin-Dufresne, P., & Solnik, B. (2001). On the Term Structure of Default Premia in the Swap and LIBOR Markets. *The Journal of Finance*, *56*(3), 1095–1115. https://doi.org/10.1111/0022-1082.00357

Cortes, F. (2003). *Quarterly Bulletin Winter 2003*.

Cortes, F. (2005). *Understanding and Modelling Swap Spreads* (SSRN Scholarly Paper

706830). https://papers.ssrn.com/abstract=706830

Driessen, J. (2005). Is default event risk priced in corporate bonds? *Review of Financial Studies*, *18*(1), 165–195. Scopus. https://doi.org/10.1093/rfs/hhi009

Duffie, D., & Singleton, K. J. (1997). An Econometric Model of the Term Structure of Interest-Rate Swap Yields. *The Journal of Finance*, *52*(4), 1287–1321. https://doi.org/10.1111/j.1540-6261.1997.tb01111.x

Fang, V., & Muljono, R. (2003). An empirical analysis of the Australian dollar swap spreads. *Pacific-Basin Finance Journal*, *11*(2), 153–173. https://doi.org/10.1016/S0927-538X(02)00112-9

Feldhütter, P., & Lando, D. (2008). Decomposing swap spreads. *Journal of Financial Economics*, *88*(2), 375–405. https://doi.org/10.1016/j.jfineco.2007.07.004

Fontaine, J.-S., & Garcia, R. (2012). Bond liquidity premia. *Review of Financial Studies*, *25*(4), 1207–1254. Scopus. https://doi.org/10.1093/rfs/hhr132

Grinblatt, M. (2001). An Analytic Solution for Interest Rate Swap Spreads. *International Review of Finance*, *2*(3), 113–149. https://doi.org/10.1111/1468-2443.00022

He, Z., Nagel, S., & Song, Z. (2022). Treasury inconvenience yields during the COVID-19 crisis. *Journal of Financial Economics*, *143*(1), 57–79. https://doi.org/10.1016/j.jfineco.2021.05.044

Hébert, B., & Li, W. (2023). Intermediary balance sheets and the treasury yield curve. *Journal of Financial Economics*, *150*(3), 103722. https://doi.org/10.1016/j.jfineco.2023.103722

In, F., Brown, R., & Fang, V. (2003). Modeling volatility and changes in the swap spread. *International Review of Financial Analysis*, *12*(5), 545–561. https://doi.org/10.1016/S1057-5219(03)00067-X

Jermann, U. J. (2020). Negative Swap Spreads and Limited Arbitrage. *Review of Financial Studies*, *33*(1), 212–238. Scopus. https://doi.org/10.1093/rfs/hhz030

Johannes, M., & Sundaresan, S. (2007). The impact of collateralization on swap rates. *Journal of Finance*, *62*(1), 383–410. Scopus. https://doi.org/10.1111/j.1540-6261.2007.01210.x

Jun, L., Longstaff, F. A., & Mandell, R. E. (2006). The market price of risk in interest rate swaps: The roles of default and liquidity risks. *Journal of Business*, *79*(5), 2337–2359. Scopus. https://doi.org/10.1086/505237

Kambhu, J. (2006). *Trading Risk, Market Liquidity, and Convergence Trading in the Interest Rate Swap Spread* (SSRN Scholarly Paper 918493). https://papers.ssrn.com/abstract=918493

Kempf, A., Korn, O., & Uhrig-Homburg, M. (2012). The term structure of illiquidity premia.

Journal of Banking and Finance, 36(5), 1381–1391. Scopus. https://doi.org/10.1016/j.jbankfin.2011.12.003

Klein, D., Nikitina, E., & Curtillet, J.-C. (2016). Forecasting Swap Spreads: A Bayesian Approach. *The Journal of Fixed Income*, *26*(2), 40–53. https://doi.org/10.3905/jfi.2016.26.2.040

Klingler, S., & Sundaresan, S. (2019). An Explanation of Negative Swap Spreads: Demand for Duration from Underfunded Pension Plans. *The Journal of Finance*, *74*(2), 675–710. https://doi.org/10.1111/jofi.12750

Kobor, A., Shi, L., & Zelenko, I. (2005). *What Determines U.S. Swap Spreads?* The World Bank. https://doi.org/10.1596/978-0-8213-6338-6

Krishnan, C. N. V., Ritchken, P. H., & Thomson, J. B. (2010). Predicting credit spreads. *Journal of Financial Intermediation*, *19*(4), 529–563. https://doi.org/10.1016/j.jfi.2009.02.004

Lang, L. H. P., Litzenberger, R. H., & Luchuan Liu, A. (1998). Determinants of interest rate swap spreads. *Journal of Banking & Finance*, *22*(12), 1507–1532. https://doi.org/10.1016/S0378-4266(98)00068-5

Lekkos, I., & Milas, C. (2001). Identifying the Factors that Affect Interest-Rate Swap Spreads: Some Evidence from the United States and the United Kingdom. *Journal of Futures Markets*, *21*(8), 737–768. https://doi.org/10.1002/fut.1803

Longstaff, F. A., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *Journal of Finance, 60*(5), 2213–2253. Scopus. https://doi.org/10.1111/j.1540-6261.2005.00797.x

Dr. Sohini Chowdhury, & Dr. Martin A. Wurm. (2017, October). *Modeling and Forecasting Interest Rate Swap Spreads*. https://www.moodysanalytics.com/risk-perspectives-magazine/managing-disruption/principles-and-practices/modeling-and-forecasting-interest-rate-swap-spreads

Smith, C. W., Smithson, C. W., & Wakeman, L. M. (1988). The Market for Interest Rate Swaps. *Financial Management*, *17*(4), 34–44. https://doi.org/10.2307/3665765

Sorensen, E. H., & Bollier, T. F. (1994). Pricing Swap Default Risk. *Financial Analysts Journal*, *50*(3), 23–33. https://doi.org/10.2469/faj.v50.n3.23

Sun, T., Sundaresan, S., & Wang, C. (1993). Interest rate swaps: An empirical investigation. *Journal of Financial Economics*, *34*(1), 77–99. https://doi.org/10.1016/0304-405X(93)90041-9

Sundaresan, S., & Wang, Z. (2009). Y2K Options and the Liquidity Premium in Treasury Markets. *The Review of Financial Studies*, *22*(3), 1021–1056.

https://doi.org/10.1093/rfs/hhn005

Turnbull, S. M. (1987). Swaps: A Zero Sum Game? *Financial Management*, *16*(1), 15–21. https://doi.org/10.2307/3665544