Aalto University
School of Business

YIELD CURVE ARBITRAGE IN THE USD INTEREST RATE SWAP MARKET

Returns and risks in 2012-2021

Master's Thesis
Miro Lammi
Aalto University School of Business
Finance
Spring 2022

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Miro Lammi	
Title of thesis Yield curve arbitrage in the USD interest rate swap market: returns and risks in 2012-2021	
Degree Master of Science	
Degree programme Finance	
Thesis advisor(s) Michael Ungeheuer	
Year of approval 2022	**Number of pages** 119 **Language** English

Abstract

In my master's thesis, I study the performance of yield curve arbitrage in the USD interest rate swap market in 2012-2021. To my knowledge, yield curve arbitrage has been extensively studied only in the article that I replicate in my thesis. The article finds strong performance for yield curve arbitrage in the USD interest rate swap market in 1988-2004. Yield curve arbitrage aims at finding and taking advantage of relative mispricings on a given rate curve.

I apply monthly data and I model the USD swap rate curve with the two-factor Vasicek short rate model and with the two-factor Cox-Ingersoll-Ross short rate model. The models assume that the development of a rate curve is explained by two stochastic factors. Yield curve arbitrage aims at being hedged against changes in the factors. My modelling indicates mispricings on the swap rate curve and I open monthly arbitrage trades for the largest mispricings. I calculate the market values for the open arbitrage trades and I calculate the returns by following the market values. I close the arbitrage trades when the mispricings disappear or when the 12-month time limit is exceeded

I try to study whether yield curve arbitrage has provided better returns in 2012-2021 than in the study period of 1988-2004 of my reference article. I also try to observe whether there are differences in the performance of the short rate models and if my data and methodology affect my results.

The means of returns of yield curve arbitrage are mostly positive but the confidence intervals of the means are wide. The distributional measures of the returns, skewness and kurtosis, indicate good performance in some arbitrage implementations and bad performance in others. I can not conclude that my returns would be different from the returns of my reference article or that one of the short rate models would provide better performance than the other. The data and methodology appear to have effect on the arbitrage performance.

Yield curve arbitrage aims at having limited exposure to the market risks. My multifactor regressions indicate that the arbitrage does not provide market risk corrected positive returns and that the arbitrage is exposed to market risks. On the other hand, the regression results suggest that the arbitrage has exposure to exploiting the mispricings on the swap rate curve. I compare my returns to the returns of a comparable hedge fund index. The performance of the hedge fund index has been strong in 2012-2021 and the correlation between my returns and the returns of the index is low. The hedge fund index is exposed to some of the same market risks as my yield curve arbitrage returns.

My analysis does not indicate strong support for yield curve arbitrage having been profitable in the USD interest rate swap market in 2012-2021. However, my results are inconsistent and inconclusive. My results are likely due to the problems in my methodology and due to my difficult study period in the 2010s when USD rate markets were extraordinary.

Keywords interest rate swap, yield curve, arbitrage, short rate model

Aalto-yliopisto
Kauppakorkeakoulu

Aalto-yliopisto, PL 11000, 00076 AALTO
www.aalto.fi

Maisterintutkinnon tutkielman tiivistelmä

Tekijä Miro Lammi

Työn nimi Korkokäyrän arbitraasi USD-koronvaihtosopimusten markkinalla: tuotot ja riskit vuosina 2012-2021

Tutkinto Kauppatieteiden maisteri

Koulutusohjelma Rahoitus

Työn ohjaaja(t) Michael Ungeheuer

Hyväksymisvuosi 2022 **Sivumäärä** 119 **Kieli** Englanti

Tiivistelmä

Tutkin opinnäytteessäni korkokäyrän arbitraasia USD-koronvaihtosopimusten markkinalla vuosina 2012-2021. Korkokäyrän arbitraasia on tietojeni pohjalta aiemmin tutkittu kattavasti vain artikkelissa, jota opinnäytteessäni replikoin. Kyseisessä artikkelissa korkokäyrän arbitraasin todettiin olleen tuottoisaa USD-koronvaihtosopimusten markkinalla vuosina 1988-2004. Korkokäyrän arbitraasissa pyrkimyksenä on löytää ja hyödyntää suhteellisia väärinhinnoitteluja korkokäyrällä.

Käytän kuukausittaista dataa ja mallinnan USD-koronvaihtosopimusten kiinteiden korkojen käyrää lyhyen koron malleilla. Hyödyntämissäni kahden faktorin Vasicekin mallissa ja kahden faktorin Cox-Ingersoll-Rossin mallissa korkokäyrän kehitystä selitetään kahdella stokastisella faktorilla, joiden muutoksilta korkokäyrän arbitraasi pyrkii olemaan suojautunut. Mallinnukseni osoittaa korkokäyrällä väärinhinnoitteluja, joista suurimpiin avaan kuukausittain arbitraasipositioita. Lasken itse markkina-arvot avatuille positioille ja lasken tuotot seuraamalla markkina-arvoja. Suljen arbitraasipositiot kun väärinhinnoittelut häviävät tai kun 12 kuukautta position avaamisesta on kulunut.

Pyrin tutkimaan, tarjoaako korkokäyrän arbitraasi vuosina 2012-2021 parempia tuottoja kuin referenssiartikkelini tutkimusperiodilla. Pyrin myös havainnoimaan, onko lyhyen koron mallien suoriutumisessa eroja ja onko mallinnuksessa hyödynnettävällä datalla ja metodologialla merkitystä.

Tuottojeni keskiarvot ovat pääsääntöisesti positiivisia, mutta tuottojen keskiarvoilla on leveät luottamusvälit. Tuottojen jakaumaan liittyvät tunnusluvut, vinous ja huipukkuus, osoittavat joissain arbitraasin toteutuksissa hyvää suoriutumista mutta toisissa taas heikompaa. En voi luotettavasti sanoa tuottojeni poikkeavan referenssiartikkelini tuotoista tai todeta toisen lyhyen koron mallin suoriutuvan paremmin kuin toinen. Datalla ja metodologialla vaikuttaa olevan merkitystä.

Korkokäyrän arbitraasi pyrkii olemaan rajatusti markkinariskeille altistunut. Monimuuttujaregressioni osoittavat, että arbitraasi ei tarjoa markkinariskeillä korjattuja positiivisia tuottoja ja että arbitraasi on markkinariskeille altistunut. Toisaalta regression tuloksista voi päätellä arbitraasin olevan kohdistunut korkokäyrän väärinhinnoittelujen hyödyntämiseen. Vertaan tuottojani myös vertailukelpoisen vipurahastoindeksin tuottoihin. Vipurahastoindeksin suoriutuminen on ollut vahvaa 2010-luvulla ja korrelaationi vipurahastoindeksin tuottoihin on pientä. Toisaalta vipurahastoindeksi on altistunut joillekin samoille markkinariskeille kuin arbitraasin toteutukseni.

Analyysini ei osoita vahvaa tukea sille, että korkokäyrän arbitraasi olisi ollut tuottoisaa USD-koronvaihtosopimusten markkinalla vuosina 2012-2021. Toisaalta tulokseni ovat epäjohdonmukaisia. Tulokseni johtuvat todennäköisesti ongelmista metodologiassani ja hyödyntämästäni tutkimusperiodista 2010-luvulla, jolloin USD-korkojen markkinat ovat olleet poikkeukselliset.

Avainsanat koronvaihtosopimus, korkokäyrä, arbitraasi, lyhyen koron mallit

Yield Curve Arbitrage In The USD Interest Rate Swap Market: Returns And Risks In 2012-2021

Miro Lammi

Master's Thesis

Deparment of Finance

Aalto University School of Business

Thesis advisor: Michael Ungeheuer

July 2022

Contents

1 Introduction **7**

2 Literature review **8**

2.1 The concept of arbitrage . . . 8

2.1.1 Market neutrality . . . 9

2.1.2 Limitations of arbitrage . . . 10

2.2 Yield curve arbitrage . . . 11

2.2.1 Sources of risks and returns in yield curve arbitrage . . . 15

2.2.2 Master's theses on yield curve arbitrage . . . 18

3 Short rate modelling **20**

3.1 Short rate modelling in a nutshell . . . 20

3.2 One-factor and multifactor models . . . 23

3.3 Equilibrium short-rate models . . . 25

3.4 Exogenous short-rate models . . . 26

3.5 The applied models . . . 27

3.5.1 Two-factor Vasicek model . . . 27

3.5.2 Two-factor Cox-Ingersoll-Ross model . . . 30

3.5.3 Calibration . . . 32

4 Contributions, research questions and hypotheses **34**

5 Data **37**

5.1 Data analysis . . . 39

5.2 Possible issues in data . . . 44

6 Methodology **45**

6.1 Background on swap mechanics . . . 45

6.2 Modelling . . . 49

6.2.1 Optimization . . . 53

6.3 Trading . . . 54

6.4 Valuation . . . 58

6.5 Returns . . . 63

7 Trading performance **67**

7.1 Full trading period 01/2012 - 12/2021 67

7.2 Subperiod 01/2015 - 12/2021 79

7.2.1 Subperiod performance with methodology modifications 83

8 Connection to hedge fund returns and risk factor analysis **89**

9 Conclusions **95**

10 References **99**

11 Appendices **102**

11.1 Python code implementation 102

11.2 Additional master's theses on yield curve arbitrage 102

11.3 Other trading on term structure 103

11.4 Interest rate swap counterparties and arbitrageurs 104

11.5 The concept of term structure of interest rates 106

11.6 Other possible issues in methodology 108

11.7 Additional trading performance analysis 112

11.8 Examples 116

List of Tables

1 Descriptive statistics for USD swap rates in 01/2007-12/2021 (%) 40

2 Summary statistics for the monthly total ex-closing cost excess returns in 01/2012 - 12/2021. 72

3 Additional summary statistics for the monthly total ex-closing cost excess returns in 01/2012 - 12/2021. 72

4 Summary statistics for the monthly total ex-closing cost excess returns in 01/2015 - 12/2021. 79

5 Additional summary statistics for the monthly total ex-closing cost excess returns in 01/2015 - 12/2021. 79

6 Summary statistics for the monthly total ex-closing cost excess returns with methodology modifications in 01/2015 - 12/2021. 84

7 Additional summary statistics for the monthly total ex-closing cost excess returns with methodology modifications in 01/2015 - 12/2021. 85
8 Summary statistics of the returns of Credit Suisse (CS) Fixed Income Arbitrage Index in 01/2012 - 12/2021. 90
9 Multifactor regression of full trading period returns. 92

List of Figures

1 USD swap rates in 01/2007-12/2021 39
2 Chosen USD OIS rates and all USD LIBOR rates in 09/2011 - 12/2021 . . 41
3 USD swap rate curves with cubic interpolation in 01/2007 - 12/2021 42
4 USD LIBOR rate curves with cubic interpolation in 09/2011 - 12/2021 . . 43
5 USD OIS rate curves with cubic interpolation in 09/2011 - 12/2021 43
6 Mispricings with two-factor Vasicek model in 01/2012 - 12/2021 67
7 Mispricings with two-factor CIR model in 01/2012 - 12/2021 68
8 Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in 01/2012 - 12/2021, 5bps mispricing limit . . . 74
9 Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in 01/2012 - 12/2021, 20bps mispricing limit . . . 75
10 Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in 01/2012 - 12/2021, 5bps mispricing limit 75
11 Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in 01/2012 - 12/2021, 20bps mispricing limit 76
12 Cumulative returns with two-factor Vasicek model in 01/2012 - 12/2021, 5bps mispricing limit 77
13 Cumulative returns with two-factor CIR model in 01/2012 - 12/2021, 5bps mispricing limit 78
14 Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit 80
15 Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit 81

16 Cumulative returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit 82
17 Cumulative returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit 82
18 Mispricings with two-factor Vasicek model in subperiod 01/2015 - 12/2021 with modified methodology 83
19 Mispricings with two-factor CIR model in subperiod 01/2015 - 12/2021 with modified methodology 84
20 Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology 86
21 Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology 86
22 Cumulative returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology 87
23 Cumulative returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology 87
24 Cumulative returns for Credit Suisse Fixed Income Arbitrage Index in 01/2012 - 12/2021. 89
25 Probability distribution of monthly excess returns of Credit Suisse Fixed Income Arbitrage Index in 01/2012 - 12/2021. 91
26 Portfolio $-value in full trading period 01/2012 - 12/2021 with two-factor Vasicek model, 5 bps mispricing limit 112
27 Portfolio $-value in full trading period 01/2012 - 12/2021 with two-factor CIR model, 5 bps mispricing limit 113
28 Monthly total ex-closing cost excess returns with two-factor Vasicek model in the full trading period 01/2012 - 12/2021, 5 bps mispricing limit 114
29 Monthly total ex-closing cost excess returns with two-factor CIR model in the full trading period 01/2012 - 12/2021, 5 bps mispricing limit 115
30 Cumulative returns with two-factor Vasicek model in 01/2012 - 12/2021, 20bps mispricing limit 115

31 Cumulative returns with two-factor CIR model in 01/2012 - 12/2021, 20bps mispricing limit . . . 116

1 Introduction

Fixed income arbitrage has historically been popular in hedge funds and proprietary trading desks[1]. This category of arbitrage includes various strategies that have done well for long periods of time but then have faced occasional heavy losses and are therefore described as "picking nickels in front of a steamroller". Many fixed income arbitrage strategies are not well covered in academia. An example is the yield curve arbitrage.

Duarte et al. (2007) presents yield curve arbitrage as a fixed income arbitrage strategy aiming at trading the curvature of the term structure of interest rates and taking advantage of relative mispricings. By applying a short rate model for modelling the USD swap rate curve in 1988-2004, Duarte et al. (2007) open arbitrage trades in mispriced swaps and hedge the trades based on the short rate model factor sensitivities. Duarte et al. (2007) presents strong performance for yield curve arbitrage, as the strategy provides positive means of excess returns and positively skewed return distributions. Duarte et al. (2007) do not find support for yield curve arbitrage being heavily exposed to market risk or tail risk, thus discarding previous findings on fixed income arbitrage.

By combining the methodologies of Duarte et al. (2007) and Smith (2013) and by making my own methodological modifications, I study yield curve arbitrage in the USD interest rate swap market in 2012-2021. I find contradictory results compared to Duarte et al. (2007) and I do not get conclusive results suggesting that yield curve arbitrage would have been lucrative strategy in the USD interest rate swap market in the 2012-2021. My confidence intervals for the means of excess returns are wide and the positive skewness in return distributions varies depending on the arbitrage implementation. My multifactor regression indicates that yield curve arbitrage would have been exposed to market risk factors in the USD swap market in the 2012-2021. My performance is likely affected most by my methodology and by the difficult USD rate environment in the 2010s.

The rest of the thesis is structured as follows: chapter 2 presents my literature review and chapter 3 presents short rate modelling. In chapter 4, I present my contributions, research questions and hypotheses. Chapters 5 and 6 present extensively my data and methdology. In chapters 7 and 8, I present yield curve arbitrage performance and the results of a multifactor regression. Finally, in chapter 9, I conclude my thesis by discussing my performance and by connecting my results to my hypotheses.

[1] Trading on a financial institution's own account.

2 Literature review

In this chapter, I present arbitrage in various forms. I begin by defining the concept of arbitrage and then move on to presenting market neutrality and yield curve arbitrage in detail.

2.1 The concept of arbitrage

The concept of arbitrage is widely used in the financial markets. Market participants often use the concept incorrectly, as according to Tuckman and Serrat (2012) the concept is broadly used for trades, that promise large profits relative to the risks involved, but could still lose significantly money. As roughly stated in Brigo and Mercurio (2007), "the absence of arbitrage is equivalent to the impossibility to invest zero today and receive tomorrow a nonnegative amount that is positive with positive probability." Shreve (2004) provides a formal definition:

Definition 2.1 (Arbitrage)**.** An arbitrage is a portfolio value process X(t) satisfying X(0) = 0 and also satisfying for some time T > 0:

$$P\{X(T) \geq 0\} = 1$$

$$P\{X(T) > 0\} > 0$$

Hull (2017) defines arbitrage in the context of trading by stating, that "arbitrage is a trading strategy, that takes advantage of two or more securities being mispriced relative to each other". Arbitrage could e.g. involve simultaneously entering into transactions in two or more markets and locking in a guaranteed profit[1]. These kind of pure arbitrage opportunities are rare in practice, though. Hull (2017) states that because of arbitrageurs, there are only very small arbitrage opportunities observed in the prices in most financial markets.

According to Tuckman and Serrat (2012), "absent confounding factors (e.g. liquidity, financing, taxes, credit risk), identical sets of cash flows should sell for the same price", which is a rough definition for the so called law of one price. Duarte et al. (2007) states the same concept as "two portfolios having the same payoff at a given future date must

[1] See Appendices section 11.8 for an example.

have the same price today". A deviation from the law of one price implies the existence of an arbitrage opportunity. Because of arbitrage activity, the law of one price usually describes security prices quite well.

The most important takeaway from this first section is the formal definition of arbitrage by Shreve (2004) that should be applied, when talking about arbitrage strategies. As I will clarify in the next sections and throughout my thesis, yield curve arbitrage does not match the formal definition of arbitrage.

2.1.1 Market neutrality

Arbitrage is not to be confused with market neutrality. Following Duarte et al. (2007), market neutral strategies could be defined theoretically as:

Definition 2.2 (Market neutral strategy)**.** Purely market neutral strategies do not have directional market risk and the returns of market neutral trading strategies cannot be explained by market risk factors. In other words, purely market neutral trading strategies should be hedged against market risk factors and such strategies should provide positive returns in all market environments.

In reality, purely market neutral strategies seldom exist and market neutral strategies are seldom neutral to all market risk factors. The idea in practical market neutral strategies is to be exposed to certain risk or risks from which risk premia is to be obtained, while being hedged againts key market risk factors.

The line between pure market neutrality and partial market neutrality is often vague and many trading strategies claiming to be market neutral are often in reality exposed to market risk factors. Patton (2009) applies data on monthly returns of different hedge fund categories and studies the dependence of the returns of hedge funds claiming to be market neutral and the returns of S&P500 index. He finds that around one quarter of the hedge funds claiming to be market neutral are not purely market neutral.

Another major risk component related to systematic market risk factors is the so called tail risk, which refers to having exposure to the negative tail of a return distribution. The tail risk can be considered as the risk of some rare and negative event. Kelly and Jiang (2012) find that hedge funds are often exposed to the tail risk. They find that tail risk explains a large part of both cross-sectional and time series variation in hedge fund returns and hedge funds that experience heavy losses during tail event periods earn on average higher annual returns than funds that are tail risk hedged. Purely market neutral strategies should not be exposed to tail risk and it is often not desired for partially market neutral strategies either.

When talking about market neutral strategies, it is important to know whether pure market neutrality or partial market neutrality is referred to. It is also important to know the risk factors to which a partially market neutral strategy is exposed to.

2.1.2 Limitations of arbitrage

As already noted, pure arbitrage opportunities rarely occur. When such opportunities occur, they are usually small and instantly arbitraged away. On the other hand, arbitrage and market neutral trading strategies are in reality limited by various reasons that help to explain why sometimes arbitrage opportunities are not instantly traded away.

Duarte et al. (2007) emphasizes the importance of intellectual capital in implementing demanding arbitrage strategies. The required intellectual capital naturally limits the number of market participants in demanding arbitrage strategies. Arbitrage opportunities may persist, if the limited number of market participants are unable or unwilling to arbitrage away all the existing arbitrage opportunities.

Limited market liquidity and limited leverage funding are other limits to arbitrage and market neutral trading strategies. Limited market liquidity and limited leverage funding may e.g. restrict the trading strategies aiming at trading convergences and divergences in relative values. For example, Fontaine and Nolin (2019) provide a relative value measure that measures a bond's value relative to other similar bonds. The measure indicates persistent pricing deviations which implies that there are limits to arbitrage in the fixed income markets. They find that the limits to arbitrage are mostly related to funding limits.

Gromb and Vayanos (2010) study market anomalies and mispricings caused by demand shocks. The shocks may stem from either institutional or behavioral considerations and institutions that are commonly seen as arbitragerus, such as hedge funds, may be causing the shocks. Gromb and Vayanos (2010) note that the costs faced by arbitrageurs can prevent them from providing liquidity and correcting the mispricings. They find that the costs faced by arbitrageurs are related to e.g. short selling costs and constraints on leverage.

The nature of demand shocks is controversial, as on one hand large, unexpected demand shocks may force arbitrageurs and market neutral traders to close their positions or react in some other unbeneficial way. On the other hand, demand shocks provide possibilities for trading divergence and convergence and such shocks are key in explaining arbitrage activity e.g. in Vayanos and Vila (2009).

2.2 Yield curve arbitrage

Yield curve arbitrage has not been widely covered in academia or in general. This is partly because there is no clear definition for yield curve arbitrage. There are rather many different trading strategies that share the same basic idea of trading mispricings in rate curves in a market neutral way but the strategies have different focus areas and names. I will later in this subsection explicitly define yield curve arbitrage as it is implemented in my main reference paper Duarte et al. (2007).

Yield curve arbitrage is fundamentally nothing more than a butterfly trade on a chosen rate curve. The details for implementing butterfly trades vary and this is where the conceptual difference between yield curve arbitrage and more general butterfly trades lies. In a conventional butterfly trade on a chosen rate curve, a trader has a view on how the curvature of the chosen rate curve will develop in the future and the trader takes long and short positions along the rate curve according to the view.

Example 2.1. A trader may think, that the 5-year swap rate will rise in the future, whereas the 1- and 10-year swap rates will fall. In this case, the curvature of the swap rate curve would change, so the trader might take benefit of it by "going short" (opening a payer swap) in the 5-year swap rate and "going long" (opening receiver swaps) in the 1- and 10-year swap rates[1]. The directions of the trades are somewhat counter-intuititive, i.e. "going short" in the rising swap rate and the opposite for the falling swap rates. This is because a fixed-for-floating interest rate swap includes both a floating leg and a fixed leg.

As the example suggests, a butterfly trade on a chosen rate curve includes opening a position in a given maturity X in one direction (long or short) and positions in maturities

[1]I will present the categories of interest rate swaps in section 6.1.

$X - a$ and $X + b$ in the opposite direction. This total portfolio can be DV01[1] hedged by choosing the weights for the swaps in $X - a$ and $X + b$ maturities such that the DV01 of the total portfolio is zero. By implementing the DV01 hedging in the butterfly trade fashion, the DV01 hedge provides automatically hedge against changes in level and slope of the swap curve, as explained in Pedersen (2015). Kakushadze and Serur (2018) and Pedersen (2015) cover these kind of risk metric[2] driven butterfly trades. The butterfly trades can also be implemented without any precise hedging by just taking the positions in butterfly trade fashion without calculating exact risk metrics.[3]

Trading the curvature means effectively trading mispricings in the rate curve. The directions of the long/short trades depend on the view of the trader on the curvature of the rate curve, which is usually based on some model or analysis. This provides the directions of the mispricings. The trader may e.g. use rigorous backtesting or just observe and analyze some anomalous events in the markets if e.g. a financial institution is forced to buy strongly a swap of certain maturity.

The mispricings on the given rate curve are relative mispricings, i.e. the rates are not compared to their intrinsic values. In other words, the trader does not usually have a clear view on the "correct" rate values but may e.g. think that the 5-year swap rate is mispriced relative to 1- and 10-year swap rates. In this case, the 1- and 10-year swap rates would be benchmark rates on the swap curve. The benchmark rates are rates that are liquid and efficiently priced and set a level on the rate curve to which other rates are compared. If a model is applied, such as a short rate term structure model, the model can be calibrated to such benchmark rates. In this case, the modelling provides the model rates to which market rates are compared but the model rates depend indirectly on the chosen benchmark rates.

To my knowledge, Duarte et al. (2007) is the only notable recent academic study that explicitly talks about yield curve arbitrage. According to Duarte et al. (2007), yield curve arbitrage was widely adopted by a number of fixed income arbitrage hedge funds

[1]DV01 indicates the approximate \$-change in mark-to-market valuation of a swap given a parallel shift of 1 basis point (0.01 %) in the underlying swap curve. A close relative of DV01 is duration that in the general fixed income context refers to the %-change in value of the fixed income instrument given small change in interest rates.

[2]Common risk metrics in fixed income world are duration and DV01.

[3]Good references for the concepts of butterfly trade, rate curve curvature, DV01 and duration are e.g. Pedersen (2015), Tuckman and Serrat (2012) and Hull (2017).

in the 1990s and the 2000s and the arbitrage was also adopted by Long-Term Capital Management in the 1990s[1]. Within hedge fund investment strategy categories, yield curve arbitrage belongs to the relative value investing category.

Duarte et al. (2007) study yield curve arbitrage in the USD interest swap market in 1988-2004. They use month-end quotes of midmarket fixed-for-floating ("vanilla") swap rates[2] for 1-, 2-, 3-, 5-, 7- and 10-year maturities. They apply the two-factor Vasicek short-rate model to find mispricings in the swap curve and implement the arbitrage in the butterfly trade fashion. The two-factor Vasicek model is perfectly calibrated to 1- and 10-year swap rates providing the model values of the other swap rates relative to the 1- and 10-year swap rates. They implement hedging by hedging out the two factors of the Vasicek short-rate model which can be interpreted as the level and slope of the swap rate curve and they are left exposed to the curvature of the swap rate curve.[3] They implement the hedging in practice by calculating the model rate sensitivities to the model factors in the mispriced swap rate and in the 1- and 10-year swap rates that are used for hedging, and take the long/short positions in the butterfly trade fashion such that the sensitivities to the factors zero out. They implement the yield curve arbitrage trading separately for different swap maturities[4] and open each month at most one butterfly trade that includes the swap rate of largest mispricing and the hedge rates, if the mispricing exceeds the given mispricing limit of 10 basis points. They close the trades after 12 months or after the original mispricing has decreased to less than 1 basis point. They also report results for equally-weighted (notional amount) portfolio of individual maturity yield curve strategies.[5]

The modelling of the swap curve with short rate model and hedging based on the

[1]Long Term Capital Management was a hedge fund that faced heavy losses in the late 1990s after the Russian Sovereign default and was later liquidated and dissolved.

[2]Duarte et al. (2007) refer to constant maturity swap rates in one section in talking about the swap rates but to my knowledge they apply vanilla swap rates. This conclusion is based on comparing their implementation of swap spread arbitrage to the general implementation of swap spread arbitrage and also on the general definition of swap spreads.

[3]The two-factor short rate models will be presented in subsection 3.5 and the methodology section 6 presents e.g. the calibration of the models.

[4]Only for 2-,3-,5- and 7-year maturities, because 1- and 10-year swap rates are perfectly modelled and used in hedging.

[5]I will go through the implementation of yield curve arbitrage by Duarte et al. (2007) in detail in section 6, because I mostly replicate their methodology.

factors of the short rate model makes this yield curve arbitrage implementation special. According to Duarte et al. (2007), this was the way most hedge funds and proprietary trading desks implemented yield curve arbitrage in their study period. Thus, as the implementation described above is still possibly used by hedge funds and as Duarte et al. (2007) is my main reference paper, I will now define yield curve arbitrage, as it is implemented in Duarte et al. (2007). I will use this definition throughout my thesis but this definition is neither a formal definition of yield curve arbitrage nor does it include all the technicalities related to e.g. mispricing limits and such things that are covered in later sections. This definition of yield curve arbitrage separates it from more general butterfly trades with modelling and hedging that are based on short rate models.

Definition 2.3 (Yield curve arbitrage)**.** Yield curve arbitrage trades mispricings that are related to the curvature of a given rate curve. The curvature of the rate curve is modelled with an endogenous two-factor short rate model that provides the model rate values and the corresponding model rate curve curvature. Based on the model rate values, mispricings are observed and long/short positions are taken in the rate curve in a butterfly trade fashion that zeros out risk related to the sensitivities of the model rates to the short rate model factors. The hedging against model factor sensitivities simultaneously provides imperfect hedging against changes in level and slope of the rate curve. The used model should be calibrated to observed market data and the model should be perfectly calibrated to some liquid and efficiently priced benchmark rates that are applied in hedging the factor sensitivities.

Duarte et al. (2007) find results that are somewhat contrary to previous beliefs on fixed income arbitrage trading. Namely, they find that yield curve arbitrage provides statistically significant positive excess returns. The Sharpe ratios and gain-to-loss ratios suggest strong performance. The skewness coefficients are positive and the means of excess returns exceed the medians of excess returns, which implies that yield curve arbitrage would not be a strategy providing mostly small returns and occasional big losses. There is more kurtosis in the distribution of excess returns than normal distribution would indicate. Duarte et al. (2007) also regress the excess returns on different equity- and bond market risk factors and get positive alphas, even after deducting hedge fund fees from the returns.

The main takeaway from section 2.2 is that yield curve arbitrage is fundamentally a butterfly trade on a rate curve but Duarte et al. (2007) and I apply specific short rate model driven version of yield curve arbitrage that separates it from general butterfly trades. Hedging and curvature modelling is based on the short rate model and I define yield curve arbitrage accordingly.

2.2.1 Sources of risks and returns in yield curve arbitrage

The next step is to link yield curve arbitrage to market neutrality and to consider the risks that yield curve arbitrage is possibly exposed to.

Vayanos and Vila (2009) build an exogenous term structure model, in which they explain the construction of the term structure with the trading of preferred-habitat investors and arbitrageurs. In their model, there are investors who demand zero-coupon bonds only for certain maturities in accordance with the preferred-habitat view. The arbitrageurs can invest in zero coupon bonds with different maturities and also in an exogenous short rate[1]. When the investors cause demand shocks (positive or negative) for certain zero coupon bond maturities, the arbitrageurs will react to the shocks and smoothen the shocks, if the expected excess return (over short rate) from buying or shorting the zero coupon bonds is large enough. Vayanos and Vila (2009) name this excess return as bond risk premia. The demand shocks would cause certain rates on the term structure to be mispriced relative to the nearby rates and therefore the term structure would be inconsistent. The short rate and the investor demand for zero coupon bonds of certain maturities are assumed to be stochastic.

The arbitrageurs change their investments between the short rate and zero coupon bonds based on the expected short rate, demand shocks and bond risk premia. The arbitrageurs face risks in trading the term structure and the higher the risks they face the higher bond risk premia they require. The demand shock effects become more localized the more there are risks for the arbitrageurs.

Although the model of Vayanos and Vila (2009) was mainly built to study the effect of demand shocks and the effects of monetary policy decisions, the model provides a theoretical framework for yield curve arbitrage. As the yield curve arbitrage is about trading relative mispricings on a rate curve[2], the framework of Vayanos and Vila (2009)

[1]The zero coupon bonds and short rates are presented more closely in the next chapter 3.

[2]More generally, in a term structure of interest rates. See Appendices section 11.5 for more

explains the mispricings by the demand shocks and also provides reasoning on when to trade the relative mispricings, i.e. when the bond risk premia is large enough.

As already noted, Duarte et al. (2007) regress the returns of yield curve arbitrage to several different factors, controlling equity-market risk, bond market risk and default risk[1]. In their regressions, market risk and in particular the default risk appears in the negative relation of yield curve arbitrage excess returns with the excess returns of general industrial corporate bonds. The default risk exposure is somewhat expected based on previous fixed income arbitrage literature. As Duarte et al. (2007) apply a two-factor model and base their hedging on the model, the negative market risk exposure could be interpreted as more than two factors driving the swap term structure. Otherwise, there would be no market risk exposure as the two-factor Vasicek model assumes that market risk is hedged away by being hedged against the model factor sensitivities. Duarte et al. (2007) state at their conclusions that "yield curve arbitrage returns are related to a combination of Treasury returns that mimic a 'curvature factor'". It is unclear to me what they refer to with this but I interpret it to mean that yield curve arbitrage is observed to target relative mispricings, as the yield curve arbitrage returns have negative and positive exposures to the returns of Treasury bond portfolios of different maturities in their regressions. The exposure to the curvature factor could explain the positive market risk adjusted returns[2]. In other words, Duarte et al. (2007) find that yield curve arbitrage is exposed to market risk through default risk and is also targeting the returns and risks related to the curvature of the rate curve. With the framework of Vayanos and Vila (2009), this could be interpreted as yield curve arbitrage targeting pricing inconsistencies in the swap curve that are possibly caused by demand shocks and that would possibly provide the bond risk premia. Duarte et al. (2007) do not directly link their curvature factor to the preferred-habitat view or demand shocks but trading the curvature in general matches the trading of mispricings caused by demand shocks in the study of Vayanos and Vila (2009).

Pedersen (2015) notes that almost all the rates and bond yields depend on risk-free interest rates in one form or another and therefore there is likely a lot of co-movement

information.

[1]Default risk is usually considered as the subcategory of credit risk but some sources generalize these concepts to be the same.

[2]The other market risk factors would be assumed not to control for this source of risk and return.

in the rates and bond yields. This likely raises relative pricing differences in closely related securities and provides the kind of trading opportunities described for yield curve arbitrage. However, according to Pedersen (2015) the relative pricing differences are likely small which requires applying leverage to take advantage of the pricing differences. This exposes the relative mispricing trading strategies to the tail risk. Duarte et al. (2007) do not find support for yield curve arbitrage being exposed to such tail risk, as the return distribution of yield curve arbitrage indicates strong performance and there is statistically significant exposure to only one systematic market risk factor.

Duarte et al. (2007) also state that yield curve arbitrage requires intellectual capital for implementation, which could explain the positive market risk adjusted excess returns. The authors find a positive relation between the increased capital amount invested in the fixed income arbitrage hedge funds and yield curve arbitrage returns, which is interesting, as usually increased amount of capital should eventually drive the returns down. They explain this connection with increased liquidity and faster convergence in the arbitrage that the increased capital provides.

Previous studies on the development of swap rates in general, such as Liu et al. (2006) and Duffie and Singleton (1997), have found that liquidity and credit risks are important in explaining the development of swap rates. Duarte et al. (2007) finds support for credit/default risk factor, as the exposure to the industrial bonds in risk factor regressions is linked to default risk. According to Fung and Hsieh (2002), many fixed income trading strategies are exposed to credit risk and credit risk affects demand for instruments, i.e. may cause demand shocks. Liquidity as itself is also connected to possible demand shocks. Therefore, these two risk factors could also be linked to the framework of Vayanos and Vila (2009) and to demand shocks and the curvature factor or bond risk premia explaining yield curve arbitrage returns.

The key takeaway from this subsection is that the yield curve arbitrage is exposed to market risk and in particular to the default risk as presented in Duarte et al. (2007). Yield curve arbitrage is also targeting the returns and risks from exposure to the curvature factor which can be seen similar to the bond risk premia in Vayanos and Vila (2009). Therefore, yield curve arbitrage returns could be obtained from targeting pricing inconsistencies in the term structure and the inconsistencies could be explained with demand shocks.

It is somewhat unclear to me, whether the exposure to the curvature factor of Duarte et al. (2007) or to the demand shock risk and short rate risk of Vayanos and Vila (2009) should be interpreted as market risk exposure. I interpret that these exposures could be interpreted as the isolated sources of market risk that partially market neutral trading strategies commonly target to. In my opinion, it is not that important to categorize the curvature factor, demand shock risk or short rate risk to being market/non-market risks. What is important is that by the results of Duarte et al. (2007), yield curve arbitrage appears to be hedged against most market risk factors while being exposed to the presented isolated sources of risk.

2.2.2 Master's theses on yield curve arbitrage

While the number of published academic studies about yield curve arbitrage is practically limited to Duarte et al. (2007), there are a couple of high-quality master's theses on yield curve arbitrage, namely the theses of Stark (2020) and Karsimus (2015). I have familiarized myself closely with these theses as they provide good benchmarks to compare my methodological choices and yield curve arbitrage performance.

Karsimus (2015) studies yield curve arbitrage in EUR constant maturity interest rate swap markets[1] with monthly data in 2002-2015 by applying two equilibrium short rate models, the two-factor Cox-Ingersoll-Ross model and the two-factor Longstaff-Schwartz model. He finds, that yield curve arbitrage produces lucrative risk-adjusted returns and return distributions. He employs enhanced methodology compared to Duarte et al. (2007), as he e.g. uses out-of-sample period for trading and in-sample period for model calibration. Duarte et al. (2007) optimize the parameters in the same in-sample period as they do their trading. Karsimus (2015) also compares high-level and style-speficic hedge fund index returns to the replicated strategy returns but finds no statistically meaningful connection. He also studies "high-noise" periods in the markets and finds that high noise coincides with large model-implied mispricings when the measure of the mispricings is smoothed. He finds no evidence in support of the idea that yield curve arbitrage alpha is compensation for carrying tail risk.

Stark (2020) analyzes the out-of-sample trading performance of yield curve arbitrage with EUR interest rate swaps in 2010-2019. He also basically replicates the paper of Duarte et al. (2007) but also introduces a novel hybrid neural network approach which

[1] Karsimus clearly states that he applies constant maturity swap data.

uses the factors of the two-factor Vasicek model as inputs. He uses an in-sample period from 2000-2010 to train the neural network model. Contrary to Duarte et al. (2007) and Karsimus (2015), Stark (2020) uses daily data as opposed to month-end data and he applies OIS discounting in valuation as in Smith (2013). He compares the performance of neural network implementation to the performance of the traditional Vasicek model implementation. Stark (2020) finds that neural network implementation provides better multifactor alpha, Sharpe ratio and cumulative returns than the traditional two-factor Vasicek model implementation and positively skewed return distribution when transaction costs are reasonable. However, the neural network implementation also has more exposure to risk factors. Both implementations of yield curve arbitrage still produce statistically significant positive returns.

3 Short rate modelling

Yield curve arbitrage could as well be referred to as 'term structure of interest rates arbitrage'. Yield curve arbitrage is not limited to only yield curve trading but in general to trading the term structure of interest rates. The concept of term structure of interest rates is more closely presented in the Appendices section 11.5 and there is clear conceptual difference between the yield curve and the term structure of interest rates. With the different term structure modelling frameworks, all interest rates and prices of interest rate products can be solved. Therefore, yield curve arbitrage can be applied for various instruments.

In this section, I will go through short rate modelling that is the term structure modelling framework that I apply. First, I present short rate modelling in general. Then, I present the two categories of short rate models. At the end of the section, I will present the term structure models that I apply in practice in the arbitrage implementation. The mathematical notation and most of the definitions in this section is based on Brigo and Mercurio (2007) and Tuckman and Serrat (2012). Other major references in this section are Hull (2017), Duarte et al. (2007), Vasicek (1977) and Cox et al. (1985).

3.1 Short rate modelling in a nutshell

The presentation of key concepts in this section is based on Brigo and Mercurio (2007) with some additions of my own. Term structure modelling was initially based on modelling the so-called short rate, a theoretical instantaneous spot rate.

Definition 3.1 (Short rate)**.** Short rate ($r(t)$) is the risk-free, annual, instantaneous spot rate that is accumulated in an infinitesimally short amount of time at time t. It is purely theoretical and economy-dependent, such that the short rate can not be directly observed and the short rate may differ between e.g. Eurozone and the US. Although the short rate is not directly observed, it can be approximated as the limit of the observed spot rates, such that:

$$r(t) = \lim_{T \to t^+} L(t,T) = \lim_{T \to t^+} Y(t,T) = \lim_{T \to t^+} R(t,T)$$

where $L(t,T)$ denotes the simply compounded spot rates, $Y(t,T)$ denotes the annually compounded spot rates and $R(t,T)$ denote the continuously compounded spot rates.[1]

Modelling the short rate allows to derive the prices of fundamental quantities, such as bond prices and other spot rates, as the expectation of a functional of the stochastic short rate process. Applying risk-neutral measure[2], the arbitrage-free price at time t of a contingent claim with payoff H_T at time T is given by:

$$H_t = E_t\{D(t,T)H_T\} = E_t\{e^{-\int_t^T r(s)ds} H_T\} \tag{3.1}$$

Here $D(t,T)$ refers to a discount factor from time t to time T that is a random variable.

Definition 3.2 (Zero coupon bond)**.** A T-maturity zero coupon bond (pure discount bond) is a contract, that guarantees its holder the payment of one unit of currency at time T, with no intermediate payments. The contract value at time $t < T$ is denoted by $P(t,T)$. Clearly, $P(T,T) = 1$ for all T.

Zero coupon bond can be understood as an expectation of the random variable discount factor $D(t,T)$ under a certain probability measure. There are actual zero coupon bonds at the markets these days, but in this context the zero coupon bond refers to a theoretical concept, whose price can be solved from the market prices of liquid base fixed income instruments with bootstrapping. As such, zero coupon bonds are dealt here as theoretical instruments that are not directly observable in the market.

The zero coupon bond price at time t for the maturity T is characterized by a unit amount of currency available at time T, so that $H_T = 1$ and:

$$P(t,T) = E_t\{e^{-\int_t^T r(s)ds}\} \tag{3.2}$$

Whenever I am able to derive zero coupon bond prices, I am able to derive all the

[1]The term structure of interest rates consists of simply compounded and annually compounded spot rates for different maturities. See the Appendices section 11.5 for more information.

[2]The change of measure from objective (real) world measure to risk-neutral world measure is one of the cornerstones in mathematical finance. I will not present this topic more closely, but there is a good introduction to the topic in Brigo and Mercurio (2007).

interest rate dynamics applying the law of one price and bootstrapping.[1] The zero coupon bond price is arguably one of the most important concepts in term structure modelling.

Therefore, there is a connection between the short rate and the zero coupon bond prices. In short rate modelling, the short rate follows a stochastic process and therefore the distribution of short rate is important. If the distribution of $e^{-\int_t^T r(s)ds}$ can be characterized in terms of a chosen dynamics for $r(t)$ conditional on the information available at time t, I am able to compute zero coupon bond prices P as presented above. In other words, the distribution of $e^{-\int_t^T r(s)ds}$ is dictated by the stochastic process and distribution of short rate $r(t)$. The distributions of $e^{-\int_t^T r(s)ds}$ and $r(t)$ refer to the final distributions, i.e. if the short rate is modelled iteratively with its stochastic process for a given period, then some final distribution of $r(t)$ is observed.

There are differences between the short rate models whether the distribution of the short rate and the stochastic short rate process allow for analytically computing zero coupon bond prices. Even if I am able to define the distribution of the short rate and define the stochastic short rate process, I may not be able to analytically solve the expected value of this process or the expected values of derived processes. Some short rate models require tree/lattice based pricing or Monte Carlo simulation for solving the expected values.

For example, models in which the distribution of short rates is a normal distribution, are called normal or Gaussian models. It was generally thought, that one problem with these models is that the short rate can become negative. A negative short rate would not make much economic sense, because people would never lend money at a negative rate when they can hold cash and earn a zero rate instead. These thoughts were somewhat revoked after the introduction of first negative nominal interest rates. It is also be important to consider e.g. the volatility structures implied by the short rate model and whether the short rate process is mean reverting.

An example of the stochastic process for the short rate is of the form

$$dr(t) = m(r)dt + s(r)dz \tag{3.3}$$

[1] I provide an example of bootstrapping in the Appendices section 11.8. The term structure of interest rates is also derived based on the zero coupon bond prices in the Appendices section 11.5.

Here the process of short rate is dictated by drift term $m(r)$ related to the short rate, standard deviation $s(r)$ related to the short rate and a Brownian motion z. This is a very simplified stochastic, time-independent, risk-neutral process presentation for the short rate process. The models that share the presented structure differ in how the drift and volatility are modelled. Usually there is a set of parameters related to the drift and volatility and these may differ between models.

The short rate process can be described either in risk-neutral world, where expectations for prices happen or in the real objective world, where different rates and proxies for short rate are observed. The short rate processes in these two worlds are formulated differently and the separating thing is the market price of risk. All my applications of the short rate models will happen in the risk-neutral world.

3.2 One-factor and multifactor models

This section provides some interpretations for the factors in the short rate models. It was difficult to find any clear answers to questions such as 'what are the factors' or 'how to interpret the factor values'. Therefore, this section includes many of my own interpretations for the factors together with explanations from Brigo and Mercurio (2007).

The short rate models can generally be categorized to one-factor models and multifactor models. The example of the stochastic short rate process presented in the previous section was a one-factor model. A factor refers to a stochastic state variable to which a source of randomness is linked. In the one-factor model presented in the previous section, there is only one source of randomness stemming from the Brownian motion z and the short rate itself is directly the stochastic state variable.

In the one-factor models, the stochastic short rate process can be presented as a function of the initial short rate $r(0)$ and a set of parameters α that commonly compose the drift and volatility of the short rate process. If the drift and volatility in the one-factor model are time-independent, as in equilibrium models, then the whole term structure of interest rates modelled with such model is fully correlated.

In the multifactor models, there are more than one stochastic state variable and source of randomness. An example structure for two-factor model is:

$$\begin{aligned} r(t) &= x(t) + y(t) \\ dx(t) &= m(x)dt + s(x)dz \\ dy(t) &= m(y)dt + s(y)dz \end{aligned} \tag{3.4}$$

For the two-factor model, the stochastic short rate process can be presented as a function of the initial factors $x(0)$ and $y(0)$ and two sets of parameters α_x and α_y. In the multifactor models, there is no perfect correlation between rates in the term structure and the state variables are usually assumed to be independent.

In multifactor models, the factors are the state variables and it can be difficult to provide interpretations for the factors. Principal component analysis and respective interpretations have been given for the factors and in two-factor short rate models the two factors could be interpreted as level and slope of the term structure. As already mentioned in section 2.2, Duarte et al. (2007) applies the two-factor Vasicek model and base their hedging on the assumption, that the model factors can be considered as level and slope of swap curve. Therefore, the modelling is left unexposed to the third major element of rate curve principal components, i.e. the curvature[1]. A two-factor model assumes that hedging against the two factors and the corresponding factor risks should provide a complete hedge against all risk.

A two-factor model could alternatively be linked to the term structure directly, such that one of the factors could be interpreted to be a fastly changing short maturity rate[2] and the other factor some slowly changing long maturity rate. These rates can be modelled with separate stochastic rate processes, one for the short maturity rate and the other for the long maturity rate and the factors would basically be these two rates. As such, the interpretation about the factors presenting the level and slope of term structure becomes more clear.

At the core of short rate modelling, it is though the case that the interpretations for the factors are neither unambiguous nor particularly important. The factors are just pure abstractions, some unknown stochastic state variables with linked sources of randomness

[1] Two-factor models assume that there are only two stochastic state variables driving the term structure of interest rates. Otherwise the third factor would be included in modelling. The interpretation of the PCA analysis for the factors is just one alternative.

[2] Not the short rate, but some other short maturity rate.

and the factors are behind the short rate evolvement. These abstract state variables become quantified in the form of the factors in short rate modelling. Only with the objective world calibration, it is necessary to interpret the factors as short maturity rates and long maturity rates so that some proxies can be used for parameter estimation. This is not necessary in the risk-neutral world with calibration to market prices, as I will later explain.

The number of factors also affects the ability to fit market data to the model values and to present realistic correlation and covariance structures. On the other hand, too many factors affect the numerically-efficient implementation of short rate modelling.

3.3 Equilibrium short-rate models

The short rate models can be divided to equilibrium models and exogenous models. This division is based on whether the model takes the market term structure of interest rates as an input and relies on time-dependence. The following presentation and discussion about the equilibrium models and the exogenous models is based on Hull (2017), Brigo and Mercurio (2007) and Tuckman and Serrat (2012).

Equilibrium models commonly make assumptions about the dynamics of the term structure and economic variables and then derive the process for the short rate r. A common form for the one-factor risk-neutral process of r in equilibrium model is the already presented:

$$dr(t) = m(r)dt + s(r)dz \tag{3.5}$$

In the equilibrium models, there is no time-dependence and as such the drift and standard deviation in the example structure above are functions of r but independent of time. The example structure for a two-factor equilibrium short rate model could be the already presented:

$$\begin{aligned} r &= x + y \\ dx &= m(x)dt + s(x)dz \\ dy &= m(y)dt + s(y)dz \end{aligned} \tag{3.6}$$

Common examples of two-factor equilibrium models are the Cox-Ingersoll-Ross model and the two-factor Vasicek model, both of which I will apply in this thesis.

Equilibrium models produce a term structure as output given the chosen parameters and the market term structure is not provided to the equilibrium models as an input. The resulting model term structure may not match observed market term structure. Equilibrium models are not particularly suitable for e.g. valuing interest rate derivatives, where it is important that the initial term structure observed in the market is matched. Equilibrium models suit well for situations where the purpose is to e.g. value bonds or swaps relative to one another and thus they suit for yield curve arbitrage.

3.4 Exogenous short-rate models

The exogenous short rate models are also known as no-arbitrage models. They take the initial term structure as an input and the models are made to fit exactly to the market values. An example structure of exogenous model could be of the form:

$$dr = m(r,t)dt + s(r)dz \tag{3.7}$$

Therefore, in exogenous models there is time dependence as can be seen in the drift of the example structure. This is because the shape of the initial term structure governs the average path taken by the short rate in the future and it is therefore necessary to have time-dependence.

Exogenous models formulate stochastic short rate processes based on the observed term structures. In this sense, exogenous models are less theoretical models and more models of data fitting. Exogenous models are referred to as no-arbitrage because these models consider the observed term structure values as correct. However, all the term structure models are arbitrage-free in the sense that the models are internally arbitrage-free.

Multifactor exogenous models can be made to perfectly fit the observed term structures by using as many time-dependent factors/state variables as there are data points. Usually a reasonable amount of time-dependent state variables is enough to provide a good enough fit to the observed term structure. Thus, the no-arbitrage models could in theory be used in arbitrage or relative value trading, if not all price or rate observations are used as data points. In general, both the equilibrium and the no-arbitrage models share some aspects of data fitting, as e.g. the practical application of the two-factor Vasicek equilibrium model also requires data calibration. An example application of exogenous models is to quote the prices of non-actively traded securities based on the prices of more liquid securities. In this sense, exogenous models can be used as an alternative to pure mathematical curve-fitting methods to obtain the prices of less liquid securities. The exogenous models are also commonly used for pricing and valuating interest rate derivatives.

In some situations, the assumption of exogenous models about fair market prices may not be reasonable, if there has been e.g. strong supply and demand shocks. In these situations, building the exogenous short rate process based on market prices makes the exogenous model flawed.

The two categories of short rate models can be blended in various ways. A model might e.g. take the prices of certain liquid and fairly-priced swaps as given, while allowing the model to valuate other securities.

3.5 The applied models

Now that I have presented all the necessary background on short rate modelling, it is time to move on to the actual models that I apply in my thesis. I apply two equilibrium short-rate models: the two-factor Vasicek model and the two-factor Cox-Ingersoll-Ross model. The mathematical notation in this section is completely based on Brigo and Mercurio (2007). I will not present all the mathematical steps for deriving the zero coupon prices from short rate processes but I will focus on the key notions about the models.

3.5.1 Two-factor Vasicek model

Vasicek (1977) presented one-factor short rate model, in which the short rate evolves as an Ornstein-Uhlenbeck process with constant coefficients under the objective world measure. When the market price of risk is chosen properly, the short rate evolves as an Ornstein-Uhlenbeck process with constant coefficients under the risk-neutral world measure as well, that is:

$$dr(t) = k[\theta - r(t)]dt + \sigma dW(t), r(0) = r_0 \tag{3.8}$$

where r_0, k, θ and σ are positive constants. $k[\theta - r(t)]dt$ defines drift with mean reversion. The constants can be interpreted as speed of mean reversion (k), long run value of the short rate (θ) and volatility of the short rate (σ). $W(t)$ is a Brownian motion and the stochastic source of randomness $\sigma dW(t)$ is normally distributed. Integrating [3.8], I obtain for each $s \leq t$,

$$r(t) = r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) + \sigma \int_s^t e^{-k(t-u)} dW(u) \tag{3.9}$$

so that $r(t)$ conditional on $\mathcal{F}_s$[1] is normally distributed with mean and variance given respectively by

$$\begin{aligned} E\{r(t)|\mathcal{F}_s\} &= r(s)e^{-k(t-s)} + \theta(1 - e^{-k(t-s)}) \\ Var\{r(t)|\mathcal{F}_s\} &= \frac{\sigma^2}{2k}\left[1 - e^{-2k(t-s)}\right] \end{aligned} \tag{3.10}$$

One-factor Vasicek model is an endogenous term structure model, in which the short rate follows Gaussian (normal) distribution. Therefore, at each time t, the rate $r(t)$ can be negative with positive probability, which was previously considered to be a drawback of the Vasicek model. Other drawbacks of the model include that the model implies flat volatility and because the modelled term structure is perfectly correlated, there are parallel shifts in the modelled term structure.

On the other hand, as the model relies on Gaussian distribution for the short rate, it has good analytical tractability. The short rate r is mean reverting, since the expected rate tends, for t going to infinity, to the long run value θ.

As the short rate follows Gaussian distribution, the price of a zero coupon bond based on the model can be directly computed with the expectation [3.2] and I obtain:

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)} \tag{3.11}$$

where

[1]Filtration that in probability theory and stochastic processes roughly corresponds to a collection of sigma algebras. To put simply, filtration corresponds to all historical but not future information available about the stochastic process.

$$
\begin{aligned}
A(t,T) &= exp\Big\{\Big(\theta - \frac{\sigma^2}{2k^2}\Big)\Big[B(t,T) - T + t\Big] - \frac{\sigma^2}{4k}B(t,T)^2\Big\} \\
B(t,T) &= \frac{1}{k}\Big[1 - e^{-k(T-t)}\Big]
\end{aligned}
\tag{3.12}
$$

As far as I understand, the two-factor Vasicek model does not exist in theory, as Vasicek (1977) did not derive any two-factor short rate model. However, two-factor short rate models relying on Gaussian distribution have been developed based on the one-factor Vasicek model. A formulation of this kind of hypothetical Gaussian two-factor Vasicek model (G2-model)[1] is presented as:

$$
\begin{aligned}
r(t) &= x(t) + y(t) \\
dx(t) &= k_1(\theta_1 - x(t))dt + \sigma_1 dW_1(t) \\
dy(t) &= k_2(\theta_2 - y(t))dt + \sigma_2 dW_2(t)
\end{aligned}
\tag{3.13}
$$

with instantaneously correlated sources of randomness $dW_1 dW_2 = \rho dt$. The constants $(k_1, \theta_1, \sigma_1, k_2, \theta_2, \sigma_2)$ are positive. Duarte et al. (2007) assume that the two Brownian motions W_1 and W_2 are independent and that $dW_1 dW_2 = 0$. To be precise, Duarte et al. (2007) only mention, that the correlation is zero, which does not imply independence of the Brownian motions, but I interpret, that this is what they must have meant. Based on the additional assumption, the zero coupon bond price based on the G2-model is presented as:

$$
P(t,T) = A(t,T,k_1,\theta_1,\sigma_1)A(t,T,k_2,\theta_2,\sigma_2)exp\Big\{ - B(t,T,k_1)x_t - B(t,T,k_2)y_t\Big\} \tag{3.14}
$$

where

$$
\begin{aligned}
A(t,T,k,\theta,\sigma) &= \Big\{\big(\theta - \frac{\sigma^2}{2k^2}\big)(B(t,T,k) - T + t) - \frac{\sigma^2}{4k}B(t,T,k)^2\Big\} \\
B(t,T,k) &= \frac{1}{k}\Big[1 - e^{-k(T-t)}\Big]
\end{aligned}
\tag{3.15}
$$

Considering, that W_1 and W_2 are independent allows me to form the zero coupon bond price as the product of the zero coupon bond prices, that are based on factors x

[1]Throughout the thesis, referring to "two-factor Vasicek model" means referring to this G2-model.

and y. The two mean reverting factor processes are modelled separately with one-factor Vasicek models, where all the assumptions of the one-factor Vasicek model apply. The factor processes are combined, when the short rate is calculated as sum of the factors and when the zero coupon bond price is solved as described above.

Duarte et al. (2007) apply different construction for the factor processes and use different notation with (α, β, σ) and (μ, γ, η) as constants in the two factor processes. I can link their factor processes to these factor processes presented by Brigo and Mercurio (2007) with basic algebra and I will present this link in section 6.2, where I go through the methodology and modelling in practice.

In summary, I apply the hypothetical two-factor Vasicek model, with the additional assumption by Duarte et al. (2007) about correlation, i.e. $dW_1 dW_2 = 0$. This model has two mean reverting processes for its factors x and y and these mean reverting processes are modelled with the one-factor Vasicek model. The short rate and zero coupon bond prices are obtained by simple additive or multiplicative combination of the results of the two factor processes.

3.5.2 Two-factor Cox-Ingersoll-Ross model

The Cox-Ingersoll-Ross (CIR) model also has its one-factor and two-factor versions. Duarte et al. (2007) do not apply the two-factor CIR model in their paper, but I will apply it as it is a very popular model and I want to get comparable results for the two-factor Vasicek model.

The one-factor CIR model was originally derived as an extension to the one-factor Vasicek model. CIR introduced a "square-root" term in the diffusion coefficient[1] of the short rate dynamics proposed by Vasicek (1977). In the CIR model, short rate is always positive. Under the risk-neutral measure, the model is presented as

$$dr(t) = k[\theta - r(t)]dt + \sigma\sqrt{r(t)}dW(t), r(0) = r_0 \tag{3.16}$$

with r_0, k, θ, σ positive constants. The constants have the same interpretations as in the Vasicek model. There is also an additional condition that $2k\theta > \sigma^2$, which ensures that the short rate remains positive all the time.

[1] Coefficient of the differential of Brownian motion $dW(t)$, i.e. $\sigma\sqrt{r(t)}$.

The short rate follows noncentral chi-squared distribution. The mean and variance of $r(t)$ conditional on $\mathcal{F}_s$ are given by

$$
\begin{aligned}
E\{r(t)|\mathcal{F}_s\} &= r(s)e^{-k(t-s)} + \theta\Big(1 - e^{-k(t-s)}\Big) \\
Var\{r(t)|\mathcal{F}_s\} &= r(s)\frac{\sigma^2}{k}\Big(e^{-k(t-s)} - e^{-2k(t-s)}\Big) + \theta\frac{\sigma^2}{2k}\Big(1 - e^{-k(t-s)}\Big)^2
\end{aligned}
\tag{3.17}
$$

The price $P(t,T)$ at time t of a zero-coupon bond with maturity T is solved as

$$
\begin{aligned}
P(t,T) &= A(t,T)e^{-B(t,T)r(t)} \\
A(t,T) &= \Big[\frac{2h exp\{\frac{(k+h)(T-t)}{2}\}}{2h + (k+h)(exp\{(T-t)h\} - 1)}\Big]^{\frac{2k\theta}{\sigma^2}} \\
B(t,T) &= \frac{2(exp\{(T-t)h\} - 1)}{2h + (k+h)(exp\{(T-t)h\} - 1)} \\
h &= \sqrt{k^2 + 2\sigma^2}
\end{aligned}
\tag{3.18}
$$

Therefore, the short rate process $dr(t)$, the distribution of the short rate and the analytical bond pricing formulas change when moving from one-factor Vasicek model to the one-factor CIR model.

The two-factor CIR model is more widely established than the hypothetical two-factor Vasicek model. In the basic two-factor CIR-model, the short rate is obtained by adding two independent processes under the risk-neutral measure, as in two-factor Vasicek model

$$
r(t) = x(t) + y(t) \tag{3.19}
$$

The independent processes for the factors x and y are obtained as

$$
\begin{aligned}
dx(t) &= k_1[\theta_1 - x(t)]dt + \sigma_1\sqrt{x(t)}dW_1(t) \\
dy(t) &= k_2[\theta_2 - y(t)]dt + \sigma_2\sqrt{y(t)}dW_2(t)
\end{aligned}
\tag{3.20}
$$

where W_1 and W_2 are independent Brownian motions under the risk-neutral measure and $\alpha_1 = (k_1, \theta_1, \sigma_1)$ and $\alpha_2 = (k_2, \theta_2, \sigma_2)$ consist of only positive constants, such that $2k_1\theta_1 > \sigma_1^2$ and $2k_2\theta_2 > \sigma_2^2$. Therefore, as the processes for factors x and y are modelled independently, the factors follow one-factor CIR models with all the assumptions of one-factor CIR model.

The zero-coupon bond price at time t for maturity T is obtained as:

$$P(t, T, x(t), y(t), \alpha) = P(t, T, x(t), \alpha_1)P(t, T, y(t), \alpha_2) \tag{3.21}$$

where the separate P-formulas with α_1 and α_2 are the zero coupon bond price formulas for corresponding one-factor CIR models.

The two-factor CIR model automatically assumes that W_1 and W_2 are independent and that $dW_1dW_2 = 0$. This is an obligatory choice, as the Brownian motions W_1 and W_2 can not be correlated, or else the two-factor CIR model would not be analytically tractable.

In summary, two-factor CIR model is an endogenous short rate model that models two independent mean reverting factor processes. The CIR model differs from the Vasicek model in that it adds the square root term in the diffusion coefficient, which makes the basis point volatilities of the factor processes proportional to the square root of the factor. As the short rate becomes very small, the standard deviation also becomes very small and the positive drift pushes the short rate up, thus preventing the short rate from becoming negative.

3.5.3 Calibration

A key thing to notice in the presented models and in the short rate models in general is the initial assumptions that the models have. One-factor models assume the initial short rate and the model parameters to be known at $t = 0$ whereas multifactor models assume the initial factors and model parameters to be known at $t = 0$. This naturally raises the question of how to find values for short rate, factors and parameters if they are not known.

There are broadly speaking two alternatives to find out the required values. In the objective world, historical data of proxies for short-maturity and long-maturity rates can be applied with methods such as regressions and maximum likelihood estimation. These estimation methods provide the parameter values, but the initial short rate or factor values may need to be approximated in another way.

The alternative is to do calibration in risk-neutral world applying prices of e.g. bonds or swap rates. In this case, the objective would be to minimize some function of data fit, e.g. sum of squared errors between model and market prices or similar. The idea would be to set the factors and parameters to match the prices. It is common to calibrate the models to cap/swaption volatility structures, such that model implied volatilities are minimized with respect to the market implied volatilities. I will be calibrating my models to the swap data that is not a very common calibration choice, but in the context of my thesis it is a necessary choice.

Especially the risk-neutral calibration is very delicate to different choices, starting with the calibration data. The quality of data is very important, such that preferably a sample of over 10 years of data would be used. According to Kim and Orphanides (2012), regime changes, crises or other big events during the data sample can make the data flawed. Another important choice regarding calibration is the instruments to which models are calibrated. If models are calibrated by setting perfect match to some benchmark prices, then the idea is that the models are internally risk-free and there are arbitrage opportunities if model prices differ from other market prices. However, when calibrating models in the risk neutral world by forcing some perfect matching to some benchmark prices, some of the analytical properties of short rate models may break. Brigo and Mercurio (2007) warns about this in talking about the analytical property of short rates remaining positive in the CIR model.

The key takeaway from this subsection is to remember that the short rate models assume initial short rates, factors and parameters to be known and the models do not provide clear guideline how to solve the values for these. Calibration is applied in this context, but it automatically requires many assumptions that may make the results of the applied models flawed or break some properties of the models.

4 Contributions, research questions and hypotheses

My main contribution to the literature on fixed income arbitrage is to study the yield curve arbitrage with a recent and unstudied dataset in the USD interest swap market. I also contribute by applying the two-factor CIR model and by extensively presenting my methodology. I list below the contributions and related research questions:

Contribution 1: Duarte et al. (2007) studied yield curve arbitrage in the USD interest rate swap market in 1988-2004. Since then, the world has seen a financial crisis, a prolonged period of low interest rates, unconventional monetary policies and a global pandemic, to name a few. As mentioned in section 2.2, yield curve arbitrage has not been extensively studied. To my knowledge, there has not been other noteworthy recent studies than Duarte et al. (2007) about yield curve arbitrage in the interest rate swap market. It is time to update the study of yield curve arbitrage to the 2010s and I do it by studying yield curve arbitrage in the USD interest rate swap market in 2012-2021. The research question related to my first contribution could be roughly formulated as: *does yield curve arbitrage provide positive excess returns in the USD interest rate swap market in 2012-2021 and do the returns differ from the returns presented by Duarte et al. (2007)?*

Contribution 2: My second contribution is to apply the two-factor Cox-Ingersoll-Ross model to the implementation of yield curve arbitrage. Some previous master's theses[1] on yield curve arbitrage have done this, but not to my knowledge any published academic studies. The CIR model is arguably one of the most popular short rate models and its two-factor version is also popular. As discussed in 3.5, the CIR model has good distributional and analytical properties and it prevents negative rates, which has been seen as a good feature. If the USD swap rates would have been negative in the 2010s, there would have automatically been a theoretical flaw in the two-factor CIR model. The research question related to my second contribution could be roughly formulated as: *does the yield curve arbitrage performance with the two-factor CIR model differ from the performance with the two-factor Vasicek model in the USD interest rate swap market in 2012-2021?*

[1] For example Karsimus (2015).

Contribution 3: My third contribution is the precise presentation of methodology and data behind yield curve arbitrage. I make some major modifications to methodology compared to Duarte et al. (2007) by e.g. applying the OIS discounting framework in swap valuation. There are numerous choices regarding yield curve arbitrage implementation that can affect the results significantly. I also raise possible issues in data and methodology that could cause issues in real-life implementation of yield curve arbitrage. The research question related to my third contribution could be roughly formulated as: *do the choices regarding applied data and methodology affect the performance of yield curve arbitrage in the USD interest rate swap market in 2012-2021?*

Next I present the hypotheses. Hypotheses 1, 2 and 5 are based on the findings of Duarte et al. (2007) and hypotheses 3 and 4 are based on Duarte et al. (2007), other studies and my own ideas:

Hypothesis 1: Yield curve arbitrage generates positive excess returns in the USD interest rate swap market in 2012-2021.

Hypothesis 2: The return distribution of yield curve arbitrage is positively skewed with a high kurtosis in the USD interest rate swap market in 2012-2021.

Hypothesis 3: The returns of yield curve arbitrage are lower in the USD interest rate swap market in 2012-2021, than in the study period of Duarte et al. (2007) in 1988-2004.

Hypothesis 4: The two-factor CIR model provides worse yield curve arbitrage performance than the two-factor Vasicek model in the USD interest rate swap market in 2012-2021.

Hypothesis 5: Yield curve arbitrage provides statistically significant multifactor alpha in the USD interest rate swap market in 2012-2021, when the returns are explained by the same market risk factors as in Duarte et al. (2007).

Hypotheses 1 and 3 are related. I hypothesize that yield curve arbitrage still provides positive returns in the 2010s but that the returns have decreased from the study period of

Duarte et al. (2007). I argue, that yield curve arbitrage still requires intellectual capital to implement and the arbitrage taps on the curvature factor/bond risk premia for returns. However, the amount of intellectual capital has likely increased from 1988-2004 with the introduction of new and effective technology and wider knowledge about term structure trading. Therefore, the competition in yield curve arbitrage has likely increased and could drive the returns down. The studies of Duarte et al. (2007) and Vayanos and Vila (2009) were published just before my study period of the 2010s, which could affect the performance of yield curve arbitrage in my study period although I doubt that the effect would be strong. As I will present in next section 5, the rate environment has also been very challenging in the 2010s. Concerns about liquidity and default risk have likely been significant in the 2010s and therefore yield curve arbitrage could have been negatively exposed to market risk through e.g. default risk.

With respect to hypothesis 2, I do not have a reason why the results concerning the shape of return distribution would be much different from Duarte et al. (2007). As there are not other published studies on yield curve arbitrage with interest rate swaps, it is difficult to justify these hypotheses otherwise. The return distribution basically addresses the possible issue of tail risk involved in yield curve arbitrage returns that was already quite well discarded by Duarte et al. (2007). Similarly, hypothesis 5 is related to the findings of Duarte et al. (2007) and to the market neutrality and source of returns for yield curve arbitrage that I do not believe have changed from 1988-2004.

Hypothesis 4 was somewhat justified in the explanation of contribution 2. In the rate environment of the last decade, rates have been close to zero and some nominal rates have been negative. For example, the prices of LIBOR floors with strike 0 % have had positive prices indicating that the market has seen positive risk-neutral probabilities for negative rates. Although the USD swap rates have not been negative, the underlying noncentral chi-squared distribution for CIR model that prevents negative short rate possibilities, could have underperformed relative to Vasicek model in the rate environment of recent years.

5 Data

My main dataset consists of monthly USD interest rate swap -, USD LIBOR - and USD overnight indexed swap (OIS) data. I need this data in modelling the swap rates, observing and trading potential mispricings, valuating the open arbitrage trades and calculating the returns. My main study period for trading is 01/2012 - 12/2021.

USD interest rate swap data consists of Intercontinental Exchange (ICE)[1] month-end mid-rates for USD fixed-for-floating ("vanilla") interest rate swaps, in which the fixed leg consists of semiannual payments[2] and floating leg consists of 3-month USD LIBOR payments[3]. In USD interest rate swaps, the underlying nominal is in USD and both the fixed and floating leg are paid in USD. Month-end refers to last trading days of months, i.e. not the average of rates in a month or similar. The swap data consist of 11 AM (London time) quotes, i.e. not the closing prices. The data is obtained from Refinitiv Datastream. I obtain rates for 1-10 full-year maturities in period 01/2007 - 12/2021. This expanded period, compared to my trading period 01/2012 - 12/2021, is needed for calibration in swap modelling. I will later in subsection 6.2 explain this more carefully.

USD LIBOR data consists of month-end ICE USD LIBOR rates for overnight-, 1 week-, 1 month-, 2 months-, 3 months-, 6 months- and 12 months maturities. USD LIBOR rates are the average interbank interest rates at which large banks in the London money market lend one another unsecured funds denominated in USD. The USD LIBOR rates are also for the 11 AM (London time) quotes for last trading days of months and not averages or similar. The USD LIBOR data is obtained from Refinitiv Datastream for 09/2011 - 12/2021[4]. I compare both the USD interest rate swap and USD LIBOR data from Refinitiv Datastream to the data from Federal Reserve Economic Data (FRED) database and the data seems to match.

USD OIS data is obtained from Refinitiv Eikon, i.e. not from Refinitiv Datastream. I will explain OIS more closely in subsection 6.1. I obtain Tullett Prebon[5] USD OIS data for 1 week, 2 weeks, 3 weeks, 1-24 months and various monthly maturities in maturities over 2 years. As with USD interest rate swaps, in USD OIS contracts the underlying nominal

[1] ICE operates e.g. global exchanges and provides e.g. pricing data on fixed income securities.

[2] 30/360 day count convention, see section 6.1.

[3] Actual/360 day count convention, see section 6.1.

[4] I am unable to get longer period for LIBOR and OIS data, when I merge the different pieces of data.

[5] One of the world's leading interdealer brokers.

is in USD and the fixed and floating leg are paid in USD. The USD OIS quotes are for month-end closing mid-rates in 09/2011 - 12/2021. Therefore, I use 11 AM (London time) data for swaps and LIBOR, but closing rates for OIS data. As the swap and LIBOR data are from ICE/Datastream and the OIS quotes are closing mid-rates of Tullett Prebon/Eikon, there is obviously a slight mismatch. There is no OIS data from ICE. Tullett Prebon OIS data is not very well described in Refinitiv Eikon, but it should be reliable data.

From here forwards, I will drop the 'USD'-prefix and only refer to the data as swap-, LIBOR- and OIS data. As a reminder, all the rate data is provided as annual rates, which is a standard in fixed income.

In practice, I first download daily data for swaps, LIBOR rates and OIS rates starting from 01-01-2007. Then I filter out dates, in which there are missing observations in any rate maturity. After that, I merge the 3 separate datasets[1] by date and choose the last dates of every month. In this way, my merged dataset has no missing month from 09/2011 to 12/2021, i.e. I have data for a total of 124 months. However, I will start my main trading period from 01/2012 for the calibration purposes[2] and will therefore have 120 months of data in the main trading period. If I had just downloaded monthly data for the 3 separate datasets and then merged these by dates, I would have had missing months in my merged dataset. Finally, I split my merged dataset back into separate datasets for swap, LIBOR and OIS rates.

As mentioned at the beginning of this section, I also use an expanded period of 01/2007-12/2021 for modelling and calibrating the swap rates. The process for preparing this dataset is quite similar as explained above for the merged dataset. I take the daily swap data from 01-01-2007 to 31-08-2011 and filter out dates with missing observations. Then, I choose the last dates of months and combine this "pre-trading" data to the monthly swap data from 09/2011 forwards that was splitted from the merged dataset. This way, I obtain modelling swap data for every month from 01/2007 to 12/2021, i.e. I have modelling swap data for a total of 180 months.

I also need data for my multifactor regression and hedge fund analysis but I present these in chapter 8. In the Appendices section 11.1 I refer to the GitHub repository that contains all the data and the Jupyter Notebooks used for data processing.

[1]Swap, LIBOR and OIS

[2]More about this in section 6.2

5.1 Data analysis

In this subsection I briefly plot and describe the data.

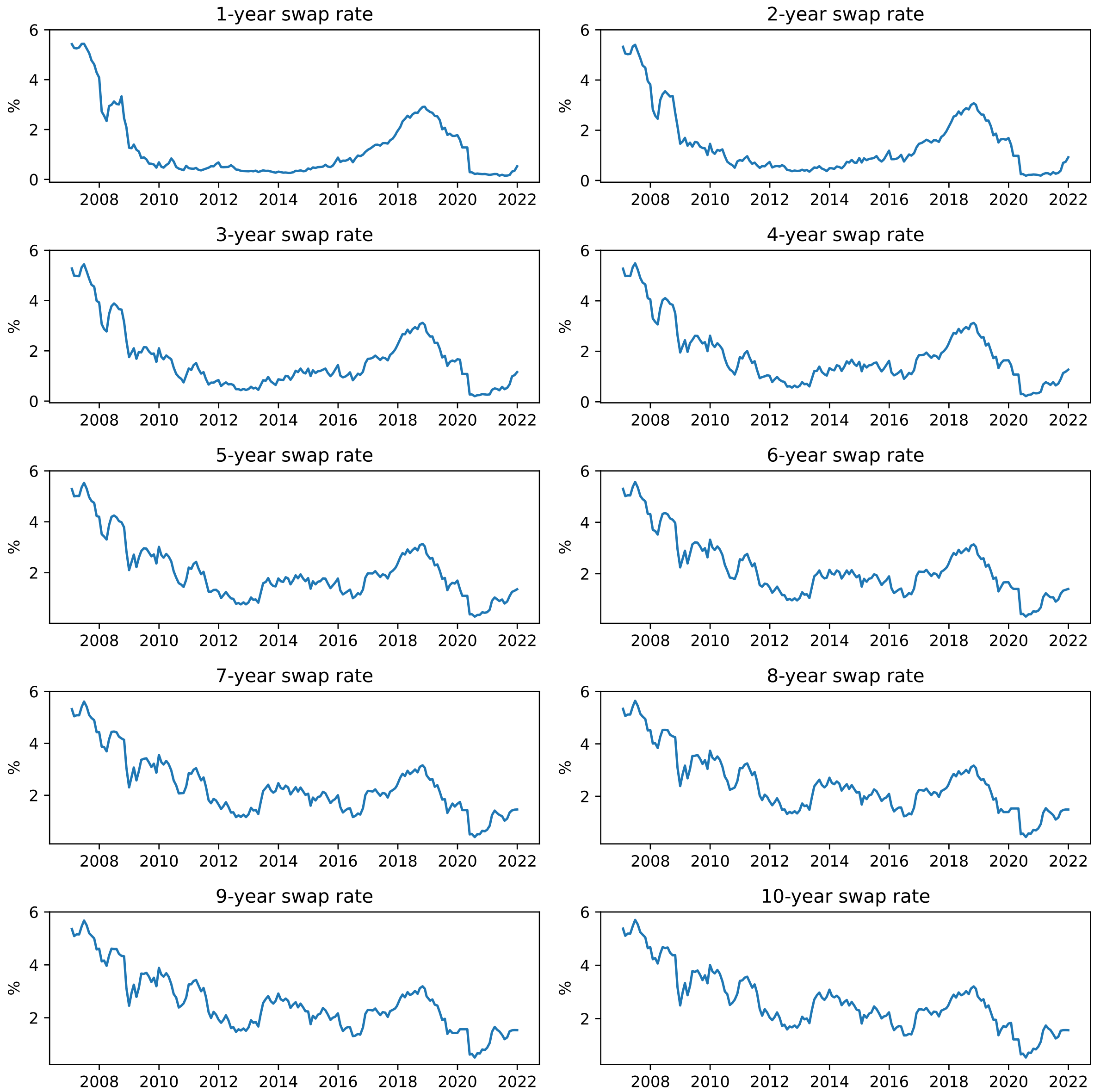

Figure 1: USD swap rates in 01/2007-12/2021

Table 1: Descriptive statistics for USD swap rates in 01/2007-12/2021 (%)

Maturity	Count	Mean	Std	Min	25%	50%	75%	Max
1-year	180	1.276	1.314	0.150	0.362	0.640	1.778	5.440
2-year	180	1.436	1.236	0.189	0.562	0.969	1.744	5.404
3-year	180	1.648	1.195	0.195	0.831	1.249	2.062	5.444
4-year	180	1.850	1.164	0.223	1.080	1.535	2.322	5.488
5-year	180	2.034	1.141	0.273	1.251	1.774	2.614	5.531
6-year	180	2.202	1.121	0.331	1.411	1.976	2.771	5.572
7-year	180	2.340	1.112	0.390	1.502	2.139	2.886	5.611
8-year	180	2.451	1.110	0.445	1.571	2.259	3.037	5.645
9-year	180	2.548	1.108	0.495	1.653	2.362	3.119	5.677
10-year	180	2.631	1.108	0.540	1.812	2.409	3.165	5.703

Figure [1] shows that all swap rates started from quite high level in 2007, as the financial crisis was ongoing at that time. After that, the rates came down in all maturities. The rates spiked again to around 3 % in 2018-2019. Rising to this level started in short maturity swap rates, such as 1-3 year swap rates, as early as in 2014. After 2019-2020 the rates have come down, but they are still quite close to each other. In early 2010s the short maturity rates were low and long maturity rates were high, as could be expected.

The latest spike in the rates is explained by the actions of Federal Reserve (the Fed), as the Fed raise interest rates in 2018, causing the short and long rates to level off to the extent that in 2019 the rate curves inverted. The pandemic caused the rates to drop and later the rates have risen again. Swap spreads were negative in the US in around 2019, so the Treasury yields were higher than the swap rates of corresponding maturity.

The descriptive statistics in Table [1] present quite expected results. One thing to notice is that the standard deviation of short maturity swap rates has been higher than the standard deviation of long maturity swap rates. This could be explained by their liquidity, high trading and thus high volatility. In high and low levels, the swap rates seem to stay close to each other, as described by the minimum and maximum of the swap rates.

The OIS rates and LIBOR rates present similar pattern as the swap rates. All rates spiked around 2019 and short and long maturities converged in around 2.5 %. It is

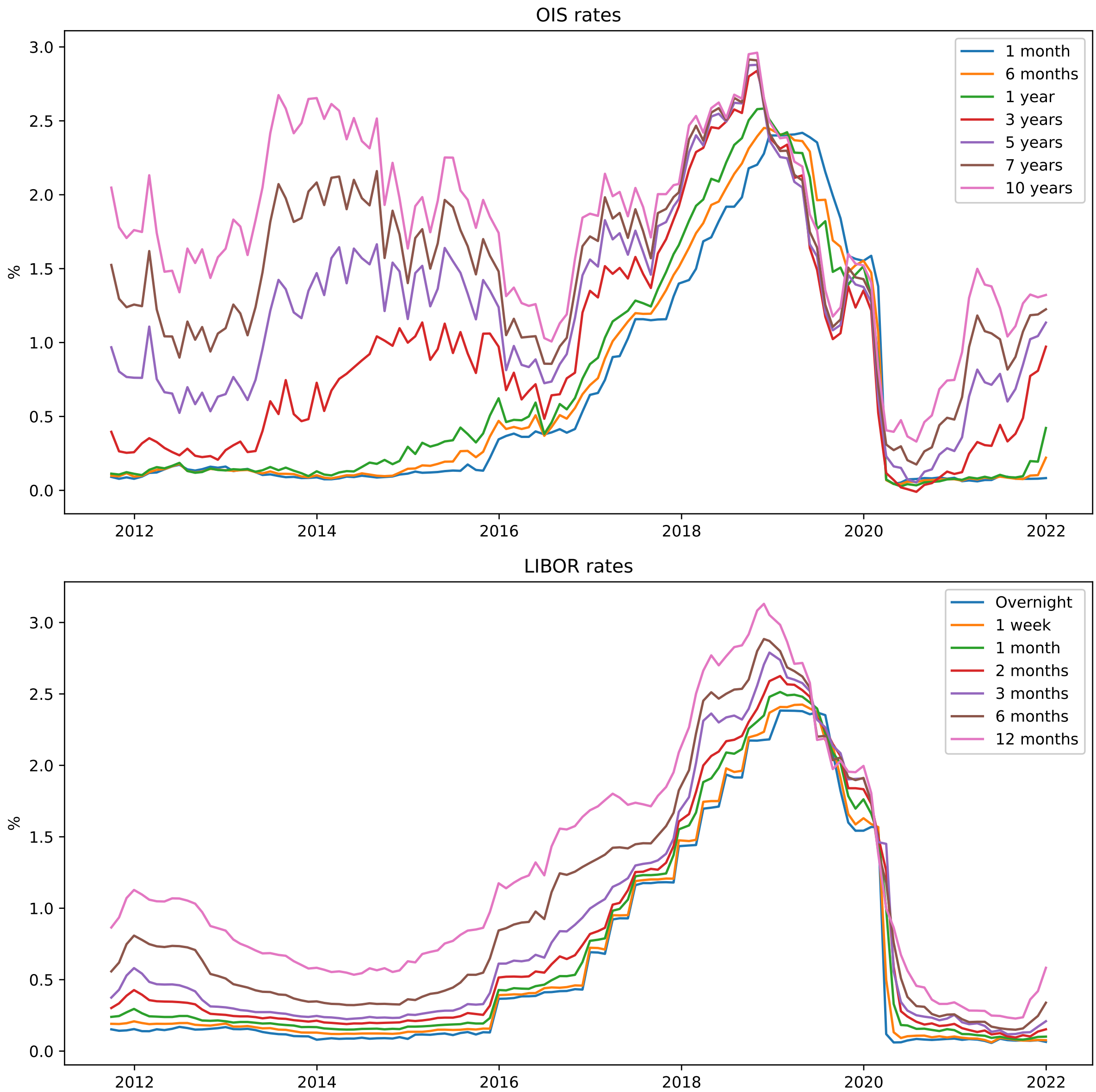

Figure 2: Chosen USD OIS rates and all USD LIBOR rates in 09/2011 - 12/2021

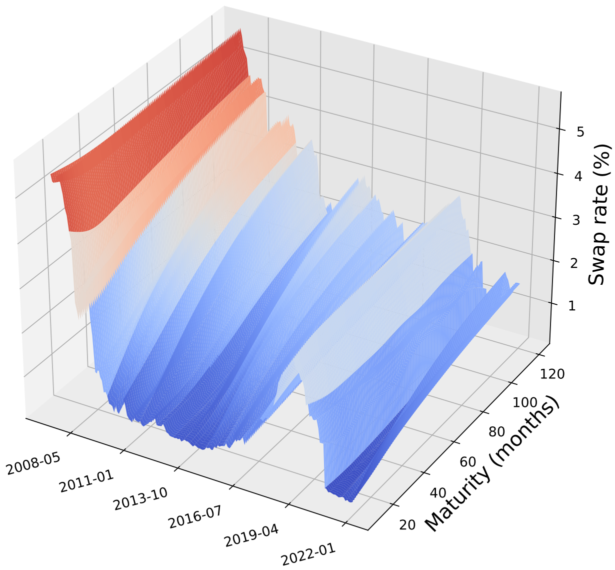

Figure 3: USD swap rate curves with cubic interpolation in 01/2007 - 12/2021

somewhat bizarre that all the rates; swap, OIS and LIBOR have spiked and converged in around 2.5 %. For instance in OIS rates, the overnight rates and 10-year rates have been very close in around 2019, whereas in the early 2010s there were big spreads between the short and long maturities. The spreads of short and long maturities in both OIS rates and LIBOR rates have somewhat returned after 2020, but as in swap rates, the rates are still quite close to each other. The explanations about Fed monetary policy in 2018 and the pandemic afterwards, apply for OIS and LIBOR rates as well.

Figures [3], [4] and [5] present cubic interpolated swap, LIBOR and OIS rate curves. The maturities are presented in months, since the interval in the interpolations is 1 month. The thing to notice is that swap and OIS rate curves seem to be upward-sloping most of the time, whereas LIBOR curves are often quite flat. LIBOR rates are for maturities less or equal to 1 year, whereas the OIS rates span up to maturity of 10 years. Therefore it is quite natural, that the OIS curves seem more upward-sloping in the figures than the LIBOR curves. The figures are not of best quality, but I would roughly say, that there has not been periods of inverted rate curves outside the previously mentioned 2018-2019 period.

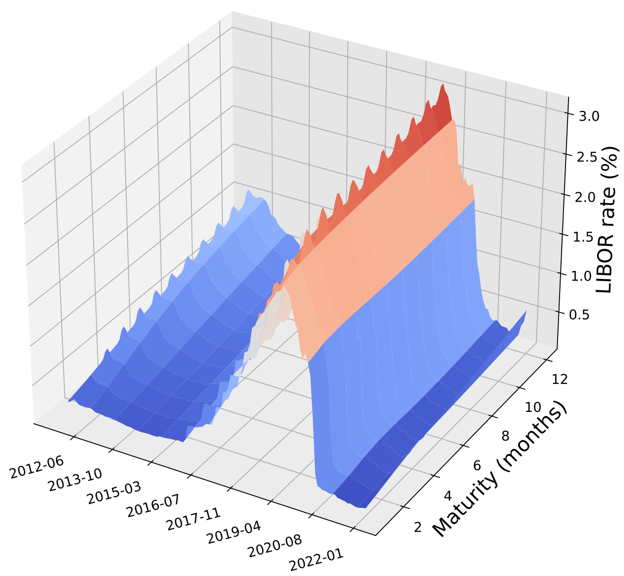

Figure 4: USD LIBOR rate curves with cubic interpolation in 09/2011 - 12/2021

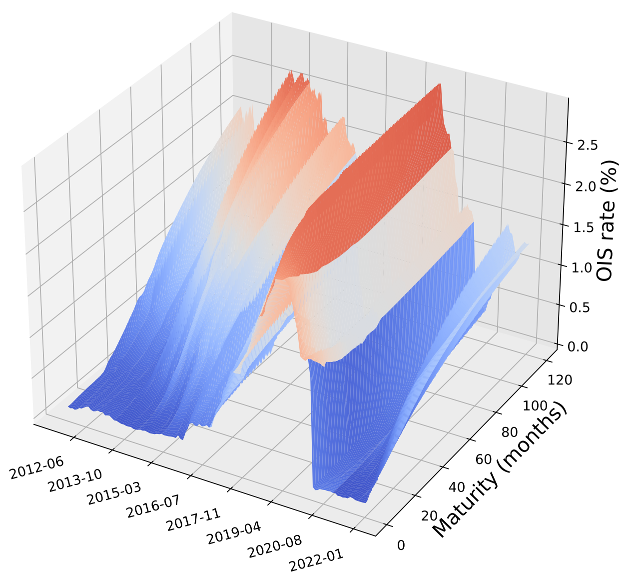

Figure 5: USD OIS rate curves with cubic interpolation in 09/2011 - 12/2021

5.2 Possible issues in data

In this section, I briefly present possible issues in my data.

First of all, my main trading period data consists of only 120 observations, which is quite low number both modelling- and trading-wise. Alternatives would have been to use weekly or daily data, but these alternatives have problems related to trading frequency and trading costs.

Another possible problem is that I combine data from different sources, as I use ICE data for swap and LIBOR rates and Tullett Prebon data for OIS rates. As mentioned before, these datasets have differences in whether the rates are closing rates or mid-day rates. These things should not have that significant impact on the quality of my results, but ideally I would have used the same data source for all the rates with same specifications for the rates.

Finally, perhaps the biggest possible issue is that my study period happens to be in a time period, when all the rates were first impacted by the financial crisis and then stayed very low for long periods of time. At the end of my study period, the rates were strongly affected by monetary policy decisions and the global pandemic. These things likely have an impact on the quality of my results.

6 Methodology

The methodology of my thesis is based on combining the methodologies of my main reference papers Duarte et al. (2007) and Smith (2013). Duarte et al. (2007) explains modelling, trading and hedging of the swap rates and applies basic LIBOR discounting in swap valuation, whereas Smith (2013) explains the swap valuation based on OIS discounting. I also want to acknowledge Stark (2020) that is very well-written and makes justified and clear choices regarding methodology. I ended up doing similar choices as Stark (2020) in many parts but I also made some major changes here and there. The notation in this section is a combination of my own notation and notations of Smith (2013) and Brigo and Mercurio (2007). Hirsa (2013) also provides a good reference on modelling the swap rates. Before moving into the methodology, I will give some background on swap mechanics.

6.1 Background on swap mechanics

By now, it is necessary to provide the definition for fixed-for-floating ("vanilla") interest rate swap that I apply throughout the thesis. Again, this definition is not formal but rather my own summary. The definition does not involve all the technical swap practicalities that I will later deal with in section 6.

Definition 6.1 (Fixed-for-floating ("vanilla") interest rate swap)**.** In a fixed-for-floating interest swap, there are two parties, A and B. Party A pays the fixed swap rate (fixed leg) to party B and party B pays the floating rate (floating leg) to party A. Party A is therefore the fixed rate payer and holds a payer swap and party B is the fixed rate receiver and holds a receiver swap. The floating leg rate is paid and reset according to the maturity of the floating leg rate, e.g. every 3 months. The fixed leg rate is fixed throughout the lifetime of the swap and payments for the fixed leg usually occur semiannually. Both the fixed leg and floating leg are paid with respect to some nominal that is not exchanged at any stage of the swap. At initiation of the swap, the fixed leg swap rate is set such that the present value of future fixed leg payments by party A is equal to the present value of the expected future floating leg payments by party B. In this case, the swap is said to be valued at par as both legs have the same present value and the fixed leg swap rates are referred to as par swap rates. The present value of the floating leg payments is equal to the nominal of the swap, if the floating leg rate is also used for discounting the future cash flows. Forward rate projections for the floating leg rate are needed for the expected future floating leg payments. The forward rates change during the lifetime of the swap and therefore the two legs of the swap may not be equally valued after the initiation.

An additional necessary assumption about fixed-for-floating interest rate swaps is that I assume that the counterparties place some initial margin collateral at the beginning of the swap and some collateral rate is paid on the collateral. I present the justification for this in the Appendices section 11.4 and in section 6.5.

The definitions for year fractions and related day count conventions are also needed to know how interest accrues over periods. Namely, as the standard is to provide interest rates as annual (per annum) rates, I need to know how the number of days between two dates affects the accrual of interest between the dates:

Definition 6.2 (Year fraction)**.** Consider the current date as date D_0. Consider also date D_i, that is i months before (denoted with negative sign, i.e. $-i$) or after D_0 and date D_j, that is j months before (denoted with negative sign, i.e. $-j$) or after D_0, such that $i < j$. The time measure for being "X months before or after some date" can be defined specifically with day count conventions. Denote by $\tau(i,j)$ the time measure (year fraction) between D_i and D_j and apply a chosen day count convention. The 3 most common day-count conventions are[1]:

Actual/365: With this convention a year is 365 days long and the year fraction between two dates is the actual number of days between them divided by 365. Denoting by $D_j - D_i$ the actual number of days between the two dates, $D_i = (d_i, m_i, y_i)$ excluded and $D_j = (d_j, m_j, y_j)$ included, the year fraction in this case is

$$\tau(i,j) = \frac{D_j - D_i}{365}$$

Actual/360: A year is in this case assumed to be 360 days long. The corresponding year fraction is

$$\tau(i,j) = \frac{D_j - D_i}{360}$$

[1]There is an example related to year fractions in the Appendices section 11.8.

30/360: With this convention, months are assumed to be 30 days long and years are assumed to be 360 days long. The year fraction between D_i and D_j is in this case given by the following formula

$$\tau(i,j) = \frac{max(30-d_i,0)+min(d_j,30)+360*(y_j-y_i)+30*(m_j-m_i-1)}{360}$$

Now I can show how the model spot swap rate at time $t=0$ for maturity T can be solved on the basis of zero coupon bond prices:

$$\widehat{s(T)} = \frac{1-P(0,T)}{\sum_{i=1}^{n}\tau_i P(0,t_i)} \tag{6.1}$$

where t_i are the times of the swap cash flows with $t_n = T$ and τ_i correspond to the year fractions between t_{i-1} and t_i. The zero coupon bond prices are solved with the short rate models presented in section 3.5. The spot swap rates for $t=0$ are the only swap rates I need in this thesis, but one could solve forward-start swap rates for $t>0$ by modifying the above formula. Brigo and Mercurio (2007) derive the above formula by considering a fixed-for-floating swap as a combination of floating rate note and a coupon bearing bond. The formula assumes that the floating rate note is valued at par and at nominal and that the floating leg and fixed leg are paid simultaneously. The formula also assumes that there is only one set of discount factors provided by the zero coupon bond prices. I also assume in solving my model swap rates that both of the legs are paid simultaneously and that the modelled zero coupon bond prices provide the correct discount factors but in my actual swap valuations I apply the real payment schedules. The formula clearly shows that once I am able to model zero coupon bond prices with short rate models, I automatically obtain the model swap rates as well.

The floating leg rate in interest rate swaps is commonly an interbank offered rate (IBOR). The floating leg rate is 3-month USD LIBOR and the fixed leg swap rate is paid semiannually in the swaps that I apply. Therefore, the floating leg and the fixed leg have different payment schedules. The two legs also have different day count conventions, as the floating leg uses Actual/360 convention and the fixed leg uses 30/360 convention. In practice, I need to deal with all the year fractions and payments separately, when I valuate the swaps. I will deal with the actual valuation of swaps in section 6.4.

The floating leg rate in interest rate swap may also be based on an overnight rate, in which case the swap is referred to as an overnight indexed swap (OIS). The following definition is not formal, but rather my own explanation for an OIS.

Definition 6.3 (Overnight indexed swap (OIS))**.** An OIS is a swap contract, in which the floating leg payments are calculated by compounding given overnight rates daily for a given period, that is based on the payment period of the fixed leg of the OIS. Common overnight rates are Federal Funds rate in USD OIS and Euro short-term rate (ESTR) in EUR OIS. Usually, the fixed leg payment occurs once at the end of the OIS maturity, if the OIS has maturity below 1 year. For maturities over 1 year, the fixed and floating leg payments usually occur in 12-month periods, apart from the possible "stub" period[1] at the beginning, if the OIS contract does not have full-year maturity. Most of the OIS contracts have full-year maturities. At the end of the given compounding period, the payment of the floating leg is based on the compounded value of the daily overnight rates. Fixed leg rate in an OIS is set at initiation such that the present value of fixed leg is equal to the projected present value of floating leg, just like in any other fixed-for-floating interest rate swap.

In traditional swap valuation, the cash flows of fixed leg and floating leg are discounted with the same rate that is used in the floating leg of the swap and the forward rates for the floating leg are bootstrapped based on these discount factors. Smith (2013) presents OIS discounting framework, where OIS-based discount factors are bootstrapped separately and then these OIS-based discount factors are used together with market LIBOR rates and with swap rates that have LIBOR as floating leg rate in bootstrapping OIS-consistent implied LIBOR forward rates. With these forward rates, the floating leg of a fixed-for-floating interest rate swap does not have present value equal to nominal at initiation. However, the fixed and floating leg of the swap should still have same present values, i.e. the swap should be valued at par. I will provide the explicit formulas for these discount factors and forward rate projections in subsection 6.4.

According to Smith (2013) and Hull (2017), the change to OIS discounting has occurred

[1]Stub periods normally occur because the interval between payments does not fit neatly into the maturity for which the instrument was issued, thus sometimes the instrument's final or first coupon period may be adjusted to make the instrument start and mature on the desired dates.

in recent years. The financial crisis of 2007-2009 left some major impacts in how interest rate swaps are valuated and traded these days. The crisis changed the conception of risk-free rates, such that interbank offered rates (e.g. USD LIBOR) are not considered risk-free anymore. During and after the crisis, the LIBOR-OIS spread jumped significantly, as the liquidity and credit risk related to the LIBOR quoting banks changed dramatically. On the other hand, in recent years new regulation on how swaps are traded through central clearing counterparties and how banks need to place collateral on their swap trades have made swaps practically risk-free[1]. As LIBOR rates are not viewed as risk-free rates anymore, OIS rates now represent risk-free rates, since the floating leg of OIS is based on risk-free overnight rates. Therefore, discount factors based on OIS rates and OIS-consistent implied LIBOR forward rates are used in the valuation of collateralized interest rate swaps. In other words, as LIBOR rates are used in the floating leg of interest rate swaps, the swap rates reflect the riskiness of the LIBOR rates, but swaps themselves are virtually risk-free. Therefore, it is justified to use OIS rates in valuation of swaps. OIS-based discount factors are used rather than e.g. discount factors based on Treasury yields, because OIS rates have the same liquidity, tax status and volatility as the interest rate swaps. This is obviously because OIS contracts are interest rate swap contracts themselves as well.

6.2 Modelling

In this section, I explain how I apply the term structure models presented in subsection 3.5. Namely, I explain how I calibrate the models to the market data and how I locate the possible mispricings of market swap rates. I will divide my data to in-sample and out-of-sample periods, such that I apply rolling in-sample periods for model calibration (calibration period) and rolling out-of-sample periods for testing the models[2]. This is contrary to Duarte et al. (2007), that use the same in-sample period for both model calibration and model testing[3].

In applying the term structure models, everything starts with calibrating the models

[1] More about this in the Appendices section 11.4.

[2] My rolling in-sample period is 60 months and the rolling out-of-sample period is 1 month.

[3] By the definition of Eurostat (2022), statistical tests of a model's forecast performance are commonly conducted by splitting a given data set into an in-sample period, used for the initial parameter estimation and model selection, and an out-of-sample period, used to evaluate forecasting performance.

to the market swap data. For this, I need the closed form expressions for zero coupon bond prices and spot swap rates that were presented in previous sections. To refresh memory, let me consider as an example, that I would want to calibrate the two-factor Vasicek model in a given month to the market data[1]. I apply the following familiar formulas:

$$
\begin{aligned}
A(t,T,k,\theta,\sigma) &= exp\Big\{(\theta - \frac{\sigma^2}{2k^2})(B(t,T,k) - T + t) - \frac{\sigma^2}{4k}B(t,T,k)^2\Big\} \\
B(t,T,k) &= \frac{1}{k}\Big[1 - e^{-k(T-t)}\Big] \\
P(t,T) &= A(t,T,k_1,\theta_1,\sigma_1)A(t,T,k_2,\theta_2,\sigma_2)exp\Big\{ - B(t,T,k_1)x_t - B(t,T,k_2)y_t\Big\} \\
\widehat{s(T)} &= \frac{1 - P(0,T)}{\sum_{i=1}^{n} \tau_i P(0,t_i)}
\end{aligned}
\tag{6.2}
$$

In applying this model, it is integral how I set the 6 parametes $(k_1, \theta_1, \sigma_1, k_2, \theta_2, \sigma_2)$ and the factors x_t and y_t in each month. This is what calibration is all about. I only need to consider the spot prices and spot swap rates, i.e. current time with $t = 0$, and therefore I am only interested in factor values x_0 and y_0 both in the in-sample calibration months and in the out-of-sample month.

First, I take the market swap data for 60 previous months (in-sample calibration period) before the current month (t=0, out-of-sample). The calibration period length can be adjusted and it is wise to test different calibration period lengths. To keep things simple, I applied as long calibration period as possible given the limits of my data. Once I have my calibration period data, I will choose the initial guess parameters for the 6 parameters. I will use the optimal parameters provided by Duarte et al. (2007) as an initial guess. As Duarte et al. (2007) uses a different structure for the two-factor Vasicek model than Brigo and Mercurio (2007), I need to derive the parameters in the model form of Brigo and Mercurio (2007) as:

[1] The process for CIR calibration is similar with differences in initial parameters, constraints for parameters and closed form expressions as presented in 3.5.

$$\begin{aligned}
k_1 &= \beta = 0.0113727 \\
\theta_1 &= \frac{\alpha}{\beta} = \frac{0.0009503}{0.0113727} = 0.083559753 \\
\sigma_1 &= 0.0548290 \\
k_2 &= \gamma = 0.4628664 \\
\theta_2 &= \frac{\mu}{\gamma} = \frac{0.0240306}{0.4628664} = 0.051916925 \\
\sigma_2 &= \eta = 0.0257381
\end{aligned} \tag{6.3}$$

As the two-factor CIR model applies same parameters as the two-factor Vasicek model, I also apply these same parameter values for the two-factor CIR model with the exception that I multiply k_1 by 1.6, so that the $2k_1\theta_1 > \sigma_1^2$ constraint of the two-factor CIR model is initially satisfied.

I will apply these parameters as initial guesses, when I move on with rolling calibration, i.e. I do not update the initial guesses. Once I have the initial parameters, I set the factors x_0 and y_0 in each month of the calibration period such that the 1- and 10-year swap rates are perfectly modelled with respect to the market 1-year and 10-year swap rates in the given month. As I do this separately for every month in the calibration period, I end up with different x_0 and y_0 values for different months in the calibration period. At this stage, I have model swap rates for maturities of 1 to 10 years for every month in the calibration period, as the 6 parameters and factors x_0 and y_0 are all set.

After this, I move on to iterating through the 6 parameters, such that the model swap rates fit the market swap rates in maturities of 2 to 9 years as well as possible in the calibration period. For this data fitting, Duarte et al. (2007) applies sum of squared differences as a measure of goodness of data fit. I apply the mean squared error, because it has a ready and robust implementation in Python but the results should be the same. The mean squared error in maturities of 2 to 9 years in an individual month can be expressed as:

$$\frac{1}{8}\sum_{T=2}^{9}(\widehat{s(T)} - s(T))^2 \tag{6.4}$$

$\widehat{s(T)}$ represents the T-year model swap rates and $s(T)$ represents the T-year market swap rates. I do not need to use the 1- and 10-year swap rates, as they are perfectly made to fit market values in the calibration period with the chosen x_0 and y_0 in the given month in calibration period. Notice, that the formula above is just for one month to give an example, but in practice I take the mean squared error over the whole calibration period. With 60-month calibration period I would be summing $60 * 8 = 480$ squared differences in total and taking the average of these 480 squared differences.

The parameters for Vasicek and CIR models need to be positive[1]. CIR also has the constraints related to the parameters, i.e. $2k_1\theta_1 > \sigma_1^2$ and $2k_2\theta_2 > \sigma_2^2$. Additionally, I bound the long-term mean θ and instant standard deviation σ parameters to range $(0, 0.15]$ but I leave the mean-reversion parameter k unbounded from above. This additional bounding of the parameters is not theoretically required, but I think it makes economically sense in real-life. As I will shortly explain, I am applying local optimization methods for which the bounding of parameters can make an important impact. I chose the 0.15 upper bounds based on the optimized parameter values of Duarte et al. (2007) and based on studying the time-series values of overnight LIBOR and 1 week LIBOR[2].

For finding the minimum of the mean squared errors, I am solving constrained nonlinear least squares problem, i.e. a constrained nonlinear optimization problem for finding the minimum. I apply the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS-B) algorithm for Vasicek calibration, as the algorithm provides the possibility to set bound constraints on the variables. For the CIR calibration, I apply the constrained optimization by linear approximation (COBYLA) algorithm[3].

Once the optimized parameters are found, the 1- and 10-year model swap rates will no longer match exactly the market swap rates in the calibration period, as the factors x_0 and y_0 were in the initial stage set with respect to initial parameter guesses. Therefore, if I want to solve the model swap rates in the calibration period, I now need to set x_0 and y_0 factors again such that the 1- and 10-year model swap rates match the market

[1]It is somewhat unclear whether Duarte et al. (2007) apply this constraint with Vasicek model, but by the model construction this constraint should be applied.

[2]This analysis happens in the objective (real) world and thus is not directly applicable to the risk-neutral world. However, I only need rough estimates for the parameter bounds and thus resorted to the objective world estimation.

[3]Both the L-BFGS-B and COBYLA algorithms have ready implementations in Python Scipy package. See the documentation of the algorithms in https://docs.scipy.org/doc/scipy/reference/optimize.html. More about the optimization aspect in calibration in the next subsection 6.2.1.

swap rates with the optimized parameters. This provides me the final model swap rates in the calibration period. However, I do not actually need the model swap rates in the calibration period. I only need the optimized parameters, that I just got as the result of the optimization of the constrained nonlinear least squares problem. I will apply these parameters now in my rolling out-of-sample period, i.e. the current month ($t = 0$). As the last step of calibration, I use the optimized parameters and set factors x_0 and y_0 such that my model swap rates match the 1- and 10-year market swap rates at $t = 0$, i.e. I do the same step that I did as first step in calibration but now only in the out-of-sample month. The way that I set the factors x_0 and y_0 relative to the 1- and 10-year swap rates demonstrates how these rates are considered the benchmark rates and how the modelling and the relative values of the 2-9 year swap rates are linked to the 1- and 10-year swap rates.

Now I have the model swap rates and market swap rates for all annual maturities from 1 to 10-years in the current out-of-sample month ($t = 0$). The 1- and 10-year model swap rates match exactly the market swap rates but the other maturities will not necessarily match. Therefore, there are potential mispricings in 2- to 9-year maturities. I will explain in the next section how to deal with these mispricings and to trade them.

This same iteration is repeated each month, i.e. the optimal parameters are not solved for the full sample or one specific subsample, but rather in a rolling manner. This rolling implementation of calibration avoids the problem of look-ahead bias and makes the arbitrage implementation more realistic.

6.2.1 Optimization

Robust calibration of the model rates to the market rates is more or less the most important step in implementing yield curve arbitrage. It is also perhaps the most difficult step and I faced many problems with this step. Duarte et al. (2007) do not specify the method, that they use for optimizing the nonlinear least squares problem. They only state that they obtain global minimum as a result of the optimization. This a major difficulty in implementing a replication of their study and may affect results quite significantly if an optimization method that does not provide global minimum would be used.

Finding the global optimum for an optimization problem is in no way a trivial issue. In general, the global optimum for optimization problems that have convex objective function over convex (constrained) set, can be found with any optimization method that finds the local optimum because in this case the local optimum is guaranteed to be global optimum. If the problem is not convex, then specific global optimization methods, such as spatial branch-and-bound methods, need to be used to find the global optimum. Most optimization methods are local optimization methods and as the convexity of an optimization problem can be difficult to prove, the global optimum can be really difficult to obtain. It is possible to obtain the global optimum for a non-convex problem with a local optimization method by finding the local optimum that happens to be also the global optimum but there is no guarantee of finding the global optimum in this case.[1]

Neither the L-BFGS-B algorithm or COBYLA algorithm that I apply are global optimization methods. Therefore, these methods are not guaranteed to provide global optimum but with well chosen initial values for the model parameters they can find a local optimum that is the global optimum. I was unable to solve the issues with global optimization and to apply global optimization method. This is a major flaw in my implementation and is likely to affect my results quite a lot. I obtain local optimum results with the above-mentioned methods but my results are not guaranteed to be global optimum.

Another possible issue related to calibration may be the length of my calibration period. My first calibration periods are also timed to the aftermath of the financial crisis, i.e. the quality of my calibration data is questionable, which likely affects my results, as I will show in chapter 7.

6.3 Trading

In this section, I explain how I open and close arbitrage trades.

Definition 6.4 (An arbitrage trade)**.** An arbitrage trade consists of 3 separate trades: a main trade for mispricing in 2-9 year swap rates and two hedge trades with 1- and 10-year swaps.

In subsection 6.2 I explained how the modelling of the swap rates moves forward in each month of market data and how possible mispricings are observed in 2 to 9-year swap

[1]My main reference for optimization related issues is the course material by Associate Professor Fabricio Oliveira in course MS-E2122 Nonlinear Optimization in Aalto University.

rates. The mispricings point out trading opportunities for arbitrage, if they are large enough. The idea in observing trading opportunities for arbitrage follows from the view that the term structure models model the swap rates correctly and the market values are incorrect. Therefore, the market should adjust to be in accordance with model values.

Definition 6.5. (Mispricing in T-year swap rate)
$\widehat{s(T)} - s(T)$, where $\widehat{s(T)}$ is the model swap rate and $s(T)$ is the market swap rate.

Example 6.1. Consider, that in a given month I have mispricing in 3-year swap rate, such that the mispricing is positive. In this case, I would enter into an arbitrage trade, in which I open payer swap in the 3-year swap (main trade) and receiver swaps in 1- and 10-year swaps (hedge trades). The view here is that the market 3-year swap rate would rise to be in the same level with the model 3-year swap rate in which case it would be profitable to have previously opened a payer swap. This is because the floating leg payments follow the market movements, whereas the fixed leg swap rate payments remain fixed throughout the lifetime of the swap. In other words, the initial market projections for forward rates are too low and the 3-year (mispriced) swap rate is set in line with the incorrect forward rate projections.

The idea in solving the hedging weights for the 1- and 10-year swaps stems from the idea, that the main trade should be hedged against the factor risk caused by the two factors of the term structure models, as explained in chapter 3. Following the methodology of Duarte et al. (2007), I solve the hedging weights analytically for T-year maturity swap by the following system of equations:

$$\begin{bmatrix} \frac{\partial \widehat{s(T)}}{\partial x_0} \\ \frac{\partial \widehat{s(T)}}{\partial y_0} \end{bmatrix} = \begin{bmatrix} \frac{\partial \widehat{s(1)}}{\partial x_0} & \frac{\partial \widehat{s(10)}}{\partial x_0} \\ \frac{\partial \widehat{s(1)}}{\partial y_0} & \frac{\partial \widehat{s(10)}}{\partial y_0} \end{bmatrix} \begin{bmatrix} w_1 \\ w_{10} \end{bmatrix} \tag{6.5}$$

w_t denotes the hedging weight. The equations basically imply that the sensitivity of the main trade to e.g. factor x_0 should equal the sum of sensitivities of hedge trades to the factor x_0. As I take the main trade and hedge trades in opposite directions[1], the

[1]Receiver swap in main trade and payer swaps in hedge trades, or vice versa.

sensitivities of the main trade and hedge trades should cancel each other. The partial derivatives can be solved quite simply analytically by applying the formula for spot swap rates with the chosen short rate model. For the two-factor Vasicek model, the partial derivative with respect to factor x_0 is solved as:

$$\frac{\partial \widehat{s(T)}}{\partial x_0} = \frac{B(0,T,k_1)P(0,T)\sum_{i=1}^{n}\tau_i P(0,t_i) + (1-P(0,T))\sum_{i=1}^{n}\tau_i B(0,t_i,k_1)P(0,t_i)}{(\sum_{i=1}^{n}\tau_i P(0,t_i))^2} \tag{6.6}$$

Correspondingly, the partial derivatives with respect to factor y_0 are obtained by replacing the parameter k_1 with k_2 in the formula above. It turns out that these same formulas for partial derivatives apply for the two-factor CIR model as well, with only the difference that corresponding A, B and P functions for the two-factor CIR model should be applied.

Notice, that receiver and payer swaps are used for implementing the arbitrage. Duarte et al. (2007) talks about taking long and short positions in swap rates but in practice I am always long in either receiver or payer swaps, i.e. I can not short the swap rates in a traditional sense. I take nominal position of \$100 dollars in the main trade and the nominals for the hedging trades are given by the hedge weights.

Example 6.2. Consider, that I would obtain $w_1 = 0.75$ and $w_2 = 1.2$ for the hedging weights. Then, I would have a nominal of \$100 in the main trade, nominal of $\$100 * 0.75 = \$75$ in 1-year swap and nominal of $\$100 * 1.2 = \$120$ in 10-year swap.

For every month of market data, I will open at most 1 arbitrage trade with main trade that corresponds to the largest mispricing observed among the 2-9 year swap rates. The mispricing must also exceed given basis point limits, i.e. the basis point difference of the model rate and the market rate. I will test separately trading for 5 basis point limit and 20 basis point limit[1]. If there are no mispricings that exceed the chosen basis point limit in a given month, I will not trade in that month but I may close my open arbitrage trades.

Even if I open at most 1 arbitrage trade each month that does not mean that there could not be more than 1 arbitrage trade open each month. The arbitrage trades are

[1]Duarte et al. (2007) tests only 10 basis point limit but I obtain so similar results with 5 and 10 basis point limits that I decided to leave 10 basis point limit out of final results.

closed, when either 12 months has passed or the market swap rate for the original maturity of the main trade deviates from the original model swap rate by less than 1 basis points or the sign (+/-) of the mispricing has changed[1].

There are no closing rules for the hedge trades, so the arbitrage trades are only closed when either of the aforementioned conditions apply for the main trade. However, in closing the arbitrage trades, I assume 1 basis point closing cost separately for each trade that the arbitrage trade consists of, i.e. the main trade and the hedge trades[2]. I do not need to apply the 1 basis point closing cost to the 1-year hedge trade, if the arbitrage trade is closed because 12 months has passed.

The closing of trades can be assumed to be made by asking for the closing price from the swap counterparty and receiving (positive value) or paying (negative value) this price if the counterparty agrees to the closing of the swap. Alternatively, I could try opening opposite swaps with maturities close to the remaining maturities of my main trade swap and hedge trade swaps, which would zero out my positions. Lastly if neither of the previously mentioned options is suitable, I could try to sell the swaps if the swap counterparty agrees to this. All these methods for closing the position may turn out to be somewhat unrealistic, as I will explain in the next subsection.

The trading can be implemented in different variations with respect to the mispricing limits and number of mispricings to be traded. I could e.g. trade more than just the largest observed mispricing in a given month by taking positions in all the observed mispricings that exceed the given mispricing limit. Duarte et al. (2007) reports results for trading separately the different swap maturities and opening a trade if there is a mispricing in the given swap maturity that exceeds their mispricing limit of 10 basis point. I follow all the different swap maturities simultaneously and open a trade in the largest mispricing among them if the mispricing exceeds the mispricing limit. Therefore, there is a difference in trading methodology between Duarte et al. (2007) and my thesis. Duarte et al. (2007) also reports the equally weighted portfolio performance in which case they would be

[1]There is an example related to the mispricings in the Appendices section 11.8.

[2]Duarte et al. (2007) applies this 1 basis point closing cost ("transaction cost") because of the bid-ask spread for swaps. It is somewhat unclear what they specifically mean by bid-ask spread, i.e. do they mean selling the open swaps or something else. It is also unclear whether Duarte et al. (2007) applies the 1 basis point closing cost separately to the hedge trades or as a whole to the arbitrage trade. I find it reasonable to apply the 1 basis point closing cost separately to the main trade and the hedge trades.

potentially opening positions in multiple mispricings each month. Trading costs and rules for closing the trades could also be adjusted but I apply the methodology of Duarte et al. (2007) in these things.

6.4 Valuation

In this section, I explain how the open trades are mark-to-market valuated based on OIS discounting methodology of Smith (2013).

The swap counterparty may not agree to the closing of the swap if that possibility is not included in the swap contract. There is also no exchange for the swaps, so it may be unrealistic to be able to sell open swap positions at a given date. Similarly, closing the swaps by taking opposite positions could also prove to be difficult, as the maturities of swaps may not match perfectly. Therefore, mark-to-market valuation is a little misleading term in a sense, but I will use it anyway to reflect the correct value for the open trades.

As interest rate swaps are OTC instruments, the valuation of an interest rate swap can be quite difficult. In theory, the valuation is nothing more than a long/short combination of a fixed coupon bond and a floating rate note as explained in Smith (2013). However, in practice the valuation requires many assumptions related to e.g. applied data and swap practicalities, some of which may turn out to be unrealistic.

To implement OIS discounting in swap valuation, I need to apply my data on OIS rates, swap rates and LIBOR rates. Starting with the OIS data, I solve the OIS discount factors for the monthly maturities n, for which I have data. m denotes the time to first payment in the OIS with maturity of over one year[1]:

$$DF_n^{OIS} = \frac{1}{1 + OIS_n * \tau_{0,n}}, n \in \{1, 2, ..., 11, 12\} \tag{6.7}$$

[1]As mentioned earlier in subsection 6.1, an OIS usually has the same payment schedule for both fixed leg and floating leg.

$$DF_n^{OIS} = \frac{1 - (OIS_n * \tau_{0,m} * DF_m + OIS_n * \sum_{j=1}^{(n-m-12)/12} \tau_{j*12+m-12,j*12+m} * DF_{j*12+m}^{OIS})}{1 + OIS_n * \tau_{n-12,n}}$$
$$n \in \{13, 14, ..., 119, 120\}$$
$$m \in \{1, 2, ..., 11, 12\} \tag{6.8}$$

These formulas and the monthly maturities are seen from the current time ($t = 0$) and $\tau_{i,j}$ denote the year fractions, as defined in section 6.1. The formulas are derived from the notion, that the OIS contract should be valued at par at initiation and the value should now equal the nominal, as the OIS rates are also used for discounting, i.e. there is no "2-curve discounting", as explained in section 6.1. Notice, that these formulas are somewhat different from Smith (2013), because I apply different dataset, I use different notation for year fractions and Smith (2013) uses indexing based on quarters but the idea remains the same.

Thus, for OIS contract maturities n below 12 months, I simply divide 1 by the OIS rate of maturity n multiplied by the year fraction from $t = 0$ until the maturity n. For maturities over 12 months, the situation is somewhat more complicated. In maturities over 12 months, there can be an irregular "stub period" for the first payment (m months), after which the payments occur in annual intervals. Therefore, the year fractions correspond to annual intervals and discount factors are iteratively bootstrapped in annual intervals after the possible first stub period payment. Once I have the discount factors for the maturities of my OIS data[1], I get the continuous discount curve with cubic spline interpolation, that has ready implementation in Python[2].

Similarly, I need continuous LIBOR and swap curves for all monthly maturities between the first and last maturities of the corresponding datasets. I apply cubic spline interpolation also for these datasets[3]. Thus, eventually I have OIS discount factors for each month between 1-120 months (i.e. 1 month to 10 years) and I have LIBOR curve for each month between 1-12 months and swap curve for each month between 12-120 months.

Once I have the continuous curves, I can move on to solving the OIS-consistent implied LIBOR forward rates for a swap. At this stage, I "rewrite" the notation and I denote by

[1]Notice that I do not have data for all monthly maturities.

[2]Python Scipy Interpolate.

[3]Cubic spline interpolated LIBOR and swap curves were plotted in chapter 5.

m the time in months to first floating leg payment of the swap. I derive the first "forward" rate separately knowing:

$$IFR_{0,m}^{OIS} = LIBOR_m \tag{6.9}$$

Therefore, the first "forward" rate is simply the market LIBOR rate for maturity of m months. At initiation and at floating leg payment/rate reset dates this is 3 months but in between it can also be 1 or 2 months. This first forward rate is not needed cash flow-wise in practice for an ongoing swap that is between the floating leg payments but it is necessary in deriving the rest of the forward rates. The rest of the forward rates for the swap for the 3-month periods that occur before 12 months are solved as follows, denoting the end month of 3-month forward period as n[1]:

$$\begin{aligned}
&IFR_{n-3,n}^{OIS} = \\
&\frac{LIBOR_n * \tau_{0,n} * DF_n^{OIS} - (IFR_{0,m}^{OIS} * \tau_{0,m} * DF_m^{OIS} + \sum_j^{n-3} IFR_{j-3,j}^{OIS} * \tau_{j-3,j} * DF_j^{OIS})}{\tau_{n-3,n} * DF_n^{OIS}} \\
&n \in \{m+3, m+6, ..., 12\} \\
&j \in \{m+3, m+6..., n-3\}
\end{aligned} \tag{6.10}$$

Let me then denote by w the time in months to the first fixed leg payment of the swap. w can be 1-6 months, depending on whether I am between fixed leg payment dates or at initiation of the swap or at payment date. The forward rates for 3-month periods, that occur after 12 months are solved as follows, denoting again by n the end month of the 3-month forward period:

[1] The assumption here is that $m \leq n$.

$$IFR^{OIS}_{n-3,n} =$$
$$\frac{SFR_n * \tau_{0,w} * DF^{OIS}_w + SFR_n * \sum_i^n \tau_{i-6,i} * DF^{OIS}_i}{\tau_{n-3,n} * DF^{OIS}_n} -$$
$$\frac{IFR^{OIS}_{0,m} * \tau_{0,m} * DF^{OIS}_m + \sum_j^{n-3} IFR^{OIS}_{j-3,j} * \tau_{j-3,j} * DF^{OIS}_j}{\tau_{n-3,n} * DF^{OIS}_n} \tag{6.11}$$
$$n \in \{m+12, m+15, m+18, ..., 120\}$$
$$i \in \{w+6, w+12, ..., n-6, n\}$$
$$j \in \{m+3, m+6, ..., n-6, n-3\}$$

The notation above is quite intense, but it should provide a good basis for deriving the forward rates. The formulas are based on Smith (2013) but I modified them to be applicable to any combination of m, n and w. In short, the sum of discounted forwards should equal the discounted LIBOR in maturities less or equal 12 months as presented in formula [6.10]. Similarly, the sum of discounted forwards should equal the sum of discounted swap payments in maturities over 12 months as presented in formula [6.11]. Again, the frequency of payments in fixed leg and floating leg is important, as now in the formulas above I handle the fixed leg payments with semiannual frequency, whereas the floating leg payments occur quarterly. Alternatively, one can modify the rates to have the same payment frequency, by solving the annualized rates, that provide same annual returns[1].

Once I have the continuos OIS discount curve and the OIS-consistent LIBOR forwards, I have everything I need for moving to the actual valuation of swaps. The floating leg of the swap having T months to maturity can be mark-to-market valuated (*clean price*) as follows, denoting nominal as N, mark-to-market value as MV and the "ongoing" LIBOR rate that was fixed $3-m$ months ago as $LIBOR_{-(3-m),m}$:

$$MV^{float}_{T,clean} = N * LIBOR_{-(3-m),m} * \tau_{0,m} * DF^{OIS}_m + N * \sum_j^T IFR^{OIS}_{j-3,j} * \tau_{j-3,j} * DF^{OIS}_j$$
$$j \in \{m+3, m+6, ..., T-3, T\}$$
$$T \in \{m+3, m+6, ..., 117, 120\} \tag{6.12}$$

[1]An example related to deriving the forward rates is presented in the Appendices section 11.8.

In the above formula, I assume that $T > m$. If $T = m$, then I can simplify the above formula by dropping the last sum. The fixed leg of the swap having T months to maturity can be mark-to-market valuated (clean price) as follows:

$$
\begin{aligned}
MV_{T,clean}^{fixed} &= N * SFR * \tau_{0,w} * DF_w^{OIS} + N * SFR * \sum_{i}^{T} \tau_{i-6,i} * DF_i^{OIS} \\
i &\in \{w+6, w+12, ..., T-6, T\} \\
T &\in \{w+6, w+12, ..., 114, 120\}
\end{aligned}
\tag{6.13}
$$

Similar to the floating leg, in the above formula I assume that $T > w$. If $T = w$, then I can simplify the above formula by dropping the last sum. Notice, that I dropped the time index from SFR, because in these formulas SFR refers to the swap rate of the specific swap contract that may have been opened in previous months and the swap rate remains fixed. In other words, I am not using the T-maturity market swap rate and the above formulas for market values can also be used after initiation of the swap. At initiation of the swap, if everything is done correctly, the clean price values of the two legs should be the same or very close to each other, i.e. the swap should be valued at par.

I will now move on to presenting the *dirty price* valuation for an ongoing swap. The dirty price values are applied if I wanted to sell the swap or close the swap by asking the (dirty) closing price from the swap counterparty. For the floating leg of the swap, I need to add the accrued interest based on the latest LIBOR fixing to obtain the dirty price value. The dirty price value corresponds to the true value of the swap and the following return calculations are based on dirty prices. Considering, that the latest LIBOR fixing was $3 - m$ months ago, the dirty price value of the floating leg for the swap having T months to maturity is obtained as follows:

$$
MV_{T,dirty}^{float} = MV_{T,clean}^{float} + N * LIBOR_{-(3-m),0} * \tau_{-(3-m),0} \tag{6.14}
$$

There is no need to discount the accrued interest, as the accrued interest does not occur in the future. For the fixed leg of the swap, the fixed swap rate does not change over the lifetime of the swap. Therefore, to obtain the accrued interest for the fixed leg, I only need to know the time that has passed since the latest payment of the fixed leg. As the payment frequencies are different for the fixed and floating leg, the accrued interests are at times calculated with respect to different latest payment date. Considering that the latest fixed leg payment was $6 - w$ months ago, the dirty price value of the fixed leg of the swap that has T months to maturity is obtained as:

$$MV^{fixed}_{T,dirty} = MV^{fixed}_{T,clean} + N * SFR * \tau_{-(6-w),0} \tag{6.15}$$

The dirty price value of a T-month maturity receiver swap is obtained as:

$$MV^{receiver}_{T,dirty} = MV^{fixed}_{T,dirty} - MV^{float}_{T,dirty} \tag{6.16}$$

The dirty price value of a T-month maturity payer swap is obtained as:

$$MV^{payer}_{T,dirty} = MV^{float}_{T,dirty} - MV^{fixed}_{T,dirty} \tag{6.17}$$

6.5 Returns

In this section, I explain how the ex-closing cost excess returns are calculated for individual open arbitrage trades that include both a main trade and hedge trades. I also consider how to calculate total ex-closing cost excess returns each month from all open arbitrage trades.

Before dwelling into the precise calculation of yield curve arbitrage returns, it is good to think for a while what return in the context of interest rate swaps actually means. Interest rate swaps (IRS)[1] should be valued at par at initiation and investing in an IRS does not require capital, apart from possible margin requirements. In addition, there is no particular marketplace for IRS as they are over-the-counter instruments and as such they do not have quoted market value. In practice, the return from an IRS investment comes from the obtained net payments of paid and received leg during the holding period of the IRS and from the positive/negative mark-to-market value of the IRS that is obtained at the end of the holding period if the IRS would not be held until maturity. The mark-to-market value is obtained by either selling the IRS in OTC-markets, by closing

[1] In this context, I refer to interest rate swaps in general and not just fixed-for-floating "vanilla" swaps.

the IRS with a counter IRS position or by asking for the closing price of the IRS from the IRS counterparty, if this is possible.

The \$-returns from an IRS investment are calculated this way and the \$-returns are what the trader would be interested in. As presented above, the development of IRS rates after the opening of an IRS positions is key in thinking about the \$-returns. The trader would also be interested in common risk metrics of IRS investments, such as DV01, that would tell how much the trader's IRS investment would change in \$-value if the fixed IRS rate or underlying projected forward rates for the remaining maturity of the swap moved by 1 basis point.

As mentioned in subsection 6.1, I assume that all the swaps require posting some initial margin collateral. Therefore, I relate the \$-returns of an IRS to the initial capital required, that means the margin collateral provided either to the central clearing counterparty or financial intermediary[1]. With reference to Gregory (2014), I assume that 5% of the swap nominal is required to be posted as collateral. In practice, the calculation of initial margin requirements is more difficult. However, Gregory (2014) provides a table of initial margins for bilateral non-centrally cleared interest rate swaps for which initial margin requirements vary between 1-4 % of notional. Gregory (2014) states that centrally cleared transactions have in general somewhat higher margin requirements, so I decided to use the rough estimate of 5 % for all the swaps.

Definition 6.6 (Excess return on yield curve arbitrage)**.** Excess return in general is the additional return on an investment over a benchmark return. The benchmark return depends on the situation but common benchmark returns include risk-free rate of return, such as the interest paid on a 3-month government Treasury bill. With yield curve arbitrage, the returns are automatically excess returns, since the strategy only involves investing capital in the form of margin requirements and I assume that risk-free returns are paid to the posted margins in the form of collateral rate. Therefore, there is no alternative benchmark, such as some cash yield, that should be deducted from the returns.

To calculate an ex-closing cost excess return in a month for an individual arbitrage trade, I first mark-to-market valuate (dirty price) the main trade and the hedge trades to have correct mark-to-market values for the swaps at month t. Then, I simply sum up

[1]More information about these concepts in the Appendices section 11.4.

the market values of the main trade and the hedge trades to get the total market value. Then, I compare the total market value to the total market value of the previous month ($t-1$). This difference is then related to the total tied initial margin, that is solved by summing the nominals of the main trade and the hedge trades and multiplying by 5%. The nominal for the main trade is \$100 and the nominals of the hedge trades are solved with the hedging weights and the nominals are time-independent, as explained in section 6.3. The excess return calculation for an open arbitrage trade is presented in formula as:

$$return_t^{arbitrage} = \frac{(MV_{t,dirty}^{main} + MV_{t,dirty}^{h_1} + MV_{t,dirty}^{h_{10}}) - (MV_{t-1,dirty}^{main} + MV_{t-1,dirty}^{h_1} + MV_{t-1,dirty}^{h_{10}})}{0.05 * (N^{main} + N^{h_1} + N^{h_{10}})} \tag{6.18}$$

As there can be multiple arbitrage trades open in a given month, the total excess return of the month is solved by taking the equally-weighted arithmetic mean of the returns of individual arbitrage trades. The possible total closing costs (CC) that occur because arbitrage trades are closed in month t are eventually deducted from the total excess returns to obtain the total ex-closing cost excess returns. Assuming there are n open arbitrage trades in month t, the total ex-closing cost excess return in a month can be presented in formula as:

$$\frac{1}{n} * \sum_{i=1}^{n} return_{t,i}^{arbitrage} - CC_t \tag{6.19}$$

Eventually, I get a time series of monthly total ex-closing cost excess returns from which I calculate the different return statistics and which I apply in other analyses. From now on, I refer to the monthly total ex-closing cost excess returns simply as 'returns'.

The monthly following of changes of values in the swaps should combine the effects of paid cash flows and obtained value of swap at the end of holding period, thus reflecting the \$-returns the swap trader would be interested in. Duarte et al. (2007) apply tied capital (i.e. provided margin) that makes the annual standard deviation[1] of returns to 10 %. This is where I make a rough simplification and I consistently apply the 5% margin. My returns have standard deviation that are quite much higher than 10 %, so in practice

[1]Annual standard deviation is obtained by multiplying the monthly standard deviation by $\sqrt{12}$.

I should de-lever my trades and use higher initial margin. As I already explained, it is difficult to define the correct margin requirements. In addition, applying margin that would make the standard deviation of returns 10 % would require knowing the returns beforehand and iteratively solving the correct margin %, which in my opinion creates a somewhat look-ahead bias. It is good to notice though that my results are not fully comparable to Duarte et al. (2007) in this sense, but on the other hand my results may better reflect real-life performance.

The returns can also be related to the standard deviation of the returns, providing Sharpe ratio as return metric. Sharpe ratio roughly depicts the amount of return obtained per unit of risk and as such it is a good metric for comparing trading strategies with different levels of returns and standard deviations. I calculate the annualized Sharpe ratio as follows, denoting the returns as $\bar{x}_{ret}$ and standard deviation of the same returns as σ_{ret}:

$$\frac{\bar{x}_{ret}}{\sigma_{ret}} * \sqrt{12} \tag{6.20}$$

Therefore, I do not follow the textbook definition of Sharpe ratio and deduct the risk-free return from the mean of returns, as there is no need to do it. I will report the Sharpe ratios of yield curve arbitrage in my return statistics.

Duarte et al. (2007) do not specifically indicate their method for calculating the returns, but I interpret that they must be calculating the returns as explained above apart from the modifications I consciously make. Thus, the method of calculating returns for open arbitrage trades may in itself be a subject of further consideration.

7 Trading performance

In this section, I present the trading performance with the two alternative two-factor short rate models. I present the performance for the full trading period and an alternative subperiod 01/2015-12/2021. I test the performance for the full trading period with the methodology described in section 6 but I also test some methodology modifications in the subperiod.

7.1 Full trading period 01/2012 - 12/2021

I begin by presenting the monthly basis point (bps) mispricings by the two models in the full trading period.

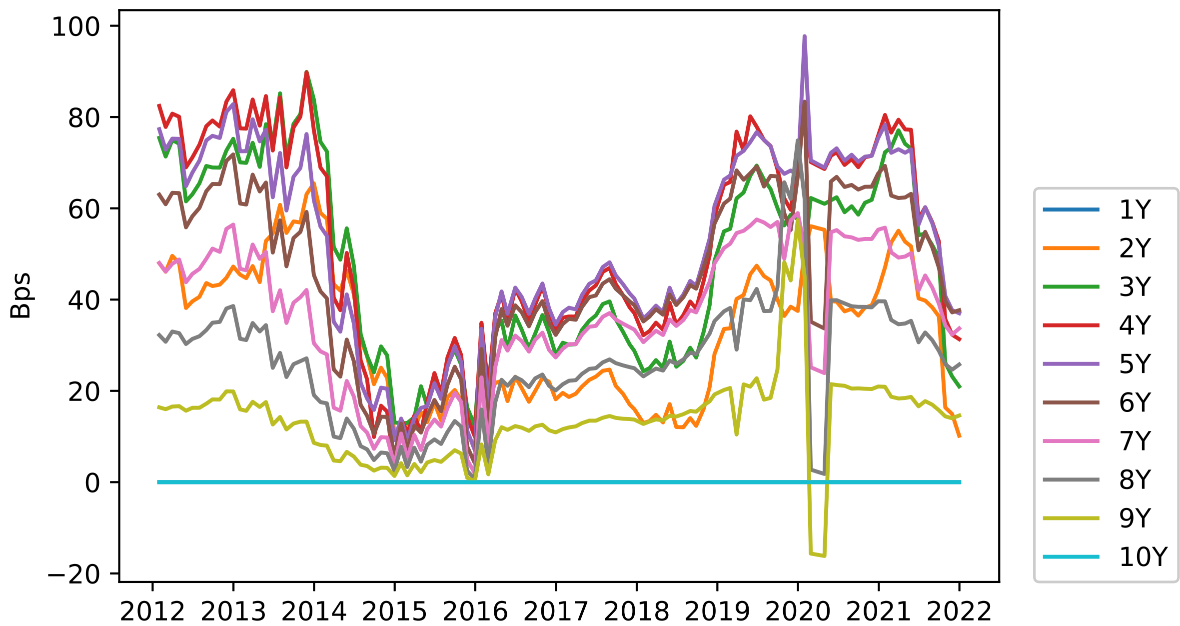

Figure 6: Mispricings with two-factor Vasicek model in 01/2012 - 12/2021

Figure [6] indicates, that with the two-factor Vasicek model the mispricings have been positive most of the 2010s. Mispricings related to 1-year and 10-year swaps are at zero, as they should be based on the modelling process. The mispricings were particularly large in 2012-2013 and in 2019-2021, but in between the mispricings were more modest. As the mispricings have exceeded the 5 bps and 20 bps mispricing limits most of the time and the largest mispricings have been consistently positive, I would have only opened payer swaps in the full trading period. As I only open the largest mispricing exceeding the basis point limit in a given month and the mispricings have been above 20 basis points most of the time, I would have opened the same arbitrage trades with both of the mispricing limits. The high level of mispricings in 2012-2013 and in 2019-2020 is most probably related first to the aftermath of financial crisis and later to the monetary policy of the Fed. In these periods, as the rates were high, the modelling likely overexaggerates the true swap rate levels. From this plot alone, it is impossible to say whether model rates have decreased or market rates increased in periods of convergence, such as in 2014. Similarly, it is impossible to say whether model rates have increased or market rates decreased in periods of divergence, such as in 2019. To do this, I can compare this mispricing plot to the plots of swap rates in Figure [1] and I can observe the convergence and divergence periods being caused by both reasons.

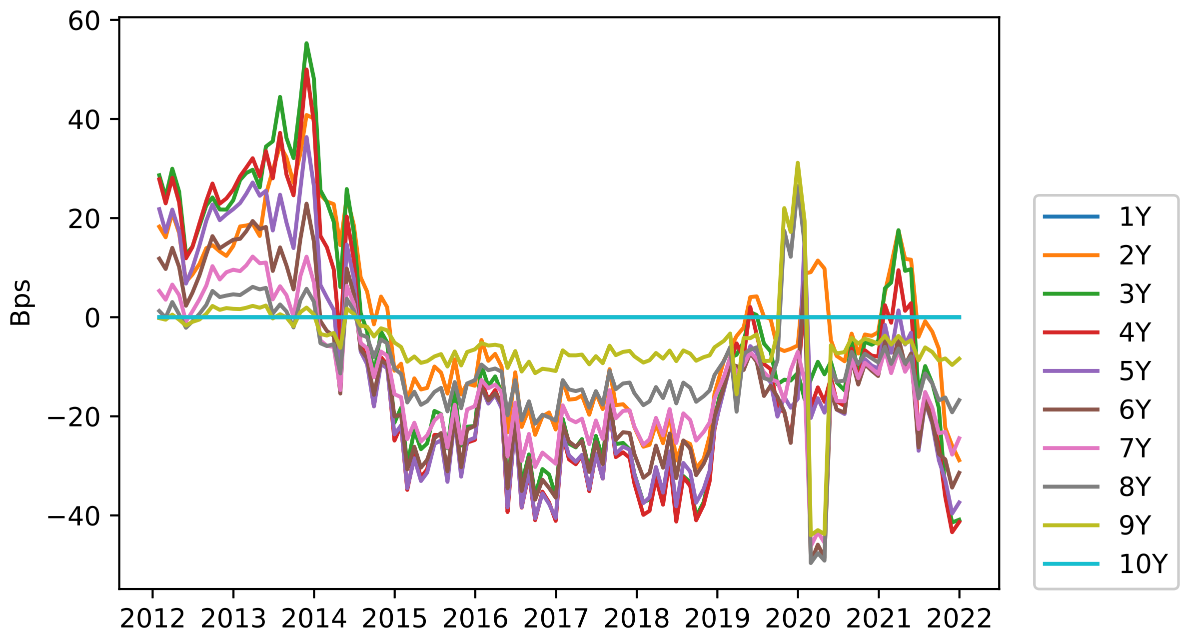

Figure 7: Mispricings with two-factor CIR model in 01/2012 - 12/2021

Figure [7] plots the mispricings by the two-factor CIR model. The mispricings are quite different compared to the Vasicek model. In 2012-2013 the mispricings were positive and quite large as with the Vasicek model but then the mispricings turn negative for the rest of the trading period. The mispricings also remain within the bounded range (-40 bps, 40 bps) whereas with the Vasicek model there are positive mispricings as large as 100 bps. Although there is fluctuation in mispricings in 2019-2021 with the CIR model, the mispricings remain surprisingly close to 0.

Overall, it is surprising that the CIR model indicates lower levels for model rates than the Vasicek model. Especially in the low rate period in 2014-2018 I would have expected the Vasicek model to provide lower levels for the model rates. In fact, I would have expected the mispricing plots to be the other way round. As far as I understand, based on the distributional properties of the short rate in both models and how they are reflected on the closed form expressions for zero-coupon bond pricing, Vasicek model should basically provide lower model rates than the CIR model. After all, I am applying the same initial parameters for both models, apart from the mean reversion parameter k that is adjusted upwards for CIR model. My expertise on the short rate models is not enough to pinpoint the reason for the observed modelling results or if the observed performance is theoretically possible. However, there are a few possible candidates that could explain the results:

- As I use different local optimization methods for Vasicek and CIR, it is most likely that the methods cause the observed differences. The models should not theoretically depend heavily on the applied optimization methods but the practical calibration procedure and e.g. the lack of global optimization likely causes the observed issues
- As mentioned in subsection 6.2.1, the two-factor models assume that the parameters and the factors x_0 and y_0 are initially known. The initial parameters are based on Duarte et al. (2007) and as such they can be considered robust estimates but the calibration of factors x_0 and y_0 based on the 1-year and 10-year swap rates may cause issues to the analytical properties of the models
- I adjust the k_1 parameter for CIR initially by multiplying it by 1.6, so that the constraint condition $2 * k_1 * \theta_1 > \sigma_1^2$ is fulfilled with initial guess parameters of Duarte et al. (2007). This may have an effect, but I consider it to be more reasonable to adjust mean-reversion rate than the long-run mean or instantaneous volatilities.

Before presenting the summary statistics for the returns, it is necessary to provide some background for the statistical analysis of the returns. Finance literature commonly makes assumption about the distribution of returns being normal distribution or some other stable distribution. Similarly, it is often assumed that the returns are independent of each other. Together, these assumptions about returns are often referred to as "i.i.d." assumptions, i.e. identically distributed and independent return observations. Such assumptions provide the possibility to study the statistical significance of returns.

However, in my case it is quite unlikely that the returns would be independent and coming from the same distribution all the time. Rather it is quite likely that e.g. the volatility of returns changes through time, thus changing the return distribution. Therefore, it is somewhat non-robust to do statistical analysis on my returns, such as to study the confidence intervals or to do statistical hypothesis testing relying on the assumptions about independence and identical distribution.

However, to provide some backbone for the returns that I just reported, I will now provide the confidence intervals and Student's t-statistics for the mean of returns and the confidence intervals for the skewness measures. I will do this by applying heteroskedasticity and autocorrelation (HAC) robust standard errors[1]. My mathematical expertise is not enough to state whether applying the HAC robust standard errors in this context is enough to completely tackle the i.i.d. assumption issues. Therefore, I may need to additionally assume at least the independence of my returns.

By the central limit theorem (CLM), the distribution of the mean of i.i.d random variables approaches normal distribution when the number of observations increases. CLM is applied in the confidence interval estimation for the expected value of a general distribution. In other words, for the confidence interval of the mean of returns, I apply the following:

$$m(x) + / - z * HAC \tag{7.1}$$

where $m(x)$ is the mean of returns and z is a random variable following standard normal distribution with $P(-z \leq Z \leq z) = \alpha$ and α is the chosen confidence level. HAC is the HAC robust standard error. I could also assume for z that it is obtained from the t-distribution, but as I have over 100 return observations, it is quite safe to rely on normality. I will use the 95 % confidence level for confidence intervals.

For the Student's t-statistic, I formulate the hypothesis testing as $H_0 = \mu = 0$ and $H_1 = \mu > 0$, where μ denotes the expected value of mean of returns[2]. The choice to use zero in null hypothesis could be argued against. I could for example assume that $H_0 = \mu < 0$ but this would make the hypothesis testing somewhat more complicated. Based on this setup for the hypothesis testing, I calculate the Student's t-statistic as:

[1] These are automatically calculated in my Python code.

[2] The mean of returns is assumed to follow normal distribution by CLM.

$$\frac{m(x) - \mu}{HAC} \tag{7.2}$$

Under the null hypothesis and with number of observations large enough, Student's t-statistic follows the standard normal distribution and I can take the critical values for the test statistic from the standard normal distribution.

Finally, I provide the HAC corrected skewness and kurtosis and the confidence intervals for skewness[1]. The confidence intervals for skewness is calculated as follows:

$$skew + / - z * \sqrt{\frac{6n(n-1)}{(n-2)(n+1)(n+3)}} \tag{7.3}$$

where *skew* is the HAC corrected skewness, z has the same previous definition and n is the number of monthly return observations. To the best of my knowledge, this is the way confidence intervals are defined for skewness but I am not fully sure. This measure of confidence interval relies on the assumption of returns stemming from stable normal distribution which is likely not the case with my returns. Thus the confidence interval for skewness is merely indicative as I do not know how to do HAC or some else correction to the margins or error. There is no easy way to calculate the correct margins of error for confidence interval estimation when the distribution is unknown and changes through time.[2]

[1] I calculate the t-statistics with HAC standard errors manually but the HAC corrected skewness and kurtosis are automatically calculated in my Python code.

[2] All in all, when the distribution of returns is unknown, it can be approximated with different boostrapping methods. The i.i.d. issues may still persist in the boostrapped distributions, though.

Table 2: Summary statistics for the monthly total ex-closing cost excess returns in 01/2012 - 12/2021. The count of observations is 119, as the first month 01/2012 is not included in return calculations. Limit is the mispricing limit for initiating trades. Mean is the mean of monthly total ex-closing cost excess returns. t-stat is the Student's t-stastistic that relies on null hypothesis $\mu = 0$ and is calculated with heteroskedasticity and autocorrelation (HAC) robust standard errors. (MLB, MUB) provide the lower and upper bound for the confidence interval of the mean of returns. HAC is the heteroskedasticity and autocorrelation (HAC) robust standard error. M. margin is the margin of error in the confidence interval for the mean of returns. Skewness (skew) and kurtosis (kurt) are calculated with HAC robust standard errors and kurtosis is based on Pearson's definition. (SLB, SUB) provide the lower and upper bound for the confidence interval of skewness. S. margin is the margin of error in the confidence interval for skewness. Limit, mean, (MLB, MUB), HAC and M.margin are in basis points.

Model	Limit	Mean	t-stat	(MLB, MUB)	HAC	M. margin	Skew	(SLB, SUB)	S. margin	Kurtosis
Vasicek	5	18.217	0.642	(-37.414, 73.848)	28.383	55.631	0.028	(-0.407, 0.463)	0.435	3.699
CIR	5	2.713	0.092	(-54.876, 60.302)	29.382	57.589	0.084	(-0.351, 0.519)	0.435	3.950
Vasicek	20	18.908	0.668	(-36.550, 74.365)	28.295	55.458	0.032	(-0.403, 0.467)	0.435	3.691
CIR	20	5.175	0.194	(-47.051, 57.402)	26.646	52.226	0.109	(-0.326, 0.544)	0.435	4.371

Table 3: Additional summary statistics for the monthly total ex-closing cost excess returns in 01/2012 - 12/2021. The count of observations is 119, as the first month 01/2012 is not included in return calculations. Limit, median, minimum, maximum, monthly standard deviation and annual standard deviation of the monthly total ex-closing cost excess returns are in basis points. Limit is the mispricing limit for initiating trades. Ratio neg is the ratio of months with negative total ex-closing cost excess returns relative to all months. Serial corr is the 1 lag autocorrelation measure. Sharpe ratios are annualized.

Model	Limit	Med	Min	Max	Stdev (m.)	Stdev (ann.)	Ratio neg	Serial corr	Sharpe
Vasicek	5	71.692	-1604.520	1702.986	590.859	2046.797	0.462	0.108	0.107
CIR	5	-18.553	-1823.045	1591.256	595.627	2063.310	0.521	0.168	0.016
Vasicek	20	71.692	-1604.520	1702.986	590.332	2044.970	0.462	0.103	0.111
CIR	20	0.000	-1814.382	1592.333	544.736	1887.020	0.429	0.149	0.033

Tables [2] and [3] finally provide the summary statistics of returns for the two models with the two mispricing limit variations. Table [2] provides the most important information. The first thing to notice is that all the trading variations provide positive mean of returns (Mean) but without statistical significance as measured by the t-statistic (t-stat). The Vasicek model appears to outperform the CIR model based on mean of returns but the conficence intervals (MLB, MUB) are very wide and they overlap between the models. The confidence intervals for the mean of returns cover 0 and also mostly the returns of Duarte et al. (2007) that were in range of about (43,62) basis points per month. Therefore, I can not say that my mean of returns would differ from the mean of returns of Duarte et al. (2007). In the same way, I can not say that my trading would provide me positive mean of returns. I would need more narrow confidence intervals for my mean of returns and I would need the interval bounds to stay on the positive. Table [3] provides the monthly (Stdev (m.)) and annualized (Stdev (ann.)) standard deviations that are around 5.9 % and 20 %, respectively. This suggests that I would need to de-lever my trades, i.e. add more initial margin as capital, to get to the 10 % annual standard deviation of Duarte et al. (2007).

Looking next at the skewness (Skew), I notice that my skewness measures are only slightly positive and they also have quite wide confidence intervals (SLB, ULB). The mean of returns exceeds the median (Med) of returns only in the CIR model which is also an indication of skewness. However, the confidence intervals of skewness remain within the bounds of symmetry ($+/-0.5$) whereas skewness of returns of Duarte et al. (2007) exceeds 0.5 in all their trading variations. Therefore, I would suggest there is no clear positive skewness in my returns. The kurtosis measures[1] exceed the limit of 3 for "leptokurtic" distribution so my return distribution appears to have fatter tails than a benchmark normal distribution. However, as I have an almost symmetric distribution there is not much value in having fatter distribution tails as I would have fatter negative tails as well. All in all, the kurtosis measures are not very big. Duarte et al. (2007) apply Fisher's definition for kurtosis and as their kurtosis measures exceed 0, it indicates there is kurtosis in their returns as well. They have particurly high kurtosis only for one trading variation of yield curve arbitrage and the other kurtosis measures are quite close to my kurtosis measures.

Looking then at the additional summary statistics in Table [3], I observe that my Sharpe ratios (Sharpe) indicate poor performance for all the trading variations and the ratios are much smaller than in Duarte et al. (2007) where the Sharpe ratios are in range (0.5, 0.7)[2]. My ratios of negative monthly returns (Ratio neg) are close to 0.5 which is much higher than the ratios of Duarte et al. (2007) that are in range (0.2, 0.3). There is small positive first-order serial correlation (Serial corr) in all the trading variations, i.e. positive returns are followed by positive returns and vice versa indicating that my returns are not independent of each other. Duarte et al. (2007) have negative serial correlations in

[1]Pearson's kurtosis.

[2]Sharpe ratio has its own problems. It relies on an assumption that the returns stem from a stable two-parameter distribution. In case the distribution is not constant or there is skewness and kurtosis, Sharpe ratio is not fully robust measure. However, as an indicative measure of the relation between risk and return, Sharpe ratio is useful.

range (-0.15, -0.10). This could suggest that the mispricings in the trading of Duarte et al. (2007) converge and diverge faster whereas in my trading the convergence and divergence is slower. The minimum (Min) and maximum (Max) of returns are in quite the same range for both the Vasicek and the CIR models and this range is very big compared to Duarte et al. (2007) who have (Min, Max) range mostly within (-6 %, 12 %). Therefore, although there appears no particular negative skewness in my returns, the negative returns are potentially very big. In the next subsections, I connect this notion to the potential tail risk exposure.

Next, I provide the return distributions for both the models and the mispricing limit variations in Figures [8], [9], [10] and [11].

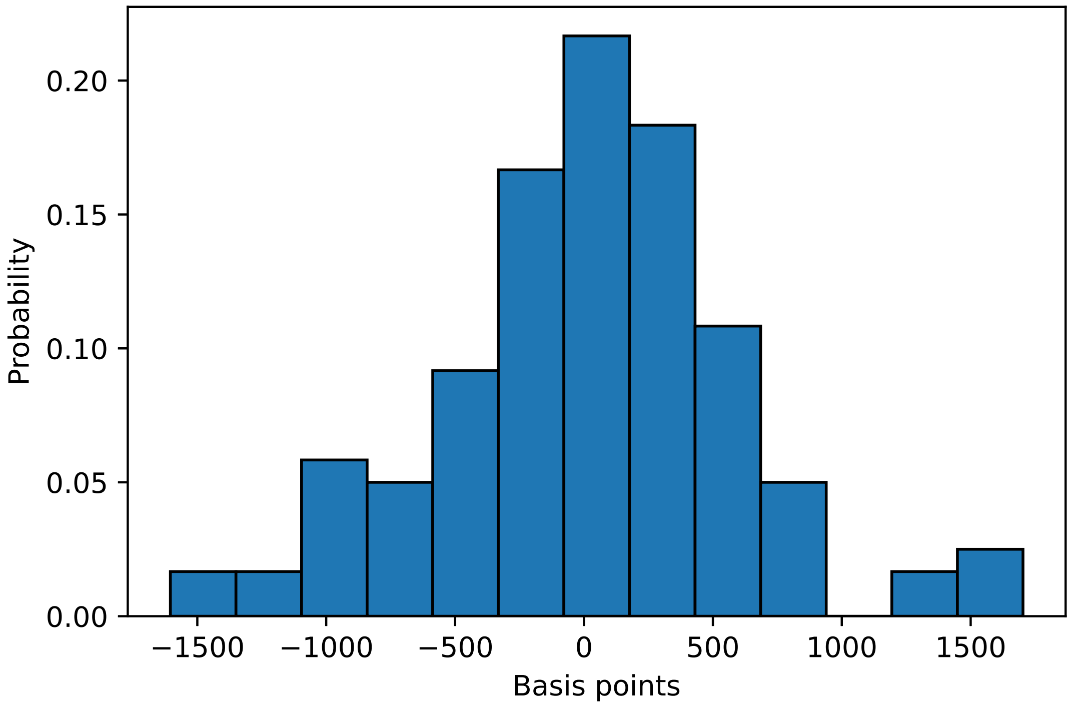

Figure 8: Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in 01/2012 - 12/2021, 5bps mispricing limit

Figures [8] and [9] confirm the results for the Vasicek model already observed in Table [2]. The return distributions are almost symmetric with only small positive skew. There is more peakness in the 20 bps mispricing limit variation than in the 5 bps variation that has fatter distribution tails. There is no clear indication of non-normal distribution with positive mean of returns, strong positive skewness and particularly fat distribution tails for the Vasicek model that would indicate a lucrative return distribution for yield curve arbitrage implementation.

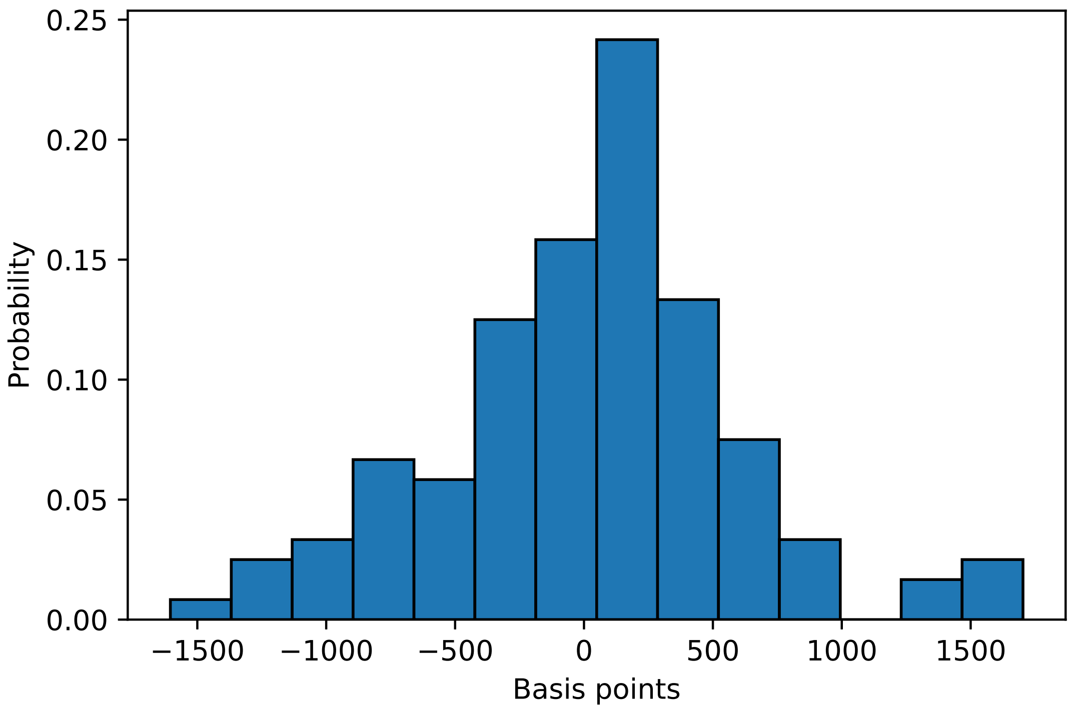

Figure 9: Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in 01/2012 - 12/2021, 20bps mispricing limit

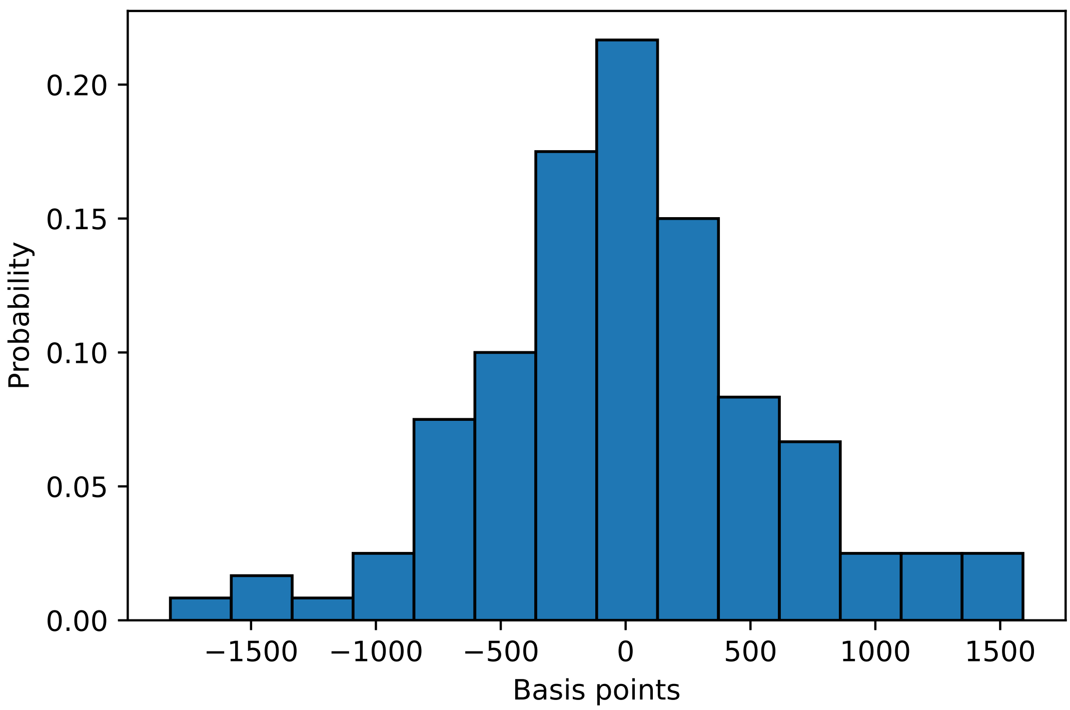

Figure 10: Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in 01/2012 - 12/2021, 5bps mispricing limit

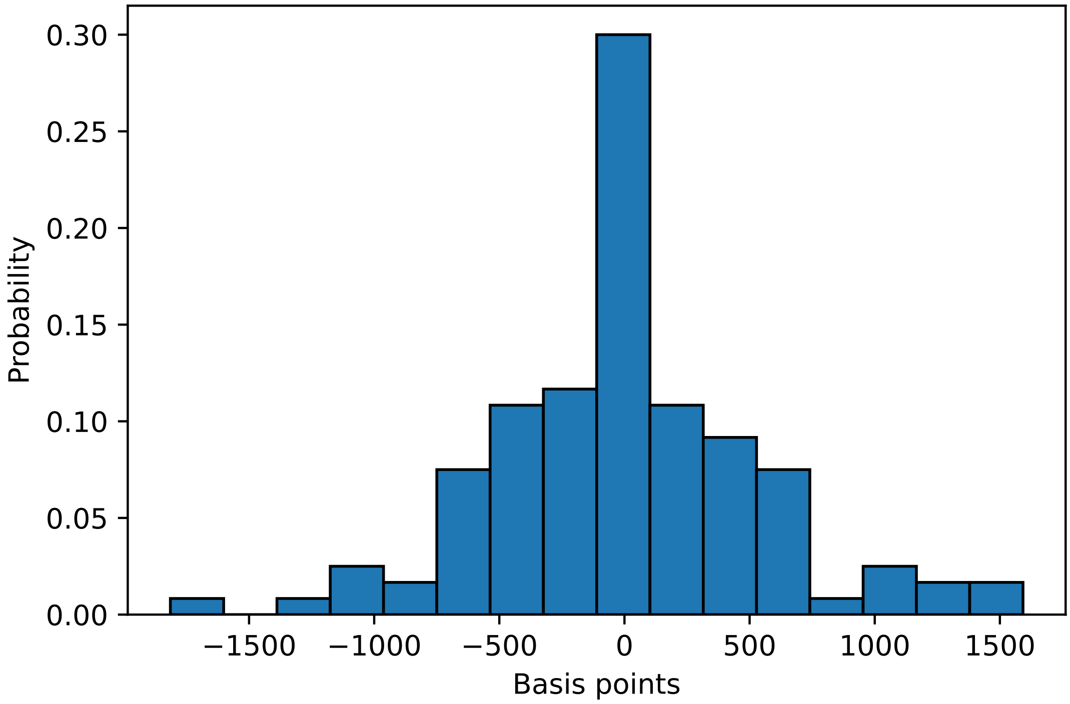

Figure 11: Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in 01/2012 - 12/2021, 20bps mispricing limit

The same goes with the CIR model as presented in Figures [10] and [11]. Positive skewness can not be observed in these distributions either but the 20 bps mispricing limit variation has particularly high peakness, i.e. kurtosis is more present through the high central peak than through fatter and longer distribution tails. This is because there are many months in this trading variation where no trades are opened as the mispricing limit is not exceeded. Therefore, the trading portfolio may be empty in certain months providing 0 excess returns. Therefore, I do not get support for lucrative return distribution for the CIR model and its mispricing limit variations either.

Finally, I provide the plots of cumulative returns in Figures [12] and [13] for the 5 bps mispricing limit trading variations. The 5 and 20 bps mispricing limit plots are quite similar for both of the models and therefore I dediced to present here only the cumulative returns for the 5 bps mispricing limit[1]. For the Vasicek model, there are large drops in cumulative returns in 2013 and in 2020-2021, but also a period of quite steady increase between 2014-2020. The cumulative return plots do not indicate good performance for the Vasicek model, as the cumulative returns remain negative for most of the trading period and there is a lot of variation. An expected and positive cumulative performance would

[1]See the Appendices section 11.7 for the 20 bps mispricing limit cumulative return plots.

have been a somewhat steady growth of cumulative returns through the trading period.

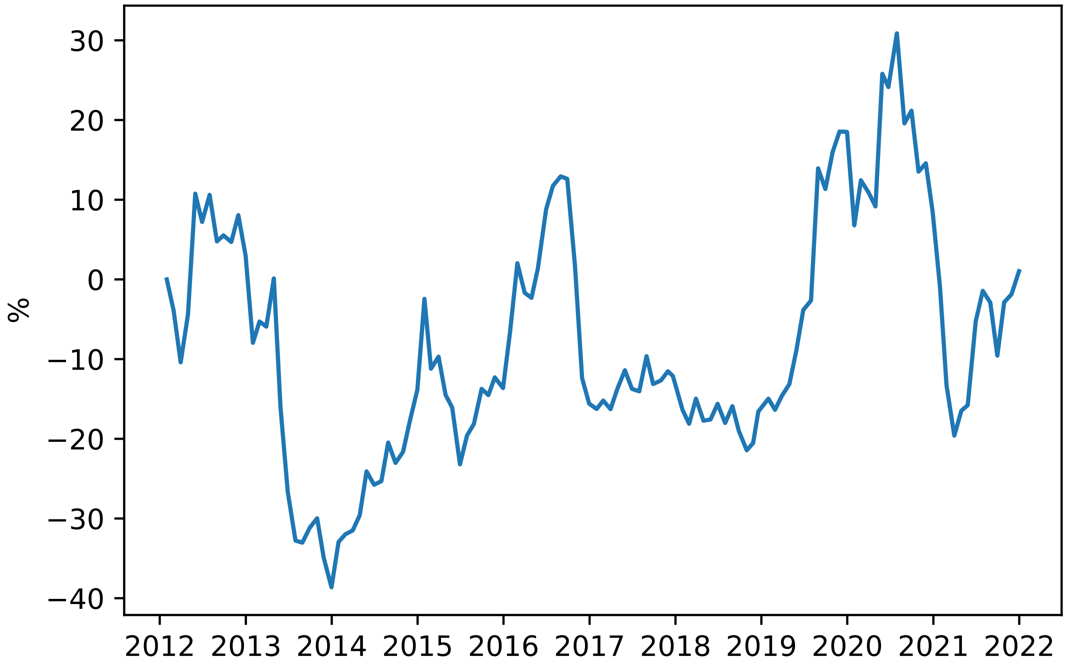

Figure 12: Cumulative returns with two-factor Vasicek model in 01/2012 - 12/2021, 5bps mispricing limit

The performance is worse for the CIR model. The cumulative performance remains negative for most of the trading period. Large negative drops and increases in cumulative returns are timed to quite the same periods as with the Vasicek model but with the CIR model the drops are bigger. For both of the models, there is too many drops and variation in returns[1].

In summary, the performance of my trading in the full trading period is in one sense positive, as I do get positive mean of returns that have confidence intervals covering the results of Duarte et al. (2007). Therefore, I can not say that yield curve arbitrage would not have provided positive returns in the full trading period. On the other hand, the return distributions do not indicate particularly good performance, as the distributions appear to be symmetric and there is quite certainly no clear positive skewness.

[1]See the Appendices section 11.7 also for plots of monthly returns.

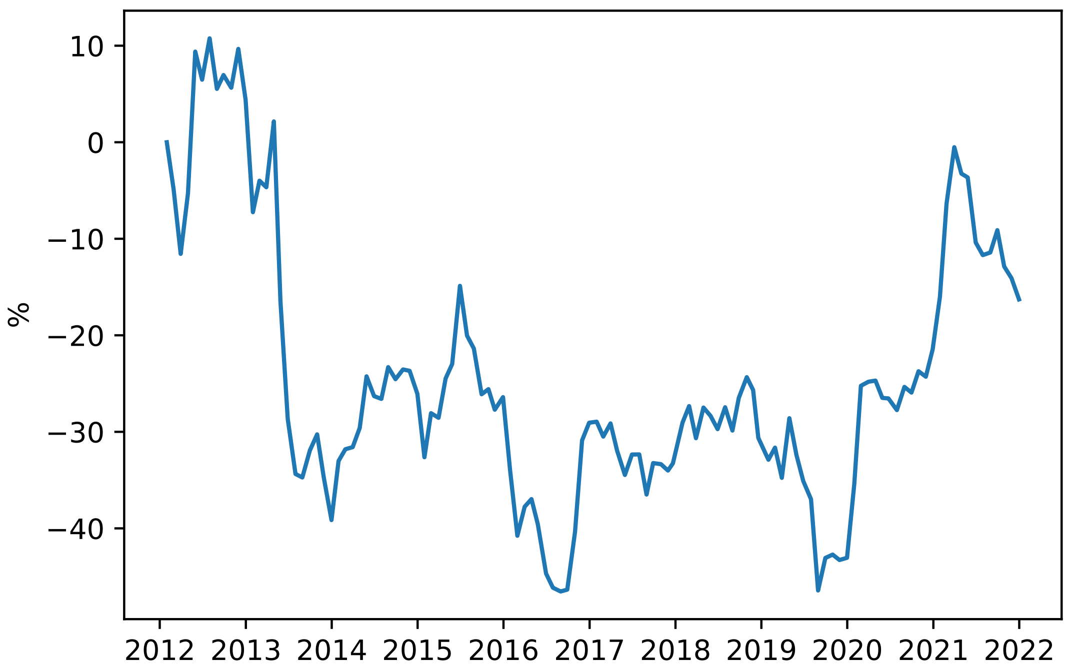

Figure 13: Cumulative returns with two-factor CIR model in 01/2012 - 12/2021, 5bps mispricing limit

7.2 Subperiod 01/2015 - 12/2021

In this subsection, I first report the return statistics for subperiod 01/2015 - 12/2021 with the same methodology that was presented in section 6. Therefore, I apply the same initial parameters and calibration period length of 60 months and I end up opening the same arbitrage trades in 01/2015 - 12/2021 as I did it in previous section 7.1, now just the first month of trading is moved from 01/2012 to 01/2015. I do this to provide a better idea how the performance changes if I drop the period 01/2012 - 12/2014 from my return statistics. I report the results only for the 5 bps mispricing limit.

Table 4: Summary statistics for the monthly total ex-closing cost excess returns in 01/2015 - 12/2021. The count of observations in this period is 83, as the first month 01/2015 is not included in return calculations. Limit is the mispricing limit for initiating trades. Mean is the mean of monthly total ex-closing cost excess returns. t-stat is the Student's t-stastistic that relies on null hypothesis $\mu = 0$ and is calculated with heteroskedasticity and autocorrelation (HAC) robust standard errors. (MLB, MUB) provide the lower and upper bound for the confidence interval of the mean of returns. HAC is the heteroskedasticity and autocorrelation (HAC) robust standard error. M. margin is the margin of error in the confidence interval for the mean of returns. Skewness (skew) and kurtosis (kurt) are calculated with HAC robust standard errors and kurtosis is based on Pearson's definition. (SLB, SUB) provide the lower and upper bound for the confidence interval of skewness. S. margin is the margin of error in the confidence interval for skewness. Limit, mean, (MLB, MUB), HAC and M.margin are in basis points.

Model	Limit	Mean	t-stat	(MLB, MUB)	HAC	M. margin	Skew	(SLB, SUB)	S. margin	Kurtosis
Vasicek	5	18.745	0.593	(-43.248, 80.737)	31.629	61.992	0.167	(-0.351, 0.685)	0.51778	4.031
CIR	5	41.174	1.246	(-23.606, 105.953)	33.051	64.779	0.440	(-0.078, 0.958)	0.51778	3.940

Table 5: Additional summary statistics for the monthly total ex-closing cost excess returns in 01/2015 - 12/2021. The count of observations in this period is 83, as the first month 01/2015 is not included in return calculations. Limit, median, minimum, maximum, monthly standard deviation and annual standard deviation of the monthly total ex-closing cost excess returns are in basis points. Limit is the mispricing limit for initiating trades. Ratio neg is the ratio of months with negative total ex-closing cost excess returns relative to all months. Serial corr is the 1 lag autocorrelation measure. Sharpe ratios are annualized.

Model	Limit	Med	Min	Max	Stdev (m.)	Stdev (ann.)	Ratio neg	Serial corr	Sharpe
Vasicek	5	52.651	-1378.186	1702.986	547.634	1897.060	0.482	0.122	0.119
CIR	5	-7.316	-1500.002	1591.256	554.476	1920.760	0.506	0.193	0.257

Tables [4] and [5] provide the summary statistics for monthly returns in subperiod 01/2015 - 12/2021. Performance somewhat improves compared to the full trading period: the mean of returns for both models improves, there is more positive skewness and the Sharpe ratios also improve. The CIR model outperforms the Vasicek model based on mean of returns, Sharpe ratio and skewness. However, the returns are still statistically insignificant by the t-statistic and the confidence intervals are wide for the mean of returns.

The confidence intervals overlap between the models so I can not conclude which model would be better than the other. The confidence intervals for the mean of returns again overlap the returns of Duarte et al. (2007), so I can not conclude that my performance would differ from Duarte et al. (2007). The monthly standard deviation of the returns is around 5.5 % and the annual standard deviation is around 19 %, suggesting that I would need to de-lever my positions. The range of minimum and maximum for returns is again very wide. The ratio of negative monthly returns and the serial correlation measures are quite the same as in the full trading period.

The skewness measures improve compared to the full trading period and now the confidence intervals for skewness is not anymore within $+/-0.5$ that would indicate symmetric distribution. For both models, the confidence intervals for skewness now cover the skewness measures of Duarte et al. (2007) and especially for the CIR model the confidence interval for skewness appears good. Therefore, I would conclude that there is positive skewness in the return distributions in the subperiod. In addition, there is again more kurtosis than in the benchmark normal distribution and in this case the kurtosis is a positive thing, as now the fatter tails would appear in the positive tail at least for the CIR model.

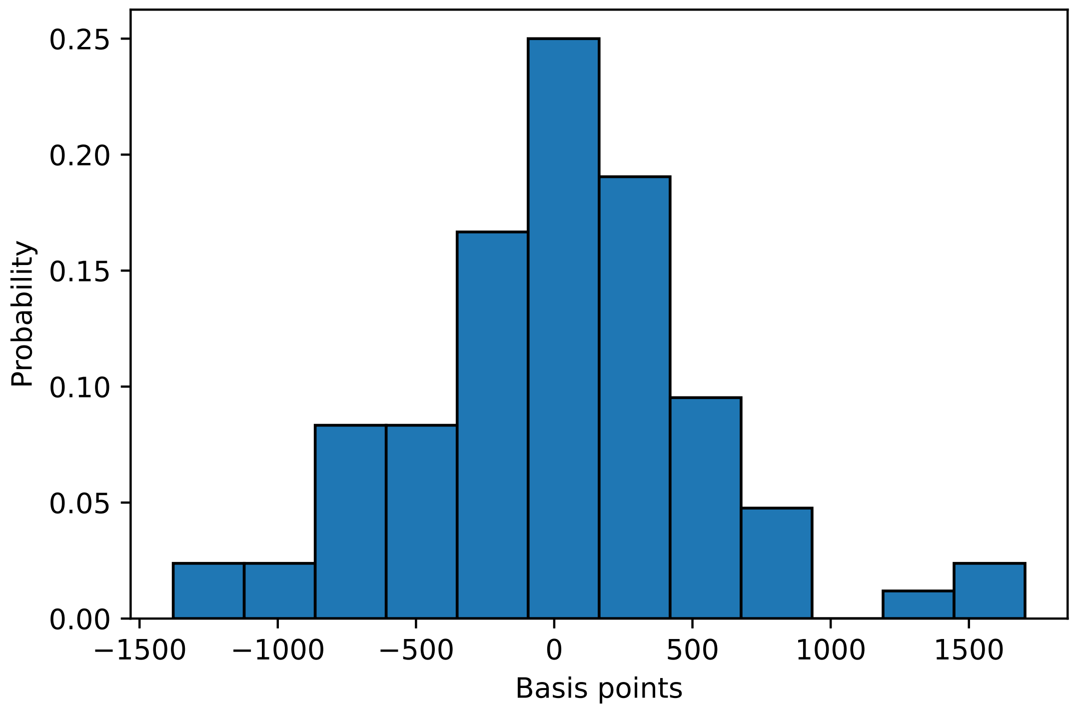

Figure 14: Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit

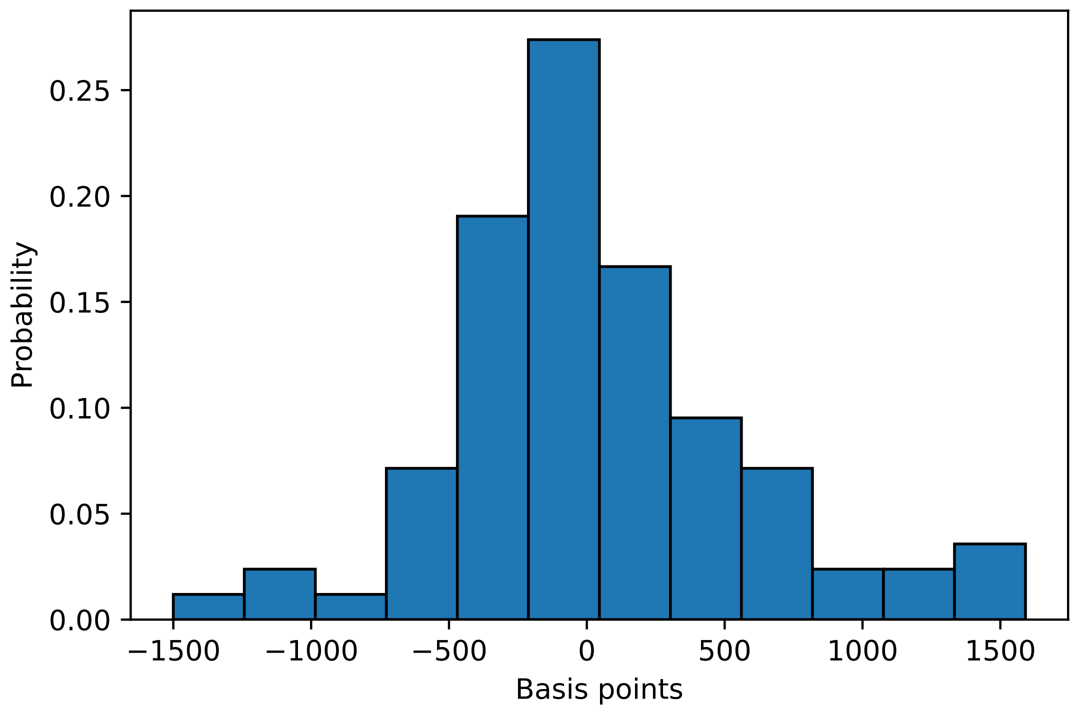

Figure 15: Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit

Figures [14] and [15] confirm the previous notions. Now the positive skewness and the fat positive distribution tails appear particularly well in the distribution of the CIR model. The positive skewness does not appear that clearly for the Vasicek model but there is still some positive skewness for the Vasicek model as well.

The cumulative performance as depicted in Figures [16] and [17] is not what would be expected. There is again large drops and increases in cumulative returns. The cumulative performance is quite similar for both models but there are also some timing differences in return drops and increases. All in all, the statistical insignificance of returns, quite poor Sharpe ratios and bad cumulative performance would suggest that yield curve arbitrage is not particularly lucrative even if 01/2012 - 12/2014 is dropped from the trading period. The same inconclusiveness regarding the mean of returns remains as in full trading period but now there appears to be positive skewness in returns.

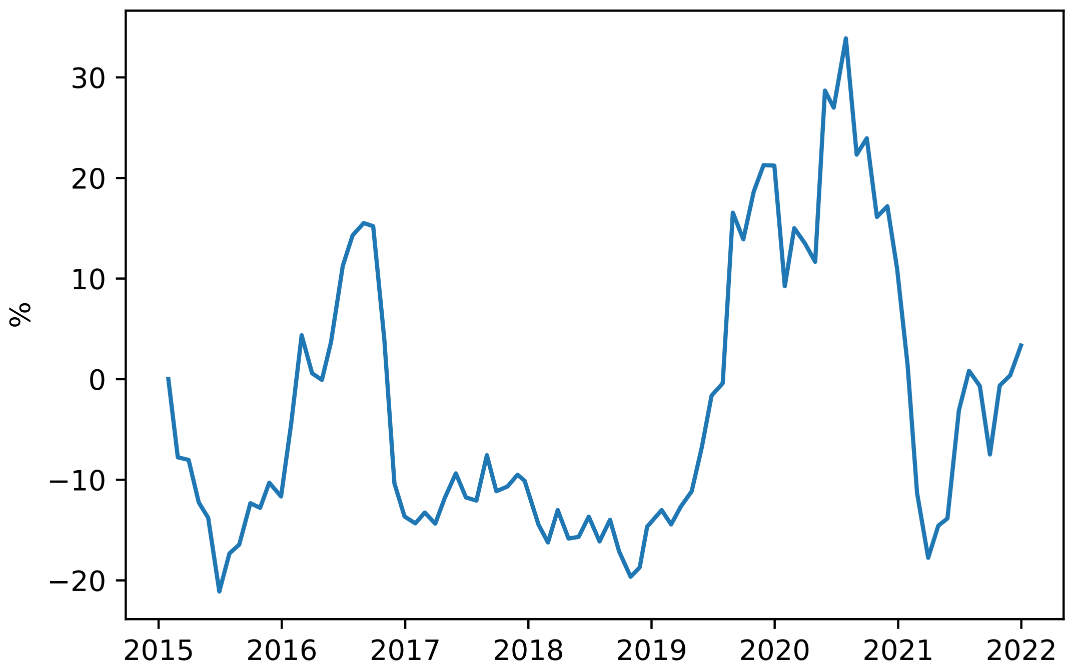

Figure 16: Cumulative returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit

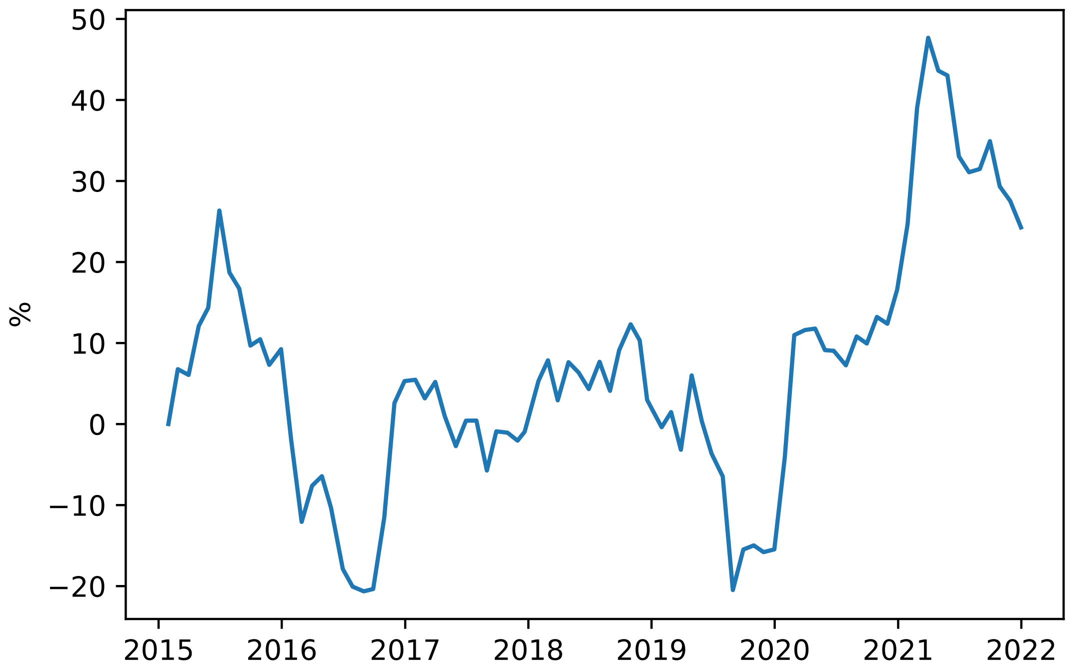

Figure 17: Cumulative returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit

7.2.1 Subperiod performance with methodology modifications

In this subsection I test the subperiod performance with methodology modifications. I want to test certain methodology modifications to see if the somewhat poor and inconsistent performance of yield curve arbitrage in my trading periods could be explained by my methodology choices. Namely, I test the following modifications:

- Upper bounds for θ and σ parameters are doubled to 0.3
- 60 month calibration period is extended to 96 months
- COBYLA optimization algorithm is applied for both the Vasicek and CIR model implementations to obtain more comparability between the models
- Parameters are updated in a rolling manner, meaning that I do not apply the same parameters of Duarte et al. (2007) as initial guesses over and over again, but rather I use the latest optimized parameters as initial guesses when moving from month to month

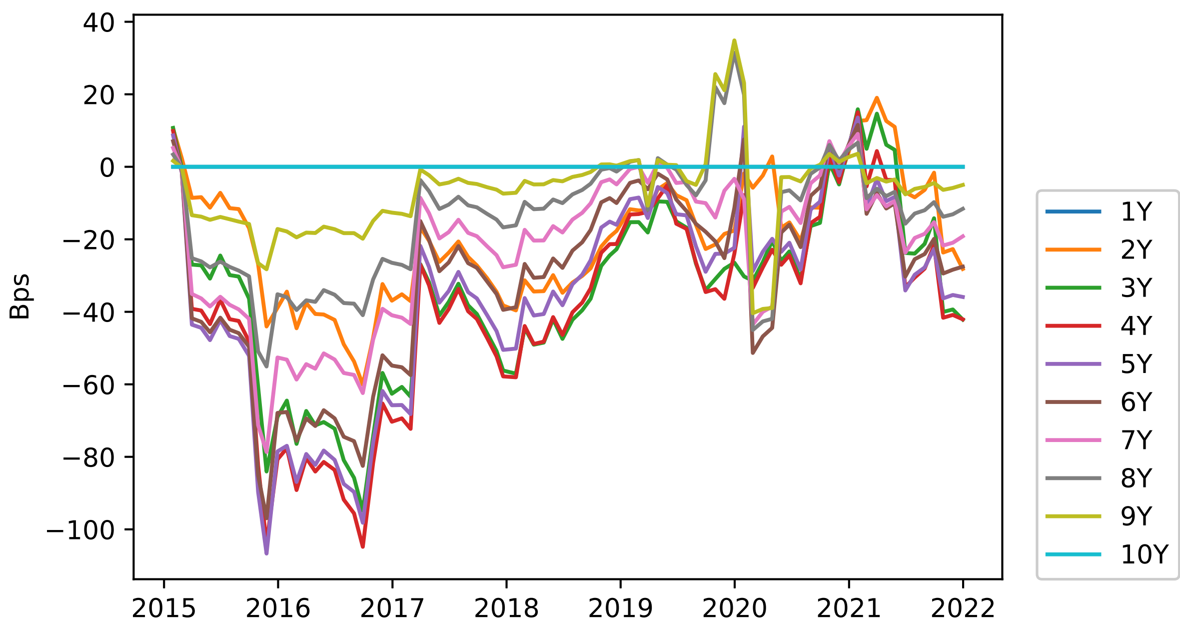

Figure 18: Mispricings with two-factor Vasicek model in subperiod 01/2015 - 12/2021 with modified methodology

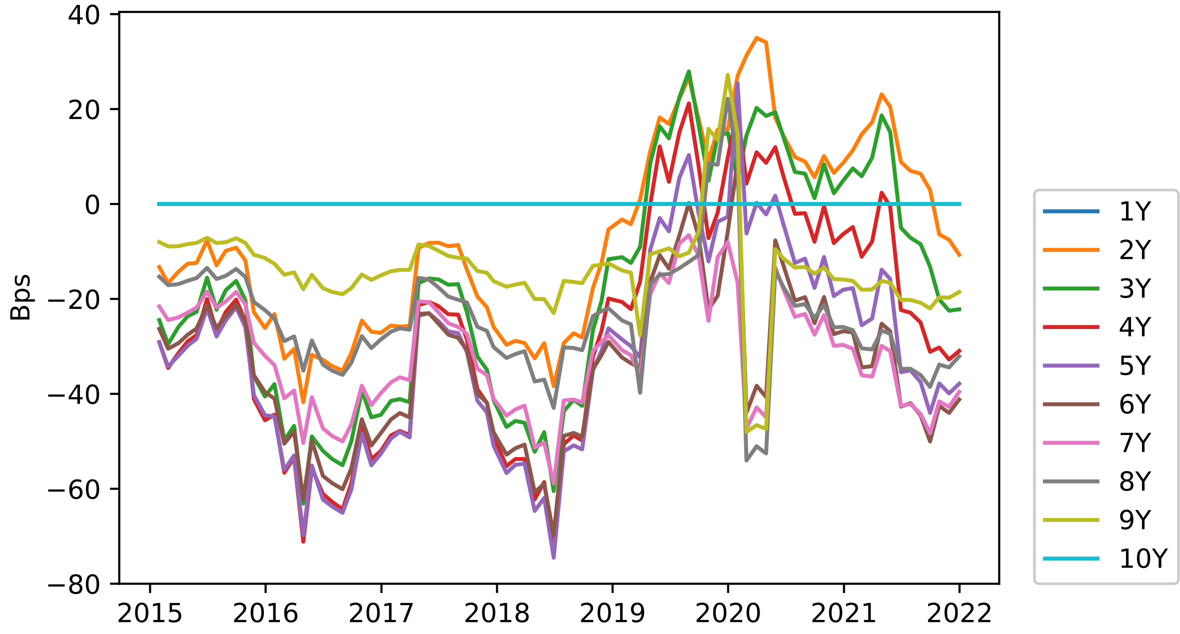

Figure 19: Mispricings with two-factor CIR model in subperiod 01/2015 - 12/2021 with modified methodology

Figure [18] presents how the mispricings modelled by the Vasicek model change totally from full trading period with unmodified methodology to the subperiod with modified methodology. Now the mispricings are negative throughout the subperiod as opposed to being positive in the full trading period. The CIR model still presents negative mispricings in Figure [19]. Now the mispricings modelled by the Vasicek model are somewhat more negative than those modelled by the CIR model, which suggests that mispricing differences observed in full trading period could indeed be explained by differences in optimization methodology. Both of the models still offer very big mispricings, although towards the end of the subperiod the mispricings decrease.

Table 6: Summary statistics for the monthly total ex-closing cost excess returns with methodology modifications in 01/2015 - 12/2021. The count of observations in this period is 83, as the first month 01/2015 is not included in return calculations. Limit is the mispricing limit for initiating trades. Mean is the mean of monthly total ex-closing cost excess returns. t-stat is the Student's t-stastistic that relies on null hypothesis $\mu = 0$ and is calculated with heteroskedasticity and autocorrelation (HAC) robust standard errors. (MLB, MUB) provide the lower and upper bound for the confidence interval of the mean of returns. HAC is the heteroskedasticity and autocorrelation (HAC) robust standard error. M. margin is the margin of error in the confidence interval for the mean of returns. Skewness (skew) and kurtosis (kurt) are calculated with HAC robust standard errors and kurtosis is based on Pearson's definition. (SLB, SUB) provide the lower and upper bound for the confidence interval of skewness. S. margin is the margin of error in the confidence interval for skewness. Limit, mean, (MLB, MUB), HAC and M.margin are in basis points.

Model	Limit	Mean	t-stat	(MLB, MUB)	HAC	M. margin	Skew	(SLB, SUB)	S. margin	Kurtosis
Vasicek	5	-22.820	-0.509	(-110.694, 65.053)	44.833	87.873	-0.158	(-0.676, 0.359)	0.518	5.556
CIR	5	14.220	0.553	(-36.170, 64.610)	25.709	50.390	0.504	(-0.014, 1.022)	0.518	5.776

Table 7: Additional summary statistics for the monthly total ex-closing cost excess returns with methodology modifications in 01/2015 - 12/2021. The count of observations in this period is 83, as the first month 01/2015 is not included in return calculations. Limit, median, minimum, maximum, monthly standard deviation and annual standard deviation of the monthly total ex-closing cost excess returns are in basis points. Limit is the mispricing limit for initiating trades. Ratio neg is the ratio of months with negative total ex-closing cost excess returns relative to all months. Serial corr is the 1 lag autocorrelation measure. Sharpe ratios are annualized.

Model	Limit	Med	Min	Max	Stdev (m.)	Stdev (ann.)	Ratio neg	Serial corr	Sharpe
Vasicek	5	-18.567	-3034.357	2554.203	779.396	2699.910	0.518	0.112	-0.102
CIR	5	16.614	-1080.173	1617.323	430.703	1492.010	0.482	0.196	0.114

Tables [6] and [7] provide the summary statistics for subperiod returns with modified methodology and the results are again quite inconclusive and inconsistent. Now the Vasicek model performs particularly bad as the mean of returns and Sharpe ratio are negative and there is negative skewness in return distribution. The CIR model provides the best performance in terms of distributional numbers, as now the model has the highest positive skewness and kurtosis compared to previous trading variations. The same notions apply for the mean of returns as in the previous subsections: I can not say that the performance would be different from Duarte et al. (2007) or that one of the models would be better than the other, as the confidence intervals for the mean of returns are wide and they overlap with each other and with the reported returns of Duarte et al. (2007). Based on the mean of returns and the confidence interval for the mean of returns, the performance appears to worsen for both of the models compared to the subperiod performance with unmodified methodology. The minimum of returns for the Vasicek model is over -30 %, which suggests particularly bad performance for some months. Sharpe ratios, ratios of months with negative returns and serial correlation measures are quite similar as in the full trading period, apart from the negative Sharpe ratio of the Vasicek model.

The skewness measures suggest negative skewness for the Vasicek model and positive skewness for the CIR model and as such these results are somewhat contradictory. The positive skewness for the CIR model is particularly clear. Figures [20] and [21] shed more light on the distributions of returns. The negative skewness for the Vasicek model is caused by an outlier, whereas there is indeed clear positive skewness in the CIR return distribution. There is kurtosis in the distributions of both models, but in the case of the Vasicek model that is observed more in the peakness of the distribution and in the case of the CIR model in the fat positive distribution tail.

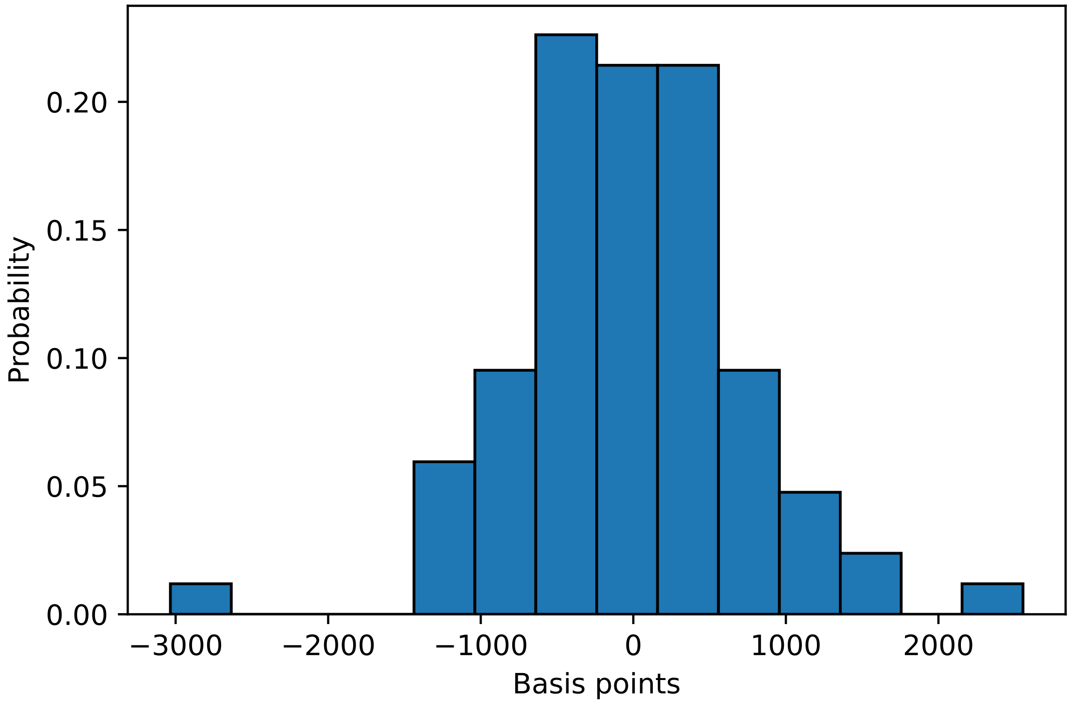

Figure 20: Probability distribution of monthly total ex-closing cost excess returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology

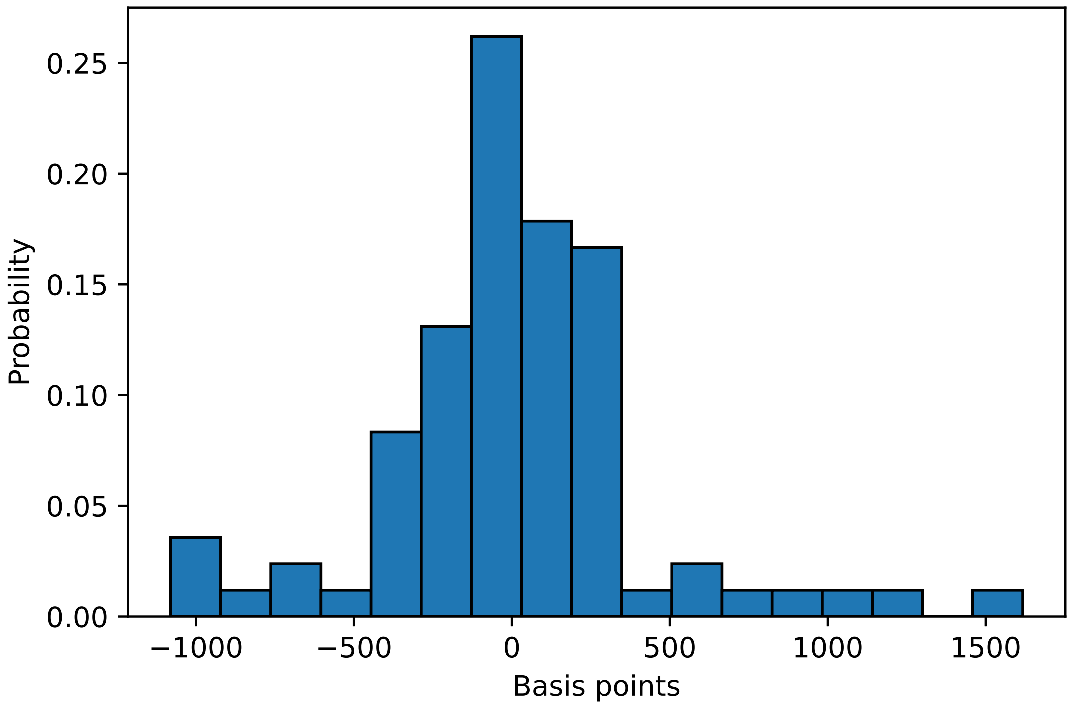

Figure 21: Probability distribution of monthly total ex-closing cost excess returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology

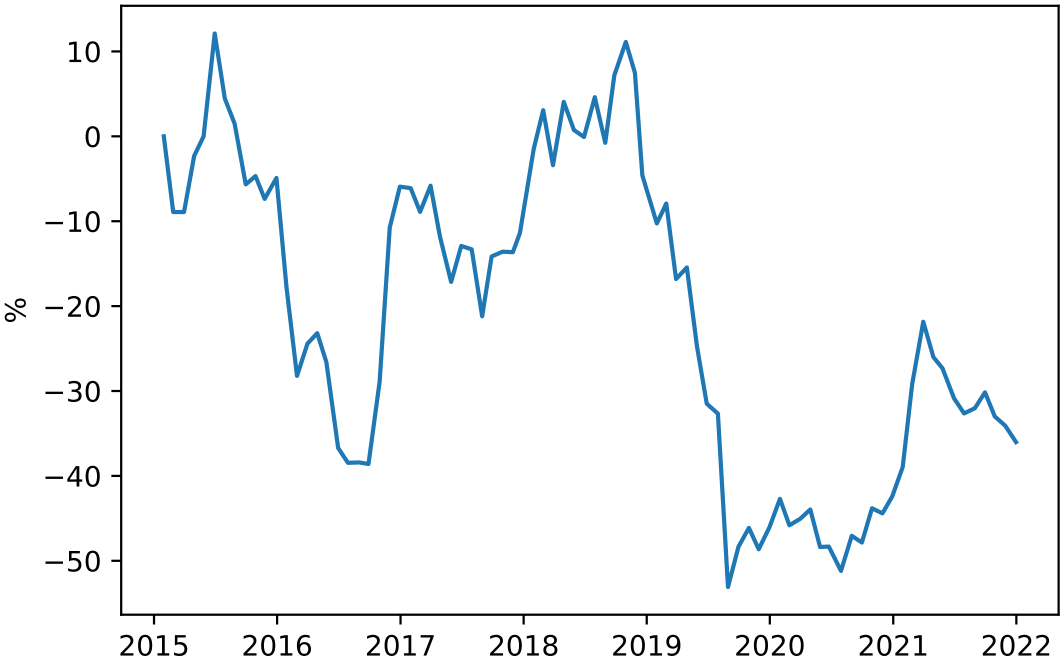

Figure 22: Cumulative returns with two-factor Vasicek model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology

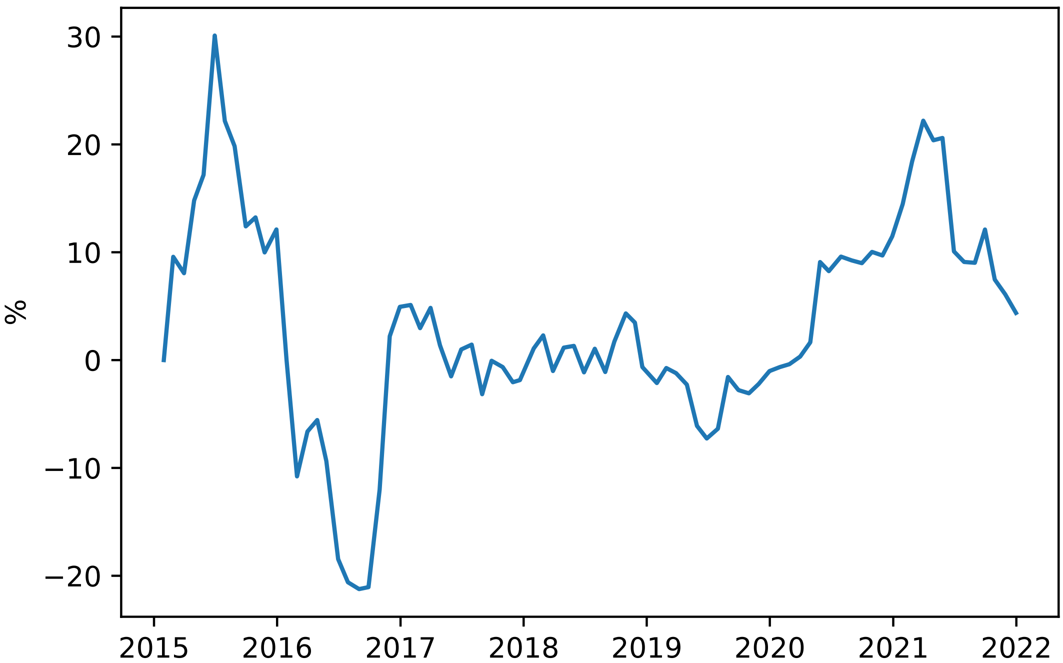

Figure 23: Cumulative returns with two-factor CIR model in subperiod 01/2015 - 12/2021, 5 bps mispricing limit and modified methodology

Finally, looking at Figures [22] and [23] it is clear that the performance of the models is still quite poor. There are still large drops and increases for the cumulative returns, although for the CIR model there is also a period of quite steady performance in 2017-2019. All in all, the CIR model appears to perform better than the Vasicek model in the subperiod with methodology modifications but the results are still inconclusive. I would conclude that the suggested methodology changes do not improve the performance of yield curve arbitrage. The conclusions about the mean of returns and skewness are the same as in the subperiod with unmodified methodology.

This subsection highlights the importance of methodology in the implementation of yield curve arbitrage. The subsection also somewhat indicates that the CIR model would have done better than the Vasicek model in the subperiod 01/2015 - 12/2021. This could be linked to the quick increase of rates in the late 2010s and the potential superiority of the CIR model to model the rates in those kind of rate environments.

8 Connection to hedge fund returns and risk factor analysis

In this section, I study the connection of my returns to the returns of fixed income arbitrage focused hedge funds. I also provide analysis on the market risk factors possibly related to yield curve arbitrage.

I apply the Credit Suisse Fixed Income Arbitrage Index for studying the returns of hedge funds focused on fixed income arbitrage[1]. This index is a subindex of Credit Suisse Hedge Fund Index and it consists of hedge funds that focus on relative value trading in global fixed income markets. Therefore, yield curve arbitrage could well be one of the strategies that the hedge funds in this index apply. Figure [24] provides the cumulative returns calculated from the index in 01/2012 - 12/2021.

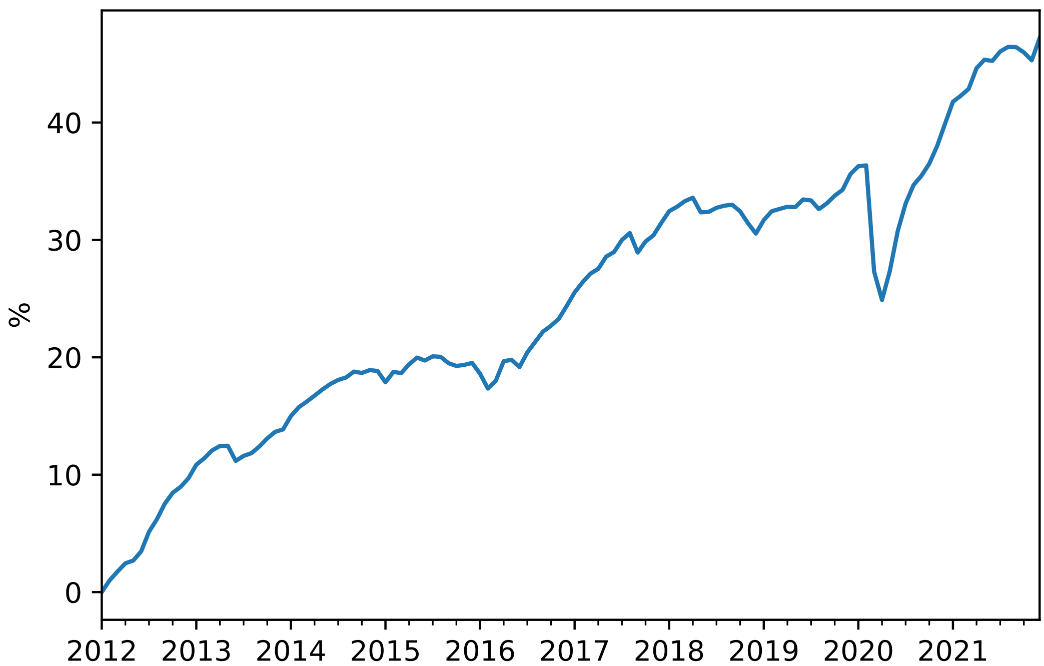

Figure 24: Cumulative returns for Credit Suisse Fixed Income Arbitrage Index in 01/2012 - 12/2021.

The cumulative returns are steadily increasing apart from the dip in around 2020, possibly caused by COVID related issues and the changes in the Fed's monetary policy. The performance of this index is quite different from the cumulative returns of my trading.

[1]This data is obtained Bloomberg Terminal.

Although this index consists of other strategies than yield curve arbitrage as well, it still indicates the lucrativeness of relative value trading in global fixed income markets in the 2010s. Table [8] provides the simplified summary statistics for the returns of the index. The mean of returns[1] is statistically significantly positive and the Sharpe ratio indicates good performance. On the other hand, the distribution of the index returns is strongly negatively skewed and has very high kurtosis. The minimum is also quite much larger than the maximum of returns. This suggests, that the index could be exposed to the tail risk which is in line with previous literature on fixed income arbitrage. The standard deviation of the returns is only around 0.9 % and the annual standard deviation is around 3.2 %. The ratio of negative monthly returns is quite low at around 20 % and there is positive serial correlation. Overall, the summary statistics for the index appear quite good, apart from the negative skewness of returns. The summary statistics are also comparable to the summary statistics of fixed income arbitrage hedge funds that Duarte et al. (2007) studied.

Table 8: Summary statistics of the returns of Credit Suisse (CS) Fixed Income Arbitrage Index in 01/2012 - 12/2021. The count of observations in this period is 119, as the first month 01/2012 is not included in return calculations. Mean is the mean of monthly excess returns calculated from the index by subtracting the risk-free return but witout subtracting any fees. t-stat is the Student's t-stastistic that relies on null hypothesis $\mu = 0$ and is calculated with heteroskedasticity and autocorrelation (HAC) robust standard errors. Skewness (skew) and kurtosis (kurt) are calculated with HAC robust standard errors. Mean, median, minimum, maximum, monthly standard deviation and annual standard deviation of the monthly excess returns are in basis points. Ratio neg is the ratio of months with negative excess returns relative to all months. Serial corr is the 1 lag autocorrelation measure. Sharpe ratios are annualized.

Index	Mean	t-stat	Med	Min	Max	Stdev (m.)	Stdev (ann.)	Skew	Kurt	Ratio neg	Serial corr	Sharpe
CS	32.961	6.797	41.013	-662.997	266.141	91.762	317.873	-3.700	29.538	0.218	0.341	0.359

Figure [25] presents the return distribution of Credit Suisse Fixed Income Arbitrage Index in 01/2012 - 12/2021. The figure reveals that there is a massive negative outlier in the return distribution of the index which explains the negative skewness of the returns. This return distribution portrays the kind of tail risk that the saying "picking nickels in front of a steamroller" refers to. The return distribution of the index indicates that the relative value investing has been lucrative but the investing has possibly been exposed to the tail risk. However, it is good to notice that this negative outlier return is still only -6 %, which is much less than the minimum returns for my reported trading variations.

[1]The returns are excess returns such that I first calculate the index return in a given month and then I subtract the risk-free return of the month. No fees or costs such as hedge fund fees are subtracted from the returns.

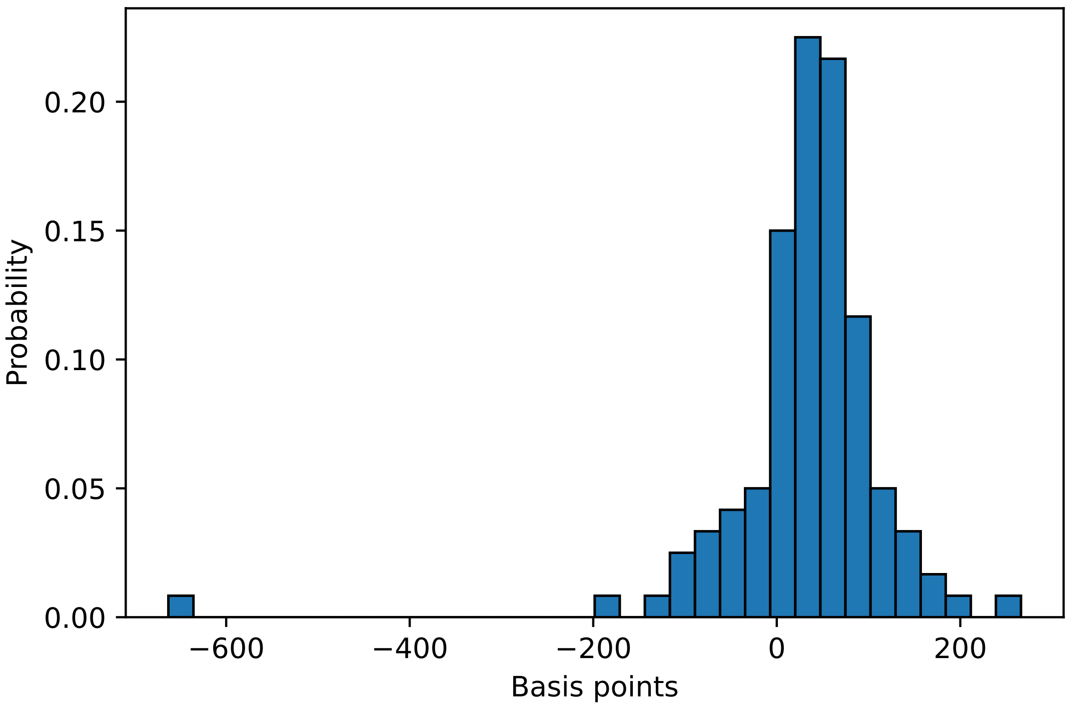

Figure 25: Probability distribution of monthly excess returns of Credit Suisse Fixed Income Arbitrage Index in 01/2012 - 12/2021.

In order to get a better idea on how my trading returns are related to the index returns, I calculate the Pearson's correlation coefficient. The correlation between the index returns and the full trading period returns with 5 bps mispricing limit and with the Vasicek modelling is -0.063 Similarly, the correlation between the index returns and the full trading period returns with 5 bps mispricing limit and with the CIR modelling is 0.145. Therefore, the correlations are somewhat small and have opposite directions. It appears that my trading returns are not particularly related to the index returns which is in line with the results of Duarte et al. (2007).

Next, I will provide the results of a multifactor regression that is formulated based on Duarte et al. (2007). I regress my returns to the common equity and bond market risk factors and control for credit and liquidity risk. I apply the excess returns of CRSP value-weighted market portfolio (R_m). SMB, HML, and MOM are the Fama-French small-minus-big, high-minus-low, and momentum market factors, respectively[1]. R_{SPBank} is the excess returns of the S&P 500 index of bank stocks. R_{INDB} is the excess total returns

[1]The excess market returns and the Fama-French factors are obtained from Kenneth French's website https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data$_library.html$.

of an iBoxx index of USD industrial AAA-BBB rated bonds. R_{BANKB} is the excess total returns of an iBoxx index of USD bank AAA-BBB rated bonds. R_{TR13}, R_{TR35}, R_{TR57} and R_{TR710} are excess total returns of the S&P/BGCantor USD index of Treasury bonds with maturities 1-3 years, 3-5 years, 5-7 years and 7-10 years, respectively[1]. Finally, I add my own measure of liquidity spread (LIQ) that is calculated as the monthly change in spread between 1-month USD LIBOR and 1-month USD OIS rates. The regressions are run for full trading period returns of the Vasicek and the CIR models with 5 bps mispricing limit and also for the Credit Suisse Fixed Income Arbitrage Index returns in the full trading period. According to Duarte et al. (2007), Fama-French factors and the bank stock index returns control for equity market risk. The industrial and bank bonds control for default/credit risk and the Treasury returns control for bond market risk. My measure of liquidity spread should control for the risk related to liquidity. Duarte et al. (2007) also state that "by including measures such as the excess returns on Treasury, banking, and general industrial bonds, or on banking stocks, we can control for the component of the fixed-income arbitrage returns that is simply compensation for bearing the risk of major (but perhaps not-yet-realized) financial events. This is because the same risk would be present, and presumably compensated, in the excess returns from these equity and bond portfolios." In other words, these provided factors should reflect the tail risk as well.

Table 9: Multifactor regression of full trading period returns. The dependent variables are the monthly total ex-closing cost excess returns of the two-factor Vasicek and the two-factor CIR models with 5 bps mispricing limit in the full trading period and the excess returns from Credit Suisse (CS) Fixed Income Arbitrage index in the full trading period. R_m is the excess returns of CRSP value-weighted market portfolio. SMB, HML, and MOM are the Fama-French small-minus-big, high-minus-low, and momentum market factors, respectively. R_{SPBank} is the excess returns of the S&P 500 index of bank stocks. R_{INDB} is the excess total returns of an iBoxx index of USD industrial AAA-BBB rated bonds. R_{BANKB} is the excess total returns of an iBoxx index of USD bank AAA-BBB rated bonds. R_{TR13}, R_{TR35}, R_{TR57} and R_{TR710} are the excess total returns of the S&P/BGCantor USD index of Treasury bonds with maturities 1-3 years, 3-5 years, 5-7 years and 7-10 years, respectively. LIQ is calculated as the monthly change in spread between 1-month USD LIBOR and 1-month USD OIS rates. The table reports the heteroskedasticity and autocorrelation (HAC) robust t-statistics, apart from the α that is provided in basis points.

Source	α	t-stat	R_m	SMB	HML	MOM	R_{SPBank}	R_{INDB}	R_{BANKB}	R_{TR13}	R_{TR35}	R_{TR57}	R_{TR710}	LIQ	R^2	Adj.R^2
Vasicek	-35.4355	-1.327	-0.176	0.725	0.409	0.054	-0.972	1.477	-0.032	2.311	-3.244	1.215	2.388	0.745	0.718	0.689
CIR	43.75749	0.553	-2.627	0.232	-0.350	-0.788	3.325	0.424	-0.637	0.776	-1.390	1.753	-1.599	-0.565	0.155	0.068
CS	31.10952	4.940	-1.721	0.012	0.013	1.021	1.630	1.521	3.234	0.576	-2.832	3.512	-4.377	0.876	0.571	0.526

Looking first at the results for the Credit Suisse Fixed Income Arbitrage Index, I observe that the index has had positive and statistically significant α in 01/2012 - 12/2021.

[1]The rest of the multifactor regression data is obtained from Refinitiv Eikon.

The R^2 measure indicates that around 50 % of the return variation in the index returns is explained by my regression model. This is quite much higher than the R^2 measures reported in Duarte et al. (2007). The index has had positive exposure to the bank bonds and to the 5-7 year Treasury bonds but also negative exposures to the 3-5 and 7-10 year Treasury bonds. These results are in line with Duarte et al. (2007). The index appears to have exposure to the credit/default risk and bond market risk. There is likely also exposure to the tail risk as suggested by the return distribution of the index but it is difficult to separate the tail risk exposure from the factors as it is by construction included in the factors. The aforementioned exposures do not take away the positive α so the index could be obtaining returns from some outside, isolated source of risk such as from the exposure to the curvature factor or demand shock risk, as suggested in Duarte et al. (2007) and in Vayanos and Vila (2009).

Looking then at the regression results for the Vasicek model, the first thing that catches my eye is the very high R^2 measure. Over 70 % of the return variation of the Vasicek model is explained by my regression model. The Vasicek model does not provide statistically significant positive α which could be expected based on the previous subsections. The Vasicek model appears to have similar exposure to the bond market risk as the Credit Suisse Fixed Income Arbitrage Index through exposure to the Treasury bond returns but no statistically significant exposure to either of the credit/default risk factors. It is somewhat surprising that the Vasicek model has similar exposure to the bond market risk as the Credit Suisse Fixed Income Arbitrage index but still the returns are that much weaker for the Vasicek model. The regression still indicates that the Vasicek model could be exposed to the curvature factor as there are positive and negative exposures to the Treasury returns as in Duarte et al. (2007).

Finally, looking at the regression results for the CIR model, I see that the CIR model does not provide statistically significant positive α and the R^2 measure is much lower than for the Credit Suisse Fixed Income arbitrage index or for the Vasicek model. Only about 15 % of the return variation of the CIR model returns is explained by the regression model, which is more in line with the results of Duarte et al. (2007). The CIR model also has the same kind of exposures to the Treasury bond returs as the Credit Suisse Fixed Income Arbitrage Index and the Vasicek model but these exposures are not statistically significant. I would interpret the regression results such that the CIR model could be modestly exposed to the curvature factor. The CIR model appears to have statistically significant negative exposure to the excess market returns and statistically significant positive exposure to the S&P 500 bank stock index returns. Therefore, CIR model appears to be exposed to the equity market risk. It is somewhat difficult to explain this exposure. This exposure could be exposure to the tail risk, as even though the return distribution for the CIR model in the full trading period was quite symmetric and even somewhat positively skewed, the range of minimum and maximum returns was very wide and the tail risk exposure could be stemming from there.

Both of the models seem to have market risk exposures. This is expected based on Duarte et al. (2007) but on the other hand, maybe the levels of market risk exposure are too high for both of the models. Especially the equity market risk exposure of the CIR model is peculiar. The hedging of neither of the models appears to provide particularly strong hedge against market risks.

9 Conclusions

In this concluding section I discuss the observed performance of yield curve arbitrage and provide possible explanations for it. I also link the performance to my hypotheses.

In general, I obtain unexpected performance. I obtain opposite performance compared to Duarte et al. (2007) and the previous master's theses. Compared to Duarte et al. (2007), I obtain lower means of returns, worse Sharpe ratios and statistically insignificant results. On the other hand, my confidence intervals for the mean of returns are wide and cover the returns reported in Duarte et al. (2007) so I can not conclude that my returns would be worse than those in Duarte et al. (2007). Some of my return distributions lack clear positive skewness, whereas others indicate positive skewness. The mispricing ranges are larger compared to Duarte et al. (2007) and the mispricings behave somewhat irregularly, i.e. there are no quick convergences and divergences but rather the mispricings remain positive or negative for long periods of time. The performance of yield curve arbitrage does not considerably change when studying the arbitrage in a subperiod and with methodology modifications. The multifactor regressions also provide somewhat inconsistent results as I observe contradictory exposures to the equity market risk and bond market risk factors between my two models. There is likely some exposure to the curvature factor of Duarte et al. (2007) and risks presented in Vayanos and Vila (2009). It is difficult to separate the possible tail risk exposure from the equity and bond market risk exposures but based on the wide range of minimum and maximum of returns, I would interpret that there is likely some tail risk exposure. All in all, I get very different results compared to Duarte et al. (2007) and I can not draw the same conclusions about the lucrativeness of yield curve arbitrage as they do. However, my results are inconclusive.

I believe that my performance is first and foremost explained by two major reasons: my methodology and the difficult interest rate environment in the 2010s. The rate environment regarding USD rates was difficult in the 2010s. First, the rates were affected by the aftermath of financial crisis, then the rates were very low for a long period of time after which the Fed started to increase the rates, causing them to converge in around 2019-2020. After that the global pandemic forced the rates down and the rates have slowly increased afterwards. In this kind of rate environment, my methodology relying on two simplistic short rate models may not be enough to hedge me against the key market risks. The short rate models that I apply assume that the term structure is driven by two risk factors and by hedging against these factors, I should be hedged against the market risk factors and level and slope changes of the term structure. These sort of assumptions have likely been too simplistic in the 2010s. My regression results provide support for this.

Another issue is in general related to the modelling, such that the short rate models may not be up-to-date enough to indicate the mispricings properly. The models rely on several assumptions regarding data, calibration and so forth that may distort the applicability of the models, especially in the rate environment where rates have evolved somewhat erratically. Also in the cases where rates have been low for long periods of time, suggesting flat term structures, it could be difficult to obtain exposure to the curvature factor and locate relative mispricings. Another explanation for the observed performance is that maybe I am trading the mispricings too seldom. A better application for the short rate models could be to apply them to daily data as in Stark (2020) and maybe observe quicker convergences and divergences.

As my performance changes when I am modifying the methodology and applying the models in the subperiod, the importance of robust methodology in yield curve arbitrage implementation is clear. It would likely require more sophisticated methodology to find the relative mispricings, react to them quickly and be hedged against market risks simultaneously. It is good to notice, that if the modelling is flawed, butterfly trades of yield curve arbitrage are constantly exposed to the risk of changes in curvature and slope of the term structure.

My regression results suggest that my trading could be exposed to the curvature factor. This connection is difficult to draw, as Duarte et al. (2007) provide very little intuition behind this idea. On the other hand, as the CIR model is exposed to the equity market risk and as the range of minimum and maximum for the returns is wide, it is likely that my trading is exposed to tail risk as well. It is difficult to extract the possible tail risk exposure from the exposure to the equity and bond market risk factors. My trading appears not to be exposed to the credit/default risk which is not that common for a fixed income trading strategy. On one hand, my factor risk hedging appears to provide quite good hedge against market risks as there are only a few significant sources of negative market risk exposures and for comparison the Credit Suisse Fixed Income Arbitrage Index has similar exposures. On the other hand, the exposures to certain market risk factors are high and my trading may have been too exposed to bond market risk and equity market risk.

Therefore, maybe yield curve arbitrage worked well in the study period of Duarte et al. (2007) because of their better methodological choices and the rate environment in the pre-financial crisis period of 1988-2004. The methodologies, technology and intellectual capital have since likely increased competition in fixed income arbitrage trading. On the other hand, the previous master's theses on yield curve arbitrage by Karsimus (2015) and Stark (2020) have found strong performance for yield curve arbitrage in EUR swap markets. Karsimus (2015) applies data from 2002-2015 and uses somewhat different methodology and Stark (2020) applies daily data. Therefore, it is difficult to directly compare my results to their results. Stark (2020) finds interesting results about applying machine learning to yield curve arbitrage suggesting that machine learning methodology could be superior to the pure short rate modelling methodology. This lends support for the idea about the importance of intellectual capital.

Although the Credit Suisse Fixed Income Arbitrage index is for global fixed income markets and covers widely different fixed income arbitrage strategies, it still indicates the lucrativeness of relative value trading in global fixed income markets in the 2010s. Hedge funds in this index likely have the best possible intellectual capital available and the hedge funds can flexibly invest in global fixed income markets. Therefore, the notion about flawed methodology in a difficult rate environment, could indeed be valid.

As to my hypotheses, my hypothesis 1 suggested that *"Yield curve arbitrage generates positive excess returns in the USD interest rate swap market in 2012-2021"*. I do not get conclusive support for the hypothesis but I can neither reject it, as my confidence intervals for the mean of returns are wide, covering positive and negative returns.

My hypothesis 2 suggested that *"The return distribution of yield curve arbitrage is positively skewed with a high kurtosis in the USD interest rate swap market in 2012-2021"*. I get partial support for this hypothesis as there are trading variations in the subperiod for which I obtain clear positive skewness and higher than normal kurtosis. Therefore, I can not reject hypothesis 2.

My hypothesis 3 suggested that *"The returns of yield curve arbitrage are lower in the USD interest rate swap markets in 2012-2021 than in the study period Duarte et al. (2007) in 1988-2004"*. Again, I get some support for this but my results are inconclusive in the sense that my confidence intervals for the mean of returns cover in most cases the reported returns of Duarte et al. (2007) so it could be the case that my trading performance does not differ from that reported by Duarte et al. (2007).

My hypothesis 4 suggested that *"The two-factor CIR model provides worse yield curve arbitrage returns than the two-factor Vasicek model in the USD interest rate swap markets in 2012-2021"*. While this is true in the full trading period, in the subperiod the CIR model beats the Vasicek model. However, the differences are quite small and the confidence intervals for the mean of returns overlap between the models in different trading variations. There are more clear differences in skewness measures and in confidence intervals for skewness but in these cases as well my results are inconclusive. Therefore, I obtain support for and against the hypothesis.

Finally, I reject my hypothesis 5 suggesting *"Yield curve arbitrage provides statistically significant multifactor alpha in the USD interest rate swap market in 2012-2021, when the returns are explained by the same market risk factors as in Duarte et al. (2007)"*. This appears not to be the case.

All in all, I have tested multiple variations in methodology and I do not get results that would be in line with Duarte et al. (2007). I am quite sure that the trading and valuation sections of my methodology are working correctly and the possible errors in my methodology are elsewhere. My implementation is open for anyone to familiarize in my GitHub repository *ycarbitrage*, username *mlammi*.

Finally, there is obviously the possibility, that my results are correct and the observed performance is the correct performance of yield curve arbitrage in the USD interest rate swap markets in the 2010s. The lack of references makes it hard to compare my results. With all this said, maybe an arbitrage strategy such as yield curve arbitrage, that faced big problems in the early 2000s and that relies on applying term structure modelling methodology, that may not be up-to-date anymore, really had bad and inconsistent performance in the USD swap markets in the 2010s that could not be fixed by improved methodology.

10 References

Ager-Wick, B. and N. Luong (2020). "Risk and Return In Yield Curve Arbitrage". Master's Thesis. Norwegian School of Economics.

Brigo, D. and F. Mercurio (2007). *Interest Rate Models-Theory and Practice: with Smile, Inflation and Credit*. Springer Science & Business Media.

Cox, J.C., J.E. Ingersoll, and S.A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 53(2): 385–407.

De Figueiredo Neto, C. (2013). "Fixed-income arbitrage strategies: swap spread arbitrage and yield curve arbitrage". Master's Thesis. University of Oulu, Business School.

Duarte, J., F. A. Longstaff, and F. Yu (2007). Risk and Return in Fixed Income Arbitrage: Nickels in Front of a Steamroller? *The Review of Financial Studies*, 20(3): 769–811.

Duffie, D. and K. J. Singleton (1997). An Econometric Model of the Term Structure of Interest-Rate Swap Yields. *The Journal of Finance*, 52(4): 1287–1321.

Eurostat (2022). *Glossary: In-sample vs. out-of-sample forecasts*. URL: `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:In-sample_vs._out-of-sample_forecasts` (visited on 03/20/2022).

Fontaine, J.-S. and G. Nolin (2019). Measuring Limits of Arbitrage in Fixed-Income Markets. *Journal of Financial Research*, 42(3): 525–552.

Fung, W. and D. A. Hsieh (2002). Risk in Fixed-Income Hedge Fund Styles. *The Journal of Fixed Income*, 12(2): 6–27.

Gregory, J. (2014). *Central Counterparties: Mandatory Central Clearing and Initial Margin Requirements for OTC Derivatives*. John Wiley & Sons Ltd.

Gromb, D. and D. Vayanos (2010). Limits of Arbitrage. *Annual Review of Financial Economics*, 2: 251–275.

Hirsa, A. (2013). *Computational Methods in Finance*. Taylor & Francis Group.

Hull, J. (2017). *Options, Futures and Other Derivatives*. 9.0. Pearson Education Limited.

Kakushadze, Z. and J.A. Serur (2018). *151 Trading Strategies*. Palgrave Macmillan.

Karsimus, L. (2015). “Yield Curve Arbitrage in the EUR Swap Rates Market”. Master’s Thesis. Aalto University School of Business.

Kelly, B. T. and H. Jiang (2012). *Tail Risk and Hedge Fund Returns*. Tech. rep. 12-44. Chicago Booth Research Paper.

Kim, D. H. and A. Orphanides (2012). Term Structure Estimation with Survey Data on Interest Rate Forecasts. *The Journal of Financial and Quantitative Analysis*, 47(1): 241–272.

Liu, J., F. A. Longstaff, and R. E. Mandell (2006). The Market Price of Risk in Interest Rate Swaps: The Roles of Default and Liquidity Risks. *The Journal of Business*, 79(5): 2337–2359.

Nelken, I. (2005). *Hedge Fund Investment Management*. Elsevier Science & Technology.

OpenGamma (2013). *Interest Rate Instruments and Market Conventions Guide*. URL: `https://quant.opengamma.io/Interest-Rate-Instruments-and-Market-Conventions.pdf` (visited on 04/24/2022).

Patton, A. J. (2009). Are Market Neutral Hedge Funds Really Market Neutral? *The Review of Financial Studies*, 22(7): 2495–2530.

Pedersen, L. H. (2015). *Efficiently Inefficient : How Smart Money Invests and Market Prices Are Determined*. Princeton University Press.

Shreve, S.E. (2004). *Stochastic Calculus for Finance II - Continuous-Time Models*. Springer Science+Business Media, Inc.

Smith, D. J. (2013). Valuing Interest Rate Swaps Using Overnight Indexed Swap (OIS) discounting. *The Journal of Derivatives*, 20(4): 49–59.

Stark, L. (2020). “Yield Curve Arbitrage In EUR Swap Rates - A Hybrid Neural Network Approach”. Master’s Thesis. Aalto University School of Business.

Tuckman, B. and A. Serrat (2012). *Fixed Income Securities: Tools for Today’s Markets*. 3.0. John Wiley & Sons, Inc.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2): 177–188.

Vayanos, D. and J-L. Vila (2009). *A Preferred-Habitat Model of the Term Structure of Interest Rates.* Tech. rep. National Bureau of Economic Research.

11 Appendices

11.1 Python code implementation

The actual code implementation of the methodology can be found from GitHub repository *ycarbitrage*, username *mlammi*. It provides separate .py-files for function-oriented code for different parts of methodology and separate Jupyter Notebooks, in which the actual arbitrage implementation is done. The original data is also included in the repository.

11.2 Additional master's theses on yield curve arbitrage

In this section I present two additional master's theses on yield curve arbitrage. I familiarized myself with these theses but I did not apply them as benchmarks the same way I applied theses of Stark (2020) and Karsimus (2015).

De Figueiredo Neto (2013) studies swap spread arbitrage, which is another fixed income arbitrage strategy, and yield curve arbitrage in USD constant maturity swap[1] markets from 1988 to 2011 replicating the paper of Duarte et al. (2007) in methodology. His sample extends the period of Duarte et al. (2007) from 2004 to 2011 and makes some changes to the multifactor regressions. He finds results similar to Duarte et al. (2007) and e.g. concludes, that the financial crisis did not have significant effects on the lucrativity of yield curve arbitrage. The idea of this thesis is very close to my basic idea of expanding the study of Duarte et al. (2007) to a more recent data set and making some modifications and additions to the methodology. However, as I am planning to add another 10 years to the data set and to make additions to the methodology e.g. in term structure modelling, our theses are similar in idea yet still different in execution. A major difference between De Figueiredo Neto (2013) and my thesis is that De Figueiredo Neto (2013) uses constant maturity swap data, whereas I use fixed-for-floating ("vanilla") swap data[2]. However, I am not fully sure if De Figueiredo Neto (2013) actually uses constant maturity swap data, or if he is just misusing the term and is actually using vanilla swap data.

Ager-Wick and Luong (2020) study yield curve arbitrage in USD and EUR swap markets in 2006-2020. They mostly replicate the paper of Duarte et al. (2007) in trading and hedging methodology, but rather than using the 2-factor Vasicek short rate model,

[1] See more about constant maturity swaps e.g. in Hull (2017).

[2] As previously mentioned, my understanding is that Duarte et al. (2007) also uses vanilla swap data.

they use an affine 3-factor term structure model. As Stark (2020), they also use daily data. Model parameters are estimated through a numerical optimization routine as in Duarte et al. (2007). Ager-Wick and Luong (2020) find similar results as Duarte et al. (2007). Ager-Wick and Luong (2020) conclude that "there is potential for risk-adjusted excess returns, but that this potential can only be unlocked through astute modeling and deep knowledge of market dynamics".

11.3 Other trading on term structure

In this section, I present two alternatives for trading the term structure. Trading the curvature of the yield curve or term structure in general is just one form of trading the term structure. For example, Pedersen (2015) also presents the trading of level and slope of the term structure.

Trading the level of the term structure means that the trader is expecting the rates to go up or down. If the trader expects the rates go down, the trader can short rates or buy bonds. Similarly, if the trader expects the rates to go up, the trader can go long rates or sell bonds. The trading of the level of the term structure is usually guided by macroeconomic views about e.g. central bank policy. Therefore, trading the level of the term structure is quite different from relative value trading such as yield curve arbitrage.

Trading the slope of the term structure also has two variations. The trader may expect the term structure to *steepen*, i.e. short rates to fall and long rates to increase, in which case the trader can short the rates in the short end of term structure and go long the rates in the long end of term structure. Alternatively the trader can buy short maturity bonds and sell long maturity bonds in trading the steepening. If the trader expects the term structure to flatten, the trader can take the trades the other way round. Trading the slope of the term structure can be duration-hedged such that the trader is not exposed to the level changes of term structure and is more exposed to the slope changes.

Nelken (2005) notes, that trading the term structure can be very difficult. The trader should have a view about the development of rates in the future but at the same time the trader needs to address whether the expectations about changes in rates are already priced into current rates. This requires studying the forward rate curve and historical yield curves and discount curves.

11.4 Interest rate swap counterparties and arbitrageurs

In this section I present extensively the possible interest rate swap counterparties and fixex income arbitrageurs.

Interest rate swaps are over-the-counter (OTC) fixed income derivative instruments and thus it is important to know the types of counterparties in interest rate swaps. In theory, any two counterparties (companies, institutions or individuals) could set up an interest rate swap. In practice though, individuals do not set up interest rate swaps, as legislation and practical swap arrangements would make it very difficult, if not impossible. Individuals and institutions/companies (financial or non-financial) neither set up interest rate swaps for the very same reasons. To take things one step further, it is also unlikely that two non-financial companies/institutions would directly arrange an interest rate swap[1] but rather according to Hull (2017) there is almost always a financial intermediary in between to facilitate the swap. The financial intermediary would facilitate the swap such that the non-financial companies/institutions are not counterparties to one another but they rather have the intermediary as their counterparty. The financial intermediary could be e.g. a commercial bank or an investment bank. Hull (2017) presents how the financial intermediary earns some profit in facilitating the swap and many large financial intermediaries are prepared to act as market makers, i.e. they can enter into swaps with counterparties without taking the usual offsetting transactions.

In the interbank markets, fixed-for-floating interest rate swaps and other standardized OTC derivatives are commonly centrally cleared through central clearing counterparties (CCPs)[2]. Two swap counterparties would still agree on the swap and have the derivative agreement between themselves but the CCP would handle the initial margin collateral and the cash flows of the swap. There is commonly a collateral also in the interest rate swaps between financial intermediary and a non-financial company/institution. There are also non-centrally cleared swaps that do not require margin collaterals but in this thesis I assume that all the swaps in some form require initial margin. The collateral is usually paid some risk-free return, i.e. collateral rate, and Tuckman and Serrat (2012) mentions

[1] To my understanding, this is not impossible though. Two non-financial companies/institutions could arrange an interest rate swap but it would require extensive legal agreements etc. Therefore, large non-financial companies/institutions, such as Fortum and Nokia, would mainly set up interest rate swaps between themselves.

[2] LCH Clearnet is the largest CCP for interest rate swaps.

that in the US the collateral rate is usually the Fed Funds rate. I also assume that in the swaps that I apply the collateral rate is paid on the collateral. Gregory (2014) provides a comprehensive guide to swap collaterals and other such practicalities.

Thus, the financial intermediaries that facilitate interest rate swaps for non-financial companies/institutions are themselves in connection with CCPs. Large non-financial companies, hedge funds and other large financial/non-financial institutions may also use CCPs for their interest rate swap transactions. The legislation in some countries even requires using CCPs, at least for the standardized interest rate swaps. However, there are still non-centrally cleared interest rate swaps, but legislation concerning them has also developed since the financial crisis, as noted in Gregory (2014). The benefit of central clearing is that it lowers the risk related to swaps, which is why the non-centrally cleared swaps agreed between two non-financial companies/institutions are not very common.

The group of possible swap counterparties limits the group of possible arbitrageurs in yield curve arbitrage. Duarte et al. (2007) talk about hedge funds and proprietary trading desks as the main category of arbitrageurs in yield curve arbitrage. In general, it is widely known that a major category of hedge fund investment styles is the so-called relative value trading, which covers yield curve arbitrage as well. Possible arbitrageurs in yield curve arbitrage include e.g. the proprietary trading desks of investment banks or large commercial banks, pension funds or fixed income mutual funds. As mentioned in Pedersen (2015), trading the term structure in general is quite common in trading financial institutions, so these institutions could be involved in yield curve arbitrage. Large corporations with their own trading desks could also be involved in yield curve arbitrage. Basically any organization that has access to interest rate swaps with reasonable terms, costs and liquidity could in theory trade yield curve arbitrage. The access to the OTC markets and the ability to find eligible counterparties for swaps or the possibility to open swaps with financial intermediaries with reasonable terms, costs and liquidity may make it difficult for some participants to implement yield curve arbitrage. Financial companies most likely have the best possibilities to implement yield curve arbitrage in the sense of finding liquid swap contracts. However, no doubt hedge funds are the most prominent operators in yield curve arbitrage since they likely have the biggest volumes and best risk tolerances for implementing yield curve arbitrage. It is also good to remember the required intellectual capital for implementing yield curve arbitrage that Duarte et al.

(2007) emphasizes: not all organizations have the required expertise to implement yield curve arbitrage successfully.

11.5 The concept of term structure of interest rates

In this section, I present the concept of term structure of interest rates. Definitions for year fractions and zero coupon bonds are assumed to be known.

In general use, the concept of term structure of interest rates is often poorly defined. It is often used analogously with yield curve, although there are clear differences between these two concepts. In the fixed income context, the concept of yield is often used for instruments for which yield-to-maturities can be derived. As Hull (2017) notes, yield curve simply depicts the yield-to-maturities for different maturities. Yield-to-maturities are internal rates of return and as such mathematically non-robust ways to express the return or yield for an instrument. Yield curves are also issuer- and instrument-specific, e.g. yield curve derived from the bond prices issued by Bank of Japan.

Term structure of spot interest rates depicts spot interest rates for different maturities. Some sources may use term structure of interest rates referring to e.g. term structure of forward interest rates, but in this thesis I refer to the spot rates. These spot rates are the required rates of return, that the markets demand for different maturities and not just the internal rates of return for some fixed income istrument as with yield-to-maturities. The spot rates are derived in a mathematically robust way, applying the previously presented bootstrapping and zero coupon bond prices. The term structure is also issuer- and instrument-specific. However, according to Brigo and Mercurio (2007) term structures of spot interest rates are commonly derived from liquid and low-risk base instruments and as such they plot the required return over different periods bypassing concerns related to liquidity or riskiness.

Let me now define the necessary spot rates.

Definition 11.1 (Simply compounded spot rates)**.** The simply compounded spot interest rate prevailing at time t for the maturity T is denoted by $L(t,T)$ and is the constant rate at which an investment has to be made to produce an amount of one unit of currency at maturity, starting from $P(t,T)$ units of currency at time t, when accruing occurs proportionally to the investment time. In formula:

$$L(t,T) = \frac{1 - P(t,T)}{\tau(t,T)P(t,T)}$$

Definition 11.2 (Annually compounded spot rates)**.** The annually compounded spot interest rate prevailing at time t for the maturity T is denoted by $Y(t,T)$ and is the constant rate at which an investment has to be made to produce an amount of one unit of currency at maturity, starting from $P(t,T)$ units of currency at time t, when reinvesting the obtained amounts once a year. In formula:

$$Y(t,T) = \frac{1}{P(t,T)^{\frac{1}{\tau(t,T)}}} - 1$$

There is also a third category of spot rates that is not directly needed in the definition of term structure of interest rates but that is important nevertheless.

Definition 11.3 (Continuously compounded spot interest rate)**.** The continuously-compounded spot interest rate prevailing at time t for the maturity T is denoted by $R(t,T)$ and is the constant rate at which an investment of $P(t,T)$ units of currency at time t accrues continuously to yield a unit amount of currency at maturity T. In formulas:

$$R(t,T) = -\frac{lnP(t,T)}{\tau(t,T)}$$

Zero coupon bond prices and the spot rates are linked as explained above. It is therefore important that the zero coupon bond prices are derived from the markets in a mathematically robust way relying on bootstrapping and no-arbitrage. According to Tuckman and Serrat (2012), "relying on the law of one price, discount factors extracted from a particular set of bonds can be used to price other bonds, outside the original set. A more complex but more convincing relative pricing methodology, known as arbitrage pricing, turns out to be mathematically identical to pricing with discount factors."

Now I have the definitions I need for defining the term structure of interest rates. Brigo and Mercurio (2007) define term structure of spot interest rates as:

Definition 11.4 (Term structure of spot interest rates)**.** The term structure of spot interest rates at time $t = 0$ is the graph of the function:

$$T \rightarrow \begin{cases} L(0,T), 0 < T \leq 1(years) \\ Y(0,T), T > 1(years) \end{cases}$$

Liquid and low-risk alternatives for deriving the term structure are e.g. interest rate swaps, eurodollar futures, interbank rates and government rates. According to Tuckman and Serrat (2012) government bonds and interest rate swaps are applied as term structure benchmarks these days. Both of these fixed income instrument categories have liquid markets and they incorporate information about interest rates that is common to all fixed income markets. Interest rate swaps are preferred over government bonds. This is because bonds are not fungible collections of cash flows. For example, on-the-run bonds may be priced at premium relative to new bonds because of their superior liquidity. It is not possible to apply one bond yield curve that is built from certain benchmark bonds to the valuation of other bonds because of these pricing inconsistencies. On the other hand, swaps are fungible and it is possible to apply one single swap curve to the valuation of entire swap book. Therefore, global fixed income markets apply interest rate swaps as benchmarks or base curves and follow the spreads relative to the swap curve.

11.6 Other possible issues in methodology

In this section I present a collection different possible issues in my methodology that I did not present in chapter 6.

A possible issue in my methodology is caused by the different dates related to OIS- and swap trading. There are roughly speaking 4 types of dates related to these rates and corresponding instruments: *transaction (trade) date*, *settlement date*, *payment dates* and *rate reset dates*. The transaction date is the date on which the actual trade occurs and generally the ownership changes. The settlement date usually happens with a small lag after the transaction date and it is the date by which margins should be provided and e.g. for swaps interest accrual starts. Payment dates are the dates, when payments are officially exchanged during the lifetime of the swap and they are agreed relative to the settlement date. Interest rate swaps and OIS contracts usually have 2 days for the *settlement lag*. There are usually also lags in payments (*payment lag*) during the lifetime of the swap, i.e. the payments would not be in practice exchanged on the payment dates but after a few days. The payment dates are the official dates agreed on the swap contract and year fractions and interest accrual are calculated based on these dates. Perhaps the most important dates during the lifetime of a swap are the rate reset dates. These are the dates when the floating rate/rates in an interest rate swap are reset. At the beginning of the swap, the first rate reset date is the same as the transaction date and during the lifetime of the swap the reset dates usually occur a few days before nearest payment date.

Example 11.1. Consider I obtain a receiver 3-year interest rate swap receiving semiannual fixed leg and paying quarterly floating leg. Let me denote the current date as t. I would obtain the ownership of the receiver swap at t, but I would obtain the first fixed leg payment after 6 months + 2 days. The 2 days are the lag for settlement date $(t + 2)$. There could be a few more days after 6 months + 2 days for me to actually obtain the first payment, because of the possible payment lag.

The settlement dates, payment dates or rate reset dates are in no way accounted for in Duarte et al. (2007). Similarly, Smith (2013) does not talk about the dates and makes quite strong simplifications regarding e.g. the payment periods of OIS contracts by assuming them to be quarterly. As I already explained in subsection 6.1 and according to OpenGamma (2013), an OIS contract usually makes one payment at the maturity of the OIS in maturities below 12 months and payments in intervals of 12 months for OIS contracts of maturities longer than 12 months. OIS contracts are the basis for OIS discount factors and as OIS contracts may have different payment- and reset dates than the swap contracts, this could affect the rigorous discounting of swap payments.

Example 11.2. Consider I have an ongoing payer swap, that has 7 months to maturity. The fixed leg of the swap would be paying in 1 months and in 7 months, whereas the floating leg of the swap would be paying in 1 months, 4 months and 7 months. Therefore, I would need forward rates for (0,1)-period, (1,4)-period and (4,7)-period and OIS discount factors for 1-, 4- and 7 months. I would obtain the OIS discount factors and OIS-consistent implied LIBOR forwards with market OIS data, but these market OIS contracts have their own settlement lags and payment dates and they may not match the payment dates of my ongoing swap.

As explained in section 5, in practice my dataset is formed such that I use the last trading days of the months for which there is data available. As I do not have specific information on the actual payment dates of swaps, I rely on simplifying assumptions:

- I valuate swaps by assuming that rate reset- and payment dates are X months in the future and I solve the dates and year fractions with Python's ready packages[1]
- I move from month to month by following my time index of dates, i.e. the last trading days of months in my data. Therefore, the payment- and rate reset dates, that I assume in valuation of swaps can be different from the dates in my time index of months
- I neglect all lags and simply assume that transaction dates are settlement dates and payment dates do not have any lags

I do not handle the date issues, but they should be considered in real life, should I actually implement the arbitrage in the markets and valuate ongoing swaps. However, the impact of these issues should be quite small in practice, such that the effect on the returns should be negligible. On the other hand, as the returns of yield curve arbitrage are not that big, if these date issues cause effects of basis points or such on the returns, the effects can be relatively quite big. Thus, the reader should be aware of these issues related to dates.

Another two major issues are the lack of valuation adjustments (XVA) and the possibly overly simplified assumptions related to margin requirements. Valuation adjustments and margin requirements are closely linked. Valuation adjustments are the different adjustments made to the theoretical dirty price value of a derivative contract and they are made because of credit risk, market risk and funding cost of margins or regulatory capital. For example, if an interest rate swap would not be sufficiently collateralized, a credit valuation adjustment (CVA) would be made to the theoretical value of the IRS to reflect the possible credit losses. On the other hand, if a collateral is posted as part of an IRS, it may be funded with external capital and if the rate of return paid on the collateral is smaller than the funding cost, this should be accounted for as a margin valuation adjustment (MVA). Overall, as highlighted in Gregory (2014), the margin- and

[1] To be precise, I use the dateutil-package and its relativedelta-module with relativedelta-function. See more about the Python-code implementation in Appendices.

valuationd adjustment issues are very case specific and they are especially important with non-centrally cleared transactions. The correct way to tackle these issues in these kind of empirical arbitrage studies is to use conservative values for costs, margins and values of derivative instruments. In this sense, my methodology may still be too optimistic.

With regard to the global optimization in short rate model calibration, I tried to apply global optimization methods in Python's Scipy package, but for some reason the methods failed to provide any reasonable results. I also tried to apply Python based Pyomo open-source optimization modeling language with BARON solver but I faced issues with this combination as well. Lastly, I tried to apply Pyomo with SCIP solver, but this combination also failed in providing any reasonable results. There was no more obtainable solutions for global optimization with Python and I would needed to either change the programming language or to obtain interfaces to some non-licensed optimization packages.

There may also be other issues in my methodology, but the aforementioned issues in the thesis and in this section are maybe the most important ones.

11.7 Additional trading performance analysis

In this section I provide some additional trading performance analysis related to the full trading period 01/2012 - 12/2021.

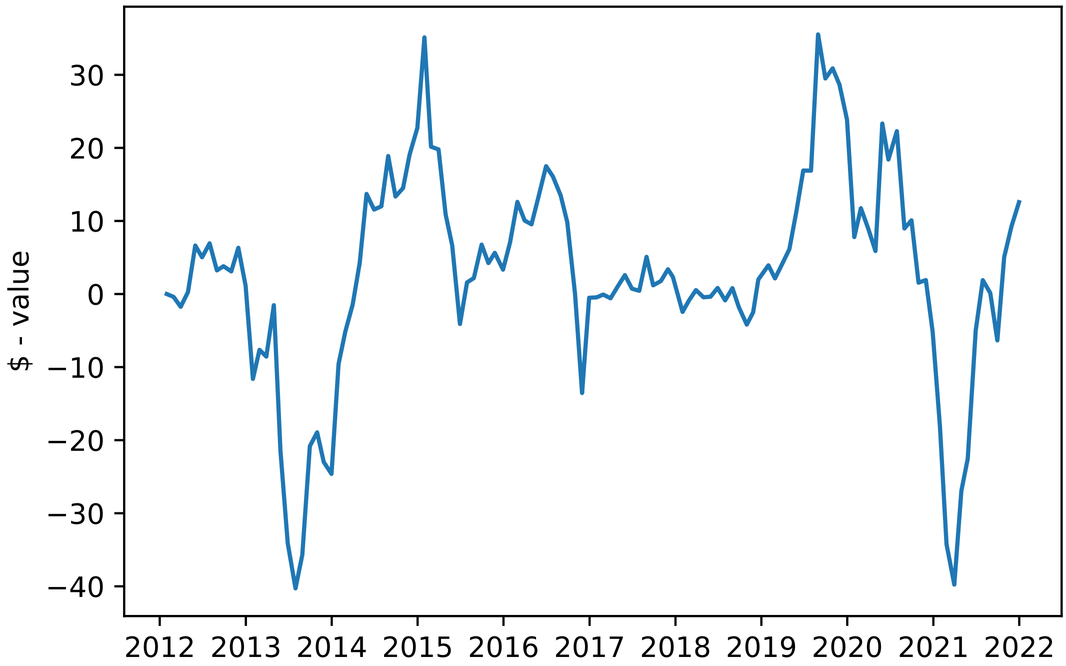

Figure 26: Portfolio \$-value in full trading period 01/2012 - 12/2021 with two-factor Vasicek model, 5 bps mispricing limit

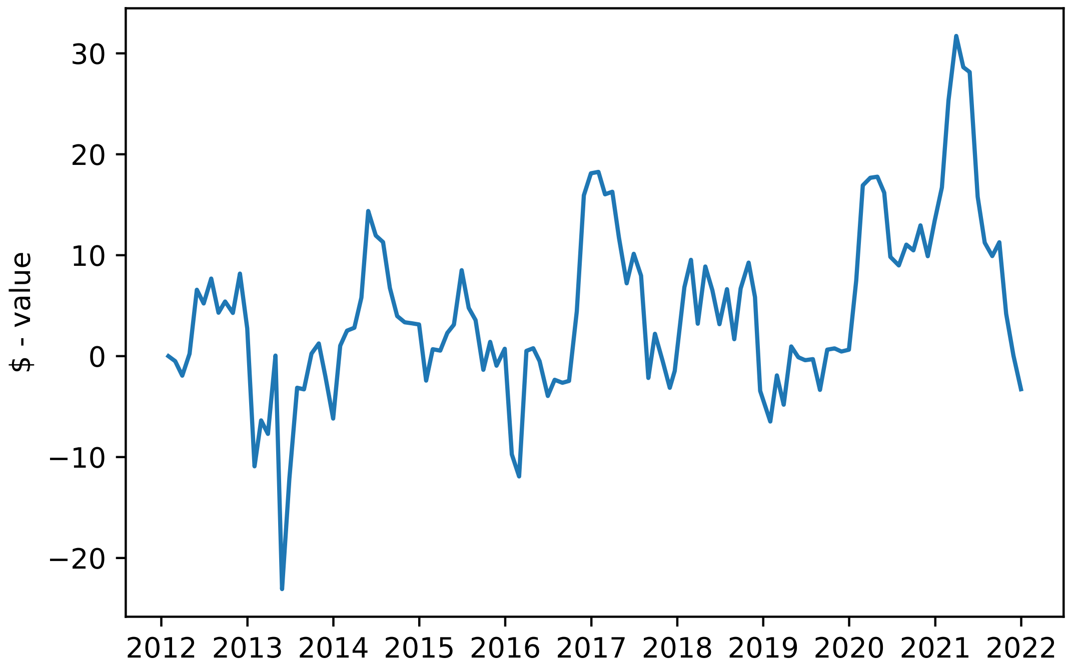

Figure 27: Portfolio \$-value in full trading period 01/2012 - 12/2021 with two-factor CIR model, 5 bps mispricing limit

Figures [26] and [27] provide the time series plots of portfolio (dirty price) \$-values with the two models. In other words, these plots present the value of all open arbitrage trades by summing them up in the given months. There have been significant drops in portfolio value with the Vasicek model in 2012-2013 and in 2020-2021 but also significant increases in value following the drops. The value variation is smaller with the CIR model but there are also periods of value drops followed by periods of value increases. The negative values with CIR have been smaller than with Vasicek and on average the CIR model has had more time of positive portfolio value than the Vasicek model.

These plots suggest, that yield curve arbitrage has not been a profitable strategy in the full trading period with the chosen strategy specifications, as the total portfolio value should be positive most of the time. As discussed in subsection 6.5, the trader of yield curve arbitrage would in reality be interested in the net payments during the holding period of the arbitrage trade and the ultimate dirty price value of the arbitrage trade at the end of the holding period. The calculation of returns based on consecutive total dirty price values combines these two return elements but the evolvement of the total dirty price value does not separate the effect of latest rate fixings and projected forwards.

Therefore, the portfolio \$-value plots alone do not tell the full story about the performance of yield curve arbitrage, as there could be drops in total portfolio \$-value if positions are closed and new positions are not opened or if the mispricings are small and they converge fast. In general though, the total portfolio \$-value should remain positive in the trading period, if the assumptions about model-market convergence were to be true, i.e. negative \$-values indicate bad perfomance. Steadily increasing \$-values are not to be expected since arbitrage trades are continuously opened and closed. The total portfolio \$-values are an indication for the yield curve arbitrage trader whether the market rates are evolving as expected and it can be utilized as a risk control metric, if the trader e.g. wonders the possible losses that would occur if the arbitrage trades were to be closed instantly.

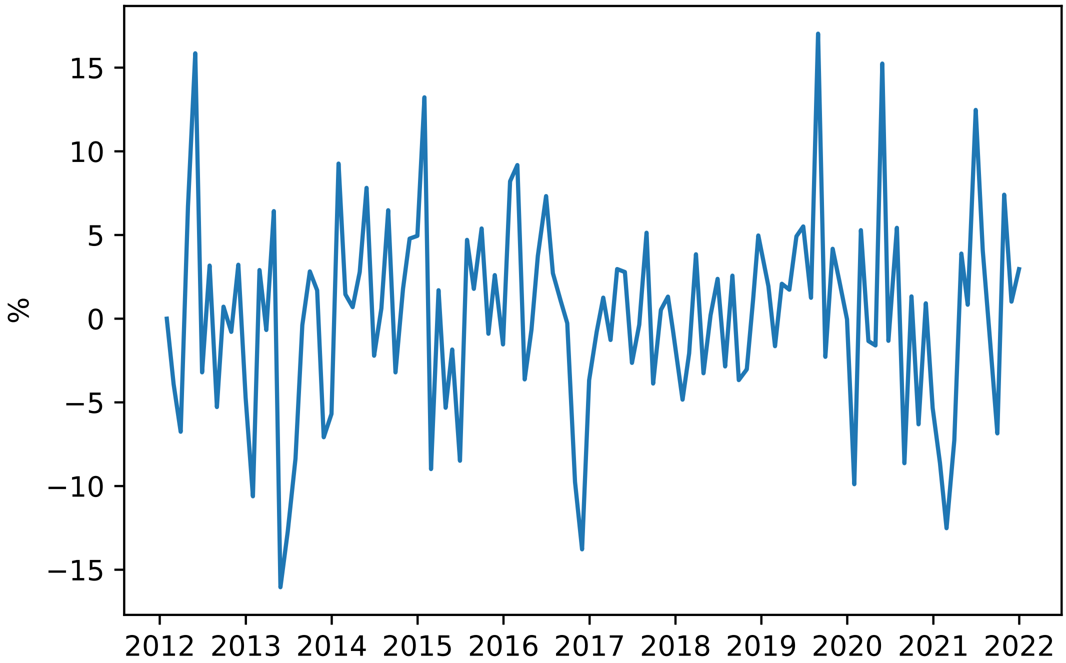

Figure 28: Monthly total ex-closing cost excess returns with two-factor Vasicek model in the full trading period 01/2012 - 12/2021, 5 bps mispricing limit

Figures [28] and [29] demonstrate how the monthly total ex-closing cost excess returns evolve in the 5 bps mispricing limit variations. These figures clearly indicate bad performance for the yield curve arbitrage. There is a lot of variation and consistently months with large negative returns with both Vasicek and CIR models. The monthly returns should remain mostly positive throughout the trading period, if the arbitrage was to behave as expected. Both the models have quite similar patterns in monthly returns, i.e. large positive and negative returns are timed to the same periods.

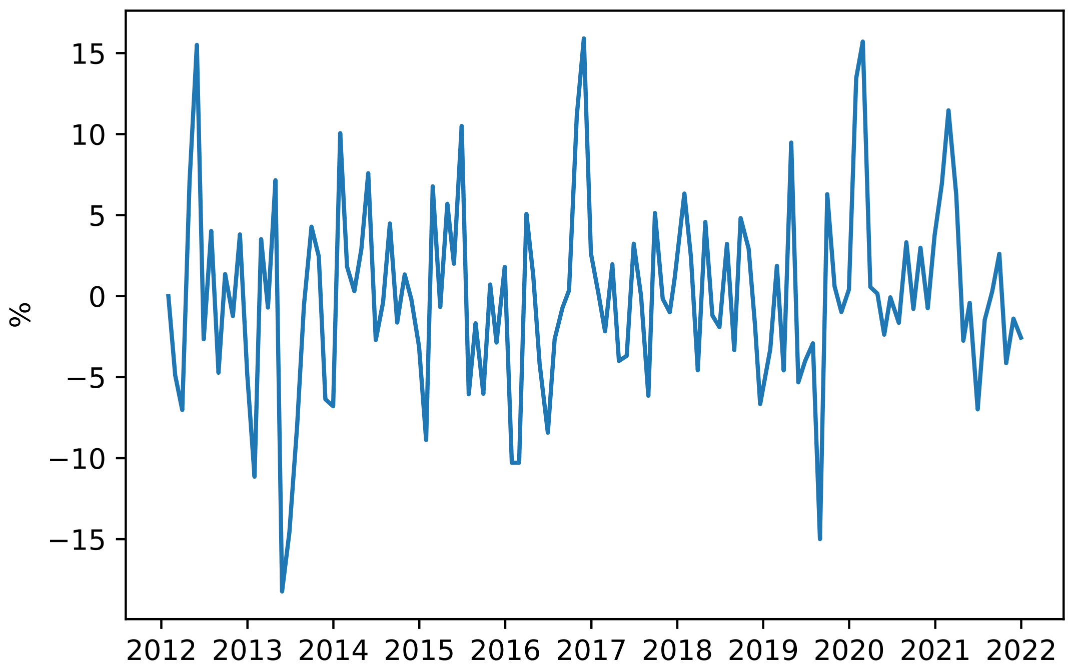

Figure 29: Monthly total ex-closing cost excess returns with two-factor CIR model in the full trading period 01/2012 - 12/2021, 5 bps mispricing limit

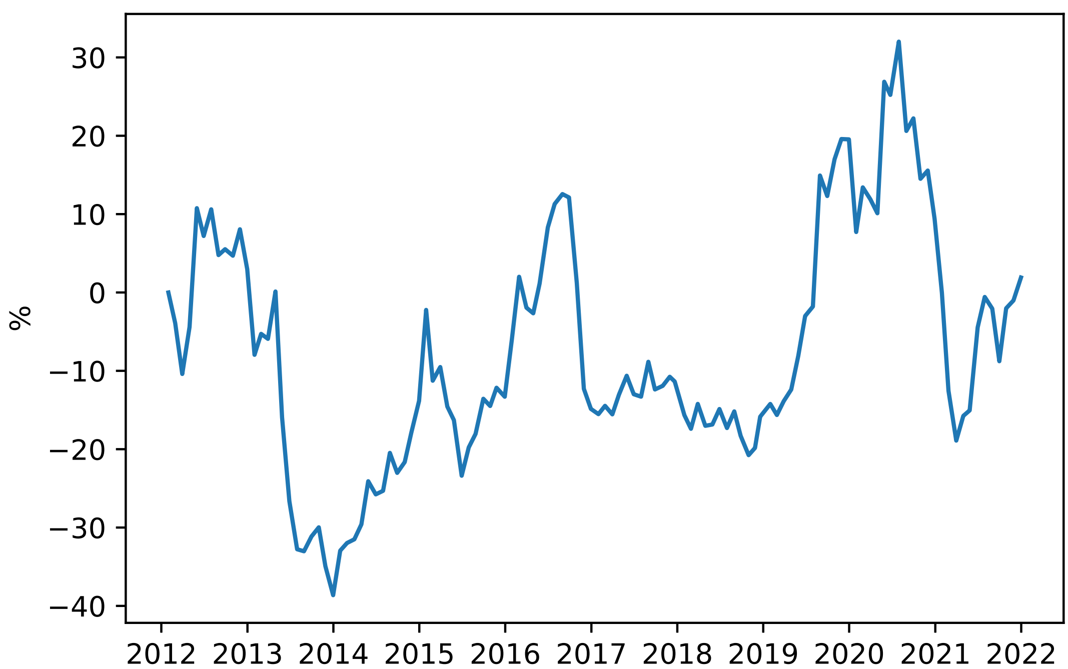

Figure 30: Cumulative returns with two-factor Vasicek model in 01/2012 - 12/2021, 20bps mispricing limit

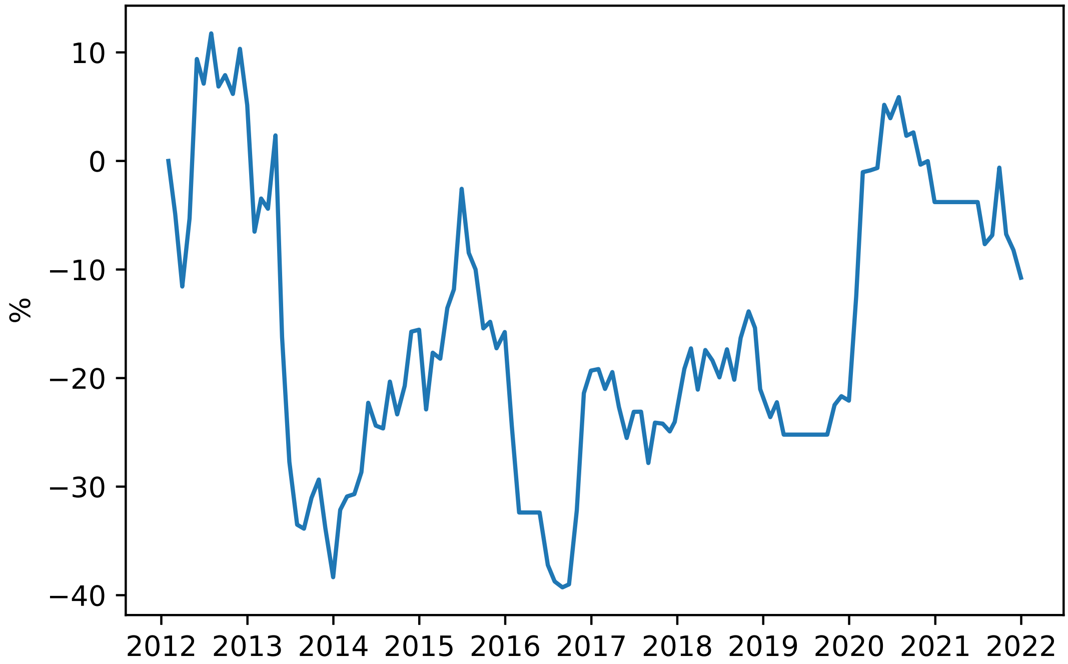

Figure 31: Cumulative returns with two-factor CIR model in 01/2012 - 12/2021, 20bps mispricing limit

Finally, Figures [30] and [31] provide the cumulative returns of the two models with 20 bps mispricing limit in the full trading period. The performance is quite similar to what was observed in the reported cumulative returns with the 5 bps mispricing limit. There are large drops and increases in the cumulative returns and the drops and increases are timed to quite the same periods for both of the models. These cumulative returns do not either indicate good performance for yield curve arbitrage in the full trading period.

11.8 Examples

In this section, I provide a collection of examples related to the topics of my thesis.

Example 11.3. Pure arbitrage trading opportunity. Consider a stock, that is traded on both the New York Stock Exchange and the London Stock Exchange. Suppose, that the stock price is \$150 in New York and £100 in London at a time when the GBP/USD exchange rate is \$1.5300 per pound. An arbitrageur could simultaneously buy 1 share of the stock in New York and sell it in London to obtain a risk free profit of

$$[(\$1.53 * 100) - \$150] = \$3$$

in the absence of transaction costs. For example, a large investment bank would probably face low transaction costs in both the stock market and the foreign exchange market and could implement this kind of arbitrage.

Example 11.4. Day count conventions. Consider, that D_0 is 31-03-2022. I want to calculate the year fraction between D_{-2}, that was 2 months before D_0 and D_3, that is 3 months after D_0. Simply adding 3 months to the month of D_0 would provide me D_3 as 31-06-2022, but this date does not exist in reality, as there are different number of days in different months. The correct D_3 is 30-06-2022. On the other hand, taking 2 months from D_0 I get D_{-2} as 31-01-2022. If I am applying the day count convention Actual/365 with these dates, the year fraction between dates D_{-2} and D_3 would be $\frac{28+31+30+31+30}{365} = \frac{150}{365}$.

Example 11.5. Bootstrapping of zero coupon bonds. Zero coupon bond prices can be bootstrapped from base instrument market prices. Consider I observe 1-year coupon bond trading at \$99.10, 2-year coupon bond trading at \$100.6 and 3-year coupon bond trading at \$98.8. All the bonds have nominals of \$100 and they pay annual coupons of 0%, 1.5 % and 1%, respectively. As the coupon bond is valued at nominal at maturity, the first zero coupon bond price is simply obtained as:

$$P(0,1) = \frac{99.10}{100} = 0.991$$

i.e. the discounted value of the nominal must equal the current price of the 1-year coupon bond. Notice here, how the zero coupon bond prices are linked to discount factors. Once I know the 1-year zero coupon bond price of 0.991, I can solve the 2-year zero coupon bond price from the equation:

$$100.6 = 1.5 * P(0,1) + 101.5 * P(0,2)$$

i.e. the 2-year coupon bond price must be obtained from its discounted cash flows. Finally, the 3-year zero coupon bond price is obtained from equation:

$$98.8 = 1 * P(0,1) + 1 * P(0,2) + 101 * P(0,3)$$

Example 11.6. An OIS. In 17-month USD OIS, the daily US Federal Funds rate is first compounded daily for the "stub" period of 5 months and then at the end of 5 months the floating leg payment is based on the compounded value. Then, the US Federal Funds rate is compounded for the first full-year period of (5,17) months and at the end of this period, the floating leg payment is again based on the compounded value. The 17-month USD OIS rate quoted in the markets is for the fixed leg of 17-month USD OIS and it is also paid at the end of the (0,5) months period and at the end of the (5,17) months period. USD OIS contracts apply Actual/360 day count convention for both the fixed leg and the floating leg.

Example 11.7. Closing swap by opening a counterposition. Consider a trader, that would open a 5-year receiver swap for speculation purposes. If a year later, the swap rates would have fallen or remained constant, such that the fixed ongoing swap rate > 4-year swap rate, the trader could open a counter IRS position in 4-year payer swap. This way, the trader would secure certain cash flows given by the difference of the fixed ongoing rate and 4-year rate and the receiver/payer swaps would cancel the floating leg payments.

Example 11.8. Mispricing of previously opened arbitrage trade. Consider there was a negative mispricing 7 months ago in 3-year swap and I opened a receiver swap. If currently the market 3-year swap rate is still larger than the original model 3-year swap rate by more than 1 basis point, I keep the trade open. Therefore, I do not compare the original model 3-year swap rate to market swap rate for maturity of 2 years and 5 months, but all the time to the new market 3-year swap rates. If currently the market 3-year swap rate would be smaller than the original 3-year model swap rate (sign of mispricing changed to positive) or the market 3-year swap rate would be larger than the original 3-year swap rate by less than 1 basis point, I would close the trade.

Example 11.9. Deriving the forward rates. Consider, that I have opened a receiver swap, that has 14 months to maturity. I need forward rates for (0,2)-, (2,5)-, (5,8)-, (8,11)- and (11,14)-periods. The first period of (0,2) is dictated by the time to first payment of floating leg. As there is LIBOR rates only up to 12-month maturity, I need to apply the 14-month market swap rate. Such swap rate does not exist in reality, but it is rather obtained from my interpolated data. As the swap contracts pay the fixed leg semiannually, I need to make a simplifying assumption about the first payment of the fixed leg. Otherwise, the semiannual payment frequency would not match with 14-month maturity. Namely, I assume that the first fixed leg payment occurs in an irregular interval[1] dictated by the time to first floating leg payment of the ongoing swap. In this case, the first fixed leg payment of the ongoing swap also occurs in 2 months, so I assume that the interpolated 14-month swap rate is paid in the fixed leg of swap in (0,2)-, (2,8)- and (8,14)-periods.

[1]The stub period.