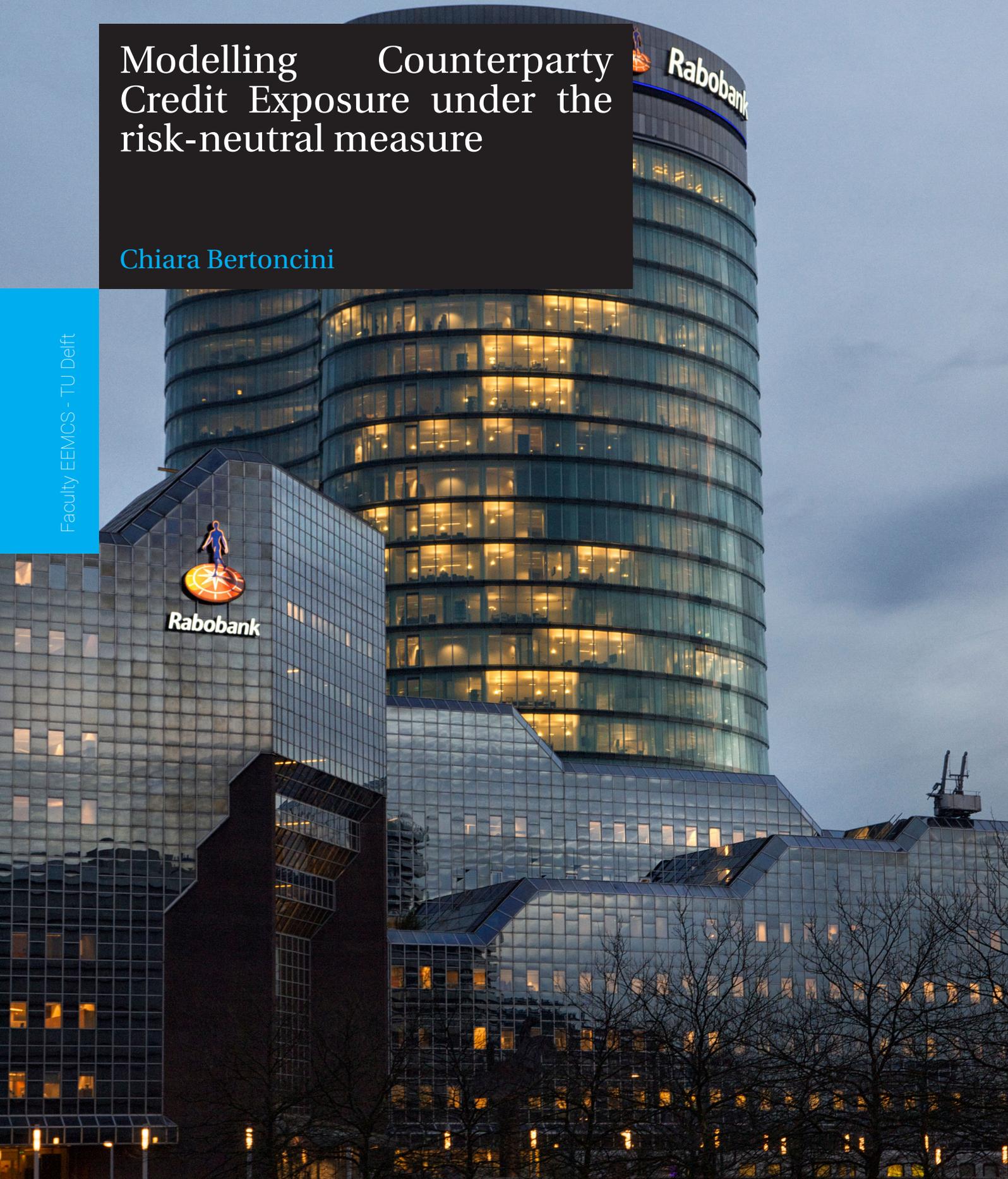


Modelling Counterparty Credit Exposure under the risk-neutral measure

Chiara Bertoncini

Faculty EEMCS - TU Delft



Modelling Counterparty Credit Exposure under the risk-neutral measure

by

Chiara Bertoncini

in partial fulfillment of the requirements for the degree of

Master of Science
in Applied Mathematics

at the Delft University of Technology,
to be defended publicly on Monday October 22, 2018 at 15:00 PM.

Student number: 4632591
Project duration: February 5, 2018 – September 30, 2018
Thesis committee: Dr. J. Cai, TU Delft, supervisor
Dr. P. Cirillo, TU Delft, co-supervisor
Dr. Ir. C. W. Oosterlee, TU Delft
Dr. Ir. M. Kolman, Rabobank

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

Three interest rate models are researched: Displaced Exponential-Vasicek, Hull-White one factor and Hull-White two factors with time-dependent volatility parameters. The motivation for this is two-fold: firstly, we would like to understand how the capital calculations would be impacted when yield curves are modelled under the three different models. This is done by looking at both magnitude and stability of the risk profiles and scalar risk-measures for three counterparties, which are highly representative for the bank. Secondly, we investigate the benefits and drawbacks of using one model and its corresponding calibration method over the others, with a special attention to the impact on yields correlations.

The first model, calibrated to historical data, is used as a nine-factors model for forward rates and is currently being used within the bank for PFE profiles and CVA regulatory capital. Historical backtest has proven the current model to perform reasonably well on real data and therefore it is used as a benchmark against which the other two models are tested. The two Hull-White models, used as short rate models, are calibrated to the risk-neutral measure (namely, to European swaptions). Precisely, a two-steps calibration procedure suited for piece-wise constant volatility functions is implemented for both. The stability analysis reveals that the variation of Exposure at Defaults is significant, which might be undesired. On the other side, the two short rate models retain the correlation structure of interest rates better than the current model. This in turn translates into higher capital impact.

Keywords: Displaced Exponential-Vasicek, Hull-White models, multi-factors models, historical calibration, risk-neutral calibration, swaption, swaps portfolios, stability analysis, risk profiles, Exposure at Default.

Acknowledgements

This thesis has been submitted in partial fulfillment of the requirements for the degree of Master of Science in Applied Mathematics at Delft University of Technology. The academic supervisors of the project has been Dr. J. Cai of the Statistics group and Dr. P. Cirillo of the Applied Probability group at the Delft Institute of Applied Mathematics. The research has been conducted under the supervision of Y. Li, M. Kolman and F. Dijk at the Rabobank office in Utrecht. This work has been promoted by the Counterparty Credit Risk team within the Risk Management Financial Markets Advisory department, with the intent of understanding the impact of risk-neutral calibration of two widely used short rate models on Rabobank portfolios.

First of all, I would like to thank my supervisors from Rabobank Yuchen Li and Marek Kolman, who kindly and very patiently introduced me to a world that was completely new to me. Considering my totally different background in statistics, it has been a great challenge to get familiar with the financial jargon and the application of mathematics to finance in such a short term. They believed in me from the very beginning and this has been a great source of motivation. In addition, I am also very thankful to the rest of the team, who highly contributed in making my experience at Rabobank very instructive and pleasant. I would also like to send my deepest thanks to Freddy van Dijk for being such a great source of inspiration and for allowing me to start my first job in his team. In addition to this, I would like to send many thanks to my supervisors at university: Juan Cai for being always very attentive and timely when it came to check my reports, and for being so kind to help me through a research whose topics are not part of her daily work. Then Pasquale Cirillo, expert in the field, whose technical knowledge has been of great help when facing the most challenging step of the thesis. Furthermore, I owe a lot to my friends and boyfriend, whose help was invaluable especially in the hardest moments. Lastly, but not less important, I would like to sincerely thank my family for being those that have believed in me the most, and for allowing me to live this incredible and unique experience in the Netherlands.

*Chiara Bertoncini
Delft, October 2018*

Contents

Abstract	iii
Acknowledgements	v
1 Introduction	1
1.1 The framework	1
1.1.1 The derivative markets	1
1.1.2 Counterparty credit risk	2
1.1.3 Metrics of credit exposure	3
1.1.4 Capital requirements and EaD modelling	5
1.2 Outline of the research	6
1.2.1 Data and Matlab codes	8
2 Basics	9
2.1 Definitions and notations	9
2.2 Core financial derivatives	12
2.2.1 Interest Rate Swap	12
2.2.2 European swaption	13
2.3 Physical and risk-neutral measure	15
2.4 Affine Term-Structure Models (ATSM)	16
2.5 $Gn++$ models (multi-factors)	17
3 The stochastic models	19
3.1 Shifted log-normal model	21
3.1.1 Introduction to the model	22
3.1.2 Modelling the forward rates	23
3.1.3 Historical Calibration	23
3.1.4 Correlation structure	24
3.1.5 Simulation	25
3.2 The Hull-White one factor model	28
3.2.1 Introduction to the model	29
3.2.2 Zero-Coupon Bond price	32
3.2.3 Dynamics under the \mathbb{Q}^T measure	32
3.2.4 Bond-option and swaption prices	34
3.2.5 Calibration to swaptions	36
3.2.6 Simulation	41
3.2.7 Perfect correlation	43
3.3 The Hull-White two factors model	48
3.3.1 Motivation and Introduction	48
3.3.2 Zero-Coupon Bond price	51
3.3.3 Dynamics under the \mathbb{Q}^T measure	51
3.3.4 Bond-option and swaption prices	52
3.3.5 Calibration to swaptions	56
3.3.6 Simulation	58

3.3.7	Imperfect correlation.	60
3.3.8	Discussion	63
4	Portfolio impact and stability analysis	65
4.1	Model and calibration impact on representative portfolios.	66
4.1.1	Revaluation of a plain vanilla swap	66
4.1.2	A real application	67
4.1.3	Historical volatility against Black's volatilities	71
4.2	Stability of risk-neutral calibration	72
4.2.1	Time-horizon and new data for calibration.	72
4.2.2	Risk-factor level.	72
4.2.3	Portfolio level	74
5	Conclusion and Outlook	77
5.1	Future research.	77
5.2	Summary and conclusion	78
A	Computation of transformation matrix $A(t)$	81
B	Estimation of mean reversion speed in HW-1F model	83
B.1	Historical Calibration	83
B.2	Swaption implied volatility approximation.	84
C	Further results	87
C.1	Simulation results of $HW - 1F$ with $a = 1.5\%$	87
C.2	Further portfolios	88
C.2.1	Single date.	88
C.2.2	Multiple dates.	89
	Bibliography	91

Introduction

1.1. The framework

In this section we would like to allow the reader to plunge inside the framework of Counterparty Credit Risk and provide him with the essentials needed to develop a clear business interpretation of the results that will come in Chapters (3) and (4). For what reported in this section we refer to [14].

1.1.1. The derivative markets

Generally speaking, derivatives are financial securities that are always referred to an underlying asset or group of assets. The derivative itself is a contract between two or more parties that agree to pay quantifiable amounts of money to one another on specified future dates. The amounts of money transferred are usually referred to as the “legs” of the derivative transaction, which can be fixed or dependent on a variable market parameter such as an interest and/or foreign exchange rate. Usually, one refers to this as the underlying risk factor of the derivative transaction and its fluctuations readily affects the price of the derivative. Two different markets exist where derivative contracts can be traded: *Exchange Market* and *Over-The-Counter market*.

Exchange-traded deals are regulated and therefore secured by some clearing entity, which may compensate losses in case a counterparty defaults. Therefore in this particular situation there is no counterparty credit risk at stake. For instance, when a future contract is traded, the actual counterparty to the contract is the exchange and the only aspect of concern is the solvency of the exchange itself. A much greater amount of derivatives are traded Over-The-Counter (OTC) since this market allows for more customization of each specific deal. Here, the number of traded instruments has grown dramatically in the last decade, especially with regard to interest rate products. Derivatives can be traded bilaterally between two parties and there is no clearing entity in between, therefore each party takes counterparty risk to the other. This is where Counterparty Credit Risk comes in. In Figure (1.1) the split of OTC derivatives by market factor type and product type as of first half of 2016 is displayed. The data have been retrieved from BIS ¹. As it is straight-forward to see, interest rate products are by far the most traded instruments in OTC, whereas foreign exchange products and credit default swaps are awarded respectively the second and third place. This is what drives us to focus the research on interest rates among all the possible risk factors.

¹Bank for International Settlements.

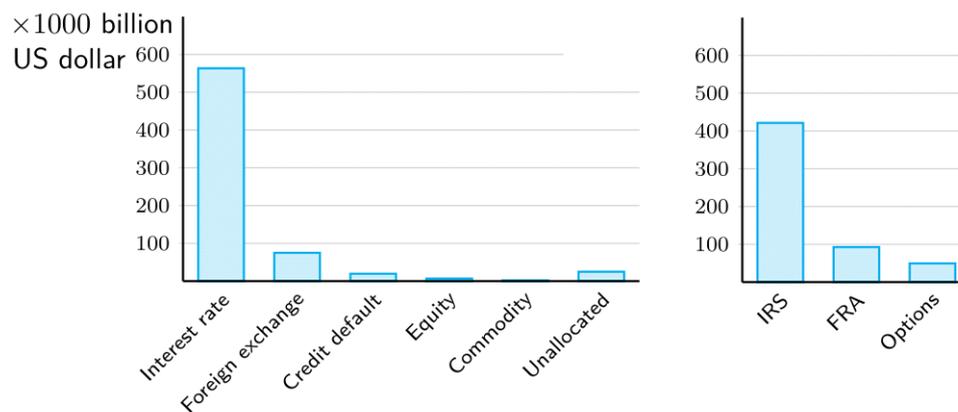


Figure 1.1: Split of OTC notional by risk factor type (left panel) and by product type (right panel) as of 2016. Source BIS.

1.1.2. Counterparty credit risk

“... probably the single most important variable in determining whether and with what speed financial disturbances become financial shocks, with potential systemic traits.” [14, p. 13]

Financial institutions face several types of risks and one of the most important activities within a bank is to develop models and methodologies in order to buffer these risks. At this stage, it is of great importance to meet regulations in the way risks are quantified. One of these is *Counterparty Credit Risk* (CCR), which can be seen as one particular form of credit risk as it arises from the credit risk in securities financing transactions such as Over The Counter derivatives. More precisely, it is the risk arising from the possibility that a counterparty may default prior to expiration of a financial contract and therefore fails to live up to its contractual obligations. Counterparty Credit Risk Management tries to measure, mitigate and ultimately control this source of risk. Counterparty Credit Risk is quantified based on *default risk* and *credit exposure*. The former is directly linked to the probability that a counterparty defaults. The latter represents the amount of capital that would be lost in the event that a counterparty defaults, and its modelling is the central focus of the research.

Before defining the exposure, it is important to clarify what is meant by “Mark-To-Market” (MtM) value with respect to a particular counterparty. This is an intrinsic value carried by a derivative contract and represents the present value of all the payments an institution is expecting to receive less than those it is obliged to make (net present value, NPV). Therefore it does not constitute an immediate liability. The MtM of a transaction is usually zero at inception as the two parties agree to exchange transaction legs of equal value. Since the future cash flows depend on the fluctuation of the underlying risk factor(s), the MtM value changes every day (even intraday) during the life of a transaction. It may then be either positive or negative, depending on whether it is an institution’s favor or not. It is also known as *replacement cost* as it defines the entry point into an equivalent transaction with another counterparty, with the assumption of no transaction costs.

Potential losses (and therefore, the credit exposure) are asymmetric with respect to the MtM value of a transaction. More precisely, credit exposure is defined as the positive value of the contract (or portfolio of contracts) with a counterparty assuming zero recovery in case of default. That is:

$$E = \max(\text{MtM}, 0) = \text{MtM}^+,$$

where E denotes the credit exposure. Indeed, there are two possible scenarios: either the value (MtM) of the contract is positive, or it is negative. In the first case, the counterparty owes us money and in the case of a default we experience a loss (positive exposure). In this situation the financial institution will close out the position and claim the positive value of the contract. Depending on the recovery rate, they may be able to recover part of the claim, however there will always be a percentage that cannot be recovered. The loss will then be quantified as the value of the contract less any recovery value. In the second case, the exposure is zero which means that we owe money and therefore no loss is experienced. What happens is that the institution will “close-out” the position by paying the market value of the contract to the counterparty, and will enter a similar contract with another counterparty in order to receive back the market value of the contract. So, depending on the development of future interest rates both parties can have an exposure, and only when the value of the contract is positive the financial institution will be affected by a credit loss in case a counterparty defaults. This easy fact translates in the so-called bilateral nature of credit risk: a contract can either be an asset or liability depending on its value. It is important to notice that the exposure is something known until today, but unknown for future dates. Therefore, this uncertainty needs then to be modelled as a stochastic process. In our case, this is implicitly done by assuming a stochastic model for the interest rate.

Two examples of OTC derivative products are Interest Rate Swap (IRS) and Forward Rate Agreements (FRA). Note that the latter is just a particular case of interest rate swap, and therefore it makes sense to only consider the more general case in which the financial institution's total exposure is driven by exposure in swaps. An interest rate swap is a contract in which the two parties agree to exchange floating interest rate payments with fixed interest rate payments based on a pre-specified notional. In this contract, notional and fixed payments are agreed in advance, but the future floating interest rate (usually the LIBOR rate) is quoted every day in the market and therefore cannot be determined in advance. The value of an interest rate will thus be uncertain for future time instants.

There are several ways with which Counterparty Credit Risk can be buffered. For instance, the parties in a derivative contract can decide to do this by means of a ISDA² master agreement. This allows the two parties entering the contract to decide on a netting agreement. As a consequence the value of all contracts within the same netting set will be netted. This means that the exposure of a set of contracts will be the net exposure, which gives legal permission to the financial institution for example to offset a contract with positive value with another contract with the same value but opposite sign, in case the counterparty defaults. Furthermore the ISDA Master Agreement might be supported by a Credit Support Annex (CSA). The two counterparties, if this option is exercised, are allowed to hold cash or securities against their exposure. More specifically, in case of a CSA one (or both) counterparties are required to pay some collateral in case the exposure exceeds some pre-specified threshold. The main advantage of this is that the exposure will then be reduced below such threshold. Note that this option will not be considered in the practical results of Chapter (4). For more on this, we refer to [30]. To conclude, also diversification and hedging with credit derivatives (e.g. Credit Default Swaps) helps to reduce the exposure and CCR.

1.1.3. Metrics of credit exposure

Here we describe the most risk-sensitive metrics that are commonly used to quantify credit exposure. These metrics can be easily reproduced from the simulated Mark-to-Market profiles and are

²ISDA is an acronym for International Swaps and Derivatives Association. This organization has created the master agreement, which is a standardized contract for derivative transactions.

set to certain limits within the bank. We rely on the Basel Committee on Banking Supervision (2005) definitions.

Potential Future Exposure (PFE)

For risk-management purposes, it is always important to assess what is the worse exposure one could have at a certain time in the future. The Potential Future Exposure is the answer to this with respect to a certain confidence level. It is an internally-used high percentile of the distribution of current exposure simulations at any particular future date before the maturity date. Also this metric can be calculated both on a risk-factor level and portfolio level. For example, in the second case the PFE at a 97.5% confidence level defines an exposure that would be exceeded with a probability of no more than 2.5%. This definition perfectly match the definition of Value At Risk (VaR), with the only two exceptions that this usually refers to a loss (instead of a gain) on a very short time horizon (whereas PFE are also calculated on quite long horizons). In Figure (1.2) a graphical representation for this metric is given. To briefly explain how we will calculate the Potential Future Exposure, let us assume that we have a derivative contract with a certain counterparty. If we call $V(t)$ the value of the contract at time t , the exposure at some time point t^* will be given by $E(t^*) = V(t^*)^+$. This value, if we stand at time $0 < t^*$ is a random variable, and it depends on the movements in interest rates (under the assumption that the interest rate is the only risk factor contributing to the value of the derivative). Given that we have chosen a certain model for the interest rate, we can simulate several scenarios of possible future interest rates up to time t^* and then calculate a range of possible future values $\overline{V}_1(t^*), \dots, \overline{V}_n(t^*)$ and $\overline{E}_1(t^*), \dots, \overline{E}_n(t^*)$. The empirical 97.5% PFE can be then easily computed from the last set of simulated values. In case there are more than one transaction at stake, the resulting PFE is calculated based on whether the two counterparties have either agreed for grossing or netting. In the first case, the total PFE will just simply be the sum of each PFE, while in the second case PFE is computed only once, posterior to the netting of the MtM values within the same netting set.

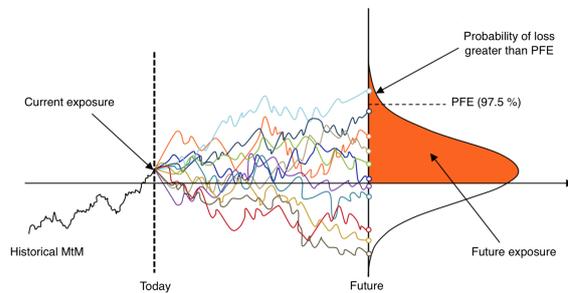


Figure 1.2: Graphical representation of simulated MtM future scenarios.

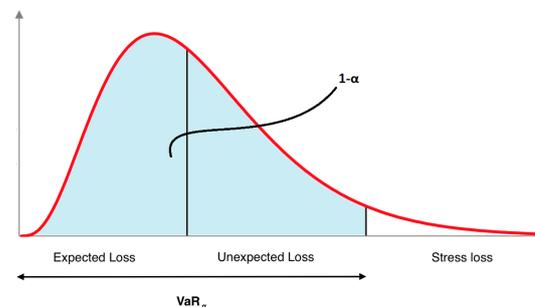


Figure 1.3: Example of yearly distribution of losses.

EE, EPE, EEPE

The Expected Exposure (EE), like the Potential Future Exposure, is a risk metric that is calculated for a given time horizon. That is, if we consider the MtM distribution at a certain future time t , the expected MtM is simply defined as the mean of such distribution, while the Expected Exposure is defined as the mean of the non-negative values of MtM. Similarly, the PFE is defined as a high quantile over the non-negative values of MtM. Of course the expected MtM is by definition always lower than the EE. On the other side, there are two important metrics that are defined through time, and not for each time step as the previous metrics: these are called Expected Positive Exposure (EPE) and Effective Expected Positive Exposure (EEPE), and represent fundamental quantities for the definition of the Exposure at Default (EaD). The EPE is simply defined as the average EE through time and is a useful single number that can represents the exposure. As EE and EPE are known to

underestimate exposure (especially for short-dated transactions), the Basel Committee on Banking Supervisions (2005) formulated two corrected versions of them. These are the Effective EE and Effective EPE (namely, the EEPE). The first one is simply a non-decreasing EE, while Effective EPE is the average of the Effective EE.

Exposure at Default

This quantity consists in a scaling of the EEPE and enters in the capital calculations formula prescribed by the regulator. More precisely, it is stated by the regulator that the Exposure at Default (EaD) must be computed as the EEPE times an “alpha factor”. If this is determined by the bank itself, then it will have to be between 1.2 and 1.3. In the case the bank does not provide its own calculation for the alpha factor, a fixed value of 1.4 has to be used. We will stick to this case. To conclude, the general simple formula to be applied is

$$\text{EaD} = \alpha \cdot \text{EEPE}.$$

1.1.4. Capital requirements and EaD modelling

Defaults are not uncommon in real life and credit trading losses, whose frequency depends on time, are not surprising to happen. Usually, the losses that a bank experiences are measured on a yearly basis. Their occurrences can be expressed with a distribution and an example is given in Figure (1.3). As summarized in the picture, there are three kind of losses a bank can experience: expected loss, unexpected loss, stressed loss. The first one constitutes the average level of credit losses that a bank can reasonably expect to experience. These are generally considered a normal cost of doing credit business, and are covered through revenues, impairments and provisions. Secondly, unexpected loss is the potential unexpected loss for which capital should be held according to the regulator. Lastly, stressed loss is the potential unexpected loss against which it is judged to be too expensive to hold capital against. Unexpected losses of this extent lead to insolvency, for which we mean that the bank will not be able to fulfill its own credit obligations by means of revenue and capital. Unexpected and stressed losses do not happen regularly but can happen to be very large. Now, of course holding capital against the worst case, that is the loss of the entire portfolio of a given year would be the safest but economically very inefficient. At the same time the probability of a stressed loss and thus the insolvency of a bank will be very small. Therefore there is a trade-off between maintaining the bank's solvency and being able to extend the credit business.

The collapse of banks such as Herstatt Bank in Germany and the Franklin National Bank in the US was, together with other reasons, linked to bad international supervision. These failures were a true symptom of serious disturbances in international currency and banking markets. This led to the establishment of the Basel Committee on Banking Supervision (BCBS) in the late 1974, mostly to enhance financial stability by improving the quality of banking supervision worldwide. Their most important works all fall under the name of *Basel accords* and are three series of banking regulations (Basel I, Basel II, Basel III). As a consequence of the Basel II framework [7], every financial institution should compute their regulatory capital for Counterparty Credit risk and such amount of capital needs to be maintained so to buffer for the default risk of all counterparties in its portfolio. How these “capital requirements” should be calculated is also established by BCBS. The most simple method is the so-called *standardized approach*, where such capital is kept at 8% of the risk-weighted assets (RWA), which can either be fixed from the regulator or derived from external ratings. This is the most simple and conservative approach. On the other side, there are the two famous *internal rating based* (IRB) approaches, which allows banks to compute capital requirements more sharply. Out of these two approaches, the less restrictive is called Advanced internal ratings based approach (A-IRB), under which the three fundamental parameters defining the capital requirements can be determined internally by the bank given that these fulfill all the regulatory requirements. These

three parameters, which enter the capital calculation formula, represent the fundamental components of credit risk for each given counterparty :

1. *Probability of Default* (PD), the probability of the counterparty defaulting within one year.
2. *Exposure at Default* (EaD), estimate of the amount outstanding with the counterparty given the event of a default.
3. *Loss Given Default* (LGD), percentage of exposure the bank will lose upon default of the counterparty.

Therefore, an accurate modelling of the Exposure at Default is needed when it comes to quantify capital requirements. In this thesis the focus will be given to the modelling of Counterparty Credit Exposure, when the portfolio is made of several different swaps and netting agreement is considered. As a consequence, all the related risk measures introduced in the previous paragraph can be calculated.

Under these approaches, the capital requirements are calculated given a bank's portfolio and its losses, in a way that the frequency of insolvencies is kept to a minimum. This minimum is recognized to be acceptable for the regulator. More precisely, the capital that must be kept apart is calculated so that the probability of (extreme) unexpected losses will not exceed a confidence level $\alpha \in (0, 1)$. In Figure (1.3) the shaded area is exactly referring to the probability of the bank remaining solvent and the corresponding threshold is known as the *Value at Risk* (VaR_α) of a given portfolio.

1.2. Outline of the research

Counterparty Credit Risk is one out of the many type of risks that a financial institution needs to deal with. This risk arises when trading derivatives in the Over-The-Counter market and its quantification is an important but complicated issue, since the loss due to a default of a counterparty is an uncertain quantity. Many market factors can contribute to the future value of a derivative, being interest rate, FX rate, credit, equity and commodity. The goal of this thesis is to research on three different interest rate modelling techniques, for which a brief summary is given in Table (1.1). These are the Displaced Exponential-Vasicek model, Hull-White one factor and two factors models, with the first being calibrated to historical data and used for the simulation of nine forward rates, and the last two being calibrated to European swaptions and used for the simulation of the unobservable short rate. The modeling of this particular kind of rate can be very convenient since all fundamental quantities are defined, by no-arbitrage arguments, as the expectation of a functional of this process. This thesis will only treat models that allow for simulation of negative interest rates, due to the current low interest rates environment. As a matter of fact, nowadays this turns out to be a desirable property as for some currencies, debt instruments have lately been experiencing negative rates for the short-term and medium-term of the yield curve. Note that the Displaced Exponential-Vasicek model assumes a log-normal distribution for the rates and therefore a shift must be applied. We will compare these three models based on: computational difficulty, easiness for implementation, performance on correlations, stability on the bank's portfolio. Stability is not an issue for the profiles implied by the historically calibrated model, while the variation of Exposure at Default over a time window of one year turns out not really on favour of the two Hull-White models. On the other side, these two outperform the other in terms of correlation of the simulated yield curves. Note that the research is not about a definite model choice, as this usually cannot be done until the models are backtested with real data. Rather, the research is driven by the interest of understanding how model choice and calibration procedure would impact the capital, and therefore provide some results that can allow to shape a model preference. This is done analyzing the implied profiles and

risk-measures on three mixed portfolios of vanilla swaps with netting agreement, which are highly representative for the bank. The reason for only dealing with plain vanilla swaps is that this represents the most traded derivative instrument in the Over-The-Counter market. Especially, Rabobank is a very conservative bank and it comes natural to believe that swaps form most of the bank's trading book.

The two Hull White models have been researched under the assumption of time-dependent volatilities, as we experienced that constant parameters were not enough to achieve good fits for certain choices of the calibration baskets. To the best of our knowledge this assumption is not much tackled in the literature, especially with regard to the two factors model. Our contribution can be summarized in four parts. First of all, we generalized some already known results, derived for constant parameters, to the case in which the parameters are time-dependent. This is done by either relying to intuition, or by deriving the results from scratch. Second contribution is in terms of calibration under the risk-neutral measure. We propose a calibration procedure to market data, with a special eye given to the calibration of the mean reversion speed(s) parameter(s). To our knowledge, in the literature there is no clear explanation on how to perform this step, especially when time-dependent parameters are considered. This assumption in fact does not allow to reduce the calibration problem to merely one optimization problem, as it could be the case when dealing with a simple one factor model with constant parameters. This result in two separate optimization problems to be solved, which are fast to be solved with one factor, whereas become more demanding with two factors. Thirdly, throughout the thesis a special attention has been given to the correlation between yield curves for all three models. We reckoned this to be necessary as capturing a wrong correlation structure between the rates can lead to underestimation of risk, which is undesired. We discovered during the research that one of the main problems of the current (benchmark) model is that the way the simulation is performed brings the correlation matrix to zero for too short simulation horizons. Whereas, the two Hull-White models exhibit a potential improvement in this direction. Especially, it turns out that the two-factors model behaves in a complete opposite manner for long (but not too long) simulation horizons, which highly impact the final risk profiles. The fourth contribution regards the following fact. Normally it is custom to use historical calibration for "risk-management-like purposes" (e.g. analyzing Potential Future Exposure profiles, capital calculations). On the other hand, for pricing of Credit Value Adjustment (CVA) one should use risk-neutral calibration. We analyzed what would be the impact (in terms of capital requirements) of adopting the risk-free measure also in the first case, driven by two main reasons: first one is very simple, that is that upcoming regulations seem to require such a change. Second reason is that the possibility of speaking the "same language" with other departments would definitely be a desirable situation, especially when it comes to deal with traders. Of course one could be quite skeptical about this. Risk-neutral calibration requires to calibrate the parameters of a model to today's market prices and it is done frequently, and therefore the exposure profiles are likely to be very volatile in time. This would translate in volatile capital requirements, which might be undesired for a financial institution. We provide the answer to this particular question, even though further research is needed in order to understand whether these answers can be considered positive or negative for the bank.

In Chapter (2) we provide the equipment of definitions and tools that are necessary for the understanding of the three models, researched in Chapter (3). Furthermore, this contains the description of how the calibration and simulation steps are performed under the three different models and the respective results. Positive and negative aspects are outlined every time a model is introduced and at the end of the chapter a conclusive discussion is addressed, merely based on how the models are designed. Posterior to this, all three models are tested on a more practical application and the procedure and results are described in Chapter (4). Here, we will present the resulting exposure dis-

Code	Calibration	SDE
DEV-MR	\mathbb{P} -measure 8 years for σ 30 years for a, b	$d \ln(f_i(t) + \gamma) = [b - a \ln(f_i(t) + \gamma)] dt + \sigma dW_{i,t}^{\mathbb{Q}}$ for $i = 1, \dots, 9$.
HW-1F G1++	\mathbb{Q} -measure 1 year for a monthly for $\sigma(t)$	$dr(t) = [\theta(t) - a(t)r(t)] dt + \sigma(t) dW_t^{\mathbb{Q}}$ $dx(t) = -ax(t) dt + \sigma(t) dW_t^{\mathbb{Q}}$
HW-2F G2++	\mathbb{Q} -measure 1 year for a, b, ρ monthly for $\sigma(t), \eta(t)$	$\begin{cases} dr(t) = [\theta(t) + u(t) - \bar{a}r(t)] dt + \sigma_1(t) dW_{1,t}^{\mathbb{Q}} \\ du(t) = -\bar{b}u(t) dt + \sigma_2(t) dW_{2,t}^{\mathbb{Q}} \\ dx(t) = -ax(t) dt + \sigma(t) dW_{1,t}^{\mathbb{Q}} \\ dy(t) = -by(t) dt + \eta(t) dW_{2,t}^{\mathbb{Q}} \end{cases}$

Table 1.1: Summary of the three models that have been investigated in the thesis.

tributions at any future time under both the historical and risk-neutral measure, for three different portfolios. After this, the stability on the risk-neutral valuation is evaluated. Lastly, in Chapter (5) we draw our conclusions and give insights about possible future research.

1.2.1. Data and Matlab codes

Data retrieval, data processing and codes writing have all been personally performed on Excel and Matlab for all three models. Regarding the data, I was given access to the Bloomberg platform so that I could retrieve the quoted volatilities for the calibration at the several dates, together with the interest rate instruments to build the discount factor curves. Furthermore, I was given access to certain packages containing Rabobank's deals and many features needed to be either selected or discarded and interpreted. The codes, which we don't provide in the thesis for privacy reasons, have been written completely from scratch for both the historical and risk-neutral calibration of the three models. Same argument applies to the simulation step (both discretized and exact have been tried in order to answer internal questions from the managers) and to the portfolios impact and stability analyses. Therefore, I take the responsibility for all the results provided in the thesis.

2

Basics

In this section we provide the mathematical framework from financial mathematics necessary to deal with the results to come. More precisely, we present and explain here the notations, definitions and some results used recurrently in the thesis. Furthermore, we reckon to be very important for our purposes to clarify the distinction between physical measure (or real-world measure) \mathbb{P} and risk-neutral measure (or risk-free measure) \mathbb{Q} . One could either chose to evaluate Potential Future Exposures and exposure profiles via the real measure, which is usually done in practice, or via the risk-neutral measure. Before we understand what are the consequences of switching to a risk-neutral measure based on our (historically calibrated) benchmark model, it is of big importance to clarify the essential distinction between such two measures.

2.1. Definitions and notations

In this section the main definitions that will be used in the remainder of the thesis are introduced. We will state the basic definition of Zero-Coupon Bond and clarify the concept of the short rate, a very important entity with which fundamental quantities can be computed, under the no-arbitrage assumption. To conclude, as many compounding and day-count conventions exist for interest rates, we will make clear our choices.

Definition 2.1.1 (Money-market account). *The money-market (or bank) account is defined as $M(t)$ and evolves according to the following differential equation:*

$$\begin{cases} dM(t) = r(t)M(t) dt, \\ M(0) = 1, \end{cases}$$

where $r(t)$ is a positive function. It follows that:

$$M(t) = \exp\left(\int_0^t r(s) ds\right).$$

What have just been introduced represents a (locally) riskless investment, where profit is accrued continuously at the risk-free rate $r(t)$, prevailing in the market at time t .

Definition 2.1.2 (Stochastic discount factor). *The (stochastic) discount factor $D(t, T)$ between time instants t and T is the amount at time t that is equivalent to one unit of currency payable at time T . It is given by:*

$$D(t, T) = \frac{M(t)}{M(T)} = \exp\left(-\int_t^T r(s) ds\right).$$

Now, it is well known that when dealing with pricing in equity markets, the Black & Scholes formula is applied with r being a deterministic function of time. In this case both money-market account and discount factor happen to be deterministic functions. However, in this thesis we will be dealing with interest rate products and we will have to deal with exposure profiles of particular portfolios for which the main variability is explained by the variability in the interest rates. Therefore, for our purposes, it is necessary to drop the deterministic setup in favour of a stochastic modeling of the process r .

Definition 2.1.3 (Zero-Coupon Bond). *A Zero-Coupon Bond with maturity T 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 $P(t, T)$. Obviously $P(T, T) = 1$.*

In this thesis we will refer to the *zero-bond curve* (or also term structure of discount factors) at time T meaning the function

$$(T - t) \mapsto P(t, T), \quad (2.1)$$

or equivalently, by calling $\tau = T - t$:

$$\tau \mapsto P(t, t + \tau).$$

It is clear that when interest rates are positive, the curve in (2.1) should be a T -decreasing function starting from $P(t, t) = 1$. Nevertheless, nowadays it is not uncommon to observe negative rates and therefore the typical shape that can be observed exhibits a slightly humped behaviour for the shortest maturities and then decreases. This typical shape is appreciable in Figure (3.12).

All interest rates can be defined in terms of Zero-Coupon Bond prices and vice versa. However, the former are directly observable in the market whilst the latter are not. There exist two main types of compounding: simple and continuous. The second one represents the convention that has been adopted in the thesis, in the sense that continuously compounded rates have been retrieved for both historical and risk-neutral calibration. The following definitions clarify the distinction between these two conventions, both in terms of forward and spot rates. After this, also their instantaneous counterparties are introduced.

Definition 2.1.4 (Simply-compounded forward interest rate). *The simply-compounded forward interest rate prevailing at time t , with expiry T and maturity S (s.t. $t < T < S$) is denoted by $F(t; T, S)$ and can be derived by rendering a prototypical FRA with same expiry and maturity a fair contract at time t . This rate will eventually read:*

$$F(t; T, S) = \frac{P(t, T) - P(t, S)}{\Delta(T, S)P(t, S)},$$

where $\Delta(\cdot)$ is the function calculating the year fraction from T to S .

Definition 2.1.5 (Simply-compounded spot interest rate). *The simply-compounded spot interest rate $L(t, T)$ at time t for maturity T is the constant rate proportional to the year fraction $\Delta(t, T)$ at which an investment of $P(t, T)$ units of currency at time t accrues to one unit of currency at maturity T . In mathematical formulas, this reads:*

$$P(t, T) + P(t, T)\Delta(t, T)L(t, T) = 1.$$

As a consequence, the simply-compounded spot interest rate and the Zero-Coupon Bond Price can be respectively expressed in terms of each other as:

$$L(t, T) = \frac{1 - P(t, T)}{\Delta(t, T)P(t, T)}, \quad P(t, T) = \frac{1}{1 + \Delta(t, T)L(t, T)}. \quad (2.2)$$

Note that $L(t, T) = F(t; t, T)$. The forward rate $F(t; T, S)$ can be seen as an estimate of the future spot rate $L(T, S)$, given the market conditions at time t . This fact can be easily accepted by intuition and becomes even more evident when deriving the pricing formula for a forward rate agreement (FRA).

Examples of these particular kind of rates are the *London Interbank Offered Rates* (LIBOR), which motivates the notation L and are fixed daily in London by high credit rating banks participating in the LIBOR panel. This is why the LIBOR rates are considered to be a good approximation of a risk-free rate and therefore represent reference rates at which banks deliver short-term loans between each other. These days, the LIBOR rates are fixed for five currencies and maturities ranging from overnight to 12 months, and are typically linked to the Zero-Coupon Bond prices by the Actual/360 day-count convention for computing $\Delta(t, T)$.

For what follows, let $\Delta(t, T) = T - t$.

Definition 2.1.6 (Continuously-compounded forward interest rate). *The continuously-compounded forward interest rate for $[T, S]$ prevailing at time t is given by:*

$$R(t; T, S) = -\frac{\log P(t, S) - \log P(t, T)}{S - T}$$

This is the quantity being modelled when performing historical calibration.

Definition 2.1.7 (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 this reads:*

$$R(t, T) := R(t; t, T) = -\frac{\log P(t, T)}{T - t}.$$

Equivalently, we can express the Zero-Coupon Bond price as a function of the continuously compounded spot interest rate:

$$P(t, T) = e^{-R(t, T)(T-t)}.$$

The *zero-coupon curve* (which most of the times we will simply refer to as the yield curve) at time t is the function:

$$(T - t) \mapsto R(t, T), \quad (2.3)$$

or equivalently, by calling $\tau = T - t$:

$$\tau \mapsto R(t, t + \tau).$$

Note that Definitions (2.1.6) and (2.1.7) implies that there is an explicit relation between $R(t, T)$ and $R(t, T, S)$. That is, thinking as if $t = 0$ it is possible to calculate the future yield on a bond starting from today's yield curve. This is summarized in the following proposition, whose result is the key that links the rates being modelled in the first model to the yield curve $R(t, T)$.

Proposition 2.1.8. *The future interest rate $R(t; T, S)$ for time period (T, S) can be calculated given the rates $R(t, T)$ and $R(t, S)$ via the following formula:*

$$R(t; T, S) = \frac{(S - t)R(t, S) - (T - t)R(t, T)}{S - T} \quad (2.4)$$

If we let the maturity of the continuously-compounded forward rate collapse towards its expiry, we have the notion of instantaneous forward rate.

Definition 2.1.9 (Instantaneous forward interest rate). *The instantaneous forward rate with maturity T prevailing at time t is defined by:*

$$f(t, T) := \lim_{S \rightarrow T} R(t; T, S) = -\frac{\partial \log P(t, T)}{\partial T}.$$

As a consequence of the definition, taking into account that $P(T, T) = 1$, one can also write the price of a Zero-Coupon Bond as $P(t, T) = e^{-\int_t^T f(t,s)ds}$.

Definition 2.1.10 (Instantaneous short rate). *The (instantaneous) short rate at time t is defined by*

$$r(t) := f(t, t) = \lim_{t \rightarrow T} R(t, T).$$

This is the quantity being modelled when performing risk-neutral calibration.

2.2. Core financial derivatives

In this section, definitions of the financial products that come into play in the thesis are given: Interest Rate Swaps and European swaptions. In addition to this, the common market practices for pricing such instruments will be outlined.

2.2.1. Interest Rate Swap

An *Interest Rate Swap* (IRS) is the natural generalization of a Forward Rate Agreement (FRA). More specifically, in an IRS cash flows of several interest-rate payments are exchanged at different time instants (while in a FRA only one interest rate payment is exchanged). As it happens with Forward Rate Agreements, there exists two kind of Interest Rate Swaps: payer and receiver. These are respectively used in practice to transform floating rate payments into fixed ones, and vice versa. If we buy a payer swap, that means we will have to pay the fixed leg (fixed rate), otherwise if we buy a receiver swap, it means we will have to pay the floating leg (floating rate, typically the LIBOR rate).

Definition 2.2.1 (Interest Rate Swap). *A Interest Rate Swap (IRS) is a contract that exchanges the payments between two legs, starting from a future time instant. At every instant T_i in a pre-specified set of dates $T_{\alpha+1}, \dots, T_{\beta}$ the fixed leg pays out the amount $N\Delta_i K$ corresponding to a fixed interest rate K , a nominal value N and year fraction $\Delta_i = \Delta(T_{i-1}, T_i)$, whereas the floating leg pays the amount $N\Delta_i L(T_{i-1}; T_{i-1}, T_i)$ corresponding to the interest rate $L(T_{i-1}; T_{i-1}, T_i)$ resetting at the previous instant T_{i-1} for maturity given by the current payment instant T_i . Clearly, the floating-leg resets at dates $T_{\alpha}, \dots, T_{\beta-1}$ and pays at dates $T_{\alpha+1}, \dots, T_{\beta}$.*

Theorem 2.2.2 (Valuation of an IRS). *: A (payer) Interest Rate Swap with nominal value N , payment dates $T_{\alpha+1}, \dots, T_{\beta}$, reset times $T_{\alpha}, \dots, T_{\beta-1}$ for the floating rates and fixed rate K , has a net value, discounted to time t , of:*

$$V^{\text{pIRS}}(t) = N(S_{\alpha,\beta}(t) - K) \sum_{i=1}^n \Delta_i P(t, T_i) \quad (2.5)$$

where $S_{\alpha,\beta}(t)$ is the forward swap rate at time t

$$S_{\alpha,\beta}(t) = \frac{P(t, T_{\alpha}) - P(t, T_{\beta})}{\sum_{i=\alpha+1}^{\beta} \Delta_i P(t, T_i)} = \frac{P(t, T_{\alpha}) - P(t, T_{\beta})}{A_{\alpha,\beta}(T_{\alpha})} \quad (2.6)$$

and $A_{\alpha,\beta}(t)$ is the present value of a basis point (PVBP).

Note that a receiver IRS is obtained by changing the sign of the cashflows c_i . Its value at time $t \leq T_{\alpha}$ is thus obtained by reversing the sign in (2.5).

Proof. Formula (2.5) easily follows from the following argument. Let $\Delta_i = \Delta(T_{i-1}, T_i)$. At every time instant T_i in the specified set of payment dates $T_{\alpha+1}, \dots, T_\beta$, the fixed leg pays out the amount $N\Delta_i K$ while the floating leg pays out the variable amount $N\Delta_i L(T_{i-1}, T_i)$, where $L(T_{i-1}, T_i)$ is the LIBOR rate at time T_{i-1} for the maturity given by the current payment instant T_i . Therefore, the net cash-flow at T_i is

$$\begin{aligned} c_i &= (L(T_{i-1}, T_i) - K)\Delta_i N \\ &= \left(\frac{1 - P(T_{i-1}, T_i)}{\Delta_i P(T_{i-1}, T_i)} - K \right) \Delta_i N \\ &= \left(\frac{1}{P(T_{i-1}, T_i)} - 1 - K\Delta_i \right) N \end{aligned}$$

where Definition (2.2) has been used. In order to determine the time t value of c_i notice that holding $\frac{1}{P(T_{i-1}, T_i)}$ at time T_i is equivalent to holding one unit of currency at time T_{i-1} , which in turn is equivalent to holding $P(t, T_{i-1})$ at time t . Equivalently, we can discount the fixed leg. The pricing formula for the swap can be obtained by summing up all the discounted cashflows:

$$\begin{aligned} V^{\text{PIRS}}(t) &= \sum_{i=\alpha+1}^{\beta} [P(t, T_{i-1}) - P(t, T_i) - K\Delta_i P(t, T_i)] N \\ &= NP(t, T_\alpha) - NP(t, T_\beta) - NK \sum_{i=\alpha+1}^{\beta} \Delta_i P(t, T_i). \end{aligned}$$

For the contract to be fair at time t , the fixed coupon rate K needs to be agreed such that $V_{\text{PIRS}}(t) = 0$. This defines the forward swap rate in (2.6). To conclude, we can rewrite the contract value as a function of the forward swap rate which ends the proof. \square

Note that the pricing of a basic Interest Rate Swap can be done without any assumptions about the underlying model. In fact, from (2.5) the instrument can be priced in $t = 0$ by simply using the interest rate instruments available in the market, that translate in today's zero-bond curve. The valuation at future time instants, that is for $t > 0$, is dependent on the model instead, since one needs to use the simulated values of the zero-coupon curve $R(t, T_i)$ (namely, the yield curve) to value $P(t, T_i)$.

In practice, the forward swap rate $S_{\alpha, \beta}(t)$ is used for the valuation of forward starting swaps. For example the 1x4 swap rate is used to value a swap entered today but starting in one year and ending in 4 years from now. This generalizes the simpler case in which we have to price a swap starting today, for which the spot swap rate

$$S_\beta(t) = \frac{1 - P(t, T_\beta)}{\sum_{i=1}^{\beta} \Delta_i P(t, T_i)}$$

is used instead.

2.2.2. European swaption

Definition 2.2.3 (European swaption). *A European swaption is the option on a Interest Rate Swap, i.e. the holder has the right, but not the obligation, to enter a swap at a future date at a given predetermined strike K . Similarly as for swaps the swaptions have payers and receivers who pay the fixed rate and receive the float leg (payer) or vice versa (receiver). We are going to assume that the swaption maturity coincides with the first reset date of the underlying IRS. For swaption maturity we intend some future date at which one has the option to enter into a swap deal, and this will happen if at maturity the contract has a positive value.*

Note that a payer swaption can be seen as a call option on a swap rate. In fact:

$$\begin{aligned} \left[V_{\alpha,\beta}^{\text{pSwpt}}(T_\alpha) \right]^+ &= \left[P(T_\alpha, T_\alpha) - P(T_\alpha, T_\beta) \right] - K \sum_{i=\alpha+1}^{\beta} \Delta_i P(T_\alpha, T_i) \Big]^+ \\ &= \left[S_{\alpha,\beta}(T_\alpha) A_{\alpha,\beta}(T_\alpha) - K A_{\alpha,\beta}(T_\alpha) \right]^+ \\ &= A_{\alpha,\beta}(T_\alpha) \left[S_{\alpha,\beta}(T_\alpha) - K \right]^+. \end{aligned} \quad (2.7)$$

In this context, it is important to make assumptions about the underlying model. There exist models that allow for analytical formulas for the pricing of swaptions in particular, and the Hull-White models we will consider are one of these. The market instead usually uses two benchmark models to quote swaptions: Black model or Normal model. The market quotes the swaptions in terms of implied volatility. In order to retrieve the price of that specific swaption (specified in terms of maturity, tenor, frequency of payments...) one then needs to plug in the quoted volatility into a pricing formula (either Black or Normal swaption pricing formula, depending on which volatility cube has been retrieved). A pricing formula can be obtained based on the distribution of the underlying swap rate, and typically the most common choice for this is to choose a log-normal distribution. Nevertheless, nowadays it is not considered strange to observe negative rates and therefore practitioners would want to consider this into their models. This problem can be overcome by either shifting the distribution (technique that leads to our first model - shifted log-normal model) or considering a normal distribution. In this thesis we are going to use Normal's (also known as Bachelier's) model as the market benchmark. This is done for two reasons: firstly, it does not assume positive rates, as we have just pointed out, and therefore all the volatility quotes are always available, for every tenor and maturity. Secondly, it assumes a normal distribution like the two Hull-White models that we will consider. In either cases, once the analytical formula for the price of a swaption has been derived, to obtain the price one only needs to plug in the volatility quoted in the market. In Subsections (3.2.4) and (3.3.2) we show how to derive the pricing formulas for European swaptions under the Hull-White model(s).

The following theorem provides the two aforementioned possible market practices for European swaptions pricing. Derivatives must be priced under risk-neutral measure to satisfy arbitrage-free condition and this assumption is adopted in both cases. In Black's formula the swap rate is assumed to follow a Geometric Brownian motion (and therefore, a log-normal process), while in Bachelier's formula it is assumed to follow a normal distribution.

Theorem 2.2.4 (Valuation of an EU swaption). *Using the annuity $A_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} \Delta_i P(t, T_i)$ as numéraire, the payoff of a swaption can be expressed as a call option on the swap rate $S_{\alpha,\beta}(T_\alpha)$. In formula, the payoff is as in (2.7).*

Now, let us call $\sigma_{\alpha,\beta}^B$ and $\sigma_{\alpha,\beta}^N$ the Black and Normal quoted volatilities respectively. If we assume a log-normal distribution for the swap rate, we end up with the so called "Black-76" formula [2]

$$\begin{aligned} V_B^{\text{Swpt}}(t_0; T_\alpha, T_\beta) &= N A_{\alpha,\beta}(t_0) (S_{\alpha,\beta}(t_0) \omega \Phi(\omega d_1) - K \omega \Phi(\omega d_2)), \\ d_1 &= \frac{\log\left(\frac{S_{\alpha,\beta}(t_0)}{K}\right) + \frac{(\sigma_{\alpha,\beta}^B)^2 (T_\alpha - t_0)}{2}}{\sigma_{\alpha,\beta}^B \sqrt{T_\alpha - t_0}}, \quad d_2 = d_1 - \sigma_{\alpha,\beta}^B \sqrt{T_\alpha - t_0}, \end{aligned}$$

where $\omega = 1$ if we are pricing a payer swaption, and $\omega = -1$ for a receiver swaption. Otherwise, the

assumption of a normal distribution for the swap rate leads to the following pricing formula [28]:

$$V_N^{\text{Swpt}}(t_0; T_\alpha, T_\beta) = NA_{\alpha,\beta}(t_0) \left(\omega(S_{\alpha,\beta}(t_0) - K)\Phi(\omega d) + \frac{\sigma_{\alpha,\beta}^N \sqrt{T_\alpha - t_0}}{\sqrt{2\pi}} e^{-\frac{d^2}{2}} \right),$$

$$d = \frac{S_{\alpha,\beta}(t_0) - K}{\sigma_{\alpha,\beta}^N \sqrt{T_\alpha - t_0}}.$$

2.3. Physical and risk-neutral measure

In mathematical finance, there exist two branches in term of measures: the “Q” area of derivatives pricing, where the main scope is to extrapolate the present, and the “P” area of quantitative risk and portfolio management, where the main goal is to model the future. Risk neutral measure can be somehow seen as a measure that represents the map of current market prices. Historical measure uses past prices to provide an estimate for future prices.

Quantitative derivatives pricing was firstly initiated by Bachelier in 1900 who applied the most basic stochastic process, the Brownian motion, to the problem of options pricing. The theory remained quite unstudied though, until Merton (1969) and Black and Scholes (1973) applied the geometric Brownian motion process to this problem. Then, another fundamental step was given by Harrison and Pliska (1981) with the Fundamental Theorem of Asset Pricing. This stated that the current price V_0 of a security (suitably normalized) is arbitrage-free if there exists a stochastic process V_t , that is a martingale and describes its future evolution. Equivalently, the theorem implies that in a complete market a derivative’s price is the discounted expected value of the future payoff under the unique risk-neutral measure, which exists if and only if the market is arbitrage-free. This means that we can create a market where assets offering the same return will share the same price in $t = 0$. Mathematically speaking, if we call H_t the price at time t of a contingent claim with payoff H_T at time $T > t$, this will be given by

$$H_t = \mathbb{E}_t \left[e^{-\int_t^T r(s) ds} H_T \right], \quad (2.8)$$

where \mathbb{E}_t denote the expectation under the risk-neutral measure \mathbb{Q} with respect to the filtration \mathcal{F}_t . Here the numéraire is the money-market account introduced in (2.1.1). For example, from Definition (2.1.3) we know that the Zero-Coupon Bond price at time t for the maturity T is characterized by a unit amount of currency available at time T . In this case, $H_T = 1$. This means that, for a Zero-Coupon Bond, Equation (2.8) reads:

$$P(t, T) = \mathbb{E}_t \left[e^{-\int_t^T r(s) ds} \right]. \quad (2.9)$$

On the other side, there is the real-world measure \mathbb{P} , which reflects the historical behaviour in the market and contrary to the risk-free measure, it can actually give the opportunity for an arbitrage, as prices today can be either cheap or expensive. If we consider the measurable space (Ω, \mathcal{F}) , where Ω is the universe of all possible outcomes and $\mathcal{F} = 2^\Omega$ the power set on Ω , then \mathbb{P} maps any event in the σ -algebra to a probability (i.e. $\mathbb{P} : \mathcal{F} \mapsto [0, 1]$). Same holds for \mathbb{Q} . Nevertheless, these two measures have very few things in common. One of these is that they agree on the unobservable sets, that is for every $A \in \mathcal{F}$, if $\mathbb{Q}(A) = 0$ then $\mathbb{P}(A) = 0$ as well.

In this thesis, we rely on the physical measure \mathbb{P} for the calibration of the benchmark model, that is the shifted log-normal model (DEV-MR) introduced in Section (3.1). Time series of log-returns of nine forward rates will be used for historical calibration. Whereas, for the two Hull-White models, a risk-free calibration will be performed and therefore we will rely on the measure \mathbb{Q} . Again, the idea is to analyze the capital impact of using the risk-free measure instead of the historical measure (which is commonly used for portfolio management, as previously mentioned).

2.4. Affine Term-Structure Models (ATSM)

In this thesis we have researched two short rate models, the Hull White one factor model and Hull White two factors model. Both of them belongs to a particular class, that is defined as follows.

Definition 2.4.1 (Affine processes class). *Suppose to have the following system of SDEs:*

$$d\mathbf{X}_t = \bar{\mu}(\mathbf{X}_t) dt + \bar{\sigma}(\mathbf{X}_t) d\widetilde{\mathbf{W}}_t,$$

where $\widetilde{\mathbf{W}}_t$ are independent Brownian motions. The processes belonging to the affine diffusion (AD) class are assumed to have drift, volatility, and interest rate components of the affine form, that means:

$$\begin{aligned}\bar{\mu}(\mathbf{X}_t) &= a_0 + a_1 \mathbf{X}_t \quad (a_0, a_1) \in \mathbb{R}^n \times \mathbb{R}^{n \times n}, \\ \bar{\sigma}(\mathbf{X}_t) \bar{\sigma}(\mathbf{X}_t)^T &= (c_0)_{ij} + (c_1)_{ij}^T \mathbf{X}_t, \quad (c_0, c_1) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times n}, \\ r(\mathbf{X}_t) &= r_0 + r_1^T \mathbf{X}_t, \quad (r_0, r_1) \in \mathbb{R} \times \mathbb{R}^n\end{aligned}$$

Example 2.4.2 ($HW - 1F$). Considering the diffusion of the short rate in $HW - 1F$ given by (3.7), it is easily proved that it belongs to the family by taking:

$$\begin{aligned}\bar{\mu}(r(t)) &= a_0 + a_1 r(t) = \theta(t) - ar(t), \\ \bar{\sigma}(r(t))^2 &= c_0 + c_1 r(t) = \sigma(t)^2 + 0r(t), \\ r(r(t)) &= r_0 + r_1 r(t) = 0 + 1r(t).\end{aligned}$$

Example 2.4.3 ($HW - 2F$). Considering the diffusion of the short rate under the $G2++$ model formulation given by (3.34), it is easily proved that it belongs to the family by taking:

$$\begin{aligned}\mathbf{X}_t &= [x(t) \ y(t)]^T, \\ a_1 &= \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}, \\ a_{12} = a_{21} = a_{22} = a_0 &= [0 \ 0]^T, \\ r_0 = \varphi(t), \ r_1 &= [1 \ 1]^T, \\ c_0 &= \begin{bmatrix} \sigma^2 + (\eta\rho)^2 & \eta^2 \rho \sqrt{1-\rho^2} \\ \eta^2 \rho \sqrt{1-\rho^2} & \eta^2 (1-\rho^2) \end{bmatrix}, \quad (c_1)_{i,j} = 0, \quad \forall i, j.\end{aligned}$$

A special property that characterizes this class is that they all provide explicit pricing formula for the Zero-Coupon Bond. which is a very useful property for our purposes as bonds are directly linkable to the yield curve and therefore to the exposure in vanilla swaps. More generally, in [10] it is proven that for affine diffusion processes the discounted characteristic function, defined as:

$$\varphi(\mathbf{X}_t, t, T, \mathbf{u}) := \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(\mathbf{X}_s) ds} e^{i\mathbf{u} \mathbf{X}_T} \middle| \mathcal{F}_t \right], \quad \text{for } \mathbf{u} \in \mathbb{R}^n,$$

with boundary condition:

$$\varphi(\mathbf{X}_T, T, T, \mathbf{u}) = e^{i\mathbf{u}^T \mathbf{X}_T},$$

has a solution of the following form:

$$\varphi(\mathbf{X}_T, t, T, \mathbf{u}) = e^{A(\mathbf{u}, t, T) + B(\mathbf{u}, t, T)^T \mathbf{X}_t}.$$

The coefficients $A(\mathbf{u}, t, T)$ and $B(\mathbf{u}, t, T)$ are found to satisfy the following system of Riccati-type ODEs:

$$\begin{cases} \frac{d}{d\tau} A(\mathbf{u}, \tau) &= -r_0 + B^T a_0 + \frac{1}{2} B^T c_0 B, \\ \frac{d}{d\tau} B(\mathbf{u}, \tau) &= -r_1 + B^T a_1 + \frac{1}{2} B^T c_1 B, \end{cases} \quad (2.10)$$

where $\tau = T - t$.

2.5. $Gn++$ models (multi-factors)

Two out of the three models that have been investigated in the thesis are part of a more general family of models. We are referring to the two Hull-White models, and the family of multi-factors models under which the spot rate r is described by

$$\begin{cases} r(t) &= \sum_{i=1}^n x_i(t) + \beta(t), \\ x_i(t) &= -a_i dt + \sigma_i dW(t), \quad \forall i = 1, \dots, n. \end{cases} \quad (2.11)$$

This means that under this particular family of models, the spot rate r is described by n Gaussian stochastic processes (factors) plus a time-dependent shift, which ensures a perfect fit to the term structure of interest rates. More precisely, it can be proven that the Hull-White one factor formulation is equivalent to (2.11) with $n = 1$ ($G1++$ model), whereas the Hull-White two factors formulation can be proven to be equivalent as well with $n = 2$ ($G2++$ model). In general, more factors are introduced in order to accommodate a better fit to market data and a more realistic description of the correlation structure between rates. Problem is that this comes at the cost of less mathematical tractability and can also result in reduced stability of the calibrated parameters, as seen in Section (3.3).

Historical analysis in [21] of the whole yield curve, based on Principal Component Analysis (PCA), suggests that under the physical measure \mathbb{P} two components can explain 85% to 90% of variations in the yield curve, whereas a third component can “only” explain from 93% to 95%. A similar but more optimistic analysis is carried out in [27] for the US market, where it is shown that one component explains 92% of the total variance. In any case, the literature tells us that it should be definitely enough to consider two factors. Three or more factors could turn out to be too computationally demanding and excessively unstable in the parameters, which is for our purposes undesired. Especially, pricing precision does not constitute much of a concern in risk-management. Despite of this, it remains still interesting to investigate a two factors model so that its more realistic description of correlations can be exploited.

3

The stochastic models

After the 2008 financial crisis, central banks have loosen their monetary policies to boost the economy. The main tool in the expansive monetary policy was lowering the interest rates to allow banks to fund cheaply and thereby motivate them to increase their lending activity. The last decade is thus characteristic for very low interest rates, across virtually all countries. In Figure (3.1) we show the evolution of interest rates from mid-2010 until mid-2018 for four different currencies: CHF, EUR, JPY and USD. A few patterns can be identified. First of all, until 2015 all rates were decreasing. In general rates are very low, getting close or oscillating around zero, with CHF having nearly hit -1% . The USD rate started to sharply increase from 2016, whereas the other rates have stabilized in 2017-2018 and are no longer decreasing. In addition, once rates are negative, these start to show a very small variation. From this we understand that the variation decreases with the level of the rates. Ideally, the interest rate model should be thus able to capture all these patterns. Note that for our purpose, only one specific asset class is concerned: the Euro interest rate, as Rabobank is mostly exposed to this currency. The models that have been studied during the research are three, and they are treated in this chapter. While two of them are by definition able to simulate negative rates as the interest rate that they model is assumed to follow a normal distribution, the other one is not as it assumes a log-normal distribution for the stochastic process. To overcome this problem, we modify it by applying a (judiciously chosen) shift.

Throughout this chapter, the interest rate models used for exposure simulation are mathematically introduced, calibrated and simulated. With regard to the Displaced Exponential-Vasicek model, we highlight the inability of properly correlating the rates after relatively short time horizons. We reckon this to constitute a very big point of weakness for our purposes. In fact, the final goal is to compute the exposure distributions of portfolios of swaps, whose value is well known to depend on several different rates. It is therefore very important to be able to well capture their correlations. Especially, this must be possible for both close and far future horizons, as Interest Rate Swaps with relatively long time-to-maturities are commonly traded within the bank. On the other hand this problem does not affect the two Hull-White models, as we will see in (3.2) and (3.3). Furthermore, in this chapter we propose a calibration method that is applicable to both the two Hull-White models, which see the calibration of the volatilities and speeds as two separate tasks. The two factors model is well known to require more time for the calibration step to be performed, and especially the tricky problem of ending up in a local minima is highly probable to happen. Nevertheless, with time-dependent volatilities it is possible to give heuristical reasoning about why our solution corresponds to the global minima. At the end of the chapter, a brief discussion prior to the real application of the models in Chapter (4) is addressed, with the main intent of making a thoughtful balance between complex and simple models and give our personal opinion up to that moment.

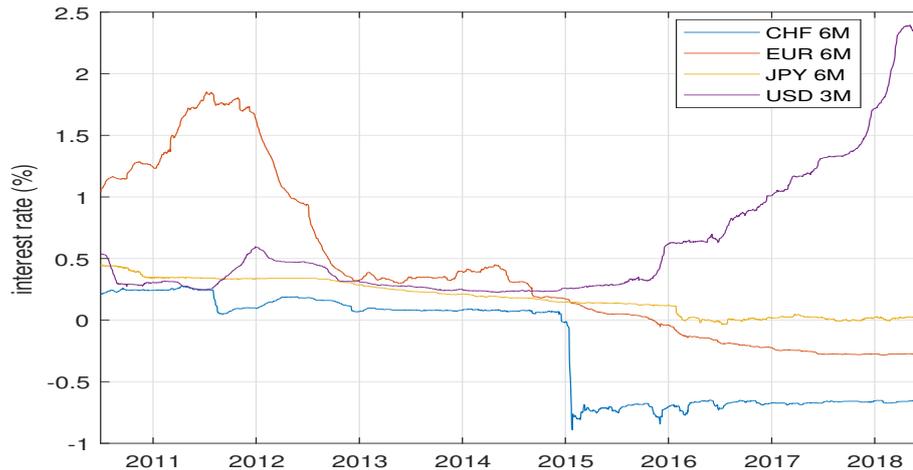


Figure 3.1: Evolution of main interest rates from 2010 to 2018.

Calibration

The choice of models for the underlying market variables is a very important issue. Nevertheless, the calibration of such models is as important as future scenarios are highly impacted by the model parameters. As already stated, two kind of calibrations have been experimented in the research: historical and risk-neutral. We will see that the mean reversion speed and volatility are two parameters that are highly linked together, when these are calibrated to the risk-free measure. That is, increase in the speed translates to increase in the volatility and vice versa. This has huge implications on the simulated rates, and as a consequence to the exposures. Models calibrated on historical data assumes that what has been observed in the past is a good indicator of what can happen in the future. Luckily, positive probability are still given to events that have not been observed yet. Nevertheless, it is well known that this type of calibration is very slow in reacting to changes in the markets, i.e. when markets become suddenly more volatile. This is not the case when dealing with risk-neutral calibration, which is always up to date as it only relies on today's market data. On the other hand, historical calibration is stable over time, which constitutes a very attractive property for risk management purposes. This is mainly due to the fact that very long time series of data are used, and therefore performing a calibration today and six months later cannot bring too serious fluctuations in the profiles.

Simulation of interest rate models can be done in different ways depending on the application. The most two popular techniques are trees and Monte-Carlo simulations. Trees describe a process as a sequence of an ever-expanding set of potential outcomes. In a few words, the state of the world today gives rise to several possible outcomes tomorrow. Each of these in turn will give rise to multiple possible outcomes the day after tomorrow, and so on. This means that the number of possible outcomes increases with the time horizon. Sometimes, the result is that the computational cost can be high. Still, this is worth it when it is important to make decisions along the trajectory of the stochastic process. Contrary to this, Monte-Carlo simulations are in nature characterized by a constant number of possible scenarios (known as "paths") throughout the whole evolution of the stochastic process. In general, trees turn out to work particularly well whenever the payoff of the derivative is not path dependent, as described in [23]. On the other hand, Monte-Carlo simulations represent a better choice when the payoff is path-dependent. Furthermore, this approach is preferable when dealing with high-dimensional problems (DEV-MR model) and in general in risk-management. This drives us in choosing the second simulation technique.

Monte-Carlo simulation

Two Monte-Carlo simulations have been approached during the research: exact simulation and a discretized simulation. While the first one is only possible when the variable distribution is known, the second one is applicable directly to the model dynamics and does not require the knowledge of an explicit solution. The first one is called “exact” in the sense that no discretization is involved, as it simply consists in drawing standard normal variables and correlating them if necessary. Regarding the second one, the simple Euler discretization schema was used. This is expected to accrue some error over time. Nevertheless, we conducted a small analysis and concluded that the accrued error is negligible. The reasons that drive us to prefer the exact simulation are two: first is that it turns out to be faster for all three models. Second reason is that for risk management purposes it is usually not necessary to simulate on uniform grids but rather on grids that are dense at the beginning and get wider for increasing time horizons. This can only be done via exact simulation, that is via sampling the distribution at the grid points of interest. Let us call $\tau = T - t$. Throughout the whole section, the following two graphical representations will be used, which we reckon to be very descriptive of both the shapes and values of the simulated yields of a given model:

1. $\tau_t \mapsto R(t, T)$: tenor τ in the x axis, simulation time t in the legend.
2. $t_\tau \mapsto R(t, T)$: simulation time t in the x axis, tenor τ in the legend.

The variable in subscript is intended to be fixed. More specifically, with the first graphical representation we will show the quantiles of the simulated yields so to understand how much fast and in which way the shapes and values change starting from today’s interest rate curve. The second graphical representation will be mainly to understand how a specific point of today’s curve will behave over time according to the given model. Note that the number of scenarios computed for this purpose is 40000, while for the results of Chapter (4) we reduce them to 2000, so not to over increase unnecessarily the computation times of the stability analysis.

3.1. Shifted log-normal model

Interest rates are the most important risk factors contributing to the bank’s portfolio. These can be simulated based on a certain model whose parameter are calibrated to a certain measure. The central goal of the research is to determine how this portfolio is impacted based on interest rate model and calibration choices. Precisely, we would like to provide an analysis of the different implications of either the physical or risk-neutral measure in the Counterparty Credit Risk framework by means of three different models. Together with this, we clarify improvements and drawbacks of each choice.

In this section we introduce the model that will serve as a benchmark with which to evaluate the performance of the two Hull-White models. This model will either be named as shifted log-normal model or Displaced Exponential-Vasicek mean-reverting model (DEV-MR). It will be used for the diffusion of nine forward rates as defined in (2.1.6). We will briefly describe the calibration procedure, which is done the time series of nine zero rates, and provide the estimated parameters. Furthermore, we will take care of the problem of modelling the correlation between these set of rates. This will be done via historically calibrating a certain correlation matrix and then defining a transformation matrix used for updating such correlation structure for future time instants. Eventually we will perform the simulations and outline the problems of the model. We believe that these problems can be solved by introducing the two Hull-White models, even though these come at the cost of other problems. Note that as the nine forward rates are governed by a particular correlation structure, we can see this model as a nine-factors model and therefore we reckon a Monte-Carlo simulation to be the best choice. As a matter of fact, this approach to simulation is well known to be

suitable to high-dimensional problems, rather than for example, trees.

This model is a variation of the Exponential-Vasicek (EV) model introduced in [5]. As the name suggests, a shift is introduced in order to give positive probability to negative values of the process. The shift is a hyper-parameter, that has to be explicitly specified before the model is calibrated. As a result, all model parameters are subject to the choice of the shift. Note that this model was firstly introduced as a short rate model but can be used as well for the simulation of other kind of rates. This is indeed our case: as previously mentioned, the model will be used for the simulation of a given set of “forward rates”. More specifically, the idea is to indirectly model the yield curve by modeling a set of points belonging to the yield curve itself, since from (2.4) we know that a straight-forward relation between forward rates and yield curve exists. This model, as well as his well-known ancestor Vasicek model, does not belong to the family of Affine Term-Structure Models and therefore an explicit formula for the Zero-Coupon Bond cannot be derived. The price of any financial product needs to be computed numerically under this model. This is of course one of its major defeats, especially when it comes to derivative pricing. Nevertheless, this is not going to touch us as this model will not be calibrated using market prices, but rather using historical prices.

3.1.1. Introduction to the model

The model find its roots in the simple Vasicek model introduced in [32], which suggests that the dynamics of $r(t)$ could be governed by a mean-reverting Ornstein-Uhlenbeck process. Therefore, it assumes that the spot rate evolves according to:

$$dr(t) = [b - ar(t)] dt + \sigma dW(t), \quad r(0) = f^M(0,0), \quad (3.1)$$

where b, a and σ are positive constants. The Exponential-Vasicek model simply assumes Equation (3.1) where $r(t)$ is replaced by $\ln[r(t)]$. At this stage, the model is not yet capable of modelling negative rates. In order to allow this, we shift the stochastic process by a constant and positive parameter γ . Therefore, the quantity that is actually being modelled is $\ln[r(t)] + \gamma$. Given the heuristic of this model, we can now formulate it more precisely.

We say that the stochastic process X_t diffuses according to the Displaced Exponential-Vasicek model if it satisfies the following stochastic differential equation:

$$d \ln(X_t + \gamma) = (b - a \ln(X_t + \gamma)) dt + \sigma dW_t, \quad (3.2)$$

where a is the mean reversion speed, $\frac{b}{a}$ is the long-run mean, γ is the shift parameter, and σ the volatility parameter. By means of the Ito's isometry, the solution to (3.2) turns out to be:

$$X_t = \exp \left[\ln X_s e^{-a(t-s)} + \frac{b}{a} (1 - e^{-a(t-s)}) + \sigma \sqrt{\frac{1 - e^{-2a(t-s)}}{2a}} \varepsilon \right] - \gamma,$$

where $s < t$ and $\varepsilon \sim \mathcal{N}(0,1)$. As a consequence, the process X_t follows a log-normal distribution. Note that this X_t is always mean reverting to a constant value and:

$$\lim_{t \rightarrow \infty} \mathbb{E}[X_t | \mathcal{F}_s] = \exp \left(\frac{b}{a} + \frac{\sigma^2}{4a} \right) - \gamma =: \xi. \quad (3.3)$$

This is well-known to be a nice property for interest rate models, as historical rates exhibit this behaviour. Note that

$$\mathbb{P}(X_t < 0) = \mathbb{P} \left(\frac{\ln \gamma - \ln X_s e^{-a(t-s)} - \frac{b}{a} (1 - e^{-a(t-s)})}{\sigma \sqrt{\frac{1 - e^{-2a(t-s)}}{2a}}} < 0 \right) \xrightarrow{t \rightarrow \infty} \Phi \left(-\frac{\frac{b}{a} - \ln \gamma}{\sqrt{\frac{\sigma^2}{2a}}} \right) \approx 0.$$

τ_0	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9
0M	1M	3M	6M	1Y	2Y	5Y	10Y	20Y	30Y

Table 3.1: Time lengths defining the 9 forward rates.

This easy fact affects the simulated term structures, which exhibit an increasing trend for increasing simulation times and for both low and high term to maturity rates, as seen in (3.1.5). Equivalently, the number of realized negative values decreases in t , which is something that does not happen in both two Hull-White models, as we will see in (3.2) and (3.3).

3.1.2. Modelling the forward rates

The DEV-MR model will be used to model the nine (continuously-compounded) forward interest rates $f(t; T_{i-1}, T_i) = f(t; t + \tau_{i-1}, t + \tau_i)$, for $i = 1, \dots, 9$. The chosen values for the τ_i 's are shown in table (3.1). In what follows, we will lighten the notation by shortening $f(t; T_{i-1}, T_i)$ to $f_i(t)$. For example, $f_4(t)$ is the forward rate at current time t for a period between 6M (expiry date) to 1Y (maturity date) from now. We then need to set $X_t = f_i(t)$ in (3.2). As a consequence each forward rate in the DEV-MR model is simulated through the following equation:

$$f_i(t) = \exp \left[\ln[f_i(t_0) + \gamma] e^{a_i(t-t_0)} + \frac{b_i}{a_i} (1 - e^{a_i(t-t_0)}) + \sigma \sqrt{\frac{1 - e^{-2a_i(t-t_0)}}{2a_i}} \varepsilon_i \right] - \gamma, \quad (3.4)$$

where ε_i is a standard normally distributed random variable and γ is a positive shift. All the nine points $f_i(t_0)$ are clearly retrievable from today's yield curve. Note that this model is not designed to fit the current term structure of interest rate, due to the constant mean reversion level. This problem is overcome by turning this formulation into a nine factors model, so that nine points of the yield curve are exactly matched. The formula to apply in order to map the forward rates back to the yields is:

$$f_i(t) = \frac{(T_i - t)R(t, T_i) - (T_{i-1} - t)R(t, T_{i-1})}{T_i - T_{i-1}},$$

which is nothing but a consequence of our notation and Equation (2.1.7).

3.1.3. Historical Calibration

For historical calibration, it is necessary to decide what history of data to use. It would make sense not to consider data that are too old and therefore meaningless, but at the same time one needs to take into account that too short histories would lead to poor statistics. Loosely speaking, the two drift parameters a and b are calibrated to match historical upper and lower bounds of rates, while the standard deviation is calibrated somehow from the standard deviation of the realized residuals, which can be made explicit from (3.4). In this case, as historical residuals are not perfectly normally distributed but rather exhibit a more fat-tailed distribution, a Kurtosis adjustment is applied to the volatility parameter. Anyhow, we reckon the technicalities of how the historical calibration is performed not to be interesting for the research and therefore such steps are omitted. We here provide only the values of the calibrated parameters so to allow for clear interpretation of the simulated yield curves and for comparisons with the other two models. In the DEV-MR model, the volatility parameter σ has been calibrated to the last 8 years' data and results in $\hat{\sigma} = 19.98\%$. On the other hand, the mean reversion parameters a and b are calibrated to the last 30 years of data and results are provided in Table (3.2), together with the long-run mean. This is calculated as $\xi = \exp\left(\frac{b}{a} + \frac{\sigma^2}{4a}\right) - \gamma$.

	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6	τ_7	τ_8	τ_9
\hat{a}	5.8	5.8	5.8	5.8	6.3	7.8	9.9	12.4	13.6
\hat{b}	-18.2	-18.2	-18.2	-18.2	-19.7	-23.5	-28.6	-34.7	-37.6
ξ	3	3.0	3.0	3.0	3.2	3.7	4.1	4.6	4.8

Table 3.2: Historically calibrated drift parameters (last 30 years of data) in percentages.

3.1.4. Correlation structure

Here we give an explanation of how the correlation structure between the nine forward rates is captured under the DEV-MR model. This procedure must be done in conjunction with the simulation step, explained in (3.1.5).

Each rate is driven by one Brownian motion and in order to compute the simulation of the nine rates in a proper way, the instantaneous correlation matrix needs to be determined. That is, if we call \mathbf{W}_t the 9-dimensional Wiener process driving the diffusion of our set of rates, the aim is to compute the matrix ρ such that:

$$\begin{aligned}\Sigma_{i,j} &= \text{Cov}(dW_t^i, dW_t^j) \\ &= \mathbb{E}(dW_t^i, dW_t^j) = \rho_{i,j} dt,\end{aligned}$$

where Σ is the covariance matrix of the multivariate-normally distributed vector $d\mathbf{W}_t$.

Consider that we want to compute 2000 scenarios per rate. Then, at each simulation time t in the set of reporting dates, we will sample simultaneously all the nine rates at that time. More precisely, we would need to sample a \tilde{Z} of size 9×2000 , where each column (corresponding to one specific scenario) is a vector of dependent normal random variables, whose dependency is given by a certain correlation matrix $\rho(t)$. Eventually formula (3.4) is used to compute $f_i(t)$. In the following we explain how this correlation matrix is computed at each t . Note that correlation of Brownian motions is a step which is not itself visible in the simulated profiles. That is, the correlation structure does not impact the marginal distribution of each forward rate, but rather their joint distribution, and thus the simulated yield curves (in both shapes and values) and the revaluation of financial instruments whose payoff is affected by rates at different maturities, which are correlated one with the other (and in our case, plain vanilla swaps).

Correlation calibration and computation of the transformation matrix A

We here explain how the correlation matrix at t_0 is calibrated and how to calculate the updated correlation matrix at time steps $t > t_0$. The idea behind the correlation calibration is to estimate pairwise correlations $\rho_{i,j}$ based on the pairwise sample correlations of the realizations of residuals. The residuals in the DEV-MR model are constructed by making ε explicit in (3.4), which (assuming Δt small enough) results in:

$$\varepsilon_i(t) = \frac{\ln\{f_i(t + \gamma)\} - e^{-\hat{a}_i \Delta t} \ln\{f_i(t - \Delta t + \gamma)\} - \frac{\hat{b}_i}{\hat{a}_i} (1 - e^{-\hat{a}_i \Delta t})}{\hat{\sigma} \sqrt{\Delta t}}.$$

Then, the calibrated correlation $\rho_{i,j}$ in t_0 is simply computed as the sample correlation $\text{Corr}(\boldsymbol{\varepsilon}_{\tau_i}, \boldsymbol{\varepsilon}_{\tau_j})$. In order to respect the assumption on Δt , in practice a time step of 10 business days is considered ($\Delta t = \frac{10}{260}$). As one could easily expect, the historically calibrated correlation matrix thus results such that the correlation between close tenors is high, while is low for tenors that are farther apart in the yield curve.

For computing the transformation matrix $A(t)$, it is handy to notice that the only source of randomness in (3.4) comes from the normal random variable ε_i . This can be easily expressed, by means of Ito's isometry, as:

$$\varepsilon_i(t) = \frac{1}{\Delta t} \int_{t-\Delta t}^t e^{a_i u} dW^i(u).$$

From this we can derive the transformation matrix $A(t)$, as explained in the following proposition.

Proposition 3.1.1. *The elements of matrix $A(t)$ should be computed in this way: $A(t)_{i,i} = 1$, for $i = 1, \dots, 9$ so that the updated matrix preserve the nature of correlation matrix. Otherwise, if $i \neq j$:*

$$A(t)_{i,j} = \frac{2\sqrt{a_i a_j}}{a_i + a_j} \frac{(1 - e^{(a_i + a_j)t})}{\sqrt{(1 - e^{2a_i t})(1 - e^{2a_j t})}}. \quad (3.5)$$

A proof of this can be found in Appendix (A). The correlation matrix at time t will then be simply computed as the element-wise product between $A(t)$ and ρ_{t_0} . This said, it is important to notice the following two aspects, which hold when $i \neq j$:

- $\rho(t)_{i,j} \approx \rho(t_0)$ when t is small, as $A(t)_{i,j} \approx 1$.
- $\rho(t)_{i,j} \rightarrow 0$ when $t \rightarrow \infty$, as $A(t)_{i,j} \approx 0$.

The drawback about the flattening quantile shapes can be now motivated by the second point. As we go on in simulation time t the correlations between the different rates (extra-diagonal elements of $\rho(t)$) tends to 0. In some way we can say that the correlation matrix brings the information about the yields' shapes. This means that that after a certain grid point t the correlations are forced to approach zero. This would not be a problem, as soon as this happens for t that is not too small. The problem is that this seems to happen in a relatively short time.

Having said that, the procedure we follow in order to properly correlate and simulate the rates can be summarized through the following three steps:

1. Calibrate historically the correlation matrix ρ_{t_0} , where with t_0 we refer to the current date.
2. Compute a certain transformation matrix $A(t)$ and then calculate the correlation matrix at time $t \geq t_0$ via the element-wise matrix product $\rho(t) = A(t)\rho_{t_0}$.
3. Simulate the correlated normal random variables via decomposition of $\rho(t)$. Which decomposition method to use shall depend on whether the matrix is positive definite or not.

3.1.5. Simulation

Simulation of correlated standard normals

Let $\tilde{Z} \sim \mathcal{N}(0, \rho)$ be a vector of correlated standard normal variables, with ρ the correlation (equivalently, the covariance) matrix. To simulate such vector, first of all simulate a vector of independent normal random variables Z . If there exists *any* decomposition such that:

$$LL^T = \rho, \quad (3.6)$$

then correlated random normals \tilde{Z} can be simulated from Z and ρ using the formula:

$$\tilde{Z} = LZ.$$

That is, any L satisfying (3.6) can be used to simulate correlated normals. The Cholesky decomposition of ρ is a popular choice for L , where L has the special property that it is a lower triangular

matrix. However, another popular choice of L is through the Principal Component Analysis (PCA). This approach offers the benefit of a possible factor reduction. Most importantly it is useful when ρ is not positive definite, as sometimes happens in practice. This is certainly something that cannot be dealt by the Cholesky decomposition. Using PCA, the L to be used in the simulation is found as

$$\rho = VDV^T = V\sqrt{D}\sqrt{D}V^T = LL^T \implies L = V\sqrt{D},$$

where any negative element in D is replaced by zero.

The results for the simulation of $f_1(t)$ and $f_9(t)$ are respectively shown in Figures (3.2) and (3.3). Note that since this is a high-dimensional problem (as nine diffusion are concerned) a Monte-Carlo simulation is suggested. At the beginning of the chapter we gave evidence of the fact that for our purposes an exact simulation is convenient. We therefore sample correlated normal variables and exploit the solution in (3.4). The red and blue lines represent respectively the simulated 97.5% and 2.5% quantiles of the rate. As expected, the model is able to simulate negative rates because of the applied shift and the lower quantiles are slowly increasing. Furthermore, the skewness of the log-

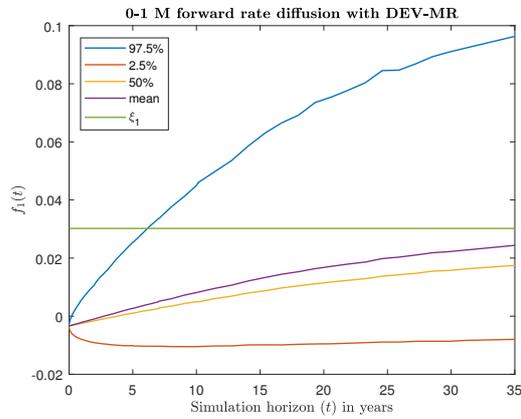


Figure 3.2: Diffusion of the forward rate with expiry date in τ_0 and maturity date in τ_1 .

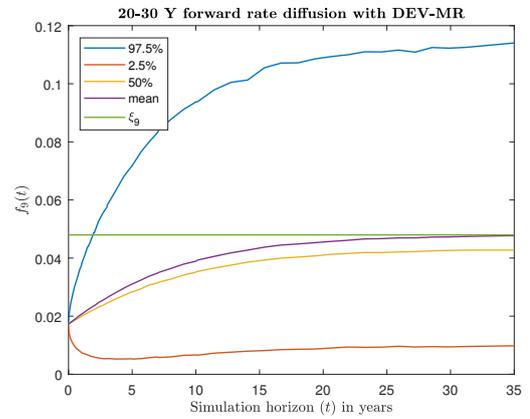


Figure 3.3: Diffusion of the forward rate with expiry date in τ_8 and maturity date in τ_9 .

normal distribution is recognizable from the non-overlapping mean and median of the simulations, drawn respectively in yellow and purple lines. The starting points of the simulations are different per simulated forward, and each of them coincide to one of the nine point of today's term structure. Note that the forwards with higher Expire-to-Maturity dates, that is in this case $f_9(t)$, exhibit bigger mean reversion speeds and higher starting points compared to the forwards with lower tenors, that is in this case $f_1(t)$. This translates in much slower convergence to the mean reversion level ξ (green line) and lower volatility in (3.2). Note that this is not surprising as $\hat{a}_9 \approx 2\hat{a}_1$. In addition, a further simulation has shown that for $i = 1$ we get a satisfactory convergence at simulation times between 50 – 60 years, that is double the time required for $i = 9$.

In Figure (3.4) the simulated forward interest rates quantiles implied by the shifted log-normal model are displayed, whereas in Figure (3.5) the yields quantiles are shown as a function of the tenor lengths. On the other hand Figure (3.6) and (3.7) displays respectively the simulated quantiles through time and the simulated term structures in three dimensions, which previews the biggest drawback of the model. That is, as the simulation time t increases, the shape of the simulated curves get very wild and are far from reproducing the current term structure, making the predicted values unreliable. This problem find its reasons in how the correlation matrix between the nine rates is updated for increasing simulation times, as explained in the previous paragraph.

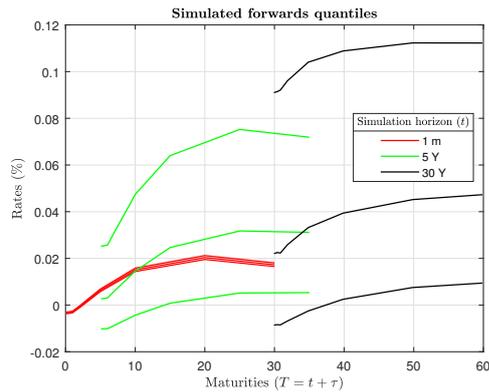


Figure 3.4: Forwards quantile shape with DEV-MR.

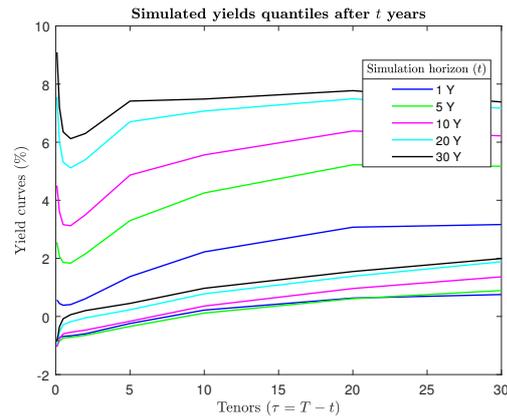


Figure 3.5: Yields quantiles as function of tenors with DEV-MR.

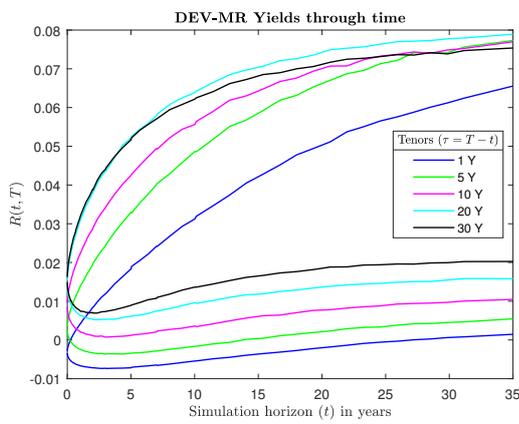


Figure 3.6: Yields quantiles as function of time with DEV-MR.

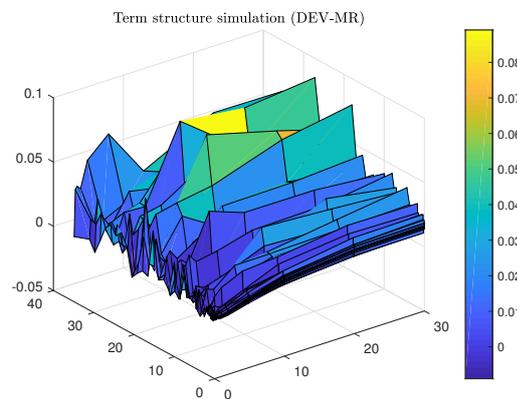


Figure 3.7: Simulated term structures under DEV-MR model in 3D.

Representative yields

The yield curve being simulated with the DEV-MR model are intersecting. As a consequence, one could easily argue that the representation of Figures (3.4) and (3.5) might be informative about the values but not the shapes, as the quantiles of simulations does not represent yield curves that have been actually simulated. To better grasp the DEV-MR’s drawback, it may be interesting to produce another graphical representation by showing two representatives of the 2.5% and 97.5% quantiles, selected out of the whole group of simulated yields. This way one could better display how the shapes develop in the simulations and from which point in time the typical yield shape is lost. There are many ways with which one could perform this selection. The simplest one would be as follow. At each simulation time t , we have our matrix of simulated yields 2000×9 . We compute the averages over rows and then rank such averages and take the quantiles from that. This way two rows out of the matrix of simulated yields have been selected, which correspond to the two representative yields. The results for this other way of assessing the quantiles are shown in Figures (3.8), (3.9), (3.10) and (3.11). The blue and red lines in the graphs are respectively the two simulated yields that “represent” the upper and lower bounds of interest. This said, we can understand that the correlation matrix governs the shapes of the predicted yield curve, which loose in this case any link to the initial yield curve starting already from 3 years simulation. Note that this won’t be the case with the two Hull-White models.

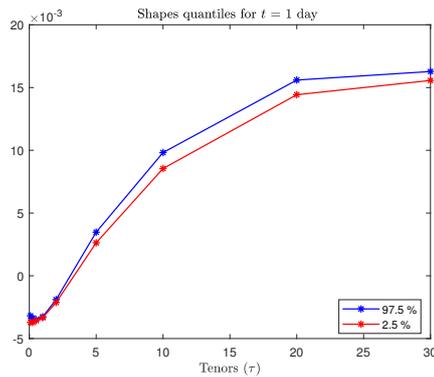


Figure 3.8: Yields quantiles shapes after 1 day simulation with DEV-MR.

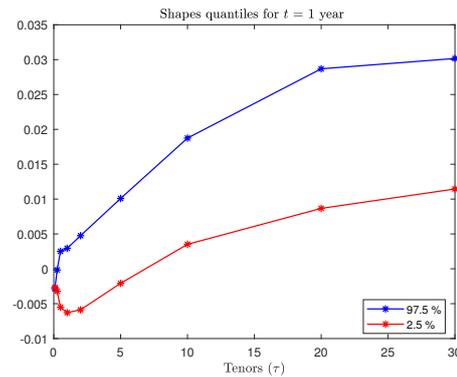


Figure 3.9: Yields quantiles shapes after 1 year simulation with DEV-MR.

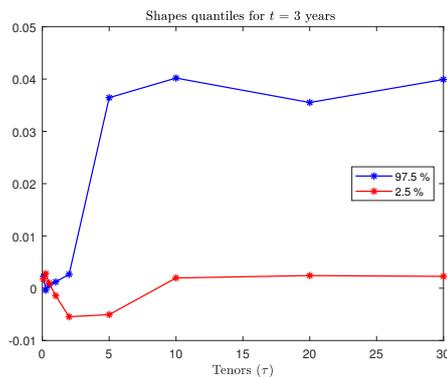


Figure 3.10: Yields quantiles shapes after 3 years simulation with DEV-MR.

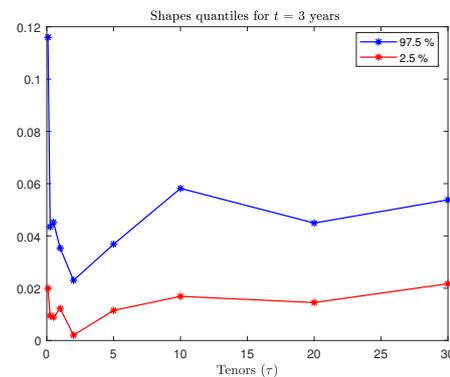


Figure 3.11: Yields quantiles shapes after 30 years simulation with DEV-MR.

3.2. The Hull-White one factor model

The Hull-White model ($HW - 1F$) was firstly introduced by J.C. Hull and A.C. White in [17] in 1990 and consists of a Gaussian Affine Term Structure model widely used in industry. Contrary to its ancestor Vasicek model, nicely allows to perfectly fit the current term structure in the market. This is a very attractive property for pricing purposes and was first introduced by Ho and Lee in [15]. In this case, it is achieved by augmenting the Vasicek model with a deterministic time-varying mean-drift. In fact, matching exactly the term structure is equivalent to solving a system with an infinite number of equations, and this is possible when an infinite number of parameters or, equivalently, a deterministic function of time is introduced. Other models of this type are the Black-Derman-Toy (BDK) [4] and the Black-Karasinski (BK) [3] models. The main drawback of these two models, as stated in [11], is that the entire forward curve cannot be expressed through analytical formulas, which is something that is met by the Hull-White model instead.

This section contains the necessary theoretical tools in order to perform a risk-neutral calibration of the model, together with the calibration method. We will introduce the model in its time-dependent volatility version and its other equivalent formulation, the Gaussian additive one-factor model ($G1++$). Main reason is that the first one gives an easier interpretation of the parameters, while the second one is preferred for proofs and give more insight about the linkage between the one-factor and two-factors model, as we will see in (3.3). After this, the formula for the mean reversion level parameter will be derived, so that the market can be perfectly replicated. As the model is nicely analytically tractable, closed-form pricing formulas can be derived for Zero-Coupon

Bond options, caps and floors. Furthermore for European swaptions semi-analytic formulas can be derived under the Hull-White model. One could chose either to calibrate the model to a set of caps/floors or to a set of swaptions. We rather prefer the second option as swaptions are known to retain more information on the correlation between interest rates. Prior to the derivation of a swaption price under the one-factor framework, it will be necessary to go through two steps: first step is to derive the dynamics of the short rate under the so-called T-forward measure, which we will refer to as \mathbb{Q}^T and in which the numéraire is the bond $P(t, T)$. Second step regards the derivation of the pricing formula for the bond option on a Zero-Coupon Bond.

3.2.1. Introduction to the model

In its most general form, the short rate is assumed to satisfy the stochastic differential equation:

$$dr(t) = [\theta(t) - a(t)r(t)] dt + \sigma(t) dW_t^{\mathbb{Q}},$$

where θ , a and σ are deterministic and W is a Brownian motion under the risk-neutral measure \mathbb{Q} . Here, $\sigma(t)$ is assumed regular enough to ensure the existence and uniqueness of a solution, while $\theta(t)$ is the time dependent parameter that replicates the current term structure in the market. This model is mean-reverting in the sense that at any time t the rate is pushed to the level $\frac{\theta(t)}{a(t)}$, which is a desirable property in interest rate modelling. In this thesis we let the mean reversion rate be modelled by a constant parameter, which does not consists of a strong approximation as the rate is known not to vary much over time. In Section (B.2) of the appendix we provide evidence to this fact. Regarding the volatility parameter, we prefer to keep it time-dependent. It is true that constant speed and volatility still allow for reasonably good fit for certain choices of the calibration basket. Still, not for all of them, especially when it comes to chose co-terminal baskets, that is when it comes to calibrate on a diagonal of the volatility surface. The reason lies in the 2008 credit crisis, when volatility surfaces became very irregular. Therefore, the safest assumption in order to well fit to the market prices is to add degrees of freedom to the model, which is done by considering a time dependent volatility $\sigma(t)$.

Under the risk neutral measure \mathbb{Q} , and under the assumption of time-dependent volatility, the instantaneous short rate r_t evolves according to the Stochastic Differential Equation:

$$dr_t = (\theta(t) - ar_t) dt + \sigma(t) dW_t^{\mathbb{Q}} = a \left(\frac{\theta(t)}{a} - r(t) \right) dt + \sigma(t) dW_t^{\mathbb{Q}}, \quad (3.7)$$

where both $\theta(t)$ and $\sigma(t)$ are deterministic functions of time. Using Ito's lemma, one gets to the following solution for the short rate process:

$$r_t = r_s e^{-a(t-s)} + \int_s^t \theta(u) e^{-a(t-u)} du + \int_s^t \sigma(u) e^{-a(t-u)} dW^{\mathbb{Q}}(u).$$

Therefore, $r(t)$ conditionally on \mathcal{F}_s is normally distributed with mean and variance given respectively by

$$\mathbb{E}[r(t)|\mathcal{F}_s] = r_s e^{-a(t-s)} + \int_s^t \theta(u) e^{-a(t-u)} du,$$

$$\text{Var}[r(t)|\mathcal{F}_s] = \int_s^t \sigma(u)^2 e^{-2a(t-u)} du.$$

Note that at each time t the short rate $r(t)$ can be negative with positive probability. More precisely, assuming constant volatility parameter, this happens with probability:

$$\mathbb{P}(r_t < 0 | \mathcal{F}_0) = \Phi \left(- \frac{r(0) e^{-at} + \int_0^t \theta(u) e^{-a(t-u)} du}{\sqrt{\frac{\sigma^2}{2a} (1 - e^{-2at})}} \right),$$

which is easily derivable from (3.2.1) and does not converge to zero. As it is appreciable in (3.2.6), this translates in simulated term structures that are not characterized by increasing trends in t , like it happens with the DEV-MR model.

The $HW-1F$ model has an equivalent formulation known as $G1++$ model, which assumes that the short rate evolves according to:

$$\begin{cases} dx(t) = -ax(t) dt + \sigma(t) dW(t), & x(0) = 0, \\ r(t) = x(t) + \beta(t), \end{cases} \quad (3.8)$$

where $\beta(t)$ is a deterministic shift having the same role of $\theta(t)$. The equivalence simply follows from Ito's lemma:

$$\begin{aligned} dr(t) &= d(x(t) + \beta(t)) \\ &= dx(t) + \frac{\partial \beta(t)}{\partial t} dt \\ &= -a(r(t) - \beta(t)) dt + \frac{\partial \beta(t)}{\partial t} dt + \sigma(t) dW(t) \\ &= \left(a\beta(t) + \frac{\partial \beta(t)}{\partial t} - ar(t) \right) dt + \sigma(t) dW(t), \end{aligned}$$

from which we can see that $\theta(t)$ and $\beta(t)$ are linked through the relation

$$\theta(t) = a\beta(t) + \frac{\partial \beta(t)}{\partial t}. \quad (3.9)$$

Fit to the initial term structure

Since the model is able to ensure an exact fit to the current term structure, then it is an arbitrage-free model. The following theorem provide the calibration formula for the mean reversion level for both the two formulations.

Theorem 3.2.1 (Calibration of mean reversion level). *The exact fit to the initial term structure of discount bonds is guaranteed if $\beta(t)$ in (3.8) is calibrated as*

$$\beta(t) = f^M(0, t) + \int_0^t \frac{\sigma(u)^2}{a} (1 - e^{-a(t-u)}) e^{-a(t-u)} du, \quad (3.10)$$

where $f^M(0, t)$ is the current instantaneous forward rate observed in the market. Equivalently, one needs to set $\theta(t) = \frac{\partial \beta(t)}{\partial t} + a\beta(t)$ in (3.7).

Proof. In order to find $\theta(t)$ so to exactly fit the initial term structure of discount bonds, consider the model as formulated in (3.8). The solution to this reads:

$$r(t) = x(t) + \beta(t) = x(s) e^{-a(t-s)} + \int_s^t e^{-a(t-u)} \sigma(u) dW(u) + \beta(t).$$

Therefore, the short rate has mean and variance under the risk-neutral measure \mathbb{Q} given by

$$\begin{aligned} \mathbb{E}[r(t)|\mathcal{F}_s] &= x(s) e^{-a(t-s)} + \beta(t), \\ \text{Var}[r(t)|\mathcal{F}_s] &= \int_s^t e^{-2a(t-u)} \sigma(u)^2 du. \end{aligned}$$

Now, the model fits the current term structure of interest rates if, for each maturity T , $P(0, T) = P^M(0, T)$, where $P^M(0, T)$ is observed in the market. To compute $P(0, T)$ we need to integrate the

process $r(t)$ over $[0, T]$, and then use Equation (2.9). In order to do so, we will use the fact that since the process $x(t)$ is normally distributed conditionally on \mathcal{F}_t , then also $\int_0^T x(t) dt$ is normally distributed. Indeed,

$$\begin{aligned} \int_0^T x(t) dt &= \int_0^T \left(\int_0^t \sigma(u) e^{-a(t-u)} dW(u) \right) dt \\ &= \int_0^T \sigma(u) e^{au} \left[\int_u^T e^{-at} dt \right] dW(u) \\ &= \int_0^T \frac{\sigma(u)}{a} (1 - e^{-a(T-u)}) dW(u), \end{aligned}$$

where we used Fubini's theorem for stochastic integral (see [13]) to exchange the two integrals. So, the integrated $x(t)$ process is normally distributed with mean zero and variance

$$\text{Var} \left[\int_0^T x(t) dt \middle| \mathcal{F}_T \right] = \int_0^T \frac{\sigma(u)^2}{a^2} (1 - e^{-a(T-u)})^2 du =: V(0, T). \quad (3.11)$$

In the case that $\sigma(t) = \sigma$ is a positive constant function, the variance reads

$$V(0, T) = \frac{\sigma^2}{a^2} \left(T - 2 \frac{1 - e^{-aT}}{a} + \frac{1 - e^{-2aT}}{2a} \right).$$

Finally, recalling the moment generating function for a normal random variable Z ¹, we have

$$P(0, T) = \mathbb{E}[e^{-\int_0^T r(t) dt} | \mathcal{F}_0] = \mathbb{E}[e^{-\int_0^T \beta(u) du - \int_0^T x(u) du} | \mathcal{F}_0] = e^{-\int_0^T \beta(u) du} e^{\frac{1}{2} V(0, T)}.$$

As a consequence, the model will give an exact fit to the market term structure if and only if for every $T > 0$:

$$P^M(0, T) = e^{-\int_0^T \beta(u) du} e^{\frac{1}{2} V(0, T)}.$$

Now, to derive the explicit expression for $\theta(t)$, we can use the definition of instantaneous forward rate and write:

$$e^{-\int_0^T f^M(0, t) dt} = e^{-\int_0^T \beta(u) du} e^{\frac{1}{2} V(0, T)},$$

which, differentiating with Leibniz integral rule², leads to:

$$\begin{aligned} f^M(0, T) &= \beta(T) - \frac{1}{2} \frac{\partial V(0, T)}{\partial T} \\ &= \beta(T) - \int_0^T \frac{\sigma(u)^2}{a} (1 - e^{-a(T-u)}) e^{-a(T-u)} du. \end{aligned}$$

From this, Equation (3.10) easily follows. Alternatively, in the $HW - 1F$ model it will suffice to calibrate $\theta(t)$ as in Equation (3.9), which, in case the volatility is assumed to be constant, translates to:

$$\theta(t) = \frac{\partial}{\partial t} f(0, t) + a f(0, t) + \frac{\sigma^2}{2a} (1 - e^{-2at}). \quad (3.12)$$

□

¹If $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$, then $M_Z(t) = \mathbb{E}[e^{tZ}] = e^{t\mu_Z + \frac{1}{2}t^2\sigma_Z^2}$.

² $\frac{\partial}{\partial \alpha} \int_{a(\alpha)}^{b(\alpha)} f(x, \alpha) dx = f(b, \alpha) \frac{\partial}{\partial \alpha} b(\alpha) - f(a, \alpha) \frac{\partial}{\partial \alpha} a(\alpha) + \int_{a(\alpha)}^{b(\alpha)} \frac{\partial}{\partial \alpha} f(x, \alpha) dx$.

3.2.2. Zero-Coupon Bond price

The Hull-White model belongs to the family of Affine Term Structure Models for whom the price of a Zero-Coupon Bond is assumed to have the following form:

$$P(t, T) = A(t, T)e^{-B(t, T)r(t)}, \quad (3.13)$$

where $A(t, T)$ and $B(t, T)$ are deterministic function of time. From the system of Riccati-type ODEs in (2.10) it is possible to prove that:

$$\begin{aligned} B(t, T) &= \frac{1 - e^{-a(T-t)}}{a}, \\ A(t, T) &= \frac{P^M(0, T)}{P^M(0, t)} \exp \left\{ -\frac{1}{2} B(t, T)^2 \Gamma(T) + B(t, T) f^M(0, t) \right\}, \end{aligned} \quad (3.14)$$

where

$$\Gamma(t) = \text{Var}\{r(t) | \mathcal{F}_0\} = \int_0^t \sigma(u)^2 e^{-2a(t-u)} du. \quad (3.15)$$

As a consequence, the explicit bond pricing formula is:

$$P(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \exp \left\{ -\frac{1}{2} B^2(t, T) \Gamma(t) - B(t, T) [r_t - f^M(0, t)] \right\}, \quad (3.16)$$

from which one can appreciate the perfect calibration to the yield curve. Indeed, for $t = 0$, the bracketed term in (3.16) becomes zero.

Equivalently, in [9] the Zero-Coupon Bond price is proven to be:

$$P(t, T) = A(t, T)e^{-B(t, T)x(t)}, \quad (3.17)$$

where $x(t)$ evolves according to (3.8), while $B(t, T)$ is as in (3.14) and $A(t, T)$ is given by:

$$A(t, T) = \frac{P^M(0, T)}{P^M(0, t)} e^{\frac{1}{2}[V(t, T) - V(0, T) + V(0, t)]}.$$

Here $V(t, T)$ is the variance of the integrated process $\int_t^T x(s) ds$, that is retrievable from (3.11). The $P(t, T)$ price under this formulation will come particularly useful in the proof of the pricing formula for a put option on a Zero-Coupon Bond.

3.2.3. Dynamics under the \mathbb{Q}^T measure

It is well-known that Girsanov's theorem provides us with a result which is fundamental when it comes to deal with pricing problems. More precisely, it is a powerful tool that allows to derive the dynamics of a stochastic process (within the Wiener-driven processes class) under a different measure. The main purpose for which it can be used is to simplify the payoff of a derivative, and in some cases helps to reach a closed form solution. In order to do so, a tradeable asset must be used as numéraire: this allows to express the value of an instrument as a function of the tradeable asset.

Even though such theorem has an incredibly wide range of applications and especially can be used to move towards any measure, for our scope it will represent the tool to move from the (spot) risk-neutral measure \mathbb{Q} towards the T-forward measure \mathbb{Q}^T , also known as risk-adjusted measure, in which the numéraire is the Zero-Coupon Bond $P(t, T)$. In the no-arbitrage setting, this implies that the price of a contingent claim with payoff H_T at maturity T will be given by

$$H_t = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} H(T) \right] = \mathbb{E}_t^{\mathbb{Q}} \left[e^{-\int_t^T r(s) ds} g(r(T)) \right] = P(t, T) \mathbb{E}_t^{\mathbb{Q}^T} [g(r(T))],$$

where we adopted the the assumption that the payoff can be written as a function of the short rate g . This is indeed the case when pricing European swaptions in the Hull-White framework. Therefore, the change of measure helps to get rid of the exponential term, whose presence is undesired as it is correlated with the payoff H_T .

In order to derive the pricing formula for a put option on a Zero-Coupon Bond, it is necessary to derive the dynamics of the process $x(t)$, defined in (3.8), under \mathbb{Q}^T by means of the Girsanov's theorem. The steps in order to derive the new dynamics are:

- Find dynamics of Zero-Coupon Bond $P(t, T)$ in term of $r(t)$ under the risk-neutral measure.
- Find dynamics of Radon-Nikodym derivative $v_{\mathbb{Q}}^T(t)$ under the risk-neutral measure.
- Compute $dW^T(t) = dW^{\mathbb{Q}}(t) - \frac{dv_{\mathbb{Q}}^T(t)}{v_{\mathbb{Q}}^T(t)} dW^{\mathbb{Q}}(t)$ (thesis of Girsanov's theorem as formulated in [1, p. 168]) and substitute this in (3.8).

The results is stated in the following proposition.

Theorem 3.2.2. *The dynamics of the process $x(t)$ under the risk-adjusted measure \mathbb{Q}^T are:*

$$dx(t) = -ax(t) dt - \sigma(t)^2 B(t, T) dt + \sigma(t) dW^T(t), \quad (3.18)$$

and from this it follows that:

$$\begin{aligned} x(t)|\mathcal{F}_s &\sim \mathcal{N}\left(x(s)e^{-a(t-s)} + M_x^T(s, t), \int_s^t \sigma(u)^2 e^{-2a(t-u)} du\right) \\ &= \mathcal{N}(\mu_{x,T}(s, t), \kappa_{x,T}^2(s, t)), \end{aligned} \quad (3.19)$$

where $M_x^T(s, t) = \int_s^t \sigma(u)^2 e^{-a(t-u)} \frac{1-e^{-a(T-u)}}{a} du$.

Proof. Here we prove the dynamics in (3.18). Ito's lemma applied to (3.13) suggest that:

$$dP(t, T) = r(t)P(t, T) dt - \sigma(t)B(t, T)P(t, T) dW^{\mathbb{Q}}(t). \quad (3.20)$$

At this point, recalling also (2.1.1), the dynamics of both numéraire for the two measure \mathbb{Q} and \mathbb{Q}^T are known. Now it is well-known that the Radon-Nikodym derivative is defined as:

$$v_{\mathbb{Q}}^T(t) := \frac{d\mathbb{Q}^T}{d\mathbb{Q}} = \frac{M(t_0) P(t, T)}{M(t) P(t_0, T)},$$

provided that $t_0 < t < T$. As a consequence, the dynamics required from the second step are:

$$\begin{aligned} dv_{\mathbb{Q}}^T(t) &= \frac{M(t_0)}{P(t_0, T)} \left(-\frac{P(t, T)}{M(t)^2} dM(t) + \frac{dP(t, T)}{M(t)} \right) \\ &= \frac{M(t_0)}{P(t_0, T)} \left(-\frac{P(t, T)}{M(t)^2} r(t)M(t) dt + \frac{1}{M(t)} (r(t)P(t, T) dt - \sigma(t)B(t, T)P(t, T) dW^{\mathbb{Q}}(t)) \right) \\ &= -\frac{M(t_0)}{P(t_0, T)} \frac{\sigma(t)B(t, T)P(t, T) dW^{\mathbb{Q}}(t)}{M(t)}. \end{aligned}$$

It is now possible to calculate the new stochastic component under the T -forward measure:

$$\begin{aligned} dW^T(t) &= dW^{\mathbb{Q}}(t) - \frac{dv_{\mathbb{Q}}^T(t)}{v_{\mathbb{Q}}^T(t)} dW^{\mathbb{Q}}(t) \\ &= dW^{\mathbb{Q}}(t) + \sigma(t)B(t, T) dt, \end{aligned}$$

where the property $dW^{\mathbb{Q}}(t) \cdot dW^{\mathbb{Q}}(t) = 1 \cdot dt$ has been used in the last equality. Substituting this in (3.8), we get to the result. \square

3.2.4. Bond-option and swaption prices

Bond-Option price

In the Hull-White model, as has been proved in [5], the price of a European call option with strike K and maturity T , written on a Zero-Coupon Bond with maturity S at time $t \in [0, T]$ is given by

$$\begin{aligned} V_{ZCB}^{\text{Call}}(t_0, T, T_F, K) &= P(t_0, T_F)\phi(h) - KP(t_0, T)\phi(h - \tilde{\sigma}), \\ V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) &= KP(t_0, T)\phi(-h + \tilde{\sigma}) - P(t_0, T_F)\phi(-h), \end{aligned} \quad (3.21)$$

where $\tilde{\sigma}$ and h are:

$$\begin{aligned} \tilde{\sigma} &= \sigma \sqrt{\frac{1 - e^{-2a(T-t_0)}}{2a}} B(T, T_F), \\ h &= \frac{1}{\tilde{\sigma}} \ln \frac{P(t_0, T_F)}{P(t_0, T)K} + \frac{\tilde{\sigma}}{2}. \end{aligned}$$

Equivalently, another formula can be found by simply computing an integral. More specifically, we proved the formula provided by the following theorem.

Theorem 3.2.3. *The price at t_0 of a European put option with strike K and maturity T , written on a Zero-Coupon Bond with maturity T_F is given, under the G1++ model, by:*

$$V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) = P(t_0, T) \left\{ K\Phi(d_1) - A(T, T_F)e^{\frac{\kappa_{\tilde{X}}^2 - 2\mu_{\tilde{X}}}{2}}\Phi(d_2) \right\}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, and $\mu_{\tilde{X}}, \kappa_{\tilde{X}}, d_1, d_2$ read

$$\begin{cases} \mu_{\tilde{X}} & := B(T, T_F)\mu_{x,T}(t_0, T), \\ \kappa_{\tilde{X}}^2 & := B(T, T_F)^2\kappa_{x,T}^2(t_0, T), \\ d_1 & := -\frac{\log A(T, T_F) - \log K - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}}, \\ d_2 & := d_1 - \kappa_{\tilde{X}}. \end{cases}$$

Proof. Consider a put option that gives the holder the right (but not the obligation) to sell a Zero-Coupon Bond with maturity T_F at a future time T . The option will be exercised if $K - P(T, T_F) > 0$. Following no-arbitrage arguments, the price at $t_0 < T$ is given by:

$$\begin{aligned} V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) &= \mathbb{E}^{\mathbb{Q}} \left[\frac{M(t_0)}{M(T)} (K - P(T, T_F))^+ | \mathcal{F}(t_0) \right] \\ &= P(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[(K - A(T, T_F)e^{-B(T, T_F)x(T)})^+ | \mathcal{F}(t_0) \right], \end{aligned}$$

where in the second equality we used Girsanov's theorem to change from the \mathbb{Q} -measure to the \mathbb{Q}^T -measure. Note that the Zero-Coupon Bond price in (3.17) has been used. From (3.19) it follows that under risk-adjusted measure

$$\tilde{X}(T, T_F) := B(T, T_F)x(T) | \mathcal{F}(t_0) \sim \mathcal{N}(B(T, T_F)\mu_{x,T}(t_0, T), B(T, T_F)^2\kappa_{x,T}^2(t_0, T)) = \mathcal{N}(\mu_{\tilde{X}}, \kappa_{\tilde{X}}^2).$$

Call $A = A(T, T_F)$ so to simplify the notation in the proof. Knowing that $\frac{\tilde{X} - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}} \sim \mathcal{N}(0, 1)$, and that

$f_{\tilde{X}}(x) = \frac{1}{\sqrt{2\pi\kappa_{\tilde{X}}}} \exp\left\{-\frac{(x - \mu_{\tilde{X}})^2}{2\kappa_{\tilde{X}}^2}\right\}$ we can continue the derivation:

$$\begin{aligned} V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) &= P(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[(K - Ae^{-\tilde{X}(T, T_F)})^+ | \mathcal{F}(t_0) \right] \\ &= P(t_0, T) \int_{-\infty}^{\infty} (K - Ae^{-\tilde{X}(T, T_F)}) \mathbb{1}(K - Ae^{-\tilde{X}(T, T_F)} \geq 0) f_{\tilde{X}}(x) dx \\ &= P(t_0, T) \int_{\log A - \log K}^{\infty} (K - Ae^{-\tilde{X}(T, T_F)}) f_{\tilde{X}}(x) dx. \end{aligned}$$

The above integral can be computed explicitly as follows.

$$\begin{aligned}
& \int_{\log A - \log K}^{\infty} (K - Ae^{-\tilde{X}(T, T_F)}) f_{\tilde{X}}(x) dx \\
&= K\mathbb{P}\left(\frac{\tilde{X} - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}} \geq \frac{\log A - \log K - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}}\right) - A \int_{\log A - \log K}^{\infty} e^{-x} \frac{1}{\sqrt{2\pi\kappa_{\tilde{X}}}} e^{-\frac{(x - \mu_{\tilde{X}})^2}{2\kappa_{\tilde{X}}^2}} dx \\
&= K\Phi\left(-\frac{\log A - \log K - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}}\right) - Ae^{\frac{(\kappa_{\tilde{X}}^2 - \mu_{\tilde{X}})^2 - \mu_{\tilde{X}}^2}{2\kappa_{\tilde{X}}^2}} \int_{\log A - \log K}^{\infty} e^{-x} \frac{1}{\sqrt{2\pi\kappa_{\tilde{X}}}} e^{-\frac{(x - \mu_{\tilde{X}} + \kappa_{\tilde{X}}^2)^2}{2\kappa_{\tilde{X}}^2}} dx \\
&= K\Phi\left(-\frac{\log A - \log K - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}}\right) - Ae^{\frac{\kappa_{\tilde{X}}^2 - 2\mu_{\tilde{X}}}{2}} \Phi\left(-\frac{\log A - \log K - \mu_{\tilde{X}} + \kappa_{\tilde{X}}^2}{\kappa_{\tilde{X}}}\right) \\
&= K\Phi(d_1) - Ae^{\frac{\kappa_{\tilde{X}}^2 - 2\mu_{\tilde{X}}}{2}} \Phi(d_2).
\end{aligned}$$

where $d_1 = -\frac{\log A - \log K - \mu_{\tilde{X}}}{\kappa_{\tilde{X}}}$ and $d_2 = d_1 - \kappa_{\tilde{X}}$. From this the theorem follows. \square

Note that the price of the corresponding call option can be obtained through the put-call parity for bond options. In fact, if the options have maturity T , strike K and are written on a Zero-Coupon Bond maturing at time T_F , their prices at time t_0 satisfy:

$$V_{ZCB}^{\text{Call}}(t_0, T, T_F, K) = V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) + P(t, T_F) - KP(t, T).$$

Swaption price

In 1989 in [20] Jamshidian introduced a method for transforming the maximum of a sum into a sum of certain maximums. The result is provided in the following lemma and it is very handy when deriving the price of an European swaption under the $HW - 1F$ model.

Lemma 3.2.4 (Jamshidian's decomposition). *Let $\varphi_i(x)$ be a finite sequence of monotone increasing (or monotone decreasing) functions, s.t. $\forall i \varphi_i(x) : \mathbb{R} \mapsto \mathbb{R}^+$. Then the following result holds:*

$$\max\left(K - \sum_i \varphi_i(x), 0\right) = \sum_i \max(\varphi_i(x^*) - \varphi_i(x), 0),$$

where x^* is determined such that:

$$K - \sum_i \varphi_i(x^*) = 0.$$

In practice, x^* is typically determined with some search algorithm like a Newton-Raphson. Thanks to this approach, it is possible to express the value of the payer (receiver) swaption as weighted sum of the values of individual puts (calls) on Zero-Coupon Bonds, as outlined in the following theorem.

Theorem 3.2.5 (Swaption price under G1++ model). *Consider an European (payer) swaption with notional N , strike K , maturity T , written on an Interest Rate Swap with first reset date T_α and payment dates $T_{\alpha+1}, \dots, T_\beta$. Assume that $T = T_\alpha$. We denote by Δ_i the year fraction from T_{i-1} and T_i , and set $c_i := K\Delta_i$ for $i = \alpha + 1, \dots, \beta - 1$ and $c_\beta := 1 + K\Delta_\beta$. The risk-free price at $t_0 \leq T$ of such instrument is given, under the G1++ model, by:*

$$V^{\text{pSwpt}}(t_0; T, T_\beta) = NP(t_0, T) \sum_{k=\alpha+1}^{\beta} c_k V_{ZCB}^{\text{Put}}(t_0, T, T_k, \hat{K}), \quad (3.22)$$

with $\hat{K} := A(T, T_k)e^{-B(T, T_k)\tilde{x}}$ where a constant \tilde{x} is chosen such that

$$1 - \sum_k c_k A(T, T_k)e^{-B(T, T_k)\tilde{x}} = 0.$$

Proof. Let us consider a payer swaption. Let T be the option maturity coinciding with T_α being the first reset date and $T_{\alpha+1}, \dots, T_\beta$ the payment dates. The value of the deal at time t_0 is given by:

$$\begin{aligned} V^{\text{pSwpt}}(t_0; T, T_\beta) &= \mathbb{E}^{\mathbb{Q}} \left[\frac{M(t_0)}{M(T)} (V^{\text{IRS}}(T))^+ \middle| \mathcal{F}(t) \right] \\ &= N \mathbb{E}^{\mathbb{Q}} \left[\frac{M(t_0)}{M(T)} \left(\sum_{k=\alpha+1}^{\beta} \Delta_k P(T, T_k) (F(T, T_{k-1}, T_k) - K) \right)^+ \middle| \mathcal{F}(t_0) \right] \\ &= NP(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[\left(\sum_{k=\alpha+1}^{\beta} \Delta_k P(T, T_k) (F(T, T_{k-1}, T_k) - K) \right)^+ \middle| \mathcal{F}(t_0) \right]. \end{aligned}$$

Now, by plugging the definition for $F(t, T_{k-1}, T_k)$ one has that

$$\sum_{k=\alpha+1}^{\beta} \Delta_k P(T, T_k) (F(T, T_{k-1}, T_k) - K) = 1 - P(T, T_n) - K \sum \Delta_k P(T, T_k) = 1 - \sum_{k=\alpha+1}^{\beta} c_k P(T, T_k),$$

with $c_i = K \Delta_i$, for $i = \alpha + 1, \dots, \beta - 1$ and $c_\beta = 1 + K \Delta_\beta$. Knowing this, the swaption price become

$$\begin{aligned} V^{\text{pSwpt}}(t_0; T, T_\beta) &= NP(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[\left(1 - \sum_{k=\alpha+1}^{\beta} c_k P(T, T_k) \right)^+ \middle| \mathcal{F}(t_0) \right] \\ &= NP(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[\left(1 - \sum_{k=\alpha+1}^{\beta} c_k A(T, T_k) e^{-B(T, T_k)x(T)} \right)^+ \middle| \mathcal{F}(t_0) \right], \end{aligned}$$

where (3.17) has been used for the last equality. Now, by means of the Jamshidian's decomposition introduced in Lemma (3.2.4), we can transform the maximum of a sum into the sum of maximums, and get to the following formula:

$$V^{\text{pSwpt}}(t_0; T, T_\beta) = NP(t_0, T) \sum_{k=1}^n c_k \mathbb{E}^{\mathbb{Q}^T} \left[(\hat{K} - A(T, T_k) e^{-B(T, T_k)x(T)})^+ \right], \quad (3.23)$$

with $\hat{K} := A(T, T_k) e^{-B(T, T_k)\tilde{x}}$ where a constant \tilde{x} is chosen such that $1 - \sum c_k A(T, T_k) e^{-B(T, T_k)\tilde{x}} = 0$. Now, the task of pricing the swaption is finished as soon as we determine the sum of expectations. We can see that each of the summands is a European put option on a Zero-Coupon Bond and pricing this instrument can be done analytically under the Hull-White model with Equation (3.2.3). The final formula easily follows. \square

3.2.5. Calibration to swaptions

Interest rate models are typically calibrated using caps, floors or swaptions, as these represents the most liquid instruments traded in the market. In this thesis, our choice goes for the calibration to European swaptions, especially because these are known to contain information on the correlation between yields of different maturities. More precisely, swaptions are sensitive to correlation of the underlying rates and therefore consist of an optimal choice for the calibration of multi-factors model.

In the calibration, the parameters to be determined are a and $\sigma(t)$. To calibrate a model under the risk-neutral measure, generally one chooses a set of calibration instruments, in this case a set of European swaptions. The parameters are chosen in such a way that the model generated prices (or model Normal implied volatilities) match the market prices (or market Normal volatilities) of the calibration instruments. Usually, when calibrating to interest rate derivatives quotes, it is better to

calibrate to the implied volatilities. This reduces the dependency on more pricey instruments (typically longer instruments are more pricey).

In general, the volatilities depend on the level of mean reversion, and therefore calibration of mean-reversion and volatility at the same time to the same market instruments should not be performed. Rather, we advise to split the calibration procedure in two separate optimization problems. This motivates us to proceed as follow: given as input any possible and reasonable mean-reversion level, we will be able to find time-dependent volatility such that the chosen set will be well-calibrated. In our case, this set is going to be a diagonal of the volatility surface (i.e. a set of co-terminal swaptions, that is a set of swaptions that have different maturities but same underlying swap with same expiry date). Afterwards, the idea is to choose the mean-reversion level such that the model fits best the quotes outside counter-diagonal. This optimization problem is formulated in detail in Equation (3.26).

Note that the swaption implied volatility matrix is constrained in the shape. The speed and volatility parameters a and $\sigma(t)$ have a direct influence on this shape. As pointed out in [26], the volatility parameter controls the level of the implied volatility curves, while the speed parameter controls the monotonicity of such curves (and the possibility of modeling humps by assuming a time dependent speed). As a consequence, under our assumptions on the parameters, we cannot expect to be successful in matching volatility humps.

Data retrieval

In order to perform a risk-neutral calibration on day t_0 , the following two objects need to be retrieved: bond curve $P(t_0, T)$ and swaptions volatility surface. More precisely for the first one, we need bonds prices at t_0 which at a later step will be interpolated by means of a cubic spline. Bloomberg³ allows you to either keep the set of interest rate instruments as default or to customize the bootstrapping procedure. For our purposes it is sufficiently accurate to go for the first option. Secondly, as we said the quoted volatility surface is needed. Also in this case the data were retrieved on Bloomberg. This platform provides the quoted volatility cube under both the (non-shifted) Black and Normal models. The former model assumes a log-normal distribution and therefore is not capable of dealing with negative values. That is the reason for why nowadays certain values are not quoted (more specifically, the values for swaptions with very short expiry and tenor). In this case, the quotes are displayed in terms of %. The latter model is capable of modelling also negative values of the underlying stochastic process, as it follows a Gaussian distribution. In this case, the quotes in Bloomberg are displayed in terms of *bps*. In order to be consistent with the negative rates that can be easily observed in the market, in this thesis we chose to calibrate to the Normal volatilities. The term structure and swaptions volatility surface dating back to 23 January 2018 is shown in Figures (3.12) and (3.13).

Specification of the volatility's functional form

In this thesis we chose to follow one of the parametrizations proposed in [26]. Precisely the one where the volatility is defined as a constant step function. Given n the number of steps we want to consider (it will turn out to coincide to the number of swaptions one consider in the calibration step), we then have the following parametrization for the volatility:

$$\sigma(t) = \sum_{j=1}^n \sigma_j \mathbb{1}_{(t_{j-1}, t_j]}(t) + \sigma_n \mathbb{1}_{(t_n, \infty)}(t), \quad (3.24)$$

³Bloomberg L.P. is a privately held financial, software, data, and media company headquartered in Midtown Manhattan, New York City.

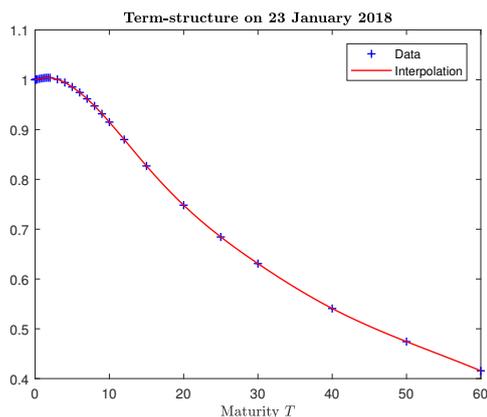


Figure 3.12: Bond curve on 23/01/18. Source: Bloomberg platform.

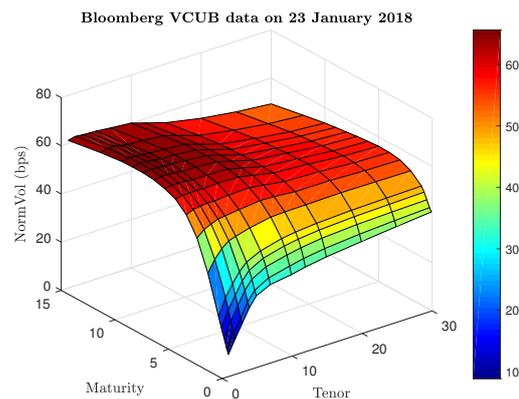


Figure 3.13: Volatility Cube on 23/01/18. Source: Bloomberg platform.

where we set $t_0 = 0$. We now define the following function:

$$\Gamma(t) = \text{Var}[r(t)|\mathcal{F}_0] = \int_0^t \sigma(u)^2 e^{-2a(t-u)} du. \quad (3.25)$$

Now, when dealing with *Exact simulation* it is important to compute this as quickly as possible. Therefore, we don't actually compute $\Gamma(t)$ as an integral but if we define the following function:

$$m(t) = \begin{cases} \inf_{1 \leq j \leq n} \{j : t_j \geq t\}, & \text{if } t \leq t_n \\ n, & \text{otherwise} \end{cases}$$

we then get the following simple sums. According to (3.24), if $t \leq t_n$ (3.25) reads:

$$\Gamma(t) = -\frac{1}{2a} \sum_{j=1}^{m(t)-1} \sigma_j^2 [e^{-2a(t-s)}]_{t_{j-1}}^{t_j} + \frac{\sigma_m^2}{2a} [e^{-2a(t-s)}]_t^{t_{m-1}},$$

and if $t > t_n$:

$$\Gamma(t) = -\frac{1}{2a} \sum_{j=1}^n \sigma_j^2 [e^{-2a(t-s)}]_{t_{j-1}}^{t_j} + \frac{\sigma_n^2}{2a} [e^{-2a(t-s)}]_t^{t_n}.$$

Bootstrap fashion when calibrating the volatility parameter

In this thesis the function $\sigma(t)$ is set piece wise constant between successive option maturities and a bootstrap fashion is followed in order to calibrate these steps. This parametrization allows for constrained and reasonable number of parameters, and it is smooth enough to ensure us with existence and uniqueness of the short rate process solution. Nevertheless, this is not the only possible way one could place the constant steps. For instance, $\sigma(t)$ could be generated from a functional form so to prevent the problem of big jumps in values from one time to the other, which instead could be recognized as a drawback of our choice. In this situation the resulting number of free parameters would be that of the functional forms. In [26] this second alternative is recognized to be a better choice in terms of stability when dealing with the pricing of exotic instruments. As we are going to price European swaptions with our models this is not definitely the case for us and therefore we recognize the first alternative to be safe enough.

Given the set of n swaptions, let us call:

- $M_1 < M_2 < \dots < M_n$ the maturities of the n swaptions in the basket

Option Maturity	Tenor	Swap Type	Strike
2Y	10Y	PAYER	ATM
5Y	7Y	PAYER	ATM
7Y	5Y	PAYER	ATM
8Y	4Y	PAYER	ATM
9Y	3Y	PAYER	ATM
10Y	2Y	PAYER	ATM

Table 3.3: 12Y-coterminal basket of ATM swaptions.

- $V_k^N\{\sigma_k^N\}$ the market Normal price of instrument $k = 1, \dots, n$
- $V_k^{\text{HW}}\{\sigma_k\}$ the Hull-White price of instrument $k = 1, \dots, n$

Having $\sigma_1, \dots, \sigma_{k-1}$, we determine σ_k by

$$V_k^{\text{HW}}\{\sigma_1, \dots, \sigma_k\} = V_k^N\{\sigma_k^N\},$$

where each V^{HW} is as in (3.22). Note that this method is pretty fast because for each step we have to solve only one equation with one unknown (σ_k).

Choice of calibration basket

The first step one should do to calibrate the model is probably the most difficult one, that is to choose the right calibration basket. Of course the choice of the calibration basket is dependent of the type of instruments one aims to price afterwards. As we are going to price portfolios that are mainly made of swaps of residual maturity between 7 to 15 years, in our opinion a good choice that would balance the availability of data and the type of instruments in the portfolio, would be to take a 12–years coterminal strip of swaptions as shown in Table 3.3. Of course with this choice we should not expect to match well the volatility surface for the shortest tenor and maturity (e.g. 1y-1y) as that is where the surface bends the most. Still, it is not of importance for us given the structure of the portfolios.

Results for the calibration of volatility parameter $\sigma(t)$

We here present the result for the calibration when the mean reversion speed, given as input, is arbitrarily set equal to 14%. This value is relatively higher than usual, but turns out to be useful when analyzing the mean reversion effects on the simulation results. In Figure (3.14) the almost perfect fit to the market volatilities is shown and Figure (3.15) displays the calibrated step-wise function for the volatility. At first sight the fit is very good. Indeed $\|\mathbf{V}^{\text{HW}} - \mathbf{V}^N\|_{L_2} \approx 6e - 08$. The jump in the volatility at the beginning is often recurrent since it is due to the irregular shape of the volatility surface for short maturities. Nevertheless, no matter how irregular it is, this extension of the Hull White model will (almost) always be able to catch the right market price because of the time-dependent assumption. We say “almost” as this statement always holds, given that a reasonable value to the mean reversion level is given as input (for instance, a choice of a in $[0.01, 0.6]$ always leads to good fit results). This tells us that the volatility is the parameter that has the biggest impact on the swaptions pricing.

Calibration of the mean reversion speed a

Of course it is also important to give a properly calibrated value for the mean reversion when it comes to use the model for pricing of other instruments or valuing the exposure with a certain counterparty. More precisely, such valuation in the future turns out to be quite sensitive to the

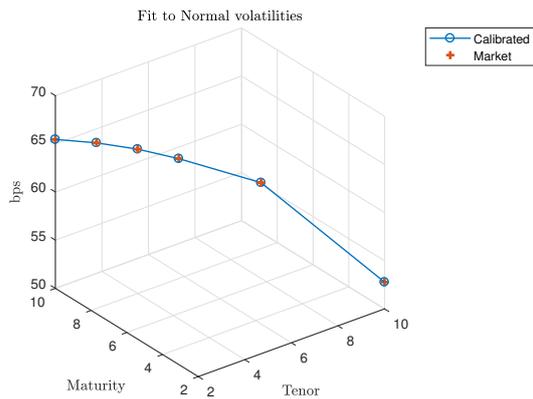


Figure 3.14: Calibrated implied volatilities.

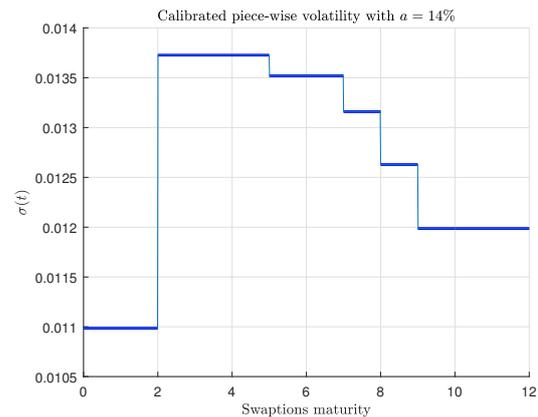


Figure 3.15: Calibrated volatility function.

mean reversion levels. The reason is that the resulting calibrated volatility increases if the input speed is increased and vice versa. If one think of it, this makes sense as mean reversion speed and volatility are usually acting in two different directions. That is, the volatility is responsible for the “wild” behaviour in the diffusion of the process, and the mean reversion speed on the other side calms the fluctuations and forces the process to stabilize after some time. Because of this, in order to fit the same market prices, if one of the two increases also the other one needs to increase so to compensate the change. At the same time it is important to point out that calibration of the mean reversion speed should not be performed as often as the volatility parameter, as $\sigma(t)$ is directly responsible for the correct pricing, and its calibrated values are stable with respect to small changes in the speed. This motivates us to calibrate the speed only once and use the same value when it comes to value exposure on different dates in Chapter (4).

Two other approaches have been experimented during the research and the results can be found in Appendix (B). Here we present our final approach, whose results are used in the subsequent portfolio revaluations. We previously noticed that the calibrated volatility depends on the speed value given as input, that is $\sigma(t) = \sigma(t, a)$. Let us call $\mathbf{A} = [-0.01, 0.6]$. The speed is calibrated as the value in \mathbf{A} that minimize the total normalized distance between the market and model prices throughout the whole volatility surface, given that the $\sigma(t)$ is calibrated to one diagonal (with maturity and tenor that always sum up to 12 years, as in table (3.3)). In mathematical formula this translates into finding \hat{a} such that:

$$\begin{aligned} \hat{a} &= \underset{a \in \mathbf{A}}{\operatorname{argmin}} g(t, a) \\ &= \underset{a \in \mathbf{A}}{\operatorname{argmin}} \min_{\sigma(t, a)} \sum_{i=1}^{n_m} \sum_{j=1}^{n_t} \left(\frac{V_{i,j}^N - V_{i,j}^{HW}(a, \sigma(t))}{V_i^N} \right)^2, \end{aligned} \quad (3.26)$$

which leads to take $a = 1.5\%$ and piece-wise $\sigma(t)$ as shown in Figure (3.16).

This calibration procedure is a particular case of the general one as follows. First we define two error functions: ERR_1 to be the sum of errors over the counter-diagonal, and ERR_2 the sum of errors over everything except for the counter-diagonal. For error any loss function measuring the distance between market or model prices (or volatilities) shall be used. Then, the general steps would be:

1. Choose mean reversion level a .
2. For a given a calibrate time-dependent volatility to counter diagonal quotes using ERR_1 .

3. Calculate ERR_2 .
4. Find a such that ERR_2 is the smallest. Or find a such $ERR = \omega ERR_1 + (1 - \omega) ERR_2$ is the smallest.

Calibration performed in (3.26) is simple done by taking $\omega = \frac{1}{2}$. We then display the error curve and

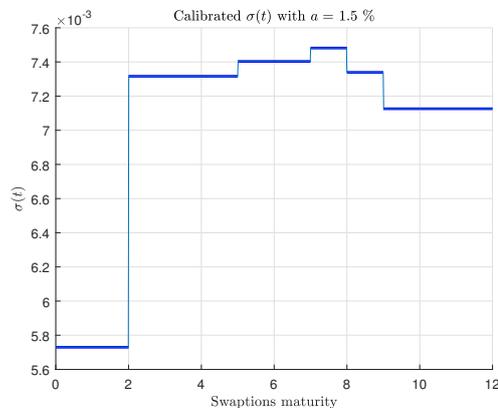


Figure 3.16: Calibrated volatility function.

match to volatility surface respectively in Figures (3.17) and (3.18). The function $g(t, a)$ is shown (3.17) as a function of a , and the minimum is effectively reached at $a = 1.5\%$.

In Figure (3.18) three volatility surfaces are displayed. Given what seen in Figure (3.13), the market implied surface is readily recognizable. Then, the implied volatility surface under the one factor Hull-White model is also easily recognizable as the farthest surface from the one observed in the market. Then, we already show here the implied surface when one more factor contributes to the short rate diffusion (that happens in the $HW - 2F$ model, introduced in Section (3.3)). The surface is closer as adding a second factor improves the pricing precision. Anyhow, in both cases two problems are clearly visible: firstly, the model is not able to fit the market volatilities for short tenors and maturities. In order to solve this “problem” it should suffice to adjust the calibration set accordingly, that is via adding the shortest swaption quoted in the market. Anyhow, this does not consist of a problem for our purposes, given the portfolio structures we will have to deal with in Chapter (4). Secondly, the volatility humps that are visible in Figure (3.13) are not caught by neither of the two models. This is due to the fact that the range of accessible swaption implied volatilities is restricted in the two Hull-White models. In fact, it is not a market model such as Libor Market Model (LMM) or Swap Market Model (SMM) for which the market implied volatilities are inputs. A possible solution so that the hump would be caught in the Hull-White framework would be to assume a time-dependent mean reversion speed, as investigated in [26]. Otherwise, other short rate models exist that are better suited for this purpose. For instance, the so-called “Cheyette short rate model”, that is a stochastic volatility model which nicely improves the fit to the market skew.

3.2.6. Simulation

During the project, two different Monte Carlo simulation have been experimented, with the first one following an exact simulation schema while the second one based on a discretized schema.

The *Exact Simulation* method was presented in [24] for the Hull-White model with one factor and constant volatility. This is possible due to the fact that under the Hull-White model, we know the distribution that r_t follows for any t . The extensions to both time-dependent volatility and the two factors model are straight-forward and the results will be presented respectively in this section

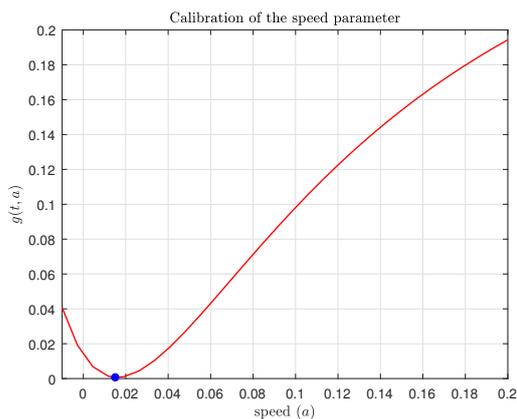


Figure 3.17: Normalized squared pricing error $g(t, a)$ over the whole swaption surface as a function of the speed.

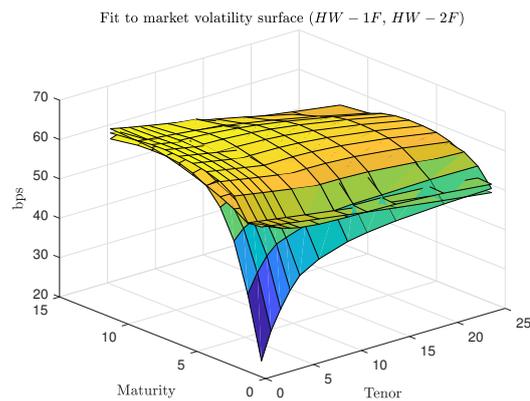


Figure 3.18: Implied volatility surface and market volatility surface.

and Subsection (3.3.6).

Driven by the interest of understanding what are the implication of the mean reversion speed parameter in terms of implied yield curves, we here provide the simulation results when a high value is given as input (14 %). The simulation results under the true calibrated value $a = 1.5\%$ are very similar to the two factors model's as their parameters all results from the calibration to the same set of market instruments. Therefore these results are put in Appendix (C).

We here briefly describe the two simulation schema that has been used for the one-factor model simulation.

Exact simulation schema

For the short rate $r(t)$, the dynamics and distribution have been already introduced in this section. We have seen that the distribution of r_t conditionally on \mathcal{F}_s is normal with mean and variance given by (3.2.1). Therefore, to simulate r_t at the reporting dates $0 = t_0 < t_1 < \dots < t_n$, we can use the following recursions:

$$x_{i+1} = x_i e^{-a\Delta_i} + \left(\sqrt{\int_{t_i}^{t_{i+1}} e^{-2a(t_{i+1}-u)} \sigma(u)^2 du} \right) Z_{i+1},$$

$$r_{i+1} = x_{i+1} + \beta(t),$$

where Z indicates a standard normal variable. Note that equivalently, in the case the parameters are constant, the factor needs to be diffused as

$$x_{i+1} = x_i e^{-a\Delta_i} + \sigma \sqrt{\frac{1}{2a}(1 - e^{-2a\Delta_i})} Z_{i+1},$$

$$r_{i+1} = x_{i+1} + \beta(t).$$

Euler simulation schema

Equivalently the short rate could be simulated using the following discretization schema:

$$x_{i+1} = x_i + a(-x_i)\Delta + \sigma(t)\sqrt{\Delta}Z_{i+1},$$

$$r_{i+1} = x_{i+1} + \beta(t),$$

where Z^1, Z^2 are correlated standard normals.

Simulation results for $a = 14\%$

Although the yield curve is not modelled directly, it can be reconstructed at each t , given the model parameters and the short rate at the grid point t . Here we provide the simulation results when the mean reversion speed is arbitrarily chosen to be 14% and the volatility parameter $\sigma(t)$ has been calibrated accordingly. The resulting short rate simulation can be found in Figure (3.19). The volatility is quite high from the beginning due to the relatively high value of the mean reversion speed and stabilizes after 10 years. Figures (3.20) and (3.21) show respectively the simulated yields from the two different perspectives introduced at the beginning of this chapter.

Note the following fact as a countercheck for the simulation profiles. When $\tau = T - t$ is taken small enough, e.g. $\tau = 1$ day (red quantiles), the yield $R(t, T)$ reads:

$$\begin{aligned} R(t, T) &= -\frac{\log P(t, T)}{T - t} \\ &= -\frac{1}{\tau} [\log A(t, t + \tau) - B(t, t + \tau)r_t] \\ &= -\frac{1}{\tau} \left[\log A(t, t + \tau) - \frac{1 - e^{-a\tau}}{a} r_t \right] \\ &\approx \frac{1}{\tau} [\tau r_t] \\ &\approx r_t + Y(t), \end{aligned}$$

where the function $Y(\cdot)$ refers to a time deterministic shift. This translates in a shape which is very similar to the short rate diffusion itself. Whereas, for bigger maturities the shape gets narrower as the factor multiplying r_t tends to 0. As a consequence, while the average value of the simulated yields stays approximately the same, the variance of $R(t, T)$ unavoidably decreases. This can be appreciated when considering the quantiles for $\tau = 30$ years, represented by the black profiles. To conclude, in Figure (3.22) the term structures simulated up to 35 years are previewed. It is clear that the problem encountered in the DEV-MR model is not present anymore, and this was expected as the correlation between rates is fixed to 1. At the same time, for this same reason, movements are quite rigid (parallel, as better explained in (3.2.7)).

Simulation results for $a = 1.5\%$

The plots for the simulation with the true calibrated parameter are provided in Appendix (C). What it is important noticing is the effect of the mean reversion rate parameter on the implied yield profiles, when this is fixed to either a high or low value. More specifically, for high values of the mean reversion speed the yields converge fast to the yield to maturity of an infinitely lived bond (see Section (3.2.7) for more explanation on this) and exhibit low quantiles range after relatively short simulation times. On the other hand, low speeds imply in comparison relatively high quantile ranges for each simulation time t . This said, we understand that mean reversion has the effect of damping the standard deviation of discount factors, and therefore the standard deviation of yields. As a matter of fact, the standard deviation of discount factors can be expressed as

$$\sigma(P(t, T)) = \sigma_{\text{HW}} \left[\frac{1 - \exp(-a(T - t))}{a(T - t)} \right].$$

as (3.20) suggests.

3.2.7. Perfect correlation

We have learned that the correlation structure between the forward rates modelled by the DEV-MR model tends to zero for increasing time horizons. This translates in uncorrelated shapes of the predicted yield curves. From this point of view the situation is completely the opposite in the $HW - 1F$

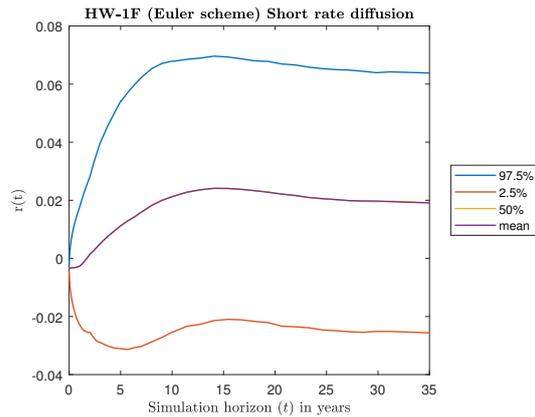


Figure 3.19: Short rate diffusion with $a = 14\%$.

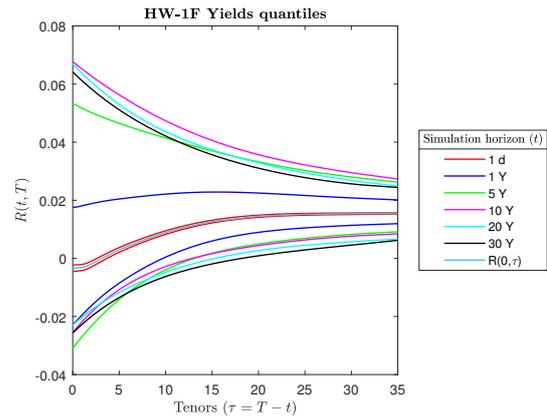


Figure 3.20: Simulated yields quantiles t years ahead.

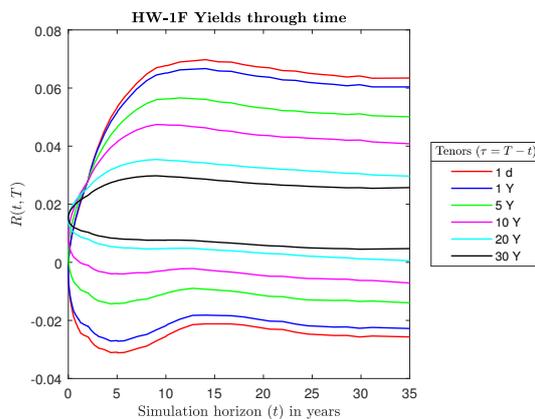


Figure 3.21: Simulated yields quantiles as a function of time t .

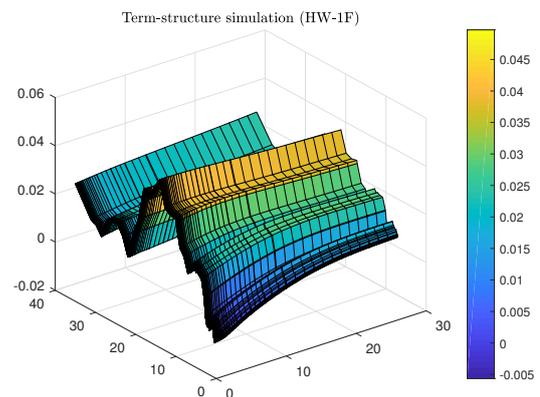


Figure 3.22: Simulated term structures under $HW - 1F$ model in 3D.

model, as it is characterized by perfect correlation between rates at different maturities independently of the time horizon t . If from one side the current model allows for quite wild and free movements of the yield curves, on the other side the one-factor Hull-White is very restrictive. In fact, in the $HW - 1F$ model a shock to the interest rate curve at time t is transmitted equally through all maturities, and the curve, when its initial point (the short rate r_t) is shocked, moves almost rigidly in the same direction.

Recall the bond price Formula (3.13), from which all rates can be computed in terms of r . We know that continuously compounded spot rates are given by an affine transformation of r , that is:

$$R(t, T) = -\frac{\log A(t, T)}{T - t} + \frac{B(t, T)}{T - t} r_t =: a(t, T) + b(t, T) r_t,$$

and therefore $\forall T_1, T_2$:

$$\text{Corr}(R(t, T_1), R(t, T_2)) = \text{Corr}(a(t, T_1) + b(t, T_1)r_t, a(t, T_2) + b(t, T_2)r_t) = 1.$$

It is true that in this case simulated yields do not lose their initial informative shape in time. Nevertheless, the perfect correlation is a strong restriction in terms of flexibility of the shapes, and the result can be that some of the risk is missed. This will be solved by means of a two factors model, as we will see in (3.3.7), which will turn out to be a good compromise between the other two in terms of correlations. Note that one undesired feature, that characterizes the Hull-White frame work in

general, is that low mean reversion speeds can make this shape very flat for relatively short tenor lengths (say, 10Y), as seen for example in Figure (C.4).

Practical implications

One can see the implication of this in both the predicted *shapes* and *values* of the yield curves. As a consequence, the model can capture very restricted movements of the yield curve and hence can miss some of the risk. For what concerns the values, at each time step t , the simulated values for two different maturities T_1 and T_2 will always lie on a straight line. This is shown in Figure (3.23) for the simulation two years ahead of the 1-year and 10-years yields. The slope and intercept values are easily retrievable from the following proposition.

Proposition 3.2.6. *Suppose X and Y have finite nonzero variances σ_X^2 and σ_Y^2 . Define ρ to be the correlation between X and Y . If $|\rho| = 1$, then Y must be an affine linear function of X (and vice versa) with probability one. In particular, $\mathbb{P}(Y = mX + b) = 1$, where:*

$$m = \rho \left(\frac{\sigma_X}{\sigma_Y} \right) \quad \text{and} \quad b = \mathbb{E}[Y] - m\mathbb{E}[X].$$

Proof. Consider the variance of $Y - mX$. By the simple variance formula, we can expand this into:

$$\begin{aligned} \text{Var}(Y - mX) &= \sigma_Y^2 + m^2 \sigma_X^2 - 2m \text{Cov}(X, Y) \\ &= \sigma_Y^2 + m^2 \sigma_X^2 - 2m\rho\sigma_X\sigma_Y \\ &= \sigma_Y^2 + \left(\rho \frac{\sigma_Y}{\sigma_X} \right)^2 \sigma_X^2 - 2 \left(\rho \frac{\sigma_X}{\sigma_Y} \right) \rho\sigma_X\sigma_Y \\ &= \sigma_Y^2 + \rho^2 \sigma_Y^2 - 2\rho^2 \sigma_Y^2 \\ &= \sigma_Y^2 + \sigma_Y^2 - 2\sigma_Y^2 \\ &= 0, \end{aligned}$$

which means that the variance of the random variable $Y - mX$ is zero. A random variable with zero variance has to be a constant with probability one, so $Y - mX = b$ with probability one for some constant b . Taking expectations of both sides shows that $b = \mathbb{E}Y - m\mathbb{E}X$. \square

This drawback characterizes all the one-factor models and can be solved by means of two or more correlated factors, as we will see. In order to give little information about what is up to come, in Figure (3.24) we show the same plot generated under the $HW - 2F$ model. This gives us some initial evidence of the capability of a multi-factors model to exhibit a more flexible behaviour in term of correlation structure.

For what concerns the shapes, there are three possible scenarios: upward sloping, downward sloping and slightly humped. Most importantly, all these predicted yields are never intersecting one with the other, as a consequence of the perfect correlation.

Yield to maturity on infinitely lived bond

From the simulation results we have understood that the yield quantiles have a tendency in converging to a certain point. This tendency is more pronounced when high speed values are concerned, while is less evident for lower values. This is of course linked to the mean-reverting property. More specifically, each upper and lower quantiles at simulation time t converge to the limiting value of the term structure at time t (equivalently, the yield to maturity on an infinitely lived bond at time t). This means that in the $HW - 1F$ model there exists one specific limit $c(t)$ per simulation time t to which the simulated yields must converge. Something similar happens with the Vasicek model,

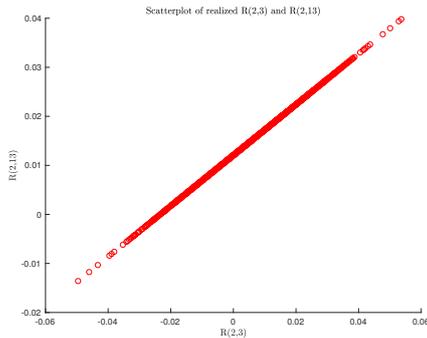


Figure 3.23: Scatterplot of yields at maturities T_1, T_2 respectively ($HW - 1F$).

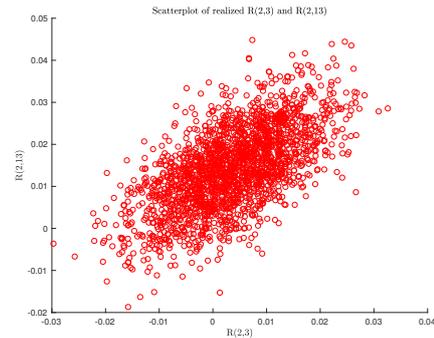


Figure 3.24: Scatterplot of yields at maturities T_1, T_2 respectively ($HW - 2F$).

with the only difference that the limit is unique, due to the fact that there is no fit to the market term structure.

Let us recall the Vasicek model, under which the Zero-Coupon Bond pricing formula has also the structure in (3.13). From [5] the expressions for $A(t, T)$ and $B(t, T)$ are easily retrievable:

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a},$$

$$A(t, T) = \exp\left\{\left(\theta - \frac{\sigma^2}{2a^2}\right)[B(t, T) - T + t] - \frac{\sigma^2}{4a}B(t, T)^2\right\}.$$

To find the yield to maturity of an infinitely lived bond, we compute the limit:

$$\begin{aligned} \lim_{(T-t) \rightarrow \infty} R(t, T) &= \lim_{(T-t) \rightarrow \infty} -\frac{1}{T-t} \{\log A(t, T) - B(t, T)r_t\} = \\ &= \lim_{\tau \rightarrow \infty} -\frac{\log A(t, t+\tau)}{\tau} = \theta - \frac{1}{2} \left(\frac{\sigma}{2a}\right)^2, \end{aligned} \quad (3.27)$$

where we used the fact that $\frac{B(t, t+\tau)}{\tau} \rightarrow 0$ for $\tau \rightarrow \infty$.

Now, the aim is to find a similar result with regard to $HW - 1F$ model. Without loss of generality, let us assume a constant volatility for the short rate dynamics. Contrary to what happens in the Vasicek model, it makes sense now to expect a limit depending on t . In fact, the Hull and White models allows a perfect fit to the term structure of interest rates, which changes every day. The Zero-Coupon Bond price in case of constant parameters a and σ reduces to (3.13), where $A(t, T)$ and $B(t, T)$ are given by:

$$B(t, T) = \frac{1 - e^{-a(T-t)}}{a},$$

$$A(t, T) = \frac{P(0, T)}{P(0, t)} \exp\left\{B(t, T)f(0, t) - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t, T)^2\right\}.$$

It is important to notice here that market bond prices are required for this bond price formula, contrary to the Vasicek model. Now, to derive the yield to maturity on an infinitely lived bond, one

needs to compute $\forall t$ the limit for $T - t = \tau$ that goes to infinity. This, from (2.3), reads:

$$\begin{aligned}
 R(t, T) &= -\frac{1}{T-t} \{ \log A(t, T) - B(t, T)r_t \} = \\
 &= -\frac{1}{T-t} \left\{ \log \frac{P(0, T)}{P(0, t)} + B(t, T)f(0, t) - \frac{\sigma^2}{4a}(1 - e^{-2at})B(t, T)^2 \right. \\
 &\quad \left. - B(t, T)r_t \right\} = -\frac{1}{\tau} \left\{ \log \frac{P(0, t + \tau)}{P(0, t)} + \frac{f(0, t)}{a}(1 - e^{-a\tau}) \right. \\
 &\quad \left. - \frac{\sigma^2}{4a^3}(1 - e^{-2at})(1 - e^{-a\tau})^2 - \frac{r_t}{a}(1 - e^{-a\tau}) \right\} \\
 &= -\frac{1}{\tau} \log \frac{P(0, t + \tau)}{P(0, t)} \xrightarrow{\tau \rightarrow \infty} c(t).
 \end{aligned}
 \tag{3.28}$$

Therefore the limit exist and is finite $\forall t$. More specifically, this corresponds to the limiting value of the term structure at time t . Nevertheless, one needs to know the whole term structure "up to ∞ " to calculate $c(t)$. This could be approximated either taking τ large enough for a given t , either by naively substituting $\theta(t)$ and $\sigma(t)$ to the constant parameters in (3.27). Both approximations lead to the same results.

In Figures (3.25) and (3.26) some of the simulated yields (dotted black lines) and the 97.5%, 2.5% quantiles (blue lines) are shown. The term structure on the 23rd January is previewed in red. While only one scenario is visible after 1 year (upward sloping), after 30 years the model exhibits all the three possible scenarios (upward, downward sloping and slightly humped). Perfect correlation is visible from the fact that there is no intersection between the yields. As a last thing, we also display the yield to maturity on a infinitely lived bond derived in (3.28) for the 1 year and 30 years ahead simulations to which the yield quantiles converge. The limit is approximated by taking τ large enough (35 years, which is possible as our bond curve is interpolated up to 70 years).

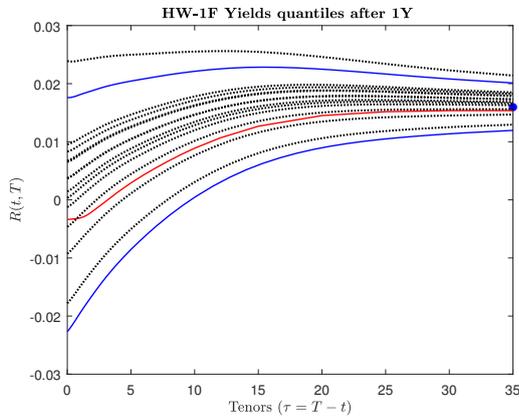


Figure 3.25: Shapes implied by the HW-1F after 1 year simulation together with quantiles and yield on infinitely lived bond (blue dot).

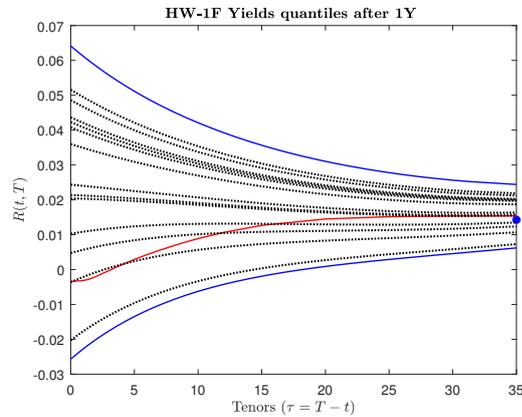


Figure 3.26: Shapes implied by the HW-1F after 30 years simulation together with quantiles and yield on infinitely lived bond (blue dot).

As a final comment for this section, the $HW - 1F$ model, together with other one-factor models has some unrealistic properties. In this case we highlighted the problem of not being able to generate all the yield curve shapes observable in the market. Only three scenarios are possible under one factor models (also for instance the $CIR++$ model exhibits this problem): an increasing curve, a decreasing curve and a curve slightly humped at the beginning. In practice, yield curves usually twist, in the sense that short-maturity yields and long-maturity yields move in opposite directions.

A further problem is represented by the inability of reflecting the imperfect correlation between rates that is observable in the market. This is due to the fact that all unexpected shocks are proportionally transmitted to the whole yield curve. On the other side, multi-factors models turn out to be definitely more able to both predict more realist shapes and correlations between rates. There exist many models in the literature but despite the one-factor model drawbacks, the one factor and two factors models are still the most used in practice as they assure the right trade-off between analytical tractability and accuracy of the results.

3.3. The Hull-White two factors model

We have seen that our way of modeling the interest rate curve in the nine factors model leads to uncorrelated interest rates in a relatively short term. We have learned that, on the contrary, with only one-factor the correlation between interest rates on different maturities is always equal to 1 for each simulation time t . This model may prove to be sufficient when the product to be priced depends only on one single rate of the whole interest curve. Of course in this case there would be no need to account for correlation. Nevertheless, in general, especially when dealing with instruments whose payoff depends on more than one rate, more realistic correlation patterns should be more appropriate. Of course if the rates are referred to maturities that are very close to each other (e.g. 6M, 1Y) then it could not be necessary, as we also know from practice that rates for adjacent maturities are well known to be highly correlated. But for example, if one needs to value a 20Y-swap in the future, one could easily argue that a model like the Hull-White with one factor is not good enough, as lots of different rates at distant maturity times enter Formula (2.5). Other reasons would simply be due to the need for a higher precision in the pricing, when performing the calibration on the swaption volatility surface.

As already done for the $HW - 1F$ model, this section aims to introduce the two-factors model and its characteristics, under the assumption that the volatilities of both stochastic factors are time-dependent. We will provide the equivalent formulation of the model named Two-Additive-Factors gaussian model ($G2 + +$), which turns out to be handier when dealing proofs. Brigo and Mercurio have already proved such equivalence in [5] and therefore we will give it for granted. After this, the formula for the mean reversion level parameter(s) will be derived, so that the yield curve is exactly matched. It is easy to understand that the nice properties of a two-factors model comes at the cost of heavier notations and computations but again, closed-form pricing formulas can be derived for Zero-Coupon Bond options and European swaptions.

3.3.1. Motivation and Introduction

J. Hull and A. White introduced in [16] a new version of the $HW - 1F$ model (1990), which dates back to 1994. This change basically involves a stochastic correction of the mean reversion level. More precisely, this new formulation involves a second Brownian motion, therefore a second source of uncertainty, that describes the diffusion of this correction. We end up modelling two factors instead of one, which solves the problem of perfect correlation. The two-factors Hull-White model ($HW - 2F$) therefore reads:

$$\begin{cases} dr(t) &= [\theta(t) + u(t) - \bar{a}r(t)] dt + \sigma_1(t) dZ_1(t), r(0) = r_0, \\ du(t) &= -\bar{b}u(t) dt + \sigma_2(t) dZ_2(t), u(0) = 0, \end{cases} \quad (3.29)$$

where (Z_1, Z_2) is a two-dimensional Brownian motion under the risk-neutral measure \mathbb{Q} with instantaneous correlation $\bar{\rho}$, so that $dZ_1(t) dZ_2(t) = \bar{\rho} dt$. We assume that $r_0, \bar{a}, \bar{b}, \sigma_1, \sigma_2$ are positive constants and $\bar{\rho} \in [-1, 1]$. Therefore the mean reversion level is not modeled as a deterministic function, but as a stochastic process $u(t)$ plus a deterministic shift $\theta(t)$. Here, $\sigma_1(t)$ and $\sigma_2(t)$ are assumed regular enough to ensure the existence and uniqueness of a solution and $\theta(t)$ again is the

time dependent parameter that replicates the current term structure in the market. Also this model is mean reverting, and precisely, the short rate $r(t)$ at any time t is pushed back to the stochastic level $\frac{\theta(t)+u(t)}{a}$. Also in this case the two speeds \bar{a} and \bar{b} will be assumed constant, for the same reasoning that applied to the one-factor model. Also in this case, the volatilities $\sigma_1(t)$ and $\sigma_2(t)$ are assumed to be time-dependent, so not to worsen the pricing precision with respect to $HW - 1F$.

Simple integration of (3.32) leads to:

$$r(t) = x(s)e^{-a(t-s)} + y(s)e^{-b(t-s)} + \int_s^t \sigma(u)e^{-a(t-u)} dW_1(u) + \int_s^t \eta(u)e^{-b(t-u)} dW_2(u) + \beta(t),$$

which implies that also under this model $r(t)$ conditional to \mathcal{F}_s is normally distributed with following mean and variance:

$$\begin{aligned} \mathbb{E}[r(t)|\mathcal{F}_s] &= x(s)e^{-a(t-s)} + y(s)e^{-b(t-s)} + \beta(t), \\ \text{Var}[r(t)|\mathcal{F}_s] &= \int_s^t e^{-2a(t-u)} \sigma(u)^2 du + \int_s^t e^{-2b(t-u)} \eta(u)^2 du + 2\rho \int_s^t \sigma(u)\eta(u)e^{-(a+b)(t-u)} du, \end{aligned} \quad (3.30)$$

where \mathbb{E} and Var denote the mean and variance operators under the risk-neutral measure \mathbb{Q} . In case the volatilities are assumed constant, the variance will read

$$\text{Var}[r(t)|\mathcal{F}_s] = \frac{\sigma^2}{2a} [1 - e^{-2a(t-s)}] + \frac{\eta^2}{2b} [1 - e^{-2b(t-s)}] + 2\rho \frac{\sigma\eta}{(a+b)} [1 - e^{-(a+b)(t-s)}].$$

Note that, for the same reasons as in $HW - 1F$ model, at each time t the short rate $r(t)$ can be negative with positive probability, which also does not converge to zero.

The $HW - 2F$ model has an equivalent formulation known as $G2++$ model, which assumes that the short rate evolves according to:

$$r(t) = x(t) + y(t) + \beta(t), \quad x(0) = 0, \quad y(0) = 0, \quad (3.31)$$

where the processes $\{x(t), t \geq 0\}$ and $\{y(t), t \geq 0\}$ satisfy

$$\begin{cases} dx(t) = -ax(t) dt + \sigma(t) dW_1(t), & x(0) = 0, \\ dy(t) = -by(t) dt + \eta(t) dW_2(t), & y(0) = 0, \end{cases} \quad (3.32)$$

with (W_1, W_2) being a two-dimensional Brownian motion with instantaneous correlation ρ , so that $dW_1(t)dW_2(t) = \rho dt$. The deterministic function $\beta(t)$ has the same role as $\theta(t)$ in (3.29), that is allowing perfect fit to the current term structure of interest rates. We assume that r_0, a, b are positive constants, $\sigma(t), \eta(t)$ are positive functions and $\rho \in [-1, 1]$. The relation between $\theta(t)$ and $\beta(t)$ is still given by (3.9), as shown in [5]. The explicit solutions are:

$$\begin{aligned} x(t) &= x(s)e^{-a(t-s)} + \int_s^t \sigma(u)e^{-a(t-u)} dW_1(u), \\ y(t) &= y(s)e^{-b(t-s)} + \int_s^t \eta(u)e^{-b(t-u)} dW_2(u). \end{aligned} \quad (3.33)$$

Furthermore one can express (3.32) in terms of two independent Brownian motion \widetilde{W}_1 and \widetilde{W}_2 by means of a Cholesky decomposition applied to their correlation matrix. In such a way we get to

$$\begin{cases} dx(t) = -ax(t) dt + \sigma(t) d\widetilde{W}_1(t), & x(0) = 0, \\ dy(t) = -by(t) dt + \rho\eta(t) d\widetilde{W}_1(t) + \eta(t)\sqrt{1-\rho^2} d\widetilde{W}_2(t), & y(0) = 0, \end{cases} \quad (3.34)$$

which is used for the simulation.

Fit to initial term structure

The following theorem provide the calibration formula for the mean reversion level for both the two formulations.

Theorem 3.3.1 (Calibration of mean reversion level). *The exact fit to the initial term structure of discount bonds is guaranteed if $\beta(t)$ in (3.31) is calibrated as*

$$\begin{aligned} \beta(t) &= f^M(0, t) + \int_0^t \frac{\sigma(u)^2}{a} (1 - e^{-a(t-u)}) e^{-a(t-u)} \, du \\ &\quad + \int_0^t \frac{\eta(u)^2}{b} (1 - e^{-b(t-u)}) e^{-b(t-u)} \, du \\ &\quad + \rho \int_0^t \sigma(u)\eta(u) \left[\frac{1 - e^{-b(t-u)}}{b} e^{-a(t-u)} + \frac{1 - e^{-a(t-u)}}{a} e^{-b(t-u)} \right] \, du, \end{aligned} \quad (3.35)$$

where $f^M(0, t)$ is the current instantaneous forward rate observed in the market. Equivalently, one needs to set $\theta(t) = \frac{\partial \beta(t)}{\partial t} + a\beta(t)$ in (3.29).

Proof. The model fits the current term structure of interest rates if, for each maturity T , $P(0, T) = P^M(0, T)$, where $P^M(0, T)$ is observed in the market. To compute $P(0, T)$ we need to integrate the process $r(t)$ over $[0, T]$. In order to do so, we will use the fact that since the process $x(t) + y(t)$ is normally distributed conditionally on \mathcal{F}_t , then also $I(t) = \int_0^T [x(t) + y(t)] \, dt$ is normally distributed with mean zero and variance given by

$$\begin{aligned} \text{Var}\{I(t)|\mathcal{F}_0\} &= \int_0^T \sigma(u)^2 \frac{(1 - e^{-a(T-u)})^2}{a^2} \, du + \int_0^T \eta(u)^2 \frac{(1 - e^{-b(T-u)})^2}{b^2} \, du \\ &\quad + 2\rho \int_0^T \sigma(u)\eta(u) \frac{1 - e^{-a(T-u)}}{a} \frac{1 - e^{-b(T-u)}}{b} \, du =: V(0, T). \end{aligned} \quad (3.36)$$

Again, this is a consequence of Fubini's theorem in [13]. Note that in the case that $\sigma(t) = \sigma$ is a positive constant function, the variance reads

$$\begin{aligned} V(0, T) &= \frac{\sigma^2}{a^2} \left(T - 2 \frac{1 - e^{-aT}}{a} + \frac{1 - e^{-2aT}}{2a} \right) + \frac{\eta^2}{b^2} \left(T - 2 \frac{1 - e^{-bT}}{b} + \frac{1 - e^{-2bT}}{2b} \right) \\ &\quad + 2\rho \frac{\sigma\eta}{ab} \left(T - \frac{1 - e^{-aT}}{a} \frac{1 - e^{-bT}}{b} + \frac{1 - e^{-(a+b)T}}{a+b} \right). \end{aligned}$$

Finally, recalling the moment generating function for a normal random variable Z ⁴, we have

$$P(0, T) = \mathbb{E}[e^{-\int_0^T r(t) \, dt} | \mathcal{F}_0] = \mathbb{E}[e^{-\int_0^T \beta(u) \, du - I(T)} | \mathcal{F}_0] = e^{-\int_0^T \beta(u) \, du} e^{\frac{1}{2}V(0, T)},$$

and going through the same steps as in the proof of Theorem (3.2.1), we get to the following formula from which $\beta(t)$ in (3.35) is easily deductible:

$$\begin{aligned} f^M(0, T) &= \beta(T) - \frac{1}{2} \frac{\partial V(0, T)}{\partial T} = \beta(T) - \int_0^T \frac{\sigma(u)^2}{a} (1 - e^{-a(T-u)}) e^{-a(T-u)} \, du \\ &\quad - \int_0^T \frac{\eta(u)^2}{b} (1 - e^{-b(T-u)}) e^{-b(T-u)} \, du \\ &\quad - \rho \int_0^T \sigma(u)\eta(u) \left[\frac{1 - e^{-b(T-u)}}{b} e^{-a(T-u)} + \frac{1 - e^{-a(T-u)}}{a} e^{-b(T-u)} \right] \, du. \end{aligned}$$

⁴If $Z \sim \mathcal{N}(\mu_Z, \sigma_Z^2)$, then $M_Z(t) = \mathbb{E}[e^{tZ}] = e^{t\mu_Z + \frac{1}{2}t^2\sigma_Z^2}$.

Note that under the assumption of constant volatilities, the former equation easily reduces to calibrating $\beta(t)$ as follows:

$$\beta(t) = f^M(0, t) + \frac{\sigma^2}{2a^2}(1 - e^{-aT})^2 + \frac{\eta^2}{2b^2}(1 - e^{-bT})^2 + \rho \frac{\sigma\eta}{ab}(1 - e^{-aT})(1 - e^{-bT}). \quad (3.37)$$

As a conclusion, in either cases, it will suffice to calibrate $\theta(t)$ as in (3.9). \square

3.3.2. Zero-Coupon Bond price

The $HW - 2F$ model also belongs to the family of Affine Structure Models. Solving the systems in (2.10) leads to the following pricing formula for the Zero-Coupon Bond:

$$P(t, T) = A(t, T) \exp\{-B(a, t, T)x(t) - B(b, t, T)y(t)\}, \quad (3.38)$$

where $A(t, T)$ and $B(z, t, T)$ are respectively given by

$$A(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \exp\left\{\frac{1}{2}[V(t, T) - V(0, T) + V(0, t)]\right\},$$

$$B(z, t, T) = \frac{1 - e^{-z(T-t)}}{z}.$$

Here $V(t, T)$ is the variance of the integrated sum process easily retrievable from (3.36). Note that the time dependency of volatilities impacts only the calculation of the integrated variance $V(t, T)$, when calculating the bond price $P(t, T)$.

3.3.3. Dynamics under the \mathbb{Q}^T measure

In order to derive the explicit formula for pricing a European swaption, it is necessary to derive the dynamics of the two factors $x(t)$ and $y(t)$ under the risk-adjusted measure, that is the T -forward measure \mathbb{Q}^T . This result is reachable from Girsanov's theorem and is summarized in the following theorem.

Theorem 3.3.2. *The dynamics of the process $x(t)$ and $y(t)$ under the risk-adjusted measure \mathbb{Q}^T are:*

$$\begin{cases} dx(t) &= -[ax(t) + \sigma(t)(\sigma(t)B(a, t, T) + \rho\eta(t)B(b, t, T))] dt + \sigma(t) dW_1^T(t), \\ dy(t) &= -[by(t) + \eta(t)(\eta(t)B(b, t, T) + \rho\sigma(t)B(a, t, T))] dt + \eta(t) dW_2^T(t), \end{cases}$$

where $dW_1^T(t) = dW_1(t) + (\sigma(t)B(a, t, T) + \rho\eta(t)B(b, t, T)) dt$ and $dW_2^T(t) = dW_2(t) + (\eta(t)B(b, t, T) + \rho\sigma(t)B(a, t, T)) dt$. From this it follows that under \mathbb{Q}^T , the vector $\mathbf{Z}_t = (x(t), y(t))$ conditioned on \mathcal{F}_s is normally distributed with mean and covariance matrix:

$$\begin{aligned} \mathbb{E}[\mathbf{Z}_t | \mathcal{F}_s] &= [x(s)e^{-a(t-s)} - M_x^T(s, t), y(s)e^{-b(t-s)} - M_y^T(s, t)], \\ \text{Cov}(\mathbf{Z}_t | \mathcal{F}_s) &= \begin{bmatrix} \int_s^t \sigma(u)^2 e^{-2a(t-u)} du & \rho \int_s^t \sigma(u)\eta(u) e^{-(a+b)(t-u)} du \\ \rho \int_s^t \sigma(u)\eta(u) e^{-(a+b)(t-u)} du & \int_s^t \eta(u)^2 e^{-2b(t-u)} du \end{bmatrix} \\ &= \begin{bmatrix} \kappa_{x,T}^2(s, t) & \varrho_{x,y,T}(s, t) \\ \varrho_{x,y,T}(s, t) & \kappa_{y,T}^2(s, t) \end{bmatrix} \end{aligned}$$

Here $M_x^T(s, t)$ and $M_y^T(s, t)$ are respectively given by:

$$M_x^T(s, t) = \int_s^t \left[\frac{\sigma(u)^2}{a} (1 - e^{-a(T-u)}) e^{-a(t-u)} + \rho \frac{\sigma(u)\eta(u)}{b} (1 - e^{-b(T-u)}) e^{-b(t-u)} \right] du,$$

$$M_y^T(s, t) = \int_s^t \left[\frac{\eta(u)^2}{b} (1 - e^{-b(T-u)}) e^{-b(t-u)} + \rho \frac{\sigma(u)\eta(u)}{a} (1 - e^{-a(T-u)}) e^{-a(t-u)} \right] du.$$

The previous theorem is easily proved by following the same steps done in the proof of Theorem (3.2.2), given that in this case the bond-price dynamics are given by

$$dP(t, T) = r(t)P(t, T) dt - \sigma(t)B(a, t, T)P(t, T) dW_1(t) - \eta(t)B(b, t, T)P(t, T) dW_2(t). \quad (3.39)$$

Note that from (3.3.2) it follows that

$$\begin{aligned} x(t) &= x(s)e^{-a(t-s)} - M_x^T(s, t) + \int_s^t \sigma(u)e^{-a(t-u)} dW_1^T(u), \\ y(t) &= y(s)e^{-b(t-s)} - M_y^T(s, t) + \int_s^t \eta(u)e^{-b(t-u)} dW_2^T(u), \end{aligned}$$

which differs with (3.33) only in the drift term.

3.3.4. Bond-option and swaption prices

Bond-Option price

The following lemma is a well known result from statistics and it is used in order to derive the bond-option price under the $HW - 2F$ model.

Lemma 3.3.3. *Distribution of sum of correlated random normals* If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ have correlation ρ , then their sum

$$Z = aX + bY$$

has a one-dimensional normal distribution such that

$$Z \sim \mathcal{N}(a\mu_X + b\mu_Y, a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_X\sigma_Y\rho).$$

Theorem 3.3.4. *The price at t_0 of a European put option with strike K and maturity T , written on a Zero-Coupon Bond with maturity T_F is given, under the $G2++$ model, by:*

$$V_{ZCB}^{\text{Put}}(t_0, T, T_F, K) = P(t_0, T) \left\{ K\Phi(d_1) - A(T, T_F)e^{\frac{\kappa_{\bar{Z}}^2 - 2\mu_{\bar{Z}}}{2}} \Phi(d_2) \right\}, \quad (3.40)$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, and $\mu_{\bar{Z}}, \kappa_{\bar{Z}}, d_1, d_2$ read

$$\begin{cases} \mu_{\bar{Z}} & := B(a, T, T_F)\mu_{x,T}(t_0, T) + B(b, T, T_F)\mu_{y,T}(t_0, T), \\ \kappa_{\bar{Z}}^2 & := B(a, T, T_F)^2\kappa_{x,T}^2(t_0, T) + B(b, T, T_F)^2\kappa_{y,T}^2(t_0, T) + 2B(a, T, T_F)B(b, T, T_F)\rho_{x,y,T}(t_0, T), \\ d_1 & := -\frac{\log A(T, T_F) - \log K - \mu_{\bar{Z}}}{\kappa_{\bar{Z}}}, \\ d_2 & := d_1 - \kappa_{\bar{Z}}. \end{cases}$$

Proof. Consider a put option that gives the holder the right (but not the obligation) to sell a Zero-Coupon Bond with maturity T_F at a future time T . The option will be exercised if $K - P(T, T_F) > 0$. The price at $t_0 < T$ is given, by no arbitrage arguments, by:

$$\begin{aligned} V_{ZCB}^{\text{Put}}(t_0, T, T_F) &= \mathbb{E}^{\mathbb{Q}} \left[\frac{B(t_0)}{B(T)} (K - P(T, T_F))^+ | \mathcal{F}(t_0) \right] \\ &= P(t_0, T) \mathbb{E}^{\mathbb{Q}^T} \left[(K - A(T, T_F)e^{-B(a, T, T_F)x(T) - B(b, T, T_F)y(T)})^+ | \mathcal{F}(t_0) \right]. \end{aligned}$$

Now, if one calls:

$$\begin{aligned} \tilde{X}(T, T_F) &:= B(a, T, T_F)x(T) | \mathcal{F}(t_0) \sim \mathcal{N}(B(a, T, T_F)\mu_{x,T}(t_0, T), B(a, T, T_F)^2\kappa_{x,T}^2(t_0, T)), \\ \tilde{Y}(T, T_F) &:= B(b, T, T_F)y(T) | \mathcal{F}(t_0) \sim \mathcal{N}(B(b, T, T_F)\mu_{y,T}(t_0, T), B(b, T, T_F)^2\kappa_{y,T}^2(t_0, T)), \end{aligned}$$

and notes that $\text{Corr}(\tilde{X}(T, T_F), \tilde{Y}(T, T_F)) = \frac{\rho_{x,y,T}(t_0, T)}{\kappa_{x,T}(t_0, T)\kappa_{y,T}(t_0, T)}$, it immediately follows from Lemma (3.3.3) that the random variable $\tilde{Z} = \tilde{X} + \tilde{Y}$ is also a normal random variable with parameters:

$$\begin{aligned}\mu_{\tilde{Z}} &= B(a, T, T_F)\mu_{x,T}(t_0, T) + B(b, T, T_F)\mu_{y,T}(t_0, T), \\ \kappa_{\tilde{Z}}^2 &= B(a, T, T_F)^2\kappa_{x,T}^2(t_0, T) + B(b, T, T_F)^2\kappa_{y,T}^2(t_0, T) + 2B(a, T, T_F)B(b, T, T_F)\rho_{x,y,T}(t_0, T).\end{aligned}$$

This said, going over again the same steps as in proof of Theorem (3.2.3), the thesis follows. \square

Swaption price

We here discuss the problem of pricing an European swaption under the two factors Hull-White model. Given (3.38), we have that

$$\begin{aligned}V^{\text{PSwpt}}(t_0; T, T_\beta) &= NP(t_0, T)\mathbb{E}_{t_0}^{\mathbb{Q}^T} \left[\left(1 - \sum_{k=\alpha+1}^{\beta} c_k P(T, T_k) \right)^+ \right] \\ &= NP(t_0, T) \int_{\mathbb{R}^2} \left(1 - \sum_{k=\alpha+1}^{\beta} c_k A(T, T_k) e^{-B(a, T, T_k)x(T) - B(b, T, T_k)y(T)} \right) g(z) \, dz,\end{aligned}$$

where $g(z)$ is the density function of vector \mathbf{Z}_T conditionally on \mathcal{F}_{t_0} . This said, the idea of having to compute a two-dimensional integral is not really appealing.

Now, trying to repeat the same steps as in the proof of Theorem (3.2.5), we can get to the intuition that the aforementioned two-dimensional integral can somehow be simplified to a one-dimensional integral. More precisely, if we apply the change of measure and plug in the pricing formula for $P(t, T)$ in (3.38), we get that the European swaption price at time t_0 , under the G2 + model, is given by:

$$V^{\text{PSwpt}}(t_0; T, T_\beta) = NP(t_0, T)\mathbb{E}_{t_0}^{\mathbb{Q}^T} \left[\left(1 - \sum_{k=\alpha+1}^{\beta} c_k A(T, T_k) e^{\{-B(a, T, T_k)x(T) - B(b, T, T_k)y(T)\}} \right)^+ \right]. \quad (3.41)$$

At this stage, we understand that it is not possible to apply the decomposition of Lemma (3.2.4), as there is no unique solution to

$$1 - \sum c_k A(T, T_k) e^{-B(a, T, T_k)\tilde{x} - B(b, T, T_k)\tilde{y}} = 0. \quad (3.42)$$

Anyhow, we can guess that there should be a one-dimensional integral in the pricing formula that can account for these two variables. The reason is due to the monotonicity of the function “sum”, which implies that for each value of \tilde{y} there exist a value \tilde{x} such that (3.42) is satisfied. Brigo and Mercurio proved in [5] that this is indeed the case. Precisely, this is proved for the case in which the two volatilities σ and η are assumed constant. The theorem is reported here and it is easily applicable to the time dependent case. In fact, the volatility parameters are not directly appearing in the formula as they are implicit in $A(t, T)$ and the parameters defining the distribution of vector \mathbf{Z}_t , which are provided by Theorem (3.3.2). Prior to the theorem statement, in the following proposition we describe an alternative way in order to proceed in the direction of Jamshidian’s decomposition.

Observation 3.3.5. *Note that hypothetically, there is a way so to proceed with Equation (3.41), and end up with a pricing formula similar to the one of Theorem (3.2.5). An idea would be to follow the steps:*

1. Choose $\tilde{y} = y^*$.
2. Find \tilde{x} such that (3.42) is satisfied.

3. Run the optimization over the model parameters, so that the swaption implied surface is matched.
4. Iterate over a certain range of values for \tilde{y} , and select the one such that the loss function of the previous optimization is minimized.

After this is done, we would have a model that has already been calibrated, and for which the pricing formula for an European swaption is given by

$$V^{\text{pSwpt}}(t_0; T, T_\beta) = NP(t_0, T) \sum_{k=\alpha+1}^{\beta} c_k V_{\text{ZCB}}^{\text{Put}}(t_0, T, T_k, \hat{K}),$$

where $\hat{K} = A(T, T_k)A(T, T_k)e^{-B(a, T, T_k)x^* - B(b, T, T_k)y^*}$ and $V_{\text{ZCB}}^{\text{Put}}$ as in (3.40). We say ‘‘hypothetically’’ as this method has not been tried and could turn out not to be efficient in terms of computation time.

In the following theorem we state the swaption pricing formula under the G2++ , proved by Brigo and Mercurio in [5].

Theorem 3.3.6 (Swaption price under G2++ model). *Consider an European (payer) swaption with notional N , strike K , maturity T , written on an Interest Rate Swap with first reset date T_α and payment dates $T_{\alpha+1}, \dots, T_\beta$. Assume that $T = T_\alpha$. We denote by Δ_i the year fraction from T_{i-1} and T_i , and set $c_i := K\Delta_i$ for $i = \alpha + 1, \dots, \beta - 1$ and $c_\beta := 1 + K\Delta_\beta$. The risk-free price at $t_0 \leq T$ of such instrument is given, under the G2++ model, by the one-dimensional integral:*

$$V^{\text{Swpt}}(t_0; T, T_\beta, \omega, a, b, \sigma(t), \eta(t), \rho) = N\omega P(t_0, T) \int_{-\infty}^{+\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu_x}{\sigma_x}\right)^2}}{\sigma_x \sqrt{2\pi}} \left[\Phi(-\omega h_1(x)) - \sum \lambda_i(x) e^{\kappa_i(x)} \Phi(-\omega h_2(x)) \right] dx, \quad (3.43)$$

where $\omega = 1$ ($\omega = -1$) for a payer (receiver) swaption,

$$\begin{aligned} h_1(x) &:= \frac{\tilde{y} - \mu_y}{\sigma_y \sqrt{1 - \rho_{xy}^2}} - \frac{\rho_{xy}(x - \mu_x)}{\sigma_x \sqrt{1 - \rho_{xy}^2}}, \\ h_2(x) &:= h_1(x) + B(b, T, T_i) \sigma_y \sqrt{1 - \rho_{xy}^2}, \\ \lambda_i(x) &:= c_i A(T, T_i) e^{-B(a, T, T_i)x}, \\ \kappa_i(x) &:= -B(b, T, T_i) \left[\mu_y - \frac{1}{2}(1 - \rho_{xy}^2) \sigma_y^2 B(b, T, T_i) + \rho_{xy} \sigma_y \frac{x - \mu_x}{\sigma_x} \right], \end{aligned}$$

$\tilde{y} = \tilde{y}(x)$ is the unique solution of the following equation:

$$\sum_{i=\alpha+1}^{\beta} c_i A(T, T_i) e^{-B(a, T, T_i)x - B(b, T, T_i)\tilde{y}} = 1,$$

and

$$\begin{cases} \mu_x & := -M_x^T(t_0, T), \\ \mu_y & := -M_y^T(t_0, T), \\ \sigma_x & := \kappa_{x,T}(t_0, T), \\ \sigma_y & := \kappa_{y,T}(t_0, T), \\ \rho_{xy} & := \frac{\varrho_{x,y,T}(t_0, T)}{\kappa_{x,T}(t_0, T) \kappa_{y,T}(t_0, T)}. \end{cases}$$

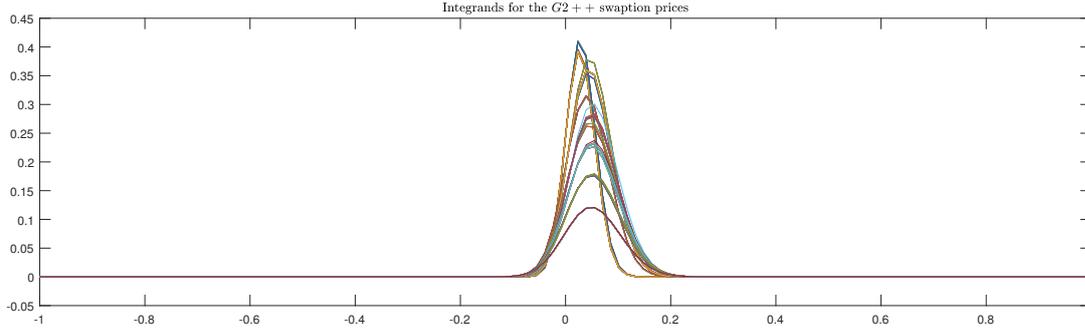


Figure 3.27: Plot of integrands from pricing formula (3.43) and different swaption parameters.

Note that integral in Equation (3.43) suffices to be computed on a smaller domain than $(-\infty, \infty)$. In our case, the integrands have proven not to exceed domains bigger than $[-0.5, 0.5]$ and therefore we reckon the interval $[-1, 1]$ to be a safe enough choice. As a support for this choice, in Figure (3.27) we displayed computed integrands as in (3.43) for several different swaptions. This integral can be calculated numerically, for example, by Gauss-Hermite quadrature⁵ to form a semi-analytical formula. In our opinion, a 12-point quadrature provides sufficient accuracy and turns out to work relatively good in terms of computation times. Anyway, note that this pricing formula requires to compute one root finding per grid point, and therefore it can turn out to be quite unfeasible when calibrating to several different dates, as it is done in Chapter (4). Therefore, a more efficient formula for that purpose is required. This is accomplished by means of the approximation formula for swaptions introduced by Schrager and Pelsser in [29, p. 173], which we report in the next paragraph. Of course there exist many other approximation formulas that are proven to perform more efficiently than (3.43). For this we refer, for example, to [6].

Approximated swaption price with time-dependent $\sigma(t)$ and $\eta(t)$

This approximation formula turns out to be very useful when calibrating a multi-factors model in general. In fact, iterating the previous reasoning for two dimensions, the swaption pricing under a k -factors model is the result of an integral over \mathbb{R}^k which is definitely computationally inefficient for large k . Schrager and Pelsser provides us with a very simple analytical pricing formula for a European swaption in the framework of gaussian short rate. This formula reads:

$$V^{\text{Swpt}}(t_0; T, T_\beta) \approx N \frac{v_{\text{Swap}}}{\sqrt{2\pi}} \sum_{j=\alpha+1}^{\beta} \Delta_j P(0, T_j) \equiv N \frac{v_{\text{Swap}}}{\sqrt{2\pi}} \mathcal{P}_T^{T_\beta}, \quad (3.44)$$

where N is the notional, K the strike, Δ_j the year fraction from T_{j-1} and T_j , $T_{\alpha+1}, \dots, T_\beta$ the payment dates, T the first reset date and v_{Swap} the approximated volatility of the forward swap rate given by:

$$v_{\text{Swap}} = \sqrt{\int_{t_0}^T \sigma(u)^2 C_x^2 e^{2au} du + \int_{t_0}^T \eta(u)^2 C_y^2 e^{2bu} du + 2 \int_{t_0}^T \sigma(u) \eta(u) \rho C_1 C_2 e^{(a+b)u} du}.$$

The two constants C_x and C_y read:

$$C_x = e^{-aT} \frac{P(t_0, T)}{\mathcal{P}_T^{T_\beta}} - e^{-aT_\beta} \frac{P(t_0, T_\beta)}{\mathcal{P}_T^{T_\beta}} - K \sum_{j=\alpha+1}^{\beta} e^{-aT_j} \Delta_j \frac{P(t_0, T_j)}{\mathcal{P}_T^{T_\beta}},$$

$$C_y = e^{-bT} \frac{P(t_0, T)}{\mathcal{P}_T^{T_\beta}} - e^{-bT_\beta} \frac{P(t_0, T_\beta)}{\mathcal{P}_T^{T_\beta}} - K \sum_{j=\alpha+1}^{\beta} e^{-bT_j} \Delta_j \frac{P(t_0, T_j)}{\mathcal{P}_T^{T_\beta}}.$$

⁵ $\int_{-\infty}^{\infty} e^{-x^2} f(x) dx \approx \sum_{i=1}^n w_i f(x_i)$, where n is the number of sample points used.

3.3.5. Calibration to swaptions

The same strip of coterminal At-The-Money swaptions used for the one factor model will be used for the calibration to swaptions under the $G2++$ model. Also in this case the two volatilities are parametrized as piece-wise constant step functions and a bootstrap method is used. The aim is to minimize a certain distance between the quoted and implied prices. As already done for the previous calibration, the chosen loss function will be the normalized squared loss function, that is:

$$l(x, y) = \left(\frac{x - y}{y} \right)^2.$$

Given the set of n swaptions, let $V_k^N\{\sigma_k^N\}$ be the Normal price of instrument $k = 1, \dots, n$ and $V_k^{\text{HW}}\{\sigma_k\}$ the two factors Hull-White price of instrument $k = 1, \dots, n$ as in either (3.43) or (3.44). Having $\sigma_1, \dots, \sigma_{k-1}, \eta_1, \dots, \eta_{k-1}$, we determine σ_k, η_k by imposing

$$V_k^{\text{HW}}\{\sigma_1, \dots, \sigma_k; \eta_1, \dots, \eta_k\} = V_k^N\{\sigma_k^N\}.$$

This method is still reasonably fast, but requires more time compared to the one factor model for mainly two reasons: the pricing formula itself requires more computation time and one equation with the two unknowns (σ_k, η_k) needs to be solved.

Problem of ending up in a local minima

As a first step, we started by calibrating the $HW - 2F$ model with the assumption of constant parameters. More specifically, we have tried to solve the following optimization:

$$\operatorname{argmax}_{a, b, \sigma, \eta, \rho} \sum \left(\frac{V_i^{\text{HW}}(a, b, \sigma, \eta, \rho) - V_i^N(\sigma_i^N)}{V_i^N(\sigma_i^N)} \right)^2,$$

subject to the linear constraints:

$$\begin{aligned} -0.001 &\leq a \leq UB_a; \\ 0.001 &\leq b \leq UB_b, \\ -1 &\leq \rho \leq 1, \end{aligned}$$

where UB_a and UB_b should be chosen in a way to allow only for realistic estimates of the mean reversion speeds (respectively of the short rate and mean reversion level correction). We set $UB_a = 0.5$ and $UB_b = 0.2$.

No stable solution could be found under this formulation. The reason is two-fold: first of all, a calibration basket of six ATM swaptions is too poor for the calibration of five parameters. As a consequence, many solutions that leads to good market fits are found from different initial conditions for the parameters. A possible solution to this problem could be to enrich the calibration basket, with maybe swaptions quoted for more extreme strikes. Secondly, the function we aim to minimize is not convex in the parameters and to our knowledge, there is no convenient choice of the loss function that would guarantee such property. As a consequence many local minima render the minimization problem even trickier. The literature tells us that different approaches could be considered to overcome such problem. One possible idea ([26]) could be to consider an adaptive lattice which reduces the danger of getting trapped near a wrong local minimum. Another option would be to perform the minimization via Simulated Annealing (main idea would be to perform a minimization for randomly generated initial points and then select the better solution, [22]).

Our winning calibration procedure follows the same line of thought used for the calibration of the one factor model, with the difference that the internal optimization is done over the two

time-dependent volatilities $\sigma(t)$ and $\eta(t)$ and the external over the remaining parameters a, b and ρ , which again are found so that the distance between model and market volatility surfaces is minimized. This is explained in more detail in the next paragraph.

Calibration to swaptions with time dependent volatilities

We have said that we would want to compare the effect of using one and two factors models for assessing the exposure profile of a portfolio of vanilla swaps. Prior to this, it makes sense to give a reasoning behind the introduction of a more complicated model, and the assumptions that are made on that model. First of all, we would like to analyze a model with two factors so that the non-perfect correlations of rates observable in the market can be better replicated. This is an important issue, as a wrong modeling of the correlations could lead to underestimation of risk, which is undesired. On the other hand, the assumption on time-dependent volatilities also in the two factor model must be preferred, as logically nobody would want a multi-factors model that worsens the pricing precision. Therefore, as we considered time dependent volatility in the $HW - 1F$ model, a comparison makes sense only if a similar assumption holds for the more complex model.

Also for the $HW - 2F$ model a bootstrap-like method for the calibration has been followed. The two volatility functions $\sigma(t)$ and $\eta(t)$ have been parametrized as piece-wise constant functions within the same intervals, where the end point of each interval corresponds to one market maturity. Let us denote with (M, T) the swaption with option exercise date M and length of underlying swap T . Given the calibration basket $\Sigma = \{(M_1, T_1), \dots, (M_n, T_n)\}$ (e.g. $n = 6$ swaptions of 12-Y coterminal as done for the one factor model), we define:

$$\sigma(t) = \sum_{j=1}^n \sigma_j \mathbb{1}_{(t_{j-1}, t_j]}(t) + \sigma_n \mathbb{1}_{(t_n, \infty)}(t),$$

$$\eta(t) = \sum_{j=1}^n \eta_j \mathbb{1}_{(t_{j-1}, t_j]}(t) + \eta_n \mathbb{1}_{(t_n, \infty)}(t),$$

where t_j is the maturity of the j -th swaption. Note that now the number of parameters that needs to be calibrated is $2n$.

The calibration process is approached in two macro steps. Firstly, we build up a set of functions for the calibration of $\sigma(t)$ and $\eta(t)$, for values of a, b and ρ that must be given as inputs. Secondly, we minimize over the remaining parameters. The algorithm is schematized through the following points:

- Choose a, b, ρ (reasonably).
- for $i = 1, \dots, n$ find (σ_i, η_i) such that a certain loss function $l(V_{i,i}^N, V_{i,i}^{HW}(\sigma_1, \dots, \sigma_i, \eta_1, \dots, \eta_i))$ is minimized, where $V_{i,i}^N$ and $V_{i,i}^{HW}$ denote the market and model price of instrument (M_i, T_i) , for $i = 1, \dots, n$.
- Calculate the total error over the counter-diagonal as

$$ERR_1 = \sum_{i=1}^n l(V_{i,i}^N, V_{i,i}^{HW}(\sigma_1, \dots, \sigma_i, \eta_1, \dots, \eta_i)).$$

- The error ERR_1 in the counter diagonal will be very small. Iterate the process and find a, b, ρ such that the error calculated outside the counter-diagonal is the smallest. Or, we can also

use a weight ω for the two error components and find $(\hat{a}, \hat{b}, \hat{\rho})$ s.t.:

$$\begin{aligned} (\hat{a}, \hat{b}, \hat{\rho}) &= \arg \min_{(a,b,\rho) \in \Omega} [\omega \text{ERR}_1(a, b, \rho) + (1 - \omega) \text{ERR}_2(a, b, \rho)] \\ &= \arg \min_{(a,b,\rho) \in \Omega} \left[\min_{\substack{\sigma(t,a), \\ \eta(t,b)}} \omega \sum_{i=j} \left(\frac{V_{i,j}^N - V_{i,j}^{\text{HW}}(a, b, \sigma(t), \eta(t), \rho)}{V_{i,j}^N} \right)^2 \right] \\ &\quad + (1 - \omega) \sum_{i \neq j} \left(\frac{V_{i,j}^N - V_{i,j}^{\text{HW}}(a, b, \sigma(t), \eta(t), \rho)}{V_{i,j}^N} \right)^2, \end{aligned}$$

with ω being the weight that the user wants to give to ERR_1 , and $(1 - \omega)$ to ERR_2 . In practice, we set $\omega = \frac{1}{2}$, so that equal weight is given over the whole surface.

Calibration of the mean reversion speeds a, b

Let us call $\mathbf{LB} = [-0.01, 0.001, -0.99]$, $\mathbf{UB} = [0.6, 0.1, 0.99]$ and $\Omega = \mathbf{LB} \times \mathbf{UB}$. The aim here is to find the minimum of a constrained nonlinear multivariable function. More precisely, a, b and ρ are calibrated as the value in Ω that minimize the total normalized distance between the market and model prices throughout the whole volatility surface, given that the $\sigma(t)$ and $\eta(t)$ are calibrated to one diagonal (with maturity and tenor that always sum up to 12 years, as in Table (3.3)).

The minimum is found at $[\hat{a}, \hat{b}, \hat{\rho}] = [0.05, 0.09, -0.97]$. In order to give more evidence to the fact that we have not been trapped in a local minima we provide the error surface $g(t, a, b)$ as a function of the two speeds in Figure (3.29), while in Figure (3.28) the final calibration for the two volatility functions is shown. As the $HW-1F$ and $HW-2F$ model, which are very similar under the Gaussian formulations, are calibrated to the same market instruments, it is logical that the calibrated short rate diffusions must be more or less the same, as it appears in Figure (4.3). Note that the calibrated correlation is by far different from being one.

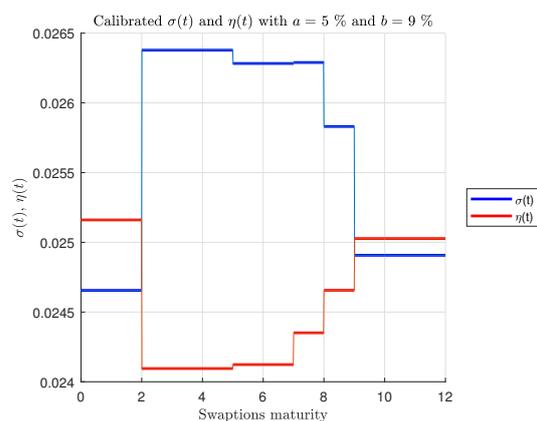


Figure 3.28: Calibrated $\sigma(t)$ and $\eta(t)$ where $\rho = -97\%$.

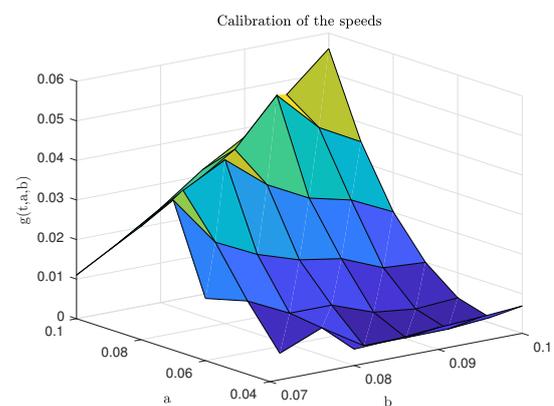


Figure 3.29: Normalized squared pricing error $g(t, a, b)$ over the whole swaption surface as a function of the speeds.

3.3.6. Simulation

In what follows the simulation results using the two factors model are presented. What has been done for the one-factor model was repeated for the two-factors model as well. Also in this case it

turns out that the exact simulation, on the line of paper [25], is faster and therefore preferable. In what follows we briefly resume how the two schema have been applied, and afterwards the result obtained by means of the exact simulation are provided.

Exact simulation schema

For the short rate $r(t)$, the dynamics and distribution have been already introduced in this section. We have seen that the distribution of r_t conditionally on \mathcal{F}_s is normal with mean and variance given by (3.30). To simulate r_t at the reporting dates $0 = t_0 < t_1 < \dots < t_n$, we can use the following recursions:

$$\begin{aligned} x_{i+1} &= x_i e^{-a\Delta_i} + \left(\sqrt{\int_{t_i}^{t_{i+1}} e^{-2a(t_{i+1}-u)} \sigma(u)^2 du} \right) Z_{i+1}^1, \\ y_{i+1} &= y_i e^{-b\Delta_i} + \left(\sqrt{\int_{t_i}^{t_{i+1}} e^{-2b(t_{i+1}-u)} \eta(u)^2 du} \right) Z_{i+1}^2, \\ r_{i+1} &= x_{i+1} + y_{i+1} + \beta(t), \end{aligned}$$

where Z^1 and Z^2 are ρ -correlated standard normal variables. Note that equivalently, in the case the parameters are constant, the two factors need to be diffused as

$$\begin{aligned} x_{i+1} &= x_i e^{-a\Delta_i} + \sigma \sqrt{\frac{1}{2a}(1 - e^{-2a\Delta_i})} Z_{i+1}^1, \\ y_{i+1} &= y_i e^{-b\Delta_i} + \eta \sqrt{\frac{1}{2b}(1 - e^{-2b\Delta_i})} Z_{i+1}^2, \\ r_{i+1} &= x_{i+1} + y_{i+1} + \beta(t). \end{aligned}$$

Euler simulation schema

Equivalently the short rate could be simulated using the following discretization schema:

$$\begin{aligned} x_{i+1} &= x_i + a(-x_i)\Delta + \sigma(t)\sqrt{\Delta}Z_{i+1}^1, \\ y_{i+1} &= y_i + b(-y_i)\Delta + \eta(t)\sqrt{\Delta}Z_{i+1}^2, \\ r_{i+1} &= x_{i+1} + y_{i+1} + \beta(t), \end{aligned}$$

where Z^1, Z^2 are correlated standard normals.

Simulation results

The simulation results under the calibrated $HW - 2F$ model are presented here. First of all, the short rate quantiles are shown in Figure (3.30). On the other hand, the two different previews of the yield curves quantiles are provided by Figures (3.31) and (3.32) respectively. Note that the same countercheck done for the one-factor can be applied in this case. This is actually linked to the Affine Term Structure of both models. That is, when $\tau = T - t$ is taken small enough, e.g. $\tau = 1$ day (red quantiles), the yield $R(t, T)$ reads:

$$\begin{aligned} R(t, T) &= -\frac{\log P(t, T)}{T - t} \\ &= -\frac{1}{\tau} [\log A(t, t + \tau) - B(a, t, t + \tau)x_t + B(b, t, t + \tau)y_t] \\ &= -\frac{1}{\tau} \left[\log A(t, t + \tau) - \frac{1 - e^{-a\tau}}{a} x_t - \frac{1 - e^{-b\tau}}{a} y_t \right] \\ &\approx \frac{1}{\tau} [\tau x_t + \tau y_t] + Y'(t) \\ &\approx r_t + Y(t), \end{aligned}$$

where the function $Y(\cdot)$ refers to a time deterministic shift. This translates in a shape which is very similar to the short rate diffusion itself. Whereas, for bigger maturities the shape gets narrower as the factor multiplying r_t tends to 0, and therefore while the average value of the simulated yields stays approximately the same, the variance of $R(t, T)$ unavoidably decreases. This can be appreciated when considering the quantiles for $\tau = 30$ years, represented by the black profiles. Note that in this case this narrowing effect is less accentuated, compared to Figure (3.21). The reason simply lies in the significantly lower speeds that have been calibrated. To conclude, the simulated 3D term structures up to 35 years are shown in Figure (3.33). More flexibility of movements is visible in comparison with Figure (3.22).

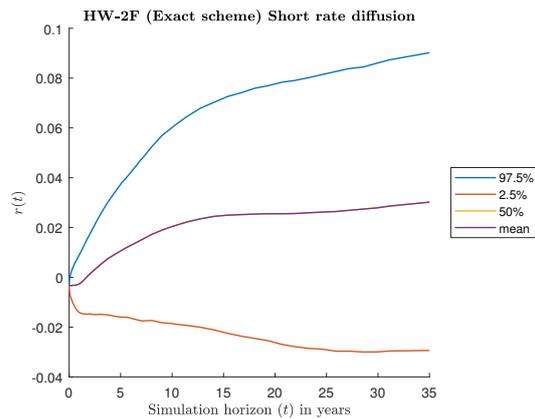


Figure 3.30: Short rate diffusion with $a = 17\%$.

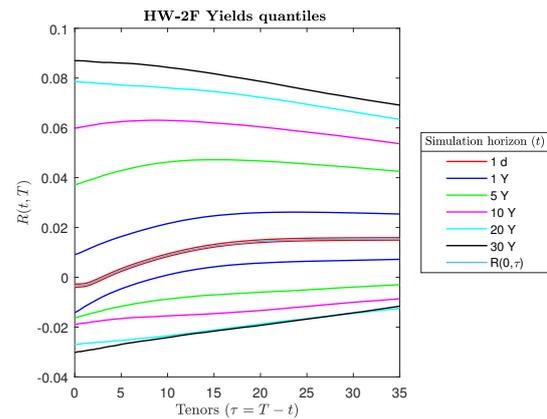


Figure 3.31: Simulated yields quantiles t years ahead.

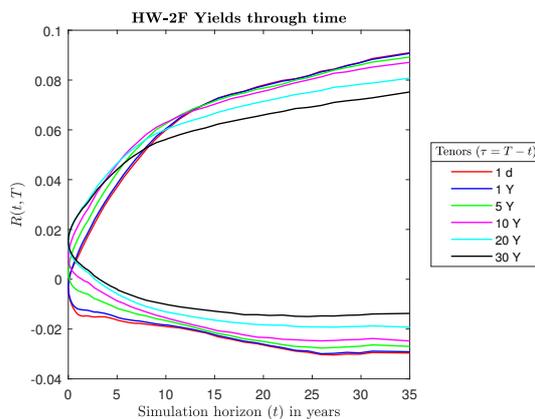


Figure 3.32: Simulated yields quantiles as a function of time t .

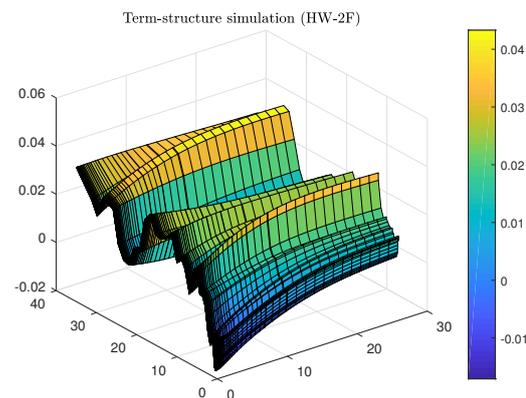


Figure 3.33: Simulated term structures under $HW - 2F$ model in 3D.

3.3.7. Imperfect correlation

In Section (3.2.7) we examined the well-known perfect correlation characteristic of the one factor Hull-White model, which holds for any simulation time as far as we like. That is, for each t we know that the following holds:

$$\text{Corr}(R(t, T_1), R(t, T_2)) = 1.$$

On the other hand, the $HW - 2F$ model performs better in these terms as it allows correlations between stochastic factors to take (theoretically) any value between $[-1, 1]$. This translates in less restrictive shapes of the predicted yield curves, which happen also to intersect for a given t . In this section we show that the asymptotic behaviour of the two factors model is completely opposite to

the one of DEV-MR model. In fact, it approaches perfect correlation when far grid points from now are concerned.

Empirical correlation between yields

In Figure (3.34) we plot the correlations for different simulation times (t days/years ahead) between $R(t, t + \tau_1)$ ($\tau_1 = 1$ day) and $R(t, t + \tau_i)$, where the τ_i 's are tenors up to 35 years. Apparently for

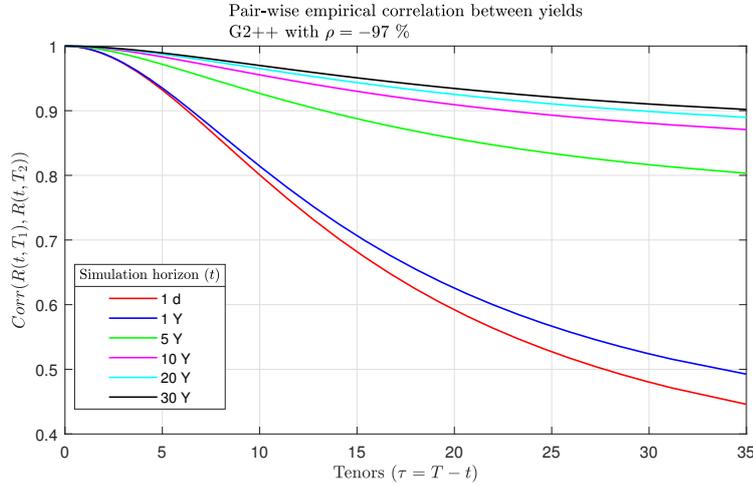


Figure 3.34: Empirical correlations between yields.

large simulation times the correlation tends to 1 and somehow we approach the same behaviour of $HW - 1F$. Whereas for small simulation times the predicted yields are far from being aligned on a line.

Theoretical correlation between yields

We here derive the theoretical formula for the computation of the correlation between yield curves implied under the $HW - 2F$ model. This served as a counter check for the results of Figure (3.34) and to motivate their asymptotic behaviour. Furthermore, we conclude that this behaviour is recurrent for any choice of the model parameters.

We know that under the risk-neutral measure \mathbb{Q} the two processes $x(t)$ and $y(t)$ are as in (3.33). Therefore the following results hold:

$$\begin{aligned} \mathbb{E}[x(t)|\mathcal{F}_s] &= x(s)e^{-a(t-s)}, \\ \mathbb{E}[y(t)|\mathcal{F}_s] &= y(s)e^{-b(t-s)}, \\ \text{Var}[x(t)|\mathcal{F}_s] &= \int_s^t \sigma(u)^2 e^{-2a(t-u)} du, \\ \text{Var}[y(t)|\mathcal{F}_s] &= \int_s^t \sigma(u)^2 e^{-2b(t-u)} du, \\ \text{Cov}[x(t), y(t)|\mathcal{F}_s] &= \rho \int_s^t \sigma(u)\eta(u)e^{-(a+b)(t-u)} du, \end{aligned}$$

where the covariance computation between the two processes $x(t)$ and $y(t)$ easily follows by applying a particular case of the Ito's Isometry. Let now the subscript s stands for "conditionally on \mathcal{F}_s " and recall the bond price formula under the $G2++$ model given in (3.38). If we introduce the shorter

notation:

$$b^x(T) = \frac{B(a, t, T)}{T - t},$$

$$b^y(T) = \frac{B(b, t, T)}{T - t},$$

the correlation between yields in the two factors model reads:

$$\begin{aligned} \text{Corr}_s[R(t, T_1), R(t, T_2)] &= \text{Corr}_s[b^x(T_1)x_t + b^y(T_1)y_t, b^x(T_2)x_t + b^y(T_2)y_t] \\ &= \frac{\text{Cov}_s[b^x(T_1)x_t + b^y(T_1)y_t, b^x(T_2)x_t + b^y(T_2)y_t]}{\sqrt{\text{Var}_s[b^x(T_1)x_t + b^y(T_1)y_t]} \sqrt{\text{Var}_s[b^x(T_2)x_t + b^y(T_2)y_t]}} \\ &= \frac{b^x(T_1)b^x(T_2)\text{Var}_s[x_t] + b^y(T_1)b^y(T_2)\text{Var}_s[y_t] + [b^x(T_1)b^y(T_2) + b^x(T_2)b^y(T_1)]\text{Cov}_s[x_t, y_t]}{\sqrt{\text{Var}_s[b^x(T_1)x_t + b^y(T_1)y_t]} \sqrt{\text{Var}_s[b^x(T_2)x_t + b^y(T_2)y_t]}}, \end{aligned} \quad (3.45)$$

where the denominator can be expanded as

$$\text{Var}_s[b^x(T)x_t + b^y(T)y_t] = b^x(T)^2\text{Var}_s[x_t] + b^y(T)^2\text{Var}_s[y_t] + 2b^x(T)b^y(T)\text{Cov}_s[x_t, y_t].$$

This ugly formula worked as a countercheck for Figure in (3.34). Precisely, from this we could understand that $\forall T > 0$ the correlation between rates is increasing in t and approaches (but never touches) 1. The parameters values affects the speed at which this happens, but not in a way that the correlations after 20-30 years do not get close to 1. Equivalently we can say that the two Hull-White models behaves similarly in terms of correlations when $t \rightarrow \infty$.

Twisted shapes

As previously mentioned, one of the advantages of using a two factors model is that this can better replicate the yield curves observed in the market. The fact that there is a correlation coefficient ρ between the two Brownian motions defining the short rate process, leads to the possibility of a wider range of scenarios in the predicted shapes of the yields. This is appreciable from Figures (3.35) and (3.36). Yields can intersect under this model as a consequence of the fact that unexpected shock are not proportionally distributed over the whole yield curve anymore.

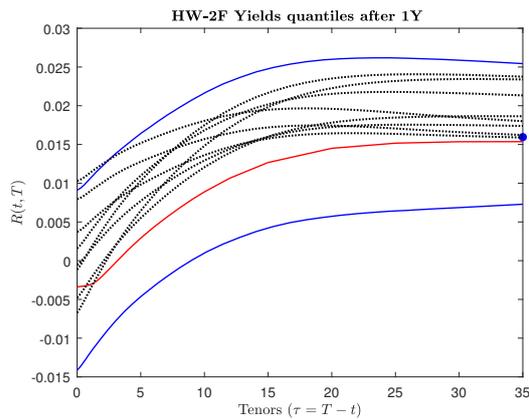


Figure 3.35: Shapes implied by the HW-2F after 1 year simulation together with quantiles and yield on infinitely lived bond (blue dot).

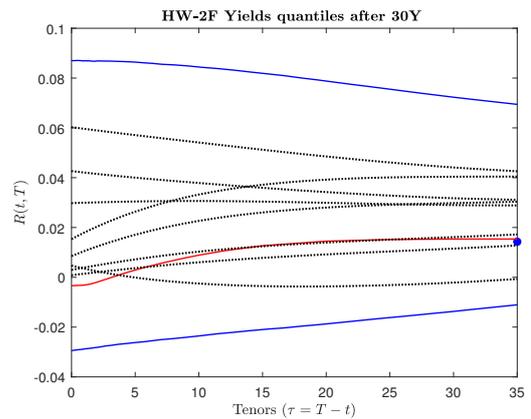


Figure 3.36: Shapes implied by the HW-2F after 30 years simulation together with quantiles and yield on infinitely lived bond (blue dot).

3.3.8. Discussion

In light of the 2007-2008 credit crisis, simple and transparent approaches are usually preferred to more complex ones. For instance, it is always preferable to maintain the number of model parameters as lower as possible, so to make sensitivity analyses of the parameters easier. In this terms, the current model and the $HW - 1F$ win over the two factors model. On the other side, more parameters could explain better the market behaviour and give more reliable estimates of the exposures. Until now, we have examined the technicalities and implications of the three models in terms of the calibration, simulation and ability in capturing the correlation between rates. We have seen that the calibration of the $HW - 2F$ model is the one requiring more effort both in terms of implementation and computational difficulty. On the other side, even though we did not write every steps of it, the historical calibration has proven itself to be the most easy to implement and understand, as it does not require much more than a properly done data cleaning. The simulation does not consist of a problem in neither of the three models while the $HW - 2F$ models the best the shape and values of yield curves. At the same time, from a correlation point of view, the two factors model behaves like a one factor model for long simulation horizon. Up to know, no preference can be shaped yet but it comes natural to be suspicious about the correlations modelled under the current model. More understanding can only be achieved by testing the models performance in a more practical application, and this is done in the next chapter.

4

Portfolio impact and stability analysis

From the risk-management side, a model is desired when it reveals to give stable credit exposure over time. Counterparty credit risk estimation of a bank is related to the exposure of individual trades, whose quantification has a direct impact on the capital that a bank is required to hold according to the regulator. If one relies on the arbitrage-free pricing theory for the computation of the risk profiles, then the problem of fluctuations in market prices needs to be faced. We reckon that it would be very inconvenient if this new approach led to high variation in capital requirements. In this section we are going to perform an analysis on a portfolio made of several types of vanilla swaps, which we reckon to be a good approximation considered that the bank is mainly exposed to these.

The analysis is done both on a risk-factor and portfolio level. Note that we can only limit ourselves to point out the consequences and differences that each of the three models brings, together with positive and negative aspects. The reason is that no final conclusion about which one is better can be drawn until all models are backtested with real data. The current model will be used as a benchmark so to understand points of strength and weakness of the two short rate models, as until now it has proven to perform reasonably well on real data. Anyhow, this is done with the awareness that this might still be not the optimal choice. In fact, there are still open issues about how to tackle issues that apply to the model, which vary from the calibration procedure to instances such as the value of the shift that should be applied. Furthermore, the model may reveal to underestimate risk, as shown later in the text. This may be either due to the fact that any risk premium embedded in market-implied volatilities leads to a conservative overestimate of risk, or simply because of the DEV-MR correlation problem described in (3.1.4).

As a first step, we will provide the implied profiles under the three different models for a representative portfolio of vanilla interest rate swaps (payers and receivers). More precisely, we will consider three different portfolios with three different counterparties (namely A, B, C) with which the bank has only EUR exposure, driven by the fact that the bank is mostly exposed to this currency. In this section we comment on the results for the most diverse counterparty (namely, Counterparty C) while the results for the other two can be found in Section (C.2). We assume that no CSA applies and only netting agreements will be considered. For simplicity we consider the whole portfolio as a unique netting set, in all three cases. The three relevant risk measures, already introduced in (1.1.3), will be provided, interpreted and commented for the counterparty with the higher number of trades. The second step regards the evaluation of the stability of short rate models on a risk-factor level and a portfolio level. A bank may not want to have a model that tells her to keep a certain amount of capital today, and double that amount for the day after. We will analyze the Potential Future Expo-

sure profiles, given that the volatility parameters are calibrated for each different date under study and for both the Hull White models. A time window of one year will be considered.

4.1. Model and calibration impact on representative portfolios

The pricing of a basic Interest Rate Swap can be done without assumptions about the underlying model. Precisely, we can price a swap at $t = 0$ by simply using the interest rate curve at $t = 0$. When, as in our case, the aim is to evaluate the future Mark-to-Market values of a swap (or a mixed portfolio of swaps), then the simulated yield curves under a specific model are needed. These future MtM's values affects banks' capital calculations and make up most of the credit exposure to its counterparties (assumed that the financial institution trades swaps in big amounts). Before dealing with real data and several swaps together, we here provide the simplest example: future MtM values (in terms of quantiles and means) of a 3Y payer Interest Rate Swap (ATM). We assume that the swaps starts today and both fixed and floating payments occur at the same date, with frequency of half a year.

4.1.1. Revaluation of a plain vanilla swap

Once the model (in this case the one factor model) has been calibrated to a specific date $t_0 = 0$, the short rate model is simulated starting from that date. Consider that we want to simulate N (2000) scenarios and our grid of (non uniform) reporting dates has length L (84 dates between today and the next 35 years). We then end up with a table \mathcal{R} of size $N \times L$ of simulated short rates which can be easily converted into bond curves by means of the formula $P(t, T) = \zeta(t, T, r(t))$, where $\zeta(\cdot)$ the Zero-Coupon Bond pricing formula under the model, that can either be (3.13) or (3.38). That is, given the set of N_τ tenor lengths τ we are interested to, for each column of \mathcal{R} (therefore for each t) we can compute the table $\mathcal{P}(t, t + \tau)$ of size $N \times N_\tau$, which contains the N simulated bond curves for a future time t at maturity dates $t + \tau$. At this point everything is known in order to use Formula (2.5), which only requires the simulated bond curves at t and the parameters defining the swaps. Note that such formula needs to be updated as soon as coupons are paid. More precisely, to compute the valuation for $t \in [0, 0.5)$ no changes should apply. Now, let $\{t_1, \dots, t_6\}$ be the set of payment dates, where $t_0 = 0$ and $t_6 = 3$ (assuming that the first reset date is today, and the first payment in half a year). When it comes to value the swap for $t \in [t_{i-1}, t_i)$, $0 < t_{i-1} < t_i \leq 3$, we need to account for the payments that have already occurred by removing the relative terms from the sum of discounted cashflows. This reflects in the typical peaked shape provided in Figure (4.1). Note that a swap with a longer maturity would have much more risk due to both a longer lifetime and more coupon payments to be exchanged. Also, note that the quantile shapes always follow the same pattern: they increase, reach a peak and then start to decrease again. This is due to the amortization of the portfolio: as time passes more coupons are paid and the bank is less exposed to the counterparty. More specifically, there is a trade off between randomness and residual time-to-maturity. At the beginning the risk-factor randomness prevails over the residual time, while after a certain time (that depends on the interest rate model being used) the valuation sensitivity to the discount factors decreases. Of course it can be easily expected the fact that, no matter how stable or unstable the risk-neutral calibration will be, the biggest variation will be found in proximity of the exposure peak.

Note that in the DEV-MR model the process for computing the future values is analogous. We only need to account for the different simulation of the yield curves. As the model is calibrated historically, only the profile relative to one day will be displayed. Calibration to any other (enough close) day would not impact the profile so much as it happens with risk-neutral calibration. In fact, historical calibration is in general quite stable as it is done on the last 8 years (30 years) of data for the volatility (mean reversion parameters).

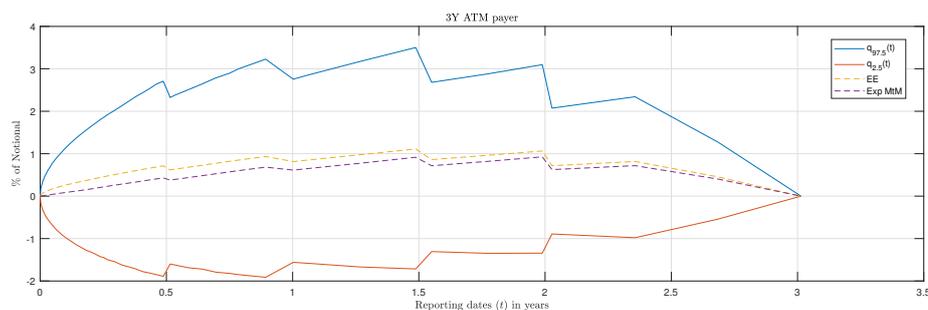


Figure 4.1: Future MtM quantiles of a payer 3Y-IRS (1F model) with calibration performed on 23/01/18.

4.1.2. A real application

A portfolio needs to be chosen and should be able to best replicate the real CCR portfolio. We will exploit the fact that Interest Rate Swaps constitutes the majority of traded instruments, and therefore an analysis on a portfolio made of only vanilla swaps should be sufficient in order to draw reliable conclusions.

Now that we made clear how to deal with one single trade, more complete analyses on combined trades can be done. Consider analyzing a single-currency (EUR) portfolio of vanilla swaps with one specific counterparty. We want to predict the future MtM values of such portfolio in order to quantify the money loss we would face in case the counterparty defaulted. This way we can understand how much capital we should keep apart in order to cover this possible loss.

The portfolios we will deal with are made of payer and receiver swaps with residual life time between 1 and 25 years. That is, such portfolios contain trades that started some date in the past and will end some date in the future. Usually, a swap is traded on par, that means that the fixed coupon rate is chosen so that the initial value of the swap is zero. Of course each swap has an history due to changes in interest rates, and by the time we start our valuation each of these will have a specific market value: either positive, consisting in a credit exposure for us, either negative, consisting in no credit exposure. So will do the netted sum. In the valuation, we will consider the case in which the financial institution and the counterparty agreed for netting. Furthermore, we assume that by the time we start the valuation, we are not in the middle of two payment dates. That is, we stand at some $t = T_s$ belonging to the set of payment dates of the whole portfolio. Note that the mean reversion speeds that will be used for the valuation are those calibrated in Chapter (3), that are $a_{1F} = 1.5\%$, $a_{2F} = 5\%$ and $b_{2F} = 9\%$. These parameters usually do not show much variation over time and it is advised to update them less frequently than the volatilities, under the risk-neutral calibration. Therefore, the speeds will be considered fixed during all the stability analysis, and therefore only the volatilities calibration will be updated for each day. This choice is also motivated by a small experiment conducted in Appendix (B).

Portfolios information

Counterparty A and B are very similar to each other, with the main differences that the bank has a higher notional with the first one while the second one is three times bigger in number of deals. These are both made of swaps (receiver type) that have relatively short time-to-maturities (between 1 and 6 years). On the other side there is Counterparty C with which the bank has the highest notional and number of trades (among the three considered). With this counterparty, deals time-to-maturities are much higher, ranging even up to 25 years. Fixed and floating payments have quarterly frequency for all three portfolios.

Aggregation

In the results to follow, each trade has been aggregated to the other as if the netting set was only one: the whole portfolio with the counterparty. In general, this agreement is used so to mitigate counterparty risk. Once the trades have all been revaluated, each of them will be represented by a matrix of values with respect to scenario number and point on the time grid. If we call $V_{i,j,k}$ the MtM value of trade i at time point j for scenario k and P the portfolio with the counterparty, then the exposure for this netting set is calculated as:

$$E_{j,k} = \left(\sum_{i \in P} V_{i,j,k} \right)^+.$$

After the revaluation and aggregation have both been completed, we have extracted the metrics introduced in (1.1.3). In what follows we report and comment on the result for the most miscellaneous portfolio, out of the three. The results for the other two portfolios are provided in Appendix (C).

Portfolio impact with Counterparty C

Graphics results are provided in Figure (4.2), for all three models. Note that each of the plots on the left refers to those on the right. This portfolio is made of 44 swaps, out of which 35 are receivers and 9 are payers. More precisely, the payer swaps retain almost 60% of the portfolio's total notional N , which amounts to 154 millions euros. Residual life times of each swaps within the portfolio varies from 1 to 25 years, which explains the profiles going to zero at $t = 25$ for all three models. In the first column of plots the Mark-to-Market simulation quantiles $q_{97.5\%}(t)$ and $q_{2.5\%}(t)$ are shown in respectively blue and red lines. Whereas, expected value of the MtM simulation and Expected Exposure are respectively shown with the purple and yellow lines. Note that the EE is always higher as it is defined as the average of the non-negative values of the MtM. The second column of plots shows the EE together with the PFE profile. Note that the PFE profiles in general results to be higher than the 97.5% quantiles of MtM values. In fact, it is the quantile calculated over the exposure simulated values, which are by definition the positive part of the MtMs. These are the two most important risk-measures that can be calculated per simulation horizon t and should retain the risks related to the portfolio. Note that for every model the risk is increasing at the beginning and then decreases (starting from different peaks) towards zero (where all coupons have been payed and each swap within the portfolio has expired). Together with this, the typical behaviour of increasing and then decreasing exposure for Interest Rate Swaps is present also here, with the peculiarity that the maximum exposure point is reached in a relatively short term, due to the initial MtM being positive and far away from zero. From this point of view, the DEV-MR model has the closest exposure peak (after around two years), whereas the Hull White models has their peaks around 5 years.

It is interesting noticing the difference in the profiles between Counterparty C and Counterparties A and B (which have similar portfolio structures). More specifically, we reckon the differences find their roots mostly in the residual life time of the swaps that contribute to each portfolio. The total revaluation of portfolio C is expected to go to zero after a relatively long time horizon t (approximately 25 Y), as the highest time-to-maturity is approximately 25 Y. This is well appreciable in (4.2). While the two Hull-White models seem to perform relatively the same, the DEV-MR clearly exhibits a big difference in shape profiles, underestimating very much the risk starting already from $t = 3$ years. A possible reason for this could be that the correlation structure between rates is lost in a relatively short term. More precisely, we know that the nine points of the yield curve that are modelled in the DEV-MR lose their correlation as the simulation time t increases. This has been explained in (3.1.4), and could translate in a fast decaying profile and therefore lower risk, compared to other two models. Anyhow, to draw a safe conclusion about this it would definitely be significant to make

this comparison when deriving the exposure profiles in the DEV-MR model under risk-neutral calibration (using it as a model for the short rate). This is not visible in terms of profile shapes when valuing portfolio A and B , as shown in Appendix (C), since those are built up with short term swaps. Still, the magnitude in PFE and EE is significantly lower.

As the two Hull-White models both assume a normal distribution for the rate, they both give much more probability to negative rates compared to the DEV-MR model and this translates in a wider range for the lower MtM quantiles. This happens for all the three counterparties and it is visible the most for counterparties A and B . In Table (4.1) the scalar risk measures EPE, EEPE and EaD introduced in Subsection (1.1.3) have been provided. These are measures that are relative to the whole life time of the portfolio. At first, we would have expected much higher values for the risk measures under the two factors model, mainly driven by the fact that a one-factor approach is restrictive in the possible yield curve movements and hence is known to “miss” some of the risk. We reckon the reason for this mainly relies on the higher mean reversion speeds a and b that have been calibrated for the two factors model. In fact, these high speeds bring the short rate simulations under the one and two factors to be very much comparable in terms of quantiles range, as visible in the first plot of Figure (4.3). In addition to this, as seen with Formula (3.45), the implied correlation between yields is not affected only by ρ , but also a and b . As a consequence, similar risk-profiles are expected.

In Table (4.1) the EaD under $HW - 2F$ is higher than the EaD under $HW - 1F$, but new runs revealed that it can also be the other way around, and therefore we can consider the two Hull-White models to give more or less the same information. As a matter of fact, the Exposure at Default is a stochastic quantity. Again, conclusion to this fact is that a two factors model, calibrated under the risk-neutral measure, does not provide significant differences from a one factor model in terms of risk profiles. We reckon this problem not to belong to the model itself (as by definition we know a second factor to be more descriptive for the correlations of rates) but rather to the kind of calibration that has been applied to the model. In general, parameters that are calibrated to market prices do not have most of the time economical interpretations, as they are merely the result of an optimization problem. In our case, we result with a very low value of mean reversion speed in the one factor model, while higher values belong to the two factors model, and this bring the two short rate diffusion to be very much comparable. If, for instance, the $HW - 1F$ calibrated speed resulted in a higher value, and this value were comparable to the two factors model speeds (in magnitude, let us say such that $a_{1F} = a_{2F} + b_{2F}$), in such a case running again the revaluation reveals that we would “benefit” of the advantage of adding a second factor to the interest rate model. In the second plot of Figure (4.3) we show such an hypothetical short rate diffusion. In this case we ended up with higher profiles in the $HW - 1F$ for counterparties A and B , while higher profiles resulted under the $HW - 2F$ model when longer time-to-maturities are concerned (Counterparty C).

This said, we can conclude that when a risk-neutral calibration is applied, the risk is for the calibrated parameters not be meaningful in their values. On the other hand the \mathbb{P} measure parameters reflect the actual, long-term behaviour of the rate and better reveal the nature of the rate. They are the description of what has happened in the past and therefore most easily admit a business interpretation. By saying this, we mean that they are more generalizable, e.g. you might deduct from historical parameters that USD is more volatile than EUR. These general conclusions are usually important for business. Under \mathbb{Q} this statement might be true today but false tomorrow; so basically under \mathbb{Q} it is very difficult to provide general statements. As a conclusion, we trust that a properly done historical calibration for both models would have suggested different speeds and risk-profiles more in line with our expectations.

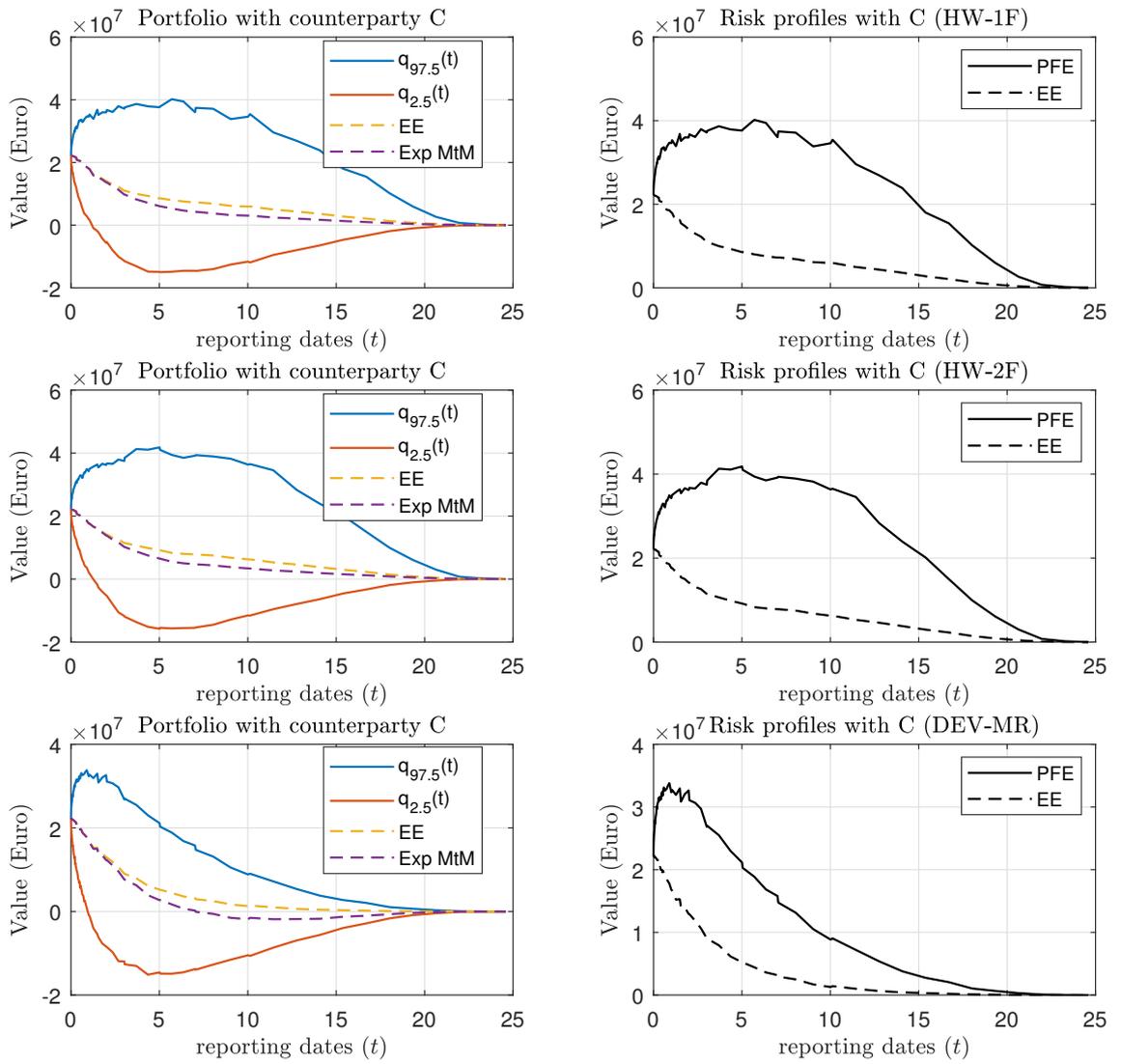


Figure 4.2: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty C implied by the DEV-MR, HW-1F and HW-2F models.

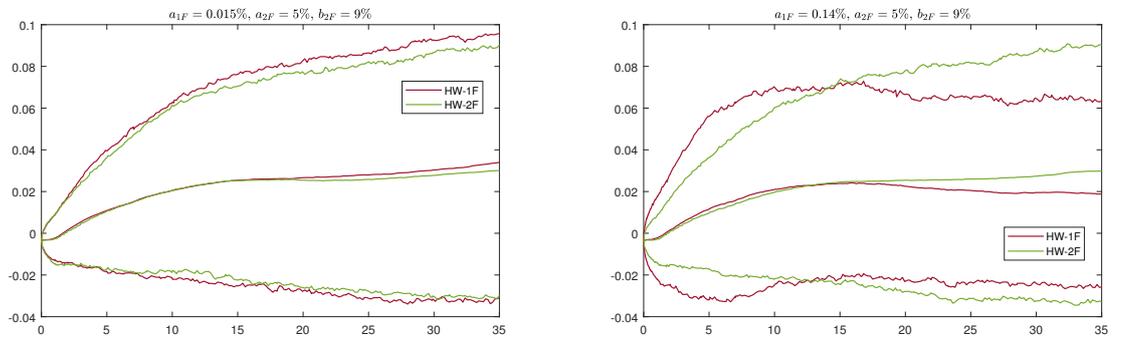


Figure 4.3: Simulated short rate quantiles with actual values of calibrated mean reversion speeds (left panel) and with more comparable values of the mean reversion speeds (right panel).

Counterparty		DEV-MR	1F	2F
A	EPE	197,252.10	261,466.41	253,299.49
	EEPE	210,632.86	280,118.93	284,445.78
	EaD	294,886.00	392,166.50	398,224.09
B	EPE	177,743.96	191,502.03	188,697.89
	EEPE	194,820.65	225,154.41	247,133.82
	EaD	272,748.91	315,216.17	345,987.35
C	EPE	13,254,239.38	14,309,114.78	14,377,411.96
	EEPE	15,517,237.92	15,779,774.75	17,359,739.31
	EaD	21,724,133.09	22,091,684.64	24,303,635.03

Table 4.1: Expected Positive Exposure (EPE), Effective Expected Positive Exposure (EEPE), Exposure at Default (EaD) for the three counterparties A, B, C under the three different models.

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	12Y	15Y	20Y	25Y	30Y
1M	NaN	NaN	314.44	104.79	73.25	48.37	33.98	30.27	27.23	25.57	25.64	26.30
3M	NaN	NaN	196.03	95.00	70.24	48.98	35.90	32.15	29.00	27.06	26.75	27.00
6M	NaN	NaN	138.44	86.84	68.39	50.08	37.78	34.05	30.91	28.81	28.29	28.38
9M	NaN	374.68	108.90	78.57	65.11	49.69	38.72	35.13	32.06	30.01	29.52	29.68
1Y	NaN	171.91	93.13	72.30	61.86	48.68	38.95	35.63	32.75	30.80	30.38	30.54
2Y	104.36	77.59	65.88	58.48	52.34	44.40	37.92	35.42	33.31	31.94	31.64	31.75
3Y	69.89	61.62	55.58	51.16	47.11	41.41	36.89	34.60	32.98	32.12	31.91	32.13
4Y	58.86	54.15	49.78	46.32	43.13	39.08	35.71	34.05	32.37	31.82	31.68	31.95
5Y	52.31	48.58	45.26	42.53	40.16	37.24	34.56	33.50	31.79	31.50	31.43	31.65
6Y	46.40	43.83	41.18	39.15	37.33	35.56	33.78	32.70	31.24	30.99	30.95	31.21
7Y	41.86	39.91	38.04	36.62	35.33	34.10	33.16	32.08	30.93	30.73	30.67	30.90
8Y	37.97	36.87	35.56	34.56	33.99	32.72	32.57	31.48	30.66	30.52	30.45	30.61
9Y	35.34	34.74	33.89	33.50	32.86	32.45	32.16	31.43	30.57	30.51	30.41	30.51
10Y	33.43	33.28	33.16	32.52	31.83	32.29	31.89	31.53	30.60	30.53	30.36	30.44
12Y	32.65	32.21	31.31	31.84	32.00	31.91	32.22	31.69	30.97	30.66	30.36	30.50
15Y	33.92	33.88	33.24	32.68	32.22	32.98	33.23	32.91	32.05	31.23	30.63	30.87

Table 4.2: Implied volatilities under non-shifted Black's model expressed in percentages. Source Bloomberg.

4.1.3. Historical volatility against Black's volatilities

One could doubt the reason we gave for the Counterparty Credit Exposure modelled in the Hull-White framework being higher than in the DEV-MR model. We said that this is due to the inability of the DEV-MR model of modelling correctly the correlation structure after some time. At the same time, one could rebut this and argue that it may simply be due to the fact that swaptions quoted volatilities are generally higher than the volatility that is observed historically, due to some risks premium embedded in the derivatives. This seems not to be the reason though. In fact, we conducted a small analysis and reached the conclusion that the opposite is true. The volatility calibrated historically under the DEV-MR model equals 19.98 %. We retrieved the log-normal volatilities on Bloomberg (which we report in Table (4.2)) and made sure that these were quoted with no shift. After this, we shifted the rates and translated the volatility accordingly with Black's formula. This results in the new swaption volatility surface of Table (4.3). This looks flatter and values are in general lower than the historical volatility (around 15%).

	1Y	2Y	3Y	4Y	5Y	7Y	10Y	12Y	15Y	20Y	25Y	30Y
1M	NaN	NaN	NaN	7.81	9.88	10.74	10.47	10.47	10.44	10.57	10.87	11.24
3M	NaN	NaN	3.89	8.94	10.62	11.47	11.36	11.33	11.26	11.27	11.40	11.59
6M	NaN	NaN	7.52	10.65	11.97	12.61	12.41	12.33	12.22	12.13	12.15	12.26
9M	NaN	2.86	9.66	11.81	12.89	13.37	13.18	13.05	12.89	12.77	12.78	12.91
1Y	NaN	8.12	11.34	12.80	13.62	13.90	13.69	13.56	13.38	13.24	13.25	13.36
2Y	12.58	14.15	14.99	15.40	15.34	15.16	14.80	14.58	14.37	14.20	14.15	14.18
3Y	16.10	16.46	16.52	16.51	16.28	15.84	15.46	15.05	14.78	14.63	14.52	14.56
4Y	17.41	17.45	17.18	16.96	16.58	16.14	15.74	15.38	14.91	14.74	14.60	14.63
5Y	17.92	17.66	17.34	17.03	16.65	16.26	15.80	15.56	14.93	14.77	14.60	14.59
6Y	17.63	17.47	17.08	16.75	16.39	16.19	15.86	15.50	14.87	14.64	14.46	14.44
7Y	17.31	17.08	16.73	16.47	16.18	16.02	15.86	15.41	14.85	14.58	14.37	14.32
8Y	16.71	16.59	16.32	16.11	16.04	15.71	15.76	15.25	14.77	14.48	14.26	14.17
9Y	16.20	16.19	16.02	16.00	15.83	15.77	15.65	15.26	14.72	14.44	14.20	14.07
10Y	15.80	15.91	15.99	15.80	15.55	15.79	15.54	15.27	14.68	14.38	14.10	13.96
12Y	15.93	15.79	15.41	15.65	15.70	15.57	15.52	15.14	14.62	14.21	13.89	13.75
15Y	16.39	16.32	15.98	15.67	15.39	15.55	15.39	15.10	14.55	14.00	13.56	13.44

Table 4.3: Implied volatilities under shifted Black's model expressed in percentages. These are calculated by bumping up the interest rates by $\gamma = 200$ basis points.

4.2. Stability of risk-neutral calibration

This analysis is addressed in two steps. Firstly we have looked at how the fluctuations in market prices impact the simulation of the short rate (and therefore the yield curve). Secondly we have looked at how these changes impact the portfolios and therefore the final exposures. It turns out that the volatility changes highly impact the changes in the risk profiles from one day to the other. That is, fluctuations in the market that are of the order of basis points translate into fluctuations of the realized values of Exposure at Default of the order of millions of Euros.

4.2.1. Time-horizon and new data for calibration

For this purpose, a time-horizon on which to perform a stability analysis needs to be chosen. Ideally the best choice would be to perform the calibrations on the longest possible time window. The adequacy of the recalibration frequency instead depends on the type of the calibration method. For instance, a quarterly historical calibration and a monthly market implied calibration may end up having comparable performance in reflecting changes in market conditions. We reckon that a risk-neutral calibration over ten different dates within one year horizon is enough to draw some conclusions. To be precise, we considered the time window from 30 June 2017 till 30 March 2018 and calibrated each model at the last business day of each month. In order to do so, bond curves and swaptions quotes from these ten days have been retrieved from Bloomberg. We provide the discount factors and volatility quotes in Figures (4.4) and (4.5) respectively. The final goal would be to understand how much sensitive are the risk profiles with respect to changes of these two sources.

4.2.2. Risk-factor level

Both Hull-White models have been calibrated to the same At-The-Money swaptions, whose Normal volatilities are identified by the black lines in (4.5).

HW – 1F model

In Figure (4.6) are displayed the calibrated function steps for $\sigma(t)$ in term of boxplots. Note that the mean reversion speed has been calibrated once in (3.2.5), and we keep this fixed for each of the ten dates. The jump from the first maturity to the second one is always recurring. This fact can

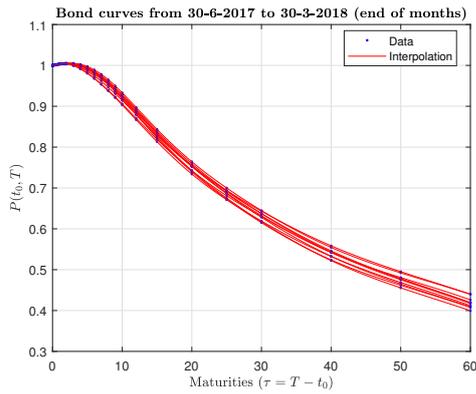


Figure 4.4: Bond curves from the set of calibration dates.

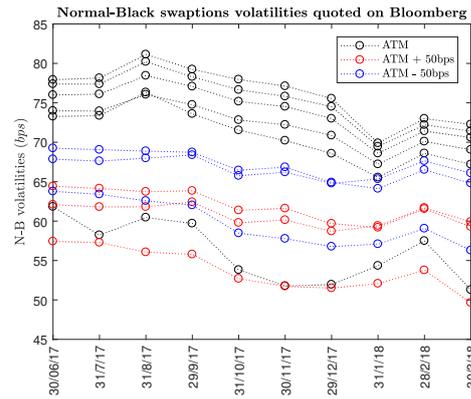


Figure 4.5: Quoted volatilities of swaptions used for calibration.

be motivated by looking at the history of EUR volatility cubes from the last year(s), where a steep shape for short maturities and tenors is exhibited. This fact is observable in (4.5), with the lowest line corresponding to the swaption with shortest maturity.

With regard to the simulation, Figure (4.7) displays the quantiles of $r(t)$ over the time horizon of 10 months. These are computed as quantiles of quantiles: for example, the two upper black lines are calculated as the upper and lower quantiles of the short rate upper quantiles realized over the ten different dates. This give us information about how the yields (and, at a later step, the valuations of vanilla swaps) are affected by the fluctuations in the market. It is undeniable that the short rates exhibits high variation for large horizons, where the upper quantiles range is higher than 2%. This constitutes a very high difference in terms of interest rates. Note that the stability of short rates and yields is very much affected by the value of the mean reversion speed. A same analysis, conducted when $a = 14\%$, reveals that such range lowers to 1% which translates in much more stable rates, especially for those related to the highest tenors.

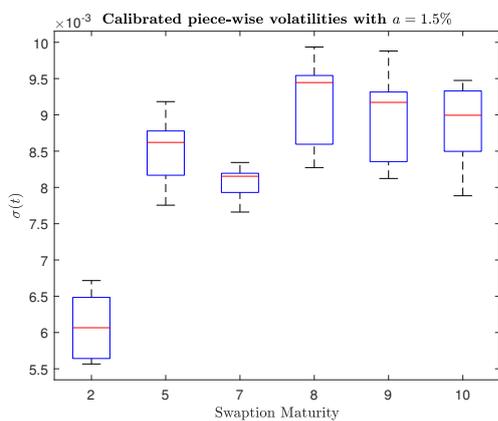


Figure 4.6: Boxplots of calibrated steps for the 1F model volatility function $\sigma(t)$ from 30-06-2017 to 30-03-2018.

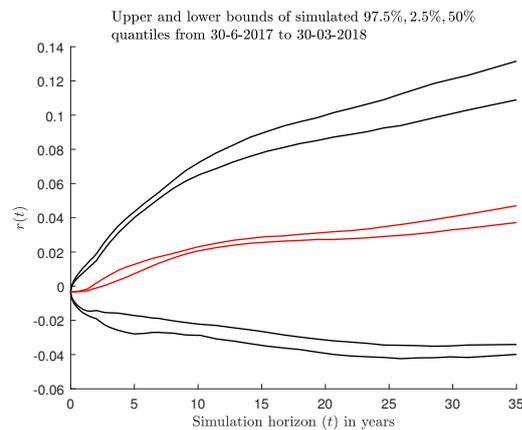


Figure 4.7: Quantiles of upper and lower quantiles of the short rate diffusion from 30-06-2017 to 30-03-2018.

HW – 2F model

Figures (4.8) and (4.9) displays the two calibrated volatilities for the two risk factors. As easily expectable, these steps are higher with respect to the one factor’s as this model comes with higher calibrated speeds. Furthermore, the first factor volatility steps move symmetrically with respect

to the second factor volatility. Note that the mean reversion speeds have been calibrated once in (3.3.5), and we keep these fixed for each of the ten dates, driven by the same motivation as for the one factor model. As it has already been said, very similar upper and lower quantiles of the short rates apply to the two factors model from 30-06-2017 to 30-03-2018. It is therefore clear that we can expect more or less same stability for the two models on a portfolio level, under these conditions.

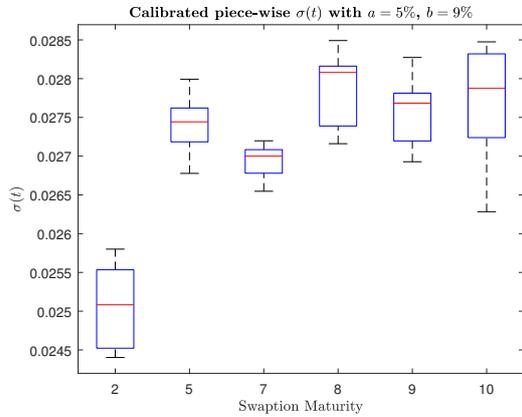


Figure 4.8: Boxplots of calibrated steps for the 2F model volatility function $\sigma(t)$ from 30-06-2017 to 30-03-2018.

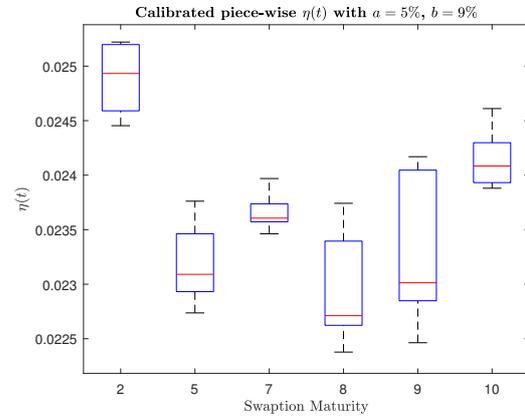


Figure 4.9: Boxplots of calibrated steps for the 2F model volatility function $\eta(t)$ from 30-06-2017 to 30-03-2018.

4.2.3. Portfolio level

To perform this step, the portfolio structure is frozen and the valuation process is repeated for the ten different days, where the calibration parameters and simulations are the ones computed in the previous paragraph. Graphic results for Counterparty A and B are put in Appendix (C). In Figure (4.10) the netted MtM quantiles for Counterparty C are shown under the two Hull-White models, while in Table (4.4) statistics measuring the stability are provided. More precisely, after normalization of each Exposure at Default ($\frac{EaD_i}{totN_i}$), the standard deviations of the realized EaDs per counterparty and model is computed and reported in the last column (these are computed over the rows of the table). The biggest impact is observed on Counterparty C as longer time-to-maturities and bigger notional are concerned. In terms of stability, the two factors model in the overall results more stable than the one factor model. It is already clearly visible from the profiles that small changes in market prices to which the model calibrates translates to huge changes in upper quantiles of netted MtMs, while measures such as EE, expected MtM, and 2.5% MtM quantile are relatively stable. For Counterparty C, the variation at the exposure peak (which is around 5-10 years for both models) is around one million Euro, which is of course a significant quantity. On the other hand in the case of Counterparty A (as seen in Subsection (C.2.2)), some calibrations bring the profiles to almost double in value, that is from one million and a half to almost 3 millions euros. This said, the bank should cautiously assess the process of setting the capital under the two Hull-White models calibrated to the market implied prices.

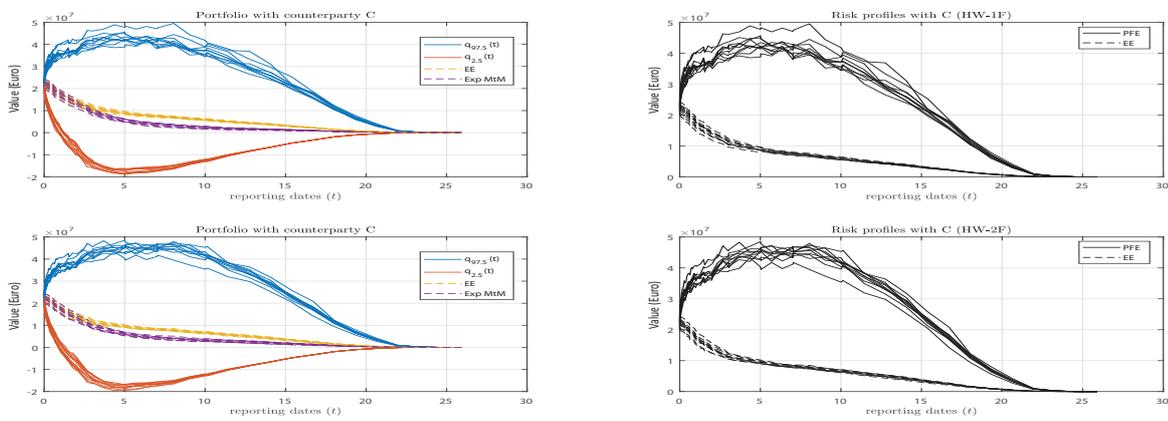


Figure 4.10: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty C implied by the $HW - 1F$ and $HW - 2F$ models over the ten different dates.

	Model	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Day 8	Day 9	Day 10	STD
A	1F	0.006	0.007	0.009	0.007	0.008	0.007	0.006	0.003	0.003	0.005	0.002
	2F	0.006	0.007	0.009	0.007	0.008	0.007	0.006	0.003	0.004	0.005	0.002
B	1F	0.016	0.018	0.020	0.017	0.017	0.019	0.015	0.012	0.011	0.014	0.003
	2F	0.015	0.016	0.020	0.018	0.018	0.019	0.014	0.012	0.013	0.012	0.003
C	1F	0.151	0.146	0.165	0.150	0.154	0.197	0.152	0.140	0.137	0.151	0.017
	2F	0.156	0.147	0.169	0.155	0.154	0.176	0.153	0.137	0.139	0.146	0.012

Table 4.4: Realized normalized EaDs over the ten calibration dates and standard deviations (STD) over rows.

5

Conclusion and Outlook

5.1. Future research

The Hull-White models, as seen, are mean reverting to a time dependent level $\theta(t)$ and because of this they allow for a perfect fit of the term structure observed in the market. Therefore, the models are designed to be calibrated to the market prices. At the same time, under the DEV-MR model explicit pricing formula for zero-coupon bonds, or options cannot be derived and these must be computed numerically. We reckon it would be informative to balance the weight of model choice and calibration choice on the realized exposures. The model impact can be measured only by employing the same calibration procedure to all the models into consideration. This said, in a future research it would be interesting to perform the same exercise done in Chapter (4) where all three models have been calibrated to the same measure, either \mathbb{P} or \mathbb{Q} . For instance, the Hull-White model could be calibrated to the historical data, as there have been problems in giving economical interpretation to the risk-neutrally calibrated parameters, which are merely the result of an optimization problem and can for this reason be very different from one model to the other. If this is done, still an open question remains about what can be the meaning (if there is) of the Hull-White level $\theta(t)$ that is calibrated historically. Anyhow, this could allow for more “realistic” values of the parameters and thus bring the exposure profiles under the two factors model to be higher with respect to the case where only one factor is considered. Otherwise, we could analyze the exposures when also the current model is calibrated to the risk-neutral measure \mathbb{Q} . In this case, we would have to numerically price European swaptions, as no explicit Zero-Coupon Bond formula could be derived.

We have shaped ourselves some ideas about how the aforementioned historical calibration could be done. First methodology requires to recall Black and Scholes (1973) and Merton (1973) assumption on the stock price dynamics, from which option prices are derived (“Black-76” formula stated in Theorem (2.2.4)). Very briefly the stock price $S(t)$ develops according to:

$$dS(t) = \mu S(t) dt + \sigma(t, S(t)) dW(t) = \mu S(t) dt + \sigma S(t) dW(t).$$

The parameter σ represents the annualized volatility of the stock price, and [18] suggests to calculate it as $\sigma = \sigma_d \sqrt{260}$, where σ_d is the daily volatility of the stock:

$$\sigma_d = \sqrt{\frac{\sum_{i=1}^n (r(t_i) - \bar{r})^2}{n-1}}, \quad \text{where } r(t_i) = \ln\left(\frac{S(t_i)}{S(t_i-1)}\right).$$

Now, given for instance the bond price dynamics under the Hull White one factor model as in (3.20), we can recognize $S(t) = P(t, t + \tau) = P_\tau(t)$ for a certain time-to-maturity τ and $\sigma = \sigma_P = -\sigma_r B(t, T)$. Therefore, one could calculate the speed a and volatility σ_r by a two-steps procedure as follow:

1. Given $r(t_1), \dots, r(t_n)$ estimate σ_r such that $dr(t) \sim \mathcal{N}(0, \sigma_r^2 dt)$. We reckon this estimate should not differ much from the volatility estimated in Section (B.1).
2. Find a such that $\sigma_d \sqrt{260} = -\sigma_r \frac{1-e^{-a\tau}}{\tau}$.

The first point is simply the $HW - 1F$ model, where the drift part is assumed to be equal to zero. In our opinion this is a reasonable approximations as in practice this term is usually very small. The problem would be here to chose an informative enough time-to-maturity τ . We leave this as an open question.

Second methodology would instead rely on the calibration steps done for the DEV-MR model. The idea would be to use the Hull-White model to simulate nine forward rates. The update of the correlation matrix would require a different transformation matrix $A(t)$, as the residuals ε would be calculated differently. In such case, nine a 's and σ 's would be calibrated. In this case, what should be understood is what kind of changes should be done to the calibration procedure in order to account for a time dependent mean reversion level, which instead we have seen to be constant in the shifted log-normal model.

The third methodology is more sophisticated and involves a particular technique for model parameters estimation called “Kalman filtering”. This requires the model to be formulated in a measurement-state transition system form. We are now assuming to start from historical data: these are collected under physical measure. For example, let us say we have a time series of zero rates $R(t, T)$ for a certain maturity $\tau = T - t$. Therefore, we already have the measurements, and the other thing that is needed is the state transition. In our case, this would be given by the short rate dynamics under the physical measure. In order to derive these, it is necessary to be able to change from the \mathbb{Q} to the \mathbb{P} measure. This is possible via the introduction of a market price of interest rate risk process λ_t , which is often assumed to be an affine function of r_t and links the Brownian motion under the risk-neutral measure \widetilde{W}_t to a Brownian motion W_t under the physical measure via the relation

$$d\widetilde{W}_t = dW_t + \lambda_t dt.$$

For more detailed explanation on this, we refer to [8, p. 117].

To conclude, if for any reason the Risk Management department will ever end up substituting the current model with the Hull-White two factors model ($HW - 2F$) for internal PFE calculations and CVA regulatory capital, we advise that more time should be spent in order to optimize the calibration, so that less computation time is required. For this, we would address the research in two directions, that is experimenting global optimization algorithms and more efficient approximation for the swaption pricing formula. For instance, this could be done by testing the efficiency of the procedure outlined in Observation (3.3.5).

5.2. Summary and conclusion

Throughout the research three models have been investigated. The first model, used for the simulation of nine forward rates, is not an Affine Term Structure Model. Its merely simulation based and does not allow for analytical solutions. For this reason, it has been calibrated to the time series of nine points of the yield curve. Whereas, for the other two short rate models the affine structure has been exploited and calibration to market implied prices has been performed. Furthermore, in order to comply with our current and future external systems, we had to perform the simulation on a Monte-Carlo basis.

The two Hull-White models have been researched under time-dependent volatilities, which is, to the best of our knowledge, an assumption that is not much tackled in the literature, especially with regard to the two factors model. We proposed a calibration procedure to market data that is easily applicable in the Hull-White framework. This is done via splitting the calibration in two nested optimizations, one for the volatilities and one for the speeds. More precisely, this procedure was designed so to comply with the time-dependency assumption. In fact, it is not possible to reduce the calibration problem to merely one optimization problem, as it could be the case when dealing with a simple one factor model with constant parameters.

We can compare these three models based on several criteria: computational difficulty, easiness for implementation, performance on correlations, stability on the bank's portfolio.

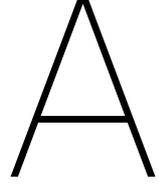
With regard to the easiness for implementation, the current model is for sure awarding the first position. The historical calibration requires many efforts in cleaning the data so that these become as most informative as possible. For instance, one needs to account for de-duplication, exclusion of observation with zero values, illiquid points, extreme observations, align dates to EUR business dates, take care of applying a Kurtosis adjustments to shifted log-returns of rates which are theoretically assumed to follow a normal distribution, but in the reality typically exhibit a more fat tailed distribution. Other precautions needs to be accounted for, but once the data have been treated in the proper way, the actual calculations to be done in order to find the calibrated parameters are very simple, and there is not much complex theoretical background to be built for this. On the other hand, the Hull-White models required much more effort both in terms of derivation of theoretical results and their implementation in Matlab. As a matter of fact, it turns out that 12 MB of codes were written for the current model, while almost 90 MB were written for the other two models. No challenging coding parts applied to the current model in particular, while reasonably tricky ones were required for the calibration of the two short rate models, especially because of the time-dependent assumption on the volatility parameters and the minor details that came with that. Constrained minimization of multi-variable loss functions needed to be tackled, which turned out to be computationally demanding in the two factors model. Therefore, the highest computational difficulty belongs to the two-factors model, for which the calibration for ten different dates required the longest computation time. On the other hand, the application to the three portfolio was not difficult from this side but rather required the codes to be efficiently organized, in a way that not many runs were required in order to retrieve the results in Chapter (4).

Throughout the thesis a special eye on the correlation between yield curves has been given when investigating all the three models, as we recognize this to have a high impact on the final results. We came to the conclusion that the current model lacks in this, while on the contrary the one factor Hull-White model performs reasonably well and the two factors is the one that capture the correlations the best. As a consequence, the two Hull-White models exhibit a potential improvement in this direction, and for this reason a change in the interest rate model may be considered in the future. It turns out that there is not much difference in the implied profiles under these two models and therefore, from this side we reckon the one factor model should be preferred (note that this is true only if a risk-neutral calibration is concerned). It is more simple to implement, and especially it requires very little time for the calibration. Of course, the two short rate models have also shown several drawbacks and therefore one needs to be cautious and further analyses should be carried out. For instance, it is true that the two factors model has potentially the ability of modelling very well the correlation between rates. At the same time, this ability is effective conditionally on the values of the parameters. The calibrated correlation between the two Brownian motion driving the two factors in $HW - 2F$ model equals -97% , which looks informative of a very different behaviour

compared to the one factor model. At the same time, the low speeds interfere considerably on the effect that this correlation can have on the actual correlation between yield rates. In fact, if we look at the shapes of the simulated term structure, we understand that the convergence to the infinitely lived bond is very slow, as the yields are slowly moving and changing shapes.

From a stability point of view, it is natural for an historical calibration to be more stable, while the risk-neutrally calibrated models appear to show high variation over one year time window. Nevertheless, this does not mean that there are not precautions that one could take in order to account for this, and therefore we do not consider appropriate to classify this as a problem which we could not dismiss. We reckon to be safer having a model that implies a more conservative estimate of risk and maybe is more unstable rather than having a model that implies stable underestimates of exposures.

To conclude, in our opinion the current model should be discarded in favour of one of the two Hull-White models. The first reason is that they are both theoretically designed in a way that underestimation of risk due to wrongly captured correlation structures between rates cannot happen. In fact, we have seen that one factor implies perfect correlation between rates and two factors imply very diverse values for the correlations. Eventually, from this side the result is that the $HW - 1F$ model behaves the opposite with respect to the current model, and so does the $HW - 2F$ asymptotically (actually, as seen, already from relatively short simulation horizons it approaches the one factor model behaviour). Nevertheless, it is undeniable that within Risk Management a model suggesting to higher the level of capital requirements due to fluctuations in the market from one day to another is not really desirable, as this would limit the business and trading activities. For this reason, we reckon that exploiting the positive aspects of a two-factors model calibrated under the historical measure represents the best option. This would probably lead to more conservative capital requirements but the historical measure would fix the problem of unstable profiles encountered in Chapter (4). In addition to this, even if the two factors model correlation tends to one, it is able to predict more flexible yield shapes for short time horizons, and therefore may reveal to provide more accurate estimates of the exposure profiles for portfolios that are made of relatively short instruments.



Computation of transformation matrix $A(t)$

We have seen in (3.1) that the correlation matrix at grid point t is updated via the element-wise product $\rho(t) = A(t)\rho_{t-\Delta t}$, where $\rho_{t-\Delta t}$ is the updated matrix at $t - \Delta t$. The transformation matrix $A(t)$ needs to have diagonal equal to one, so to preserve the nature of a correlation matrix. We here show how we computed the elements outside, that is for $i \neq j$, with $i, j = 1, \dots, 9$.

Let us consider factor i and factor j . From Equation (3.4) and Ito's isometry which tells us that $\sqrt{\frac{1-e^{-2a(t-s)}}{2a}} \varepsilon = \int_s^t e^{-a(t-u)} dW_u$ (where $s < t$), we have (conditionally on $\mathcal{F}_{t-\Delta t}$) that

$$\begin{cases} \ln[f_i(t) + \gamma] = \ln[f_i(t - \Delta t) + \gamma] e^{a_i(\Delta t)} + \frac{b_i}{a_i} (1 - e^{a_i(\Delta t)}) + \sigma \int_{t-\Delta t}^t e^{-a_i(t-u)} dW_u^i, \\ \ln[f_j(t) + \gamma] = \ln[f_j(t - \Delta t) + \gamma] e^{a_j(\Delta t)} + \frac{b_j}{a_j} (1 - e^{a_j(\Delta t)}) + \sigma \int_{t-\Delta t}^t e^{-a_j(t-u)} dW_u^j. \end{cases}$$

Therefore, $\rho(t) = \text{Corr}(f_i(t), f_j(t)) = \text{Corr}(\int_{t-\Delta t}^t e^{a_i u} dW_u^i, \int_{t-\Delta t}^t e^{a_j u} dW_u^j)$. To compute explicitly the correlation, note that

$$\begin{aligned} \mathbb{E} \left[\left(\int_{t-\Delta t}^t e^{a_i u} dW_u^i \right)^2 \right] &= \int_{t-\Delta t}^t e^{2a_i u} du = \frac{e^{2a_i t} (1 - e^{-2a_i \Delta t})}{2a_i}, \\ \mathbb{E} \left[\left(\int_{t-\Delta t}^t e^{a_j u} dW_u^j \right)^2 \right] &= \int_{t-\Delta t}^t e^{2a_j u} du = \frac{e^{2a_j t} (1 - e^{-2a_j \Delta t})}{2a_j}, \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left[\left(\int_{t-\Delta t}^t e^{a_i u} dW_u^i \right) \left(\int_{t-\Delta t}^t e^{a_j u} dW_u^j \right) \right] &= \mathbb{E} \left[\int_{t-\Delta t}^t e^{(a_i+a_j)u} dW_u^i dW_u^j \right] \\ &= \int_{t-\Delta t}^t e^{(a_i+a_j)u} \mathbb{E} \left[dW_u^i dW_u^j \right] du \\ &= \frac{e^{(a_i+a_j)t} (1 - e^{-(a_i+a_j)\Delta t})}{(a_i + a_j)} \rho_{i,j}^{t-\Delta t} \end{aligned}$$

Therefore, the correlation between $f_i(t)$ and $f_j(t)$ is given by

$$\rho(t) = \frac{e^{(a_i+a_j)t} (1 - e^{-(a_i+a_j)\Delta t}) \frac{1}{(a_i+a_j)}}{e^{(a_i+a_j)t} \sqrt{\frac{(1-e^{-2a_i\Delta t})}{2a_i}} \sqrt{\frac{(1-e^{-2a_j\Delta t})}{2a_j}}} \rho_{i,j}^{t-\Delta t} = A_{i,j}^t \rho_{i,j}^{t-\Delta t}.$$

B

Estimation of mean reversion speed in HW-1F model

In the thesis we deal with the following formulation of the $HW - 1F$ model:

$$dr(t) = (\theta(t) - ar(t)) dt + \sigma(t) dW(t), \quad (\text{B.1})$$

where $W(t)$ is a standard Brownian motion under the risk-neutral measure \mathbb{Q} . The parameter a represents the speed at which the short rate $r(t)$ is pulled back towards the mean reversion level $\frac{\theta(t)}{a}$ and a proper calibration of this parameter is very important, as the volatility calibration is directly affected by the value of a . That is, the volatility is calibrated with a given as input. Two other alternatives have been investigated for the calibration of such parameter, in addition to the one presented in (3.2.5).

- historical calibration of a
- estimating such speed independently of the volatility parameter by means of an approximated formula for the swaption implied volatility

The two approaches are described in the following two sections.

B.1. Historical Calibration

The mean reversion speed a could be historically calibrated based on the simple Vasicek model, where all parameters are assumed constant and therefore no perfect fit to the term structure can be achieved. A proxy for the short rate needed to be considered for this purpose, as suggested in [19, p. 77]. Our choice falls in the (daily) 3 months Euribor rate, whose time series can be easily retrieved from the Bloomberg platform. In figure (B.1) we display such rate from 03/01/2000 to 02/07/2018. Under the risk-neutral measure, the spot rate in the Vasicek model evolves according to:

$$dr(t) = (\theta - ar(t)) dt + \sigma dW(t).$$

There exist mainly two method for estimating the three parameters θ , a and σ : maximum Likelihood method, described in [12] and least squares regression method, as described in [31].

Results

The calibration was performed with the Maximum Likelihood method on the time series starting from 11/06/2013 in order not to take into account too diverse economic periods, or by now meaningless data. The resulting estimates are $\theta = -11$ bps, $\sigma = 3.90$ bps and $a = 6.37\%$. In figure (B.2) are shown three scenarios from the calibrated Vasicek model together with the data. The idea would be then to give $a = 6.37\%$ as input to the function that calibrates $\sigma(t)$ in (B.1).

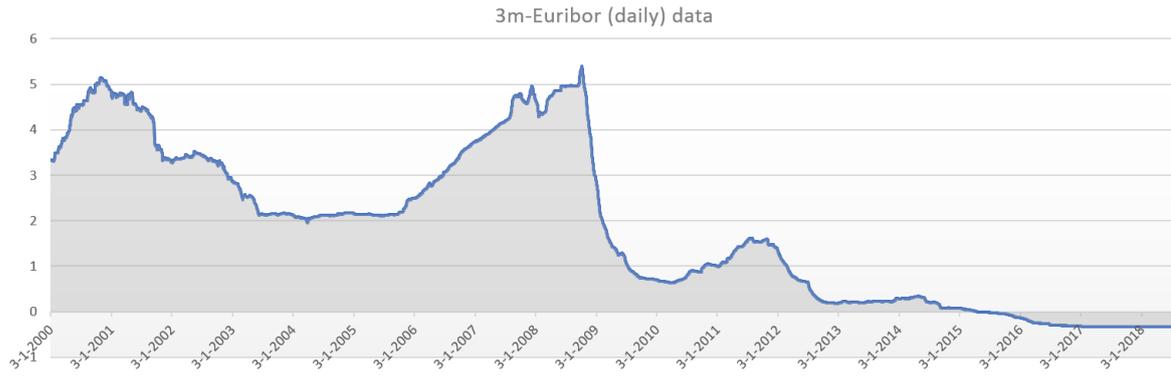


Figure B.1: Time series of daily 3m-Euribor rate (%) used for historical calibration.

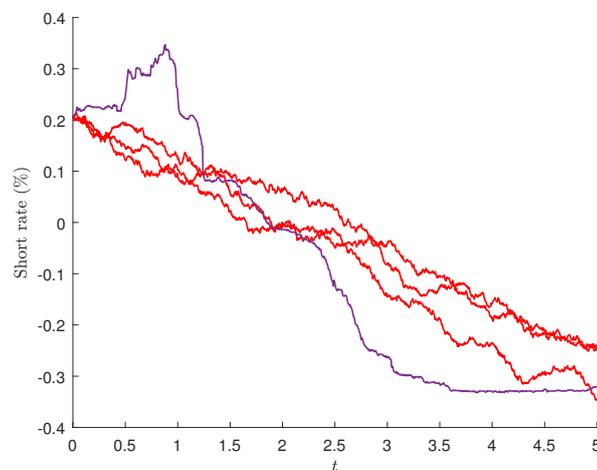


Figure B.2: Three simulations (red) of the historically calibrated Vasicek model together with Euribor (purple).

B.2. Swaption implied volatility approximation

This approach is introduced in [26]. The price of an European (payer) swaption has been proven to be a weighted sum of zero-bond (put) options and the result is stated in Theorem (3.2.5). Although this provides us with a closed form pricing formula, it is not explicit in the parameters and it may not be to handy to calibrate on this. The main idea of the paper is to derive a formula whose dependency between a and $\sigma(t)$ is decoupled, and the direct relation between market implied volatility and model parameters is more visible. More precisely, under the Hull-White model introduced in Equation (B.1), it is possible to prove that the variance of the swap rate can be approximately calculated with the following expression:

$$V_S(T_\alpha, T_\beta) \approx \left[\frac{P(0, T_\alpha)}{P(0, T_\alpha) - P(0, T_\beta)} \right]^2 V(0, T_\alpha) B(T_\alpha, T_\beta)^2, \quad (\text{B.2})$$

where $V(s, t) := \text{Var}\{r(t)|\mathcal{F}_s\} = \int_s^t e^{-2a(t-u)} \sigma(u)^2 du$ and $B(s, t) = \frac{1-e^{-a(t-s)}}{a}$. Note that this implied variance, derived under $HW - 1F$, depends on the model volatility $\sigma(t)$ only through the function $V(0, T_\alpha)$. This means that the dependencies in the swap tenor T_β and $\sigma(t)$ are decoupled in the approximation formula, as desired.

The objective function

Now, one could estimate the speed parameter independently of the volatility by taking the ratios of (B.2) for fixed maturity M_i and different tenors T_j and T_k and try to find a that allows for the best fit of the market implied volatilities ratios. That is, given the ratio:

$$\frac{V_S(M_i, T_j)}{V_S(M_i, T_k)} = \left[\frac{(P(0, M_i) - P(0, T_k))B(M_i, T_j)}{(P(0, M_i) - P(0, T_j))B(M_i, T_k)} \right]^2,$$

whose dependency on a is implicit in the expression for $B(s, t)$, the parameter can be estimated such that:

$$F_i(a) = \sum_j \frac{W_{i,j+1}}{W_{i,j}} \left(\sqrt{\frac{V_S(M_i, T_{j+1})}{V_S(M_i, T_j)}}(a) - \frac{\sigma_{i,j+1}^B}{\sigma_{i,j}^B} \right)^2, \quad (\text{B.3})$$

is minimized. Here we denoted $IV_{i,j}^B$ as the Black implied volatility for the swaption of maturity M_i and tenor T_j .

Results

The function to be minimized in (B.3) is related to one single swaption maturity M_i . In Figure (B.3) are displayed the calibrated speeds over a time period of 10 months (where calibration is done on the last business day of each month), relative to 10, 20 and 30 years maturities. The problem with this approach is that the calibrated mean reversion speeds is dependent on the chosen maturity, and therefore we rather prefer to stick to the method described in Subsection (3.2.5). Still, this small exercise gave us the motivation to calibrate the speed only once and then keep it fixed for other calibration days in Chapter (4). In fact, a quite constant behaviour over time for each of the three cases is exhibited.

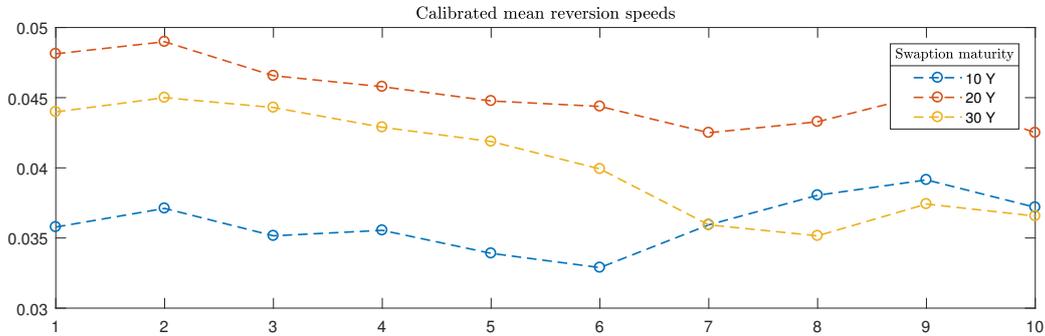


Figure B.3: Calibrated speeds for $HW - 1F$ model over ten months.

C

Further results

C.1. Simulation results of HW-1F with $a = 1.5\%$

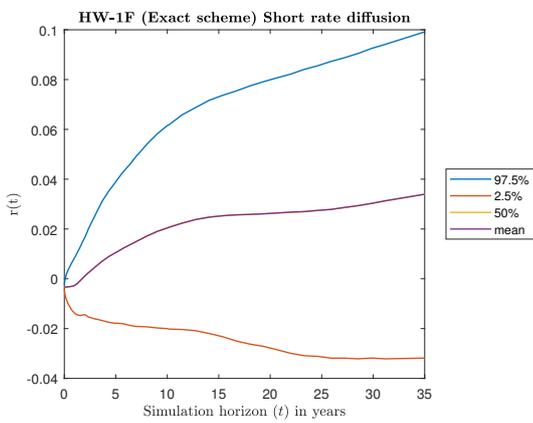


Figure C.1: Short rate diffusion with $a = 1.5\%$.

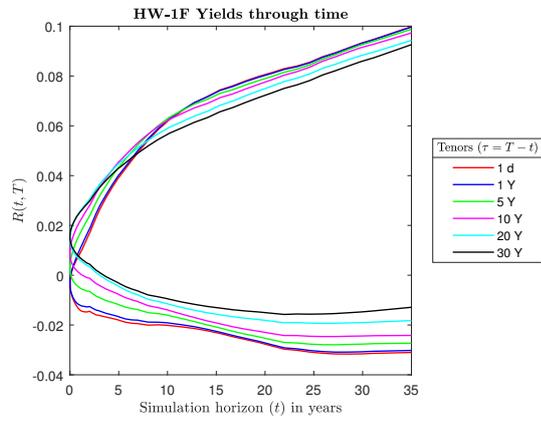


Figure C.2: Simulated yields quantiles as a function of time t .

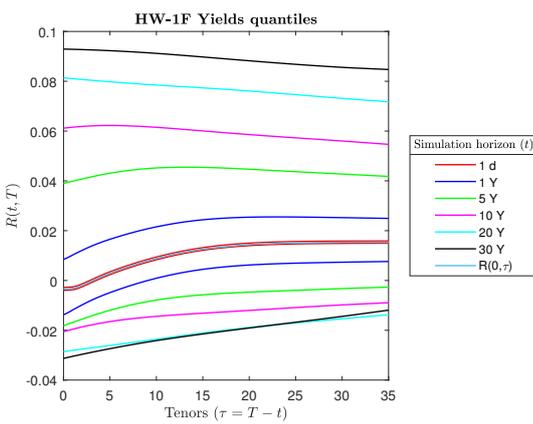


Figure C.3: Simulated yields quantiles t years ahead.

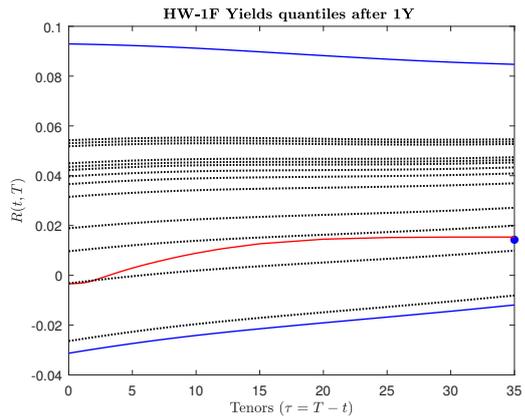


Figure C.4: Shapes implied by the HW-1F after 30 years simulation together with quantiles and yield on infinitely lived bond (blue dot).

C.2. Further portfolios

C.2.1. Single date

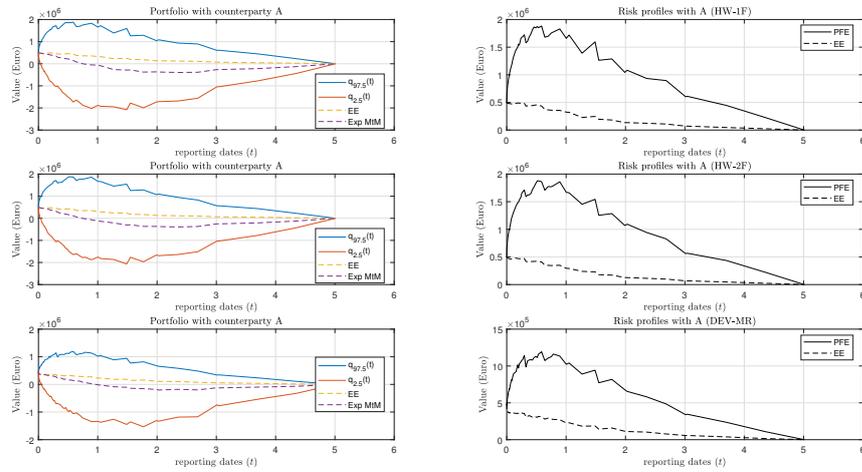


Figure C.5: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty A implied by the DEV-MR, $HW - 1F$ and $HW - 2F$ models.

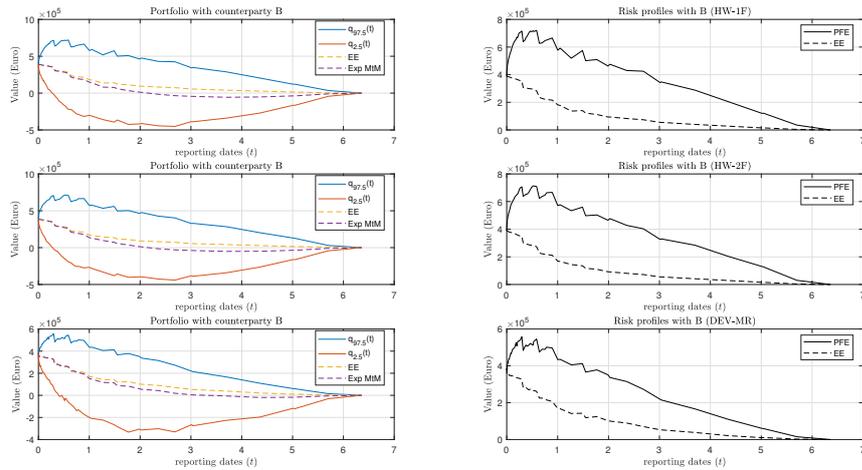


Figure C.6: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty B implied by the DEV-MR, $HW - 1F$ and $HW - 2F$ models.

C.2.2. Multiple dates

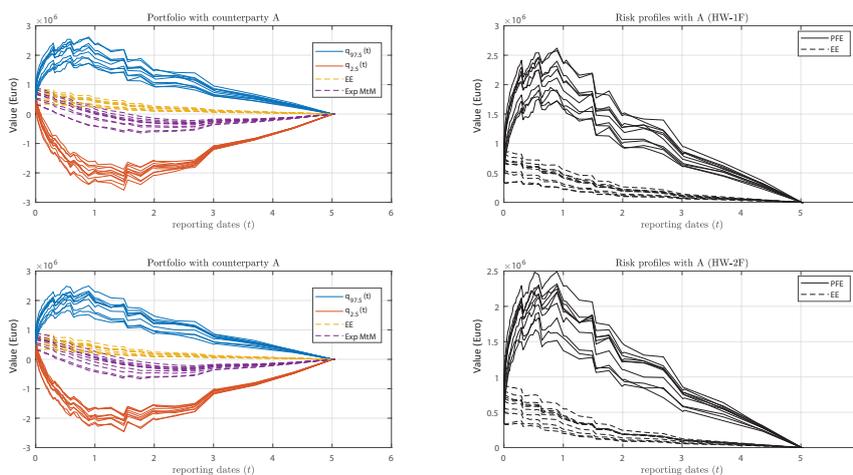


Figure C.7: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty A implied by the $HW - 1F$ and $HW - 2F$ models over the ten different dates.

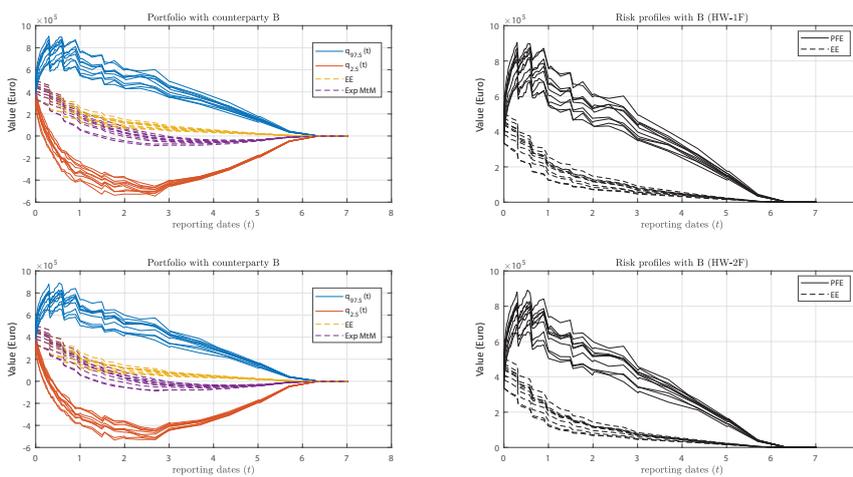


Figure C.8: MtM quantiles, PFE and EE profiles for the single-currency portfolio with Counterparty B implied by the $HW - 1F$ and $HW - 2F$ models over the ten different dates.

Bibliography

- [1] Tomas Björk. *Arbitrage theory in continuous time*. Oxford university press, 2009.
- [2] Fischer Black. The pricing of commodity contracts. *Journal of financial economics*, 3(1-2), 1976.
- [3] Fischer Black and Piotr Karasinski. Bond and option pricing when short rates are lognormal. *Financial Analysts Journal*, 1991.
- [4] Fischer Black, Emanuel Derman, and William Toy. A one-factor model of interest rates and its application to treasury bond options. *Financial analysts journal*, 1990.
- [5] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.
- [6] Jaehyuk Choi and Sungchan Shin. Fast swaption pricing in gaussian term structure models. *Mathematical Finance*, 26(4), 2016.
- [7] Basel Committee et al. An explanatory note on the basel 2 irb risk weight functions. *Bank for International Settlements*, 2005.
- [8] C.Xiong. *Introduction to interest rate models*. 2018. Available at <http://www.cs.utah.edu/cxiong/>.
- [9] Marco Di Francesco. A general gaussian interest rate model consistent with the current term structure. *ISRN Probability and Statistics*, 2012, 2012.
- [10] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6), 2000.
- [11] Frank J Fabozzi et al. *Advances in fixed income valuation modeling and risk management*, volume 18. John Wiley & Sons, 1997.
- [12] Kevin Fergusson and Eckhard Platen. Application of maximum likelihood estimation to stochastic short rate models. *Annals of Financial Economics*, 10(02), 2015.
- [13] Damir Filipovic. *Term-Structure Models. A Graduate Course*. Springer, 2009.
- [14] Jon Gregory. *Counterparty credit risk: the new challenge for global financial markets*, volume 470. John Wiley & Sons, 2010.
- [15] Thomas SY Ho and Sang-Bin Lee. Term structure movements and pricing interest rate contingent claims. *the Journal of Finance*, 41(5), 1986.
- [16] John Hull. Numerical procedures for implementing term structure models: Two-factor models. *Journal of Derivatives*, 2, 1994.
- [17] John Hull and Alan White. Pricing interest-rate-derivative securities. *The Review of Financial Studies*, 3(4), 1990.

- [18] John C Hull and Sankarshan Basu. *Options, futures, and other derivatives*. Pearson Education India, 2016.
- [19] Jessica James and Nick Webber. *Interest rate modelling*, volume 77. Wiley, 2000.
- [20] Farshid Jamshidian. An exact bond option formula. *The journal of Finance*, 44(1), 1989.
- [21] Farshid Jamshidian and Yu Zhu. Scenario simulation: Theory and methodology. *Finance and stochastics*, 1(1), 1996.
- [22] Scott Kirkpatrick, C Daniel Gelatt, and Mario P Vecchi. Optimization by simulated annealing. *science*, 220(4598), 1983.
- [23] Attilio Meucci. 'p' versus 'q': Differences and commonalities between the two areas of quantitative finance. 2011.
- [24] Vladimir Ostrovski. Efficient and exact simulation of the hull-white model. *International Journal of Financial Engineering*, 2013.
- [25] Vladimir Ostrovski. Efficient and exact simulation of the gaussian affine interest rate models. *International Journal of Financial Engineering*, 2016.
- [26] Martin Rainer. Calibration of stochastic models for interest rate derivatives. *Optimization*, 58(3):373–388, 2009.
- [27] Riccardo Rebonato. *Interest-rate option models: understanding, analysing and using models for exotic interest-rate options*. John Wiley & Sons, 1996.
- [28] Walter Schachermayer and Josef Teichmann. How close are the option pricing formulas of bachelier and black–merton–scholes? *Mathematical Finance: An International Journal of Mathematics, Statistics and Financial Economics*, 18(1), 2008.
- [29] David F Schrager and Antoon AJ Pelsser. Pricing swaptions and coupon bond options in affine term structure models. *Mathematical Finance*, 16(4), 2006.
- [30] M Pykhtin SH Zhu. A guide to modeling counterparty credit risk. 2007. Available at <http://ssrn.com/abstract=1032522>.
- [31] Thijs van den Berg. Calibrating the ornstein-uhlenbeck (vasicek) model. 2011. Available at <http://www.sitmo.com/article/calibrating-the-ornstein-uhlenbeck-model/>.
- [32] Oldrich Vasicek. An equilibrium characterization of the term structure. *Journal of financial economics*, 5(2), 1977.