

**Essays on the modeling of
risks in interest-rate and
inflation markets**

Allan Sall Tang Andersen

Advisor : Bjarne Astrup Jensen

Ph.D. dissertation

Department of Finance

Copenhagen Business School

May, 2011

Contents

Contents	iii
Acknowledgements	v
Introduction	1
Summary	11
1 Modeling stochastic skewness in a Heath-Jarrow-Morton framework	19
1.1 Introduction	20
1.2 Evidence of stochastic skewness	21
1.3 Modeling stochastic skewness	27
1.4 Specifying the Model	33
1.5 Model calibration	35
1.6 Results	38
1.7 Conclusion	43
1.8 Appendix: Proof of proposition 1	44
1.9 Appendix: MCMC details	45
2 Inflation derivatives modeling using time changed Lévy processes	49
2.1 Introduction	50
2.2 Inflation linked products	52
2.3 The Jarrow-Yildirim model	56
2.4 Time changed Lévy processes	58
2.5 An inflation HJM framework based on time changed Lévy processes	61
2.6 Pricing inflation products	68
2.7 Specification of the time-change	72
2.8 Calibration	74
2.9 Conclusion	79

2.10	Appendix: Proofs	81
2.11	Appendix: The Variance Gamma process	85
3	Inflation risk premia in the term structure of interest rates: Evidence from Euro area inflation swaps	87
3.1	Introduction	88
3.2	Data: Inflation swap rates and the nominal term structure	90
3.3	Inflation risk premia: What theory predicts	97
3.4	A no-arbitrage model of nominal and inflation swap rates	99
3.5	Model Estimation	102
3.6	Empirical results	105
3.7	Conclusion	120
3.8	Appendix: Derivation of nominal ZCB prices, real ZCB prices and Inflation expectations	122
3.9	Appendix: MCMC estimation of the model	126
4	Affine Nelson-Siegel Models and Risk Management Performance	131
4.1	Introduction	132
4.2	The Danish Government Bond Term Structure	133
4.3	Affine Term Structure Models	136
4.4	Multi-factor Cox-Ingersoll-Ross models	138
4.5	Affine Nelson-Siegel models	140
4.6	Model Estimation	143
4.7	Empirical Results: In-Sample	146
4.8	Empirical Results: Out-of-Sample	153
4.9	Conclusion	159
4.10	Appendix: Affine Nelson-Siegel models with stochastic volatility	161
4.11	Appendix: MCMC details	167
4.12	Appendix: Density Forecasts	172
4.13	Appendix: Tables with out-of-sample forecast results	177
	Conclusion	203
	Bibliography	205

Acknowledgements

With the hand-in of this thesis, I have now finished my years as a Ph.D. student at the Department of Finance at Copenhagen Business School and Danmarks Nationalbank. The years as a Ph.D. student have at times been hard work, but it has most definitely been a good experience.

First of all, I would like to thank my supervisor Bjarne Astrup Jensen for encouraging me to consider a position as a Ph.D. student, for providing me with suggestions and pin-pointing errors and unclear reasoning in my work during the process.

I would like to thank Danmarks Nationalbank for giving me the opportunity to enter into their Ph.D. programme and the Capital Markets/Financial structure division at the European Central Bank for giving me the possibility to spend 6 months in Frankfurt. The stay in Frankfurt both spurred an interest in and deepened my understanding of inflation markets. I would also like to thank professor Rama Cont for providing the opportunity to stay a semester at the Center for Financial Engineering, Columbia University.

Fellow Ph.D. students at the Department of Finance and Danmarks Nationalbank also deserve thanks for creating a good environment, both socially and professionally. From Danmarks Nationalbank I would especially like to thank my office-mate Jannick Damgaard for making the Ph.D. a pleasant time, despite non-overlapping research interests. From the Department of Finance I would like to thank Mads Stenbo Nielsen and Jens Dick-Nielsen for always having their door open if I needed to discuss a topic. Outside the world of Ph.D. students, I would like to thank Jacob Ejsing and Kasper Ahrndt Lorenzen for providing valuable insights, especially regarding inflation markets.

I also want to thank Jesper Lund and Fred Espen Benth for participating in my pre-defence and providing me with valuable comments for my thesis.

On a personal level, I want to thank my parents for supporting me my entire life and providing abstractions outside the world of finance. Finally, I would like to thank my fiancée Anne-Sofie for constant encouragement

and patience. Your insight into the life of a Ph.D. student, have made my life as a, at times slightly confused, Ph.D. student significantly better.

Copenhagen, May 2011
Allan Sall Tang Andersen

Introduction

The topic of this thesis is the modeling of risks in interest-rate and inflation markets.

Interest-rate risk is an important issue to investors. For instance, according to BIS (2010) the notional value of over-the-counter interest-rate derivatives markets is 465,260 billion US-dollar. This corresponds to 77 percent of the notional of the entire OTC derivatives market. Thus interest-rate derivatives is at the back-bone of the financial markets. According to ISDA (2009) 83 percent of Fortune 500 companies report using interest-rate derivatives in their risk management. Furthermore, many mortgage-based loans and pension contracts contain either explicit or implicit interest-rate options. Thus a better understanding of the interest-rate derivative markets, and the risk associated with the traded products is of great value, both to financial and non-financial companies as well as individuals.

The market for inflation linked products, such as bonds and swaps, is significantly smaller than the one for interest-rate derivatives. The market is also significantly newer than the nominal interest-rate market, with one of the prominent examples being US Treasury Inflation Protected Bonds (TIPS). TIPS were introduced in 1997 and with a notional outstanding of less than 50 billion US dollar. By 2010 the US Treasury has issued TIPS worth over 600 billion US dollar, see Christensen and Gillan (2011). The issuance of inflation linked bonds is not limited to the US; countries such as France, United Kingdom¹, Germany and Japan also have a significant issuance of inflation linked bonds.

With the increase in issuance and trading in inflation linked bonds derivative markets have evolved. Inflation swaps, which can be compared to interest-rate swaps, have been traded at least since 1995 (see Barclays Capital (2008)), and in recent years options on inflation have also been traded. With the increase in trading in inflation linked products academic interest

¹United Kingdom was one of the first countries to issue inflation linked bonds. UK inflation linked bonds has been issued as early as in 1981.

has spurred, but perhaps more interesting to individuals, central banks have started using the information content embedded in inflation linked products in relation to the economic analysis used to set policy rates.

The goal of this Ph.D. thesis is to add a small piece to the puzzle of understanding interest-rate and inflation markets. The thesis consists of four essays, two of which are focused on modeling interest-rate risk and two are focused on modeling inflations risks and risk premia. Each essay contributes to the literature in its field and can, of course, be read independently. In short, here follows a brief motivation for each paper.

Modeling stochastic skewness in a Heath-Jarrow-Morton framework

Several facts on interest rate behaviour are well known. First, interest-rate volatilities are obviously stochastic, and these volatilities tend to cluster in periods with low respectively high volatility (see for instance Andersen and Lund (1997)). Carr and Wu (2007) show that currency options have time-varying skewness. By using model-free estimates of the volatility and skewness priced in interest-rate options, it can be shown that interest-rate distributions also show time-varying skewness (see Trolle and Schwartz (2010)). The main purpose of the paper is to provide a consistent framework for modeling the stochastic volatility and skewness. Finally, calibrating the model to time-series of market data is interesting, as it shows the applicability of the model.

Inflation derivatives modeling using time changed Lévy processes

With the rise of inflation derivatives and more liquid markets, non-linear inflation contracts have been introduced. When considering time-series of inflation swap rates, the fact that changes in inflation swap rates show large sudden movements, i.e. jumps, is easily acknowledged. The standard method for modeling inflation derivatives is the Gaussian forward-rate framework introduced in Jarrow and Yildirim (2003). In this paper we wish to model inflation derivatives in a no-arbitrage framework which could both include time varying volatility and jumps. We also want to provide evidence of the applicability of the model framework by using market data.

Inflation risk premia in the term structure of interest rates: Evidence from Euro area inflation swaps

Market based inflation expectations such as inflation swap rates or the Break Even Inflation Rate, i.e. an inflation measure derived from nominal and inflation linked bonds, provide market participants with real-time measures of inflation expectations. However, these measures include both an inflation expectation and an inflation risk premium. The purpose of the paper is to use a no-arbitrage model to disentangle the two components by using nominal swap rates, inflation swap rates, surveys on inflation expectations and CPI data. The model output can be used to interpret whether changes in inflation swap rates correspond to changes in inflation expectations or inflation risk premia.

Affine Nelson-Siegel Models and Risk Management Performance

The ability to correctly assess the interest-rate risk one faces as an investor is a critical issue. This is obviously the case for a pension fund with a large bond portfolio, but it is also the case for government debt agencies who need to assess the risks of different issuance strategies. The importance of using a good framework for managing interest-rate risk is further emphasized by the current sovereign debt crisis. The purpose of the paper is to assess the medium- and long-term forecasts of the Affine Nelson-Siegel model-class introduced in Christensen, Diebold, and Rudebusch (2011) and Christensen, Lopez, and Rudebusch (2010). With regard to data we use Danish government bond yields from 1987 to 2010, and use a Bayesian Markov Chain Monte Carlo method to estimate the models.

A short summary of the considered markets and related literature

For the convenience of the reader we provide a brief description of the main concepts used in the essays in this thesis. For a general description of continuous time asset pricing we refer to Duffie (2001) or Björk (2004). For an introduction to Lévy processes we refer to Cont and Tankov (2004). A good introductory book on inflation markets and their institutional features is Deacon, Derry, and Mirfendereski (2004), whereas a textbook treatment on the modeling of interest-rate and inflation derivatives can be found in Brigo and Mercurio (2006).

The term structure of interest-rates

In this section we briefly describe the basics of the term structure of interest-rates. We also describe the literature on the modeling of the term structure of interest-rates.

A *zero-coupon bond* with maturity T (also called T -bond), is a contract that pays the owner of the zero-coupon bond 1 unit of currency at time T . The zero coupon bond observed at time t that matures at time T is denoted $p(t, T)$.

In the money market it is customary to consider simple compounding, i.e. we are considering the simple forward rate (*LIBOR forward rate*) $L(t; S, T)$. The LIBOR forward rate agreement is an agreement to borrow or lend between time S and T at a time t specified simple rate $L(t; S, T)$. By no-arbitrage we have the following relationship between zero coupon bond prices and LIBOR forward rates.

$$\frac{p(t, S)}{p(t, T)} = 1 + (T - S)L(t; S, T) \Leftrightarrow L(t; S, T) = -\frac{1}{T - S} \frac{p(t, T) - p(t, S)}{p(t, T)}$$

The case of the simple spot rates (*LIBOR spot rates*), is denoted $L(t, T)$, i.e. a simple compounded rate starting from today (time t) to some future point in time T . This implies that the LIBOR spot rate is the same as a LIBOR forward rate with $S = t$, i.e. $L(t, T) = L(t; t, T)$. Hence we have

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

In most models we are not considering simple rates (the market model being the exception), but instead it is more convenient to consider continuously compounded rates.

When considering how to derive the *continuously compounded forward rates* from zero-coupon bonds we will use that investors are indifferent between investing in a zero-coupon bond or an asset with continuously compounded rate

$$\frac{p(t, S)}{p(t, T)} = e^{y(t; S, T)(T - S)} \Leftrightarrow y(t; S, T) = -\frac{\log p(t, T) - \log p(t, S)}{T - S}$$

Again, we can find the *continuously compounded spot rates*, $y(t, T)$, by letting $S = t$ in the continuously compounded forward rates

$$y(t, T) = -\frac{\log p(t, T)}{T - t}$$

Finally, in most continuous time interest rate models it is customary to use instantaneous continuously compounded rates, i.e. rates that prevail over an instantaneous time interval $[T, T + dt]$. The *instantaneous forward rates* can be obtained by letting $(T - S) \rightarrow 0$ in the continuously compounded forward rates

$$f(t, T) = \lim_{(T-S) \rightarrow 0} -\frac{\log p(t, T) - \log p(t, S)}{T - S} = -\frac{\partial \log p(t, T)}{\partial T}$$

Similarly, we define the instantaneous *short rate* as

$$r(t) = f(t, t)$$

Using the definitions above we have the following relationship for $t \leq s \leq T$

$$p(t, T) = p(t, s) \exp \left\{ - \int_s^T f(t, u) du \right\}$$

and in particular

$$p(t, T) = \exp \left\{ - \int_t^T f(t, u) du \right\}$$

Term structure models can mainly be put in three categories: Short-rate models, forward-rate models and market models.

Short-rate models

Short-rate models are based on modeling the instantaneous short rate $r(t)$. The short rate is typically described as a diffusion process

$$dr(t) = \mu(t, r(t), X(t))dt + \sigma(t, r(t), X(t))dW(t)$$

where $\mu(t, r(t), X(t))$ is the drift of the process and $\sigma(t, r(t), X(t))$ is the diffusion term. $X(t)$ represents factors that are not the short rate, typically assumed to be latent factors.

An often referenced short-rate model is the Vasicek-model (see Vasicek (1977)), where the short rate follows an Ornstein-Uhlenbeck process

$$dr(t) = \kappa(\theta - r(t))dt + \sigma dW(t)$$

Using this specification yields can be shown to be affine functions of the short rate

$$y(t, T) = \frac{A(t, T)}{T - t} + \frac{B(t, T)}{T - t}r(t)$$

where $A(t, T)$ and $B(t, T)$ are functions of the model parameters.

Duffie and Kan (1996) extend the Vasicek-model to include a more general multivariate affine diffusion process. The framework of Duffie and Kan (1996) includes prominent short-rate models such as the Vasicek-model and the Cox-Ingersoll-Ross model, see Cox, Ingersoll, and Ross (1985). The results in Duffie and Kan (1996) have been generalized further in Duffie, Pan, and Singleton (2000) and Duffie, Filipovic, and Schachermayer (2003) such that interest-rate options also can be priced in a general affine model.

A part of the finance literature has focused on short-rate models and their ability to forecast yields and term premia, see Dai and Singleton (2002), Dai and Singleton (2003), Duffie (2002), Cheredito, Filipovic, and Kimmel (2007), Feldhütter (2008) and Christensen, Diebold, and Rudebusch (2011). Their results point towards that models mainly consisting of Gaussian factors provide the best forecasts. The fact that a Gaussian model is preferred can in part be linked to the structure of the affine model framework. Non-Gaussian models put more restrictions on the interest-rate dynamics due to admissibility conditions on the diffusion processes.

Recently, there has been a significant focus on capturing interest-rate volatilities, see Andersen and Lund (1997), Andersen, Benzoni, and Lund (2004), Collin-Dufresne and Goldstein (2002), Collin-Dufresne, Goldstein, and Jones (2008), Jacobs and Karoui (2009) and Christensen, Lopez, and Rudebusch (2010). Collin-Dufresne and Goldstein (2002) and Collin-Dufresne, Goldstein, and Jones (2008) argue that interest-rate volatility is not spanned by yields, i.e. that interest-rate derivatives cannot be perfectly hedged by using zero-coupon bonds. On the other hand, Jacobs and Karoui (2009) argue that the results of Collin-Dufresne and Goldstein (2002) and Collin-Dufresne, Goldstein, and Jones (2008) are driven by the considered data and sampling-period.

Forward-rate models

An alternative to modeling the short rate is the forward-rate modeling framework of Heath, Jarrow, and Morton (1992), also termed the Heath-Jarrow-Morton framework, henceforth HJM framework.

In this framework we consider the instantaneous forward rate $f(t, T)$, which is assumed to solve the stochastic differential equation

$$df(t, T) = \alpha(t, T)dt + \sigma(t, T)dW(t)$$

where $\sigma(t, T)$ is a volatility loading, which describes how forward rates with specific maturities are affected by changes in the Wiener process W . $\alpha(t, T)$

is a drift term, which ensures that the model is arbitrage-free.

Heath, Jarrow, and Morton (1992) show that, under the risk neutral measure, the drift term has the following form:

$$\alpha(t, T) = \sigma(t, T) \int_t^T \sigma(t, s) ds$$

Thus, the risk neutral dynamics are entirely determined by the volatility loading $\sigma(t, T)$. The HJM framework can be specified such that it is consistent with a short-rate model; however, with the modification that the model has a perfect fit to the initial term structure. For instance, when $\sigma(t, T) = \sigma e^{-\kappa(T-t)}$ the model is consistent with the Vasicek-model.

A strand in the mathematical finance literature has extended the HJM-framework to include innovations driven by a Lévy process, see Eberlein and Raible (1999), Raible (2000) and Kluge (2005). Even though deriving a HJM framework based on Lévy processes is an accomplishment of its own, Kluge (2005) show that the Heath-Jarrow-Morton-framework based on Lévy processes can provide a good fit to market data from single trading days.

Casassus, Collin-Dufresne, and Goldstein (2005) find that models based on the HJM framework with particular ease can generate interest-rate volatility that is not spanned by yields. In recent papers Trolle and Schwartz (2009) and Trolle and Schwartz (2010) estimate a model based on a HJM framework with stochastic volatility and find that their model is able to describe both yields and prices of interest-rate derivatives.

Market models

Market models are an alternative to short-rate and forward-rate models and are mainly used for pricing of interest-rate derivatives. The market models were introduced in Miltersen, Sandmann, and Sondermann (1997) and Brace, Gatarek, and Musiela (1997).

One of the main insights of the market model is that the LIBOR forward rate is a martingale under an appropriately chosen forward risk neutral measure:

$$dL(t; S, T) = \sigma(t, T)L(t; S, T)dW^T(t)$$

where $\sigma(t, T)$ is a volatility loading, which describes how LIBOR forward rates with specific maturities are affected by changes in the Wiener process W^T .

One consequence of the log-normal market model specified above, is that it is consistent with Black's formula, which is used as a market convention to quote cap and floor prices.

Since the market model provides a straightforward way to model interest-rate derivatives it has been adopted by many banks. Therefore a large part of the literature on the market model is based on the pricing of interest-rate derivatives of varying complexity, see Andersen and Andreasen (2000), Andersen and Brotherton-Ratcliffe (2001) and Brigo and Mercurio (2006).

The market models have also been extended to include Lévy processes, see Eberlein and Özkan (2005) and Kluge (2005). Similar to introducing Lévy processes in a forward-rate model, introducing Lévy processes in the market model improves the fit to market data on single trading days. With respect to fitting a model on time-series data, Jarrow, Li, and Zhao (2007) estimate a market model with stochastic volatility and jumps and find that jumps improve the model performance.

Inflation linked securities

In this section we describe the typical inflation linked products traded in the market. A good description of inflation derivatives can also be found in Barclays Capital (2008), which is more exhaustive than the description given in this section. Finally, we refer to the Consumer Price Index (henceforth CPI), which is the price of a consumer basket measured in Dollars, or the representative local currency.

An *Inflation Protected Zero-Coupon bond* is a bond where the payoff at maturity, T , is compounded by the CPI²

$$p_{IP}(T, t_0, T) = \frac{I(T)}{I(t_0)}$$

where $I(T)$ is the value of the CPI at time T . The denominator $I(t_0)$ normalizes the dependence of the CPI, such that the inflation indexation is initiated at the issuance of the bond. The price of an inflation protected bond will be given by the expectation

$$p_{IP}(t, t_0, T) = E_t^Q \left[\exp \left(- \int_t^T n(s) ds \right) \frac{I(T)}{I(t_0)} \right] = p_n(t, T) E_t^T \left[\frac{I(T)}{I(t_0)} \right]$$

²Typically inflation protected bonds are linked to the CPI some months prior to maturity, however we use this simpler specification to enhance the understanding of the product.

where $n(t)$ is the nominal spot rate and $p_n(t, T)$ is the nominal Zero coupon bond with maturity T , observed at time t .

Based on the observed market price of an inflation protected ZCB, we define the real ZCB as

$$p_r(t, T) = \frac{p_{IP}(t, t_0, T)I(t_0)}{I(t)}$$

such that $p_r(T, T) = 1$. Note that the relation above tells us that $p_r(t, T)$ is measured in units of the CPI-basket, and the real bond will give the investor one CPI-basket at time T . Also note that real bonds are derived quantities and thus not directly traded. Finally differences between yields from nominal and real ZCBs are termed *Break Even Inflation Rates*, as it reflects the inflation compensation required by investors.

A *Zero-Coupon Inflation Indexed Swap (ZCIIS)* is a swap agreement where one party pays the percentage change on the CPI over the period $[t, T]$ and the other party pays a fixed amount K . The payoff, at maturity, for the holder of the ZCIIS is then given by

$$\begin{aligned} ZCIIS_T(t, T, K) &= \left(\frac{I(T)}{I(t)} - 1 \right) - \left((1 + K)^{T-t} - 1 \right) \\ &= \frac{I(T)}{I(t)} - (1 + K)^{T-t} \end{aligned}$$

ZCIISs are initiated with a value of zero and are quoted in terms of the fixed payment K , and thus ZCIIS quotes offer a term structure of the expected (risk adjusted) future inflation, also known as *Swap Break Even Inflation Rates*. Although it may not appear so, the pricing of a ZCIIS is completely model independent and only depends on nominal and real zero-coupon bonds.

In terms of modeling real and nominal interest-rates there are two main strands in the literature. First, an approach which is mainly focused on the modeling of prices of inflation derivatives, i.e. both linear and non-linear derivative contracts. Secondly, a macro-finance based view where the main purpose is to extract information from the inflation linked bonds.

Jarrow and Yildirim (2003) were among the first to consider modeling inflation, nominal and real rates in a no-arbitrage framework. By using a forward-rate framework, they derive drift conditions for the CPI, real and nominal forward rates. Hinnerich (2008) extends the results of Jarrow and Yildirim (2003) to include a jump process. Three papers that use a market model framework are Mercurio (2005), Mercurio and Moreni (2006)

and Mercurio and Moreni (2009). Mercurio (2005) use a standard log-normal market model, where Mercurio and Moreni (2006) and Mercurio and Moreni (2009) introduce a volatility smile in their model by adding stochastic volatility.

In terms of the macro-finance view the modeling is typically based on the no-arbitrage relationship between real and nominal pricing kernels

$$M^R(t) = M^N(t)I(t)$$

where M^R is the real stochastic discount factor and M^N is the nominal stochastic discount factor. The no-arbitrage relationship implies, that in a no-arbitrage setting we can 1) model nominal rates and inflation and then infer real rates, 2) model real rates and inflation and infer nominal rates, and 3) model real and nominal rates and infer the CPI. These three approaches are the ones used in the literature, with no consistent pattern on which to prefer. The models based on the macro-finance view are typically based on a short-rate model, sometimes including observable factors such as GDP.

A number of papers analyse inflation markets using TIPS, see Ang, Bekaert, and Wei (2008), D'Amico, Kim, and Wei (2008), Chernov and Mueller (2008) and Christensen, Lopez, and Rudebusch (2008). With regard to Euro Area data, we are aware of three papers, namely Tristani and Hördahl (2007), Garcia and Werner (2010) and Tristani and Hördahl (2010). All papers extract real yields from inflation indexed bonds, and then estimate inflation expectations and inflation risk premia.

Overall, only a few of these studies agree on the size of the inflation risk premia. Some papers have inflation risk premia of up to 300 basis points (Chernov and Mueller (2008)), where others show more moderate fluctuations (-50 to 50 basis points, see for instance Christensen, Lopez, and Rudebusch (2008)). These differences seem to arise from small differences in data periods and the data included, e.g. for instance the inclusion of surveys or not. Finally, only Tristani and Hördahl (2007) present confidence bands on their of estimates inflation risk premia. They find that their estimate of inflation risk premia is statistically insignificant for most of the considered maturities.

Summary

English Summary

Modeling stochastic skewness in a Heath-Jarrow-Morton framework

In this paper we model the stochastic skewness present in interest-rate options by using a Heath-Jarrow-Morton framework and time-changed Lévy processes. The approach is inspired by Carr and Wu (2007) who consider modeling stochastic skewness in currency options.

Most of the term structure modeling literature is focused on capturing stochastic volatility, see for instance Casassus, Collin-Dufresne, and Goldstein (2005), Trolle and Schwartz (2009) and Jarrow, Li, and Zhao (2007). The only paper to consider stochastic skewness is Trolle and Schwartz (2010), who use a Heath-Jarrow-Morton-framework driven by Wiener processes with two stochastic volatility factors. Trolle and Schwartz (2010) are able to generate the skewness implied by the 1-year option on the 10-year swap rate. However, Trolle and Schwartz (2010) acknowledge that their model will understate volatility and skewness for short-term swaptions. Our contribution is to provide a framework which can capture skews, also for short-term interest-rate options. Our calibration to data suggests that the model provides a reasonable fit to the skewness data and that the jump components in the time-changed Lévy processes mainly affect short-term maturities.

The structure of the paper is as follows. First, we show model-free evidence of time-varying skewness in the LIBOR distributions and describe results on the relationship between volatilities and skews at different maturities. Secondly, we use the intuition from these results to specify a model based on a Heath-Jarrow-Morton-framework and time-changed Lévy processes. The model framework allows for semi-analytical solutions of caplet prices and moments of the LIBOR distribution. Finally, by using these model-based moments we calibrate a simple case of the model to time-series of

the model-free volatility and skewness and show that the model is able to capture the volatility and skewness in the data.

Inflation derivatives modeling using time changed Lévy processes

In this paper we consider a consistent no-arbitrage framework for modeling inflation which incorporates both stochastic volatility and jumps in inflation, real and nominal rates.

More precisely, we model inflation derivatives by using the time changed Lévy processes of Carr and Wu (2004) in a Heath-Jarrow-Morton framework, i.e. we consider an extension of the model framework found in Jarrow and Yildirim (2003). Incorporating stochastic volatility into an inflation derivatives model is also considered in Mercurio and Moreni (2006) and Mercurio and Moreni (2009), where Hinnerich (2008) describes the possibility of adding jumps to the model of Jarrow and Yildirim (2003). The paper adds to the existing literature on inflation modeling by providing a model framework which can incorporate both stochastic volatility and jumps, while still being analytically tractable. The modeling framework can also form a good basis for an analysis of time-series of inflation swaps and caps.

The structure of the paper is as follows. First, we briefly describe inflation linked securities, and show evidence of volatility smiles, i.e. non-Gaussian behaviour. Secondly, we describe the framework and derive drift conditions for nominal and real forward rates. Similarly, a drift condition for the consumer price index is found. Thirdly, we show how to price standard inflation derivatives by considering a complex (time dependent) measure. By specifying the subordinator as an affine process, the prices of the considered derivatives can be obtained up to ordinary differential equations and possibly Fourier inversion. Finally, we calibrate our model to market data. Our results show that even though Lévy processes can improve the fit to data, an investigation into the exact specification of the Lévy process and volatility loading is still needed.

Inflation risk premia in the term structure of interest rates: Evidence from Euro area inflation swaps

We consider the estimation of inflation risk premia in the Euro area by using inflation swaps. Our approach is based on a reduced-form no-arbitrage term structure model.

With regard to similar research, Ang, Bekaert, and Wei (2008), D'Amico, Kim, and Wei (2008), Chernov and Mueller (2008) and Christensen, Lopez, and Rudebusch (2008) analyze inflation markets by using TIPS. With regard to Euro Area data, we are aware of three papers, namely Tristani and Hördahl (2007), Garcia and Werner (2010) and Tristani and Hördahl (2010). All papers extract real yields from inflation indexed bonds, and then estimate inflation expectations and inflation risk premia. Only one other paper uses inflation swaps, namely Haubrich, Pennacchi, and Ritchken (2008), who use US inflation swap data. However, in the US, TIPS dominate the inflation linked market, thus having a negative effect on the liquidity of US inflation swaps.

Our contribution to the literature is that we are the first to derive inflation risk premia based on fairly liquid inflation swap data, namely Euro area inflation swaps. Secondly, by using a Bayesian Markov Chain Monte Carlo approach we present confidence intervals for the inflation risk premia and by using this model output we can assess the impact of using surveys on inflation expectations in the identification of inflation risk premia.

The structure of the paper is as follows. First, we examine the relationship between nominal swap rates and inflation swap rates and use this information to specify a no-arbitrage term structure model. Secondly, we estimate the model using Markov Chain Monte Carlo and find that estimates of inflation risk premia on average show an upward sloping term structure, with 1 year risk premia of 18 bps and 10 year risk premia of 43 bps; however, with fluctuations in risk premia over time. Thirdly, our estimates suggest that surveys are important in identifying inflation expectations and thus inflation risk premia. Finally, we relate estimates of inflation risk premia to agents beliefs, and find that skews in short term inflation perceptions drive short term inflation risk premia, where beliefs on GDP growth drive longer term risk premia.

Affine Nelson-Siegel Models and Risk Management Performance

In this paper we assess the ability of the Affine Nelson-Siegel model-class with stochastic volatility to match the observed distributions of Danish Government bond yields.

The ability of affine term structure models to capture interest-rate volatility has recently been discussed in Collin-Dufresne, Goldstein, and Jones (2008), Jacobs and Karoui (2009) and Christensen, Lopez, and Rudebusch (2010). Collin-Dufresne, Goldstein, and Jones (2008) argue that unspanned

stochastic volatility is needed to capture the dynamics of volatility, where Christensen, Lopez, and Rudebusch (2010) argue that their preferred 3-factor affine Nelson-Siegel model can capture the one-month interest-rate volatility reasonably well. Our contribution is to use the models in Christensen, Lopez, and Rudebusch (2010), with respect to both short-term and long-term forecasts. Most other related papers focus on short-term forecasts, i.e. one month. However, in some applications the long-term dynamics of interest-rates are of greater interest. Furthermore, in our estimation we use Danish Government bond yields.

The structure of the paper is as follows. First, we describe the data and describe the 7 different models used in the forecasting exercise. Secondly, based on data from 1987 to 2010 and using a Markov Chain Monte Carlo estimation approach we estimate the 7 different model specifications and test their ability to forecast yields (both means and variances) out of sample. We find that models with 3 CIR-factors perform the best in short term predictions, while models with a combination of CIR and Gaussian factors perform well on 1 and 5-year horizons. Overall, our results indicate that no single model should be used for risk management, but rather a suite of models.

Dansk Resumé

Modellering af stokastisk skævhed i Heath-Jarrow-Morton modelrammen

I dette papir modellerer vi stokastisk skævhed i renteoptioner. Vi modellerer stokastisk skævhed ved at benytte en Heath-Jarrow-Morton-modelramme og tidstransformerede Lévy processer. Vores fremgangsmåde er inspireret af Carr and Wu (2007), der beskriver en model med stokastisk skævhed for at beskrive priser på valutaoptioner.

Store dele af rentestruktur-litteraturen fokuserer på at fange stokastisk volatilitet, se for eksempel Casassus, Collin-Dufresne, and Goldstein (2005), Trolle and Schwartz (2009) og Jarrow, Li, and Zhao (2007). Det eneste papir, der beskriver stokastisk skævhed, er Trolle and Schwartz (2010). Her benyttes en Heath-Jarrow-Morton modelramme baseret på Wiener-processer med to faktorer der genererer stokastisk volatilitet. Trolle and Schwartz (2010) kan generere skævheder der er meget lig den implicitte skævhed fra 1-årige optioner på en 10-årig swap rente. Trolle and Schwartz (2010) bemærker dog, at deres model vil undervurdere skævheden for renteoptioner med kort restløbetid. Vores bidrag er at beskrive en modelramme, der kan beskrive denne skævhed, også for renteoptioner med kort restløbetid. En kalibrering på data viser, at vores model giver en god beskrivelse af skævheden, og at springkomponenterne i de tidstransformerede Lévy processer hovedsageligt påvirker korte løbetider.

Strukturen af papiret er som følger. Først beskriver vi modeluafhængige estimater af tidsvarierende skævhed i fordelingen af LIBOR og sammenhængen mellem volatiliteter og skævheder med forskellige løbetider. Dernæst benytter vi intuitionen fra disse resultater til at specificere en model baseret på en Heath-Jarrow-Morton modelramme og tidstransformerede Lévy processer. Modelrammen tillader semi-analytiske løsninger til priser på capletter og momenter af fordelingen af LIBOR. Endelig, ved at benytte disse modelbaserede momenter, kalibrerer vi modellen baseret på tidsrækker af modeluafhængige estimater af volatilitet og skævhed. Kalibreringen viser, at modellen er i stand til at beskrive volatiliteten og skævheden i data.

Modellering af inflationsderivater ved brug af tidstransformerede Lévy processer

I dette papir beskriver vi en konsistent arbitrage-fri modelramme med både stokastisk volatilitet og spring i inflationen samt reale og nominelle renter.

Mere præcist modellerer vi inflationsderivater ved at benytte de tidstransformerede Lévy processer fundet i Carr and Wu (2004). Disse processer anvendes i en Heath-Jarrow-Morton modelramme. Herved betragter vi en udvidelse af modelrammen i Jarrow and Yildirim (2003). Indarbejdelsen af stokastisk volatilitet i en model, der anvendes til prisning af inflationsderivater, er også beskrevet i Mercurio and Moreni (2006) og Mercurio and Moreni (2009). Hinnerich (2008) beskriver, hvorledes spring kan indarbejdes i den samme Heath-Jarrow-Morton modelramme. Dette papir bidrager til den eksisterende litteratur ved at tilbyde en modelramme, der både kan indarbejde stokastisk volatilitet og spring og samtidig være analytisk tilgængelig. Modelrammen kan også være en god base for en tidsrækkeanalyse af inflationsswaps og -caps.

Strukturen i papiret er som følger. Først beskriver vi inflationsindeksede aktiver og påviser tegn på volatilitetssmil, dvs. ikke-Gaussisk adfærd. Dernæst beskriver vi modelrammen og udleder betingelser for driften for reale og nominelle renter. Tilsvarende udleder vi en betingelse for driften for prisindekset. Vi viser også, hvorledes standard inflationsderivater kan prissættes ved at benytte et komplekst, tidsafhængigt sandsynlighedsmål. Ved at definere tidstransformationen som en affin proces kan vi finde priser ved at løse ordinære differential ligninger og muligvis Fourier inversion. Endelig kalibrerer vi modellen på markedsdata og vores resultater viser, at selv om en model baseret på Lévy processer kan forbedre beskrivelsen af data, vil en nærmere undersøgelse af volatilitetsspecifikationen kunne forbedre modellen yderligere.

Inflationsrisikopræmier i rentestrukturen: Resultater baseret inflationsswaps fra Euro-området

Vi betragter estimationen af inflationsrisikopræmier i Euro-området ved at benytte inflationsswaps. Vores fremgangsmåde er baseret på en statistisk arbitragefri rentestruktur model.

Hvad angår lignende forskning analyserer Ang, Bekaert, and Wei (2008), D'Amico, Kim, and Wei (2008), Chernov and Mueller (2008) og Christensen, Lopez, and Rudebusch (2008) inflationsmarkedet ved at anvende TIPS. Med hensyn til data på Euroområdet kender vi til tre papirer, Tristani and Hördahl (2007), Garcia and Werner (2010) and Tristani and Hördahl (2010). Alle disse papirer anvender reale renter baseret på inflationsindeksede obligationer, og estimerer derefter inflationsforventninger og -risikopræmier. Det eneste andet papir, der benytter inflationsswaps, er Haubrich, Pennacchi, and Ritchken (2008), som benytter inflationsswaps fra

USA. I USA dominerer TIPS det inflationsindekserede marked, og de har derved en negativ effekt på likviditeten i inflationsswaps.

Vores bidrag til litteraturen er, at vi er de første til at estimere inflationsrisikopræmier baseret på likvide inflationsswaps, nemlig inflationsswaps fra Euro-området. Dernæst, ved at benytte en Bayesiansk Markov Chain Monte Carlo metode til at estimere modellen, kan vi præsentere konfidensbånd for inflationsrisikopræmien. Ved at benytte disse modelresultater kan vi analysere effekten af at inkludere spørgeskemaundersøgelser omkring inflationsforventninger på identifikationen af inflationsrisikopræmier.

Strukturen på papiret er som følger. Først undersøger vi sammenhængen mellem nominelle swap-renter og inflationswap-renter, og dernæst benytter vi denne information til at specificere en arbitragefri rentestrukturmodel. Vi estimerer modellen ved at benytte en Markov Chain Monte Carlo metode og finder, at vores estimer på inflationsrisikopræmier er stigende som funktion af løbetiden. Den 1-årige inflationsrisikopræmie er på 18 bps og den 10-årige inflationsrisikopræmie er 43 bps. Begge tidsrækker udviser dog betydelig variation over tid. Vores estimer viser også, at identifikationen af inflationsrisikopræmierne forbedres ved at inkludere spørgeskemaundersøgelser omkring inflationsforventninger. Endelig relaterer vi vores estimer på inflationsrisikopræmier til agents antagelser. Vi finder, at skævheder i den kortsigtede inflationsopfattelse beskriver kortsigtede inflationsrisikopræmier, hvor forventninger til BNP vækst beskriver langsigtede inflationsrisikopræmier.

Affine Nelson-Siegel modeller og risikostyringsperformance

I dette papir analyserer vi de affine Nelson-Siegel modelleres evne til beskrive de observerede fordelinger af danske statsobligationsrenter.

Affine rentestrukturmodellers evne til at beskrive rentevolatiliteten er for nyligt blevet diskuteret i Collin-Dufresne, Goldstein, and Jones (2008), Jacobs and Karoui (2009) og Christensen, Lopez, and Rudebusch (2010). Collin-Dufresne, Goldstein, and Jones (2008) argumenterer for at såkaldt *unspanned stochastic volatility* er nødvendigt for at beskrive rentevolatilitetsdynamikken, hvorimod Christensen, Lopez, and Rudebusch (2010) argumenterer for, at deres fortrukne 3-faktor affine Nelson-Siegel model har en god beskrivelse af rentevolatiliteten én måned frem. Vores bidrag er at anvende modellerne i Christensen, Lopez, and Rudebusch (2010) både til kort- og langsigtede fremskrivninger. De fleste andre papirer fokuserer på kortsigtede fremskrivninger, dvs. én måned, men i nogle anvendelser er

langsigtodynamikken af større interesse. Endelig benytter vi danske statsobligationsrenter i vores estimation.

Strukturen på papiret er som følger. Først beskriver vi data samt de 7 forskellige modeller brugt til at fremskrive med. Dernæst, baseret på data fra 1987 til 2010, og ved at benytte en Markov Chain Monte Carlo metode, estimerer vi de 7 modeller. Disse 7 modeller anvendes til at fremskrive renter (både middelværdier og varianser) *out-of-sample*. Vi finder at modeller med 3 CIR-faktorer klarer sig bedst ved kortsigtede fremskrivninger, hvor modeller med en kombination af CIR- og Gaussiske-faktorer klarer sig bedst på 1- og 5-årige horisonter. Overordnet set indikerer vores resultater, at en samling af modeller, i modsætning til en enkelt model, bør benyttes til risikostyring.

Essay 1

Modeling stochastic skewness in a Heath-Jarrow-Morton framework

Abstract

In this paper we model the stochastic skewness present in interest-rate options. More precisely, we show model-free evidence of time-varying skewness in the LIBOR distributions and use the intuition from these results to specify a model based on a Heath-Jarrow-Morton framework and time-changed Lévy processes. The model framework allows for semi-analytical solutions of caplet prices and moments of the LIBOR distribution. By using these model-based moments, we calibrate a simple case of the model to time-series of the model-free volatility and skewness and show that the model is able to capture the volatility and skewness in the data.

Stochastic skewness, HJM framework, Time-changed Lévy processes, Markov Chain Monte Carlo

JEL Classification: G12, G13, C11, C58

¹Parts of this paper are based on the previous paper 'A tractable Heath-Jarrow-Morton framework based on time changed Lévy processes'. I would like to thank Bjarne Astrup Jensen, Fred Espen Benth, and Anne-Sofie Reng Rasmussen for useful comments.

1.1 Introduction

Several facts on interest rate behaviour are well known. First interest-rate volatilities are obviously stochastic, and these volatilities tend to cluster in periods with low respectively high volatility (see for instance Andersen and Lund (1997)). Another stylized fact is that changes in interest rates and changes in volatility tend to be positively correlated (see Andersen and Lund (1997) and Trolle and Schwartz (2009)). Finally, jumps have also been shown to be an integral part of interest rate dynamics (see Das (2002), Andersen, Benzoni, and Lund (2004) and Johannes (2004)). Furthermore, in a recent analysis of swaption prices, Trolle and Schwartz (2010) also find evidence of stochastic skewness in the probability distributions implied by swaption prices.

A model used in the pricing and risk management of interest-rate dependent assets should ideally capture all of these facts. Indeed, models based on Wiener processes have implemented the time-inhomogeneous behaviour through stochastic volatility processes. This allows for semi-analytical pricing of many interest rate derivatives, which is indeed preferable. Among papers taking this approach can be mentioned Duffie and Kan (1996), Casassus, Collin-Dufresne, and Goldstein (2005), Trolle and Schwartz (2009) and Jarrow, Li, and Zhao (2007).

In terms of capturing stochastic skewness in interest-rates, only Trolle and Schwartz (2010) model the stochastic skewness explicitly. Using a Heath-Jarrow-Morton framework driven by Wiener processes with two stochastic volatility factors, they are able to generate the skewness implied by the 1-year option on the 10-year swap rate. They also find that a model with a single stochastic volatility factor does capture the correct volatility patterns, but not the correct skewness patterns. The main objective in Trolle and Schwartz (2010) is not the modeling of stochastic skewness, but rather explaining the dynamics of medium and long-term swaption distributions. They also acknowledge that their model will understate volatility and skewness for short-term swaptions.

The model in Trolle and Schwartz (2010) is closely linked to the foreign exchange model of Carr and Wu (2007). However, where Carr and Wu (2007) use the time-changed Lévy processes of Carr and Wu (2004), Trolle and Schwartz (2010) only use a special case of the Lévy process, namely the Wiener process.

In this paper we complete the link between Carr and Wu (2007) and Trolle and Schwartz (2010) and specify a Heath-Jarrow-Morton framework with stochastic skewness, where the changes in forward rates are driven by time-

changed Lévy process. To generate stochastic skewness, we need to consider a framework where positive and negative skewness can evolve independently. When using Lévy processes with jumps, distributions showing positive and negative skewness are easily obtained as the skewness is an integral part of the jump process.² The amount of skewness generated by each factor is governed by the activity rates associated with each process.³

First, we show evidence of stochastic skewness in the distribution of Euro area LIBOR-rates. We derive model-free estimates of the standard deviation and skewness present in interest-rate caps and floors. We show that at least two factors should be used to capture the dynamics of the volatility and skewness, as there is a decoupling between short-term and long-term skews.

Secondly, we explicitly formulate a Heath-Jarrow-Morton framework driven by time-changed Lévy processes that can generate stochastic volatility and skewness. We show that we can derive a semi-analytical expression for the characteristic function of log Zero-Coupon Bond prices, which enables us to show how to price interest-rate caps and floors. Finally, we relate the characteristic function to the moments of the LIBOR distribution.

Thirdly, we calibrate our model to standard deviations and skewness measured from mid-2005 to end-2009. Our calibration also shows the decoupling of short-term and long-term skewness and also shows that the jumps in our model mainly affect short-term caps and floors. We also show that the activity rates have natural interpretations with relation to the volatility and skew in the caps and floors.

The structure of the paper is as follows. In section 1.2 we derive and describe the stochastic volatility and skewness present in caps and floors. In section 1.3 we consider our modeling framework, and in section 1.4 we describe the specific version of the model which we use for calibration. Section 1.5 presents the calibration method and section 1.6 shows the results from the calibration. Finally, section 1.7 concludes the paper.

1.2 Evidence of stochastic skewness

In this section we provide evidence of stochastic skewness in interest-rate markets. More precisely, we consider deriving variance and skewness mea-

²When only using Wiener processes, such as in Trolle and Schwartz (2010), the correlations between the forward-rate innovations and the stochastic volatility determine the skewness.

³For a Wiener process an activity rate is equivalent to a stochastic volatility factor.

tures of LIBOR distributions implied by interest-rate caps and floors. First, LIBOR is a discrete rate over the period $[T_0, T_1]$. The rate can be written as

$$L(T_0, T_1) = \frac{1}{T_1 - T_0} \left(\frac{1}{p(T_0, T_1)} - 1 \right)$$

Similarly, a forward contract initiated at time t , but paying interest over the period $[T_0, T_1]$, gives us forward LIBOR

$$L(t, T_0, T_1) = \frac{1}{T_1 - T_0} \left(\frac{p(t, T_0)}{p(t, T_1)} - 1 \right)$$

where by construction $L(T_0, T_0, T_1) = L(T_0, T_1)$.

Consider an interest rate caplet, with cap rate K , fixing at time T_0 and payment at time T_1 . By standard arbitrage arguments we arrive at:

$$\mathcal{C}(t, T_0, T_1, K) = p(t, T_1) E_t^{T_1} [(L(T_0, T_1) - K)^+]$$

Similarly, an interest rate floorlet is defined by:

$$\mathcal{F}(t, T_0, T_1, K) = p(t, T_1) E_t^{T_1} [(K - L(T_0, T_1))^+]$$

Obviously, given that prices are expectations, they contain information on the forward risk adjusted distribution. Furthermore, by construction of the T_1 -forward measure, we have that $E_t^{T_1} [L(T_0, T_1)] = L(t, T_0, T_1)$.

Following Bakshi and Madan (2000), Carr and Madan (2001) and Bakshi, Kapadia, and Madan (2003), a twice differentiable function of $L(T_0, T_1)$, $g(L(T_0, T_1))$, can be written as

$$\begin{aligned} g(L(T_0, T_1)) = & g(Z) + g'(Z)(L(T_0, T_1) - Z) + \int_Z^\infty g''(K)(L(T_0, T_1) - K)^+ dK \\ & + \int_0^Z g''(K)(K - L(T_0, T_1))^+ dK \end{aligned}$$

for any suitable choice of Z . Taking expectations under the T_1 -forward measure, and setting $Z = L(t, T_0, T_1)$, yields the result

$$\begin{aligned} E_t^{T_1} [g(L(T_0, T_1))] = & g(L(t, T_0, T_1)) + \frac{1}{p(t, T_1)} \int_{L(t, T_0, T_1)}^\infty g''(K) \mathcal{C}(t, T_0, T_1, K) dK \\ & + \frac{1}{p(t, T_1)} \int_0^{L(t, T_0, T_1)} g''(K) \mathcal{F}(t, T_0, T_1, K) dK \end{aligned}$$

Thus, by setting $g(K) = K^n$ we can obtain the n 'th non-central moment. For $n = 1, 2, 3$ the expression yields:

$$\begin{aligned} E_t^{T_1} [L(T_0, T_1)] &= L(t, T_0, T_1) \\ E_t^{T_1} [L(T_0, T_1)^2] &= L(t, T_0, T_1)^2 + \frac{2}{p(t, T_1)} \int_{L(t, T_0, T_1)}^{\infty} \mathcal{C}(t, T_0, T_1, K) dK \\ &\quad + \frac{2}{p(t, T_1)} \int_0^{L(t, T_0, T_1)} \mathcal{F}(t, T_0, T_1, K) dK \\ E_t^{T_1} [L(T_0, T_1)^3] &= L(t, T_0, T_1)^3 + \frac{6}{p(t, T_1)} \int_{L(t, T_0, T_1)}^{\infty} K \mathcal{C}(t, T_0, T_1, K) dK \\ &\quad + \frac{6}{p(t, T_1)} \int_0^{L(t, T_0, T_1)} K \mathcal{F}(t, T_0, T_1, K) dK \end{aligned}$$

Finally we can relate the non-central moments to mean, variance and skewness by using the standard relations

$$\begin{aligned} \mu &= L(t, T_0, T_1) \\ \sigma^2 &= E_t^{T_1} [L(T_0, T_1)^2] - \mu^2 \\ \text{skewness} &= \frac{E_t^{T_1} [L(T_0, T_1)^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3} \end{aligned}$$

Data

To extract the market implied variance and skewness we use cap and floor data based on 6M EURIBOR. Our data consists of weekly flat volatility surfaces (sampled on Wednesdays)⁴ and zero-coupon bonds (extracted from LIBOR and swap rates by using bootstrapping), from the June 1st 2005 to December 30th 2009.

The caps are annual caps, i.e. the 1 year cap consists of one caplet, the 2 year cap consists of three caplets, etc. Thus a cap is a portfolio of caplets:

$$\text{Cap}(t, T_N, K) = \sum_{j=1}^N \mathcal{C}(t, T_{j-1}, T_j, K)$$

Without any additional assumptions we cannot extract caplet prices (except for the 1 year cap). To obtain an estimate, we use linear interpolation to get implied volatility estimates for semi-annual maturities, i.e. a 1.5 year cap which consists of two caplets, a 2.5 year cap which consists of four caplets.

⁴A flat volatility is one single volatility which is used to price all the caplets in a cap, using Blacks formula.

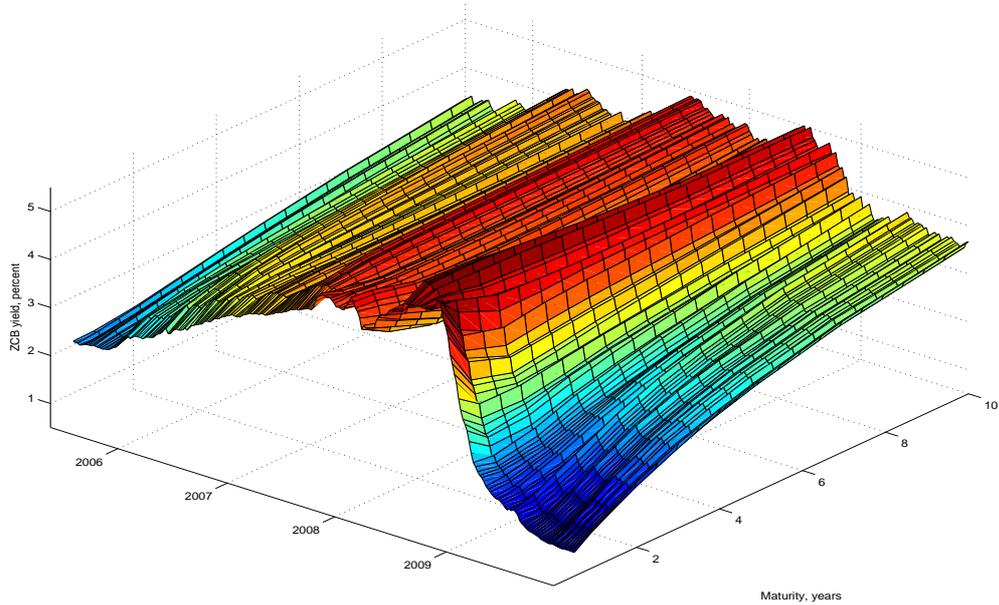


Figure 1.1: **Zero-Coupon Bonds yields.** The yields are extracted from interest rate swaps and LIBOR rates by using bootstrapping. Source: Bloomberg.

Thus using these interpolated volatilities we can get prices for each tenor in the caps, i.e. we can obtain a caplet price as

$$\mathcal{C}(t, T_{s-1}, T_s, K) = \text{Cap}(t, T_s, K) - \text{Cap}(t, T_{s-1}, K)$$

To obtain prices at cap-rates which are not quoted, but are required to obtain the variance and skewness, we use linear interpolation between the flat volatilities and then perform the interpolation in the maturity dimension to obtain the caplet prices. Outside the range of available strikes we extend the linear interpolation. We have considered keeping volatilities constant outside the range of available strikes. Doing so, the derived variance and skewness are slightly more noisy.

Figure 1.2 presents the derived standard deviations. First, as expected we find that the distributions get wider as maturity increases. Furthermore, there appears to be common patterns in the standard deviations. The patterns are confirmed when performing a principal component analysis (PCA) on standard deviations (normalized by the square root of time).⁵ The PCA shows that the first principal component (PC) explains close to 90

⁵Detailed results from the PCA are not reported here, but are available upon request.

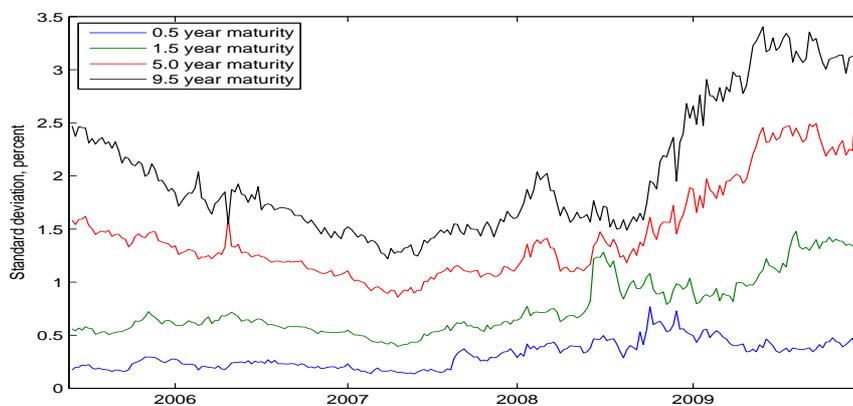


Figure 1.2: **LIBOR distribution standard deviation for different maturities.** The standard deviations are derived from interest-rate cap and floor prices using a model independent approach.

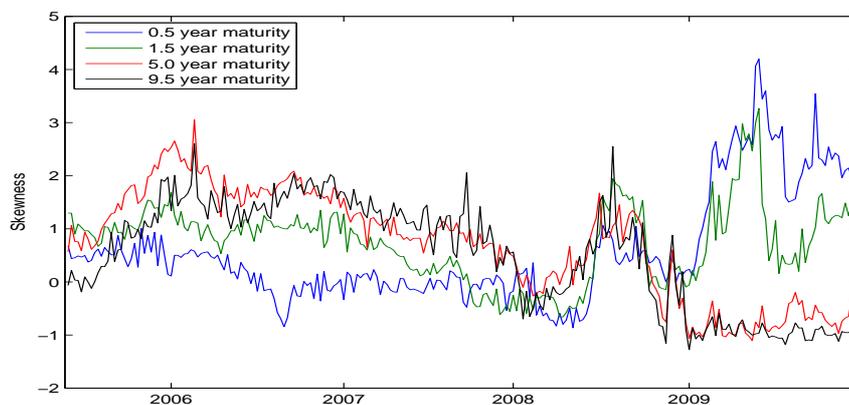


Figure 1.3: **LIBOR distribution skewness for different maturities.** The skewness measures are derived from interest-rate cap and floor prices using a model independent approach.

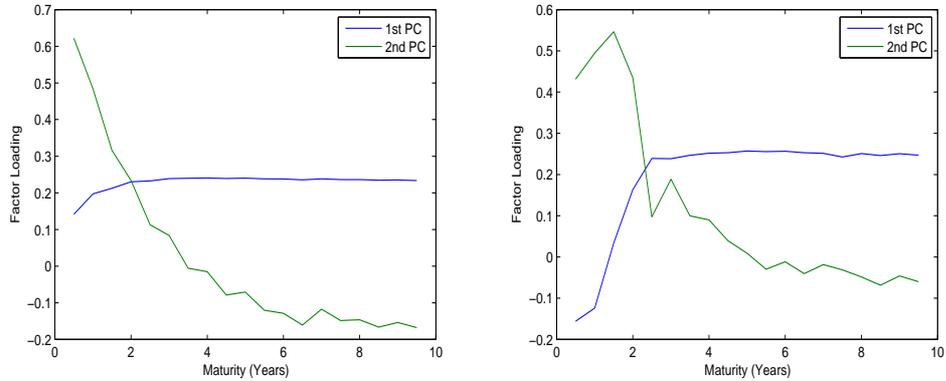


Figure 1.4: **Left:** Factor loadings for standard deviations **Left:** Factor loading for skewness

percent of the variation and the second PC explains close to 8 percent.⁶ The left-hand panel in Figure 1.4 shows the factor loadings from the PCA. The factor loadings are similar to the level and slope curvature factor loadings from a PCA on the term structure of interest-rates, albeit with the difference that the level factor shows a small drop for the short-term maturities.

Figure 1.3 presents the derived skewness measures. First, the skewness is clearly time-varying. Secondly, short- and long-term skewness measures appear to differ in their dynamics. To further examine this, we perform a PCA on the skewness measures. The PCA shows that the first PC explains 78 percent of the variation in data, whereas the second PC explain around 15 percent of the variation in data.⁷ The right-hand panel in Figure 1.4 shows the factor loadings from the PCA. They are similar to the ones from the PCA performed on the standard deviations, albeit with a more pronounced drop in the level factor loading for short-term maturities. The lower expla-

⁶The PCA is performed on the levels of the standard deviations, rather than changes. When using changes the amount of variation described by the first PC is around 53 percent, and for the second PC it is around 18 percent. The factor loadings remain similar to the ones from the PCA performed on Levels. Factor loadings related to the higher order PCs, are quite noisy, and do not show the usual patterns, i.e. level, slope or curvature.

⁷Again, performing the PCA on changes leads to a lower explanation rate from the PCs. The first PC explains 31 percent of the variation in data and the second PC explains 14 percent of the variation in data. The interpretations of the factor loadings remain the same, though they are more noisy. For higher order factor loadings, the factor loadings are quite noisy. We believe that the lower explanation rates, and more noisy factor loadings can be related to the fact than the skewness data is more noisy than the standard deviation data.

nation rates and the more pronounced drop in the factor loadings indicates that at least a two factor model should be used in order to generate the patterns in the data.

1.3 Modeling stochastic skewness

To model the stochastic skewness present in interest-rate options, we use a Heath-Jarrow-Morton framework (see Heath, Jarrow, and Morton (1992)) in combination with time-changed Lévy processes (see Carr and Wu (2004)).

In the following we consider a complete stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q})$. We assume that all processes defined below are adapted to the filtration $\{\mathcal{F}_t\}_{t \geq 0}$.

Time-changed Lévy processes provide a way to model time-inhomogeneous activity rates for Lévy processes. In case of a Wiener process the time-change is equivalent to modeling stochastic volatility and for a compound Poisson process the time-change corresponds to a stochastic intensity.

To model stochastic skewness, consider a positively skewed Lévy process, L^+ , and a negatively skewed Lévy process, L^- . By having two different time-changes, τ^+ and τ^- , both stochastic volatility and skewness can be captured, as the activity rates of the positively and negatively skewed innovations can evolve independently, i.e. when $d\tau^+(t)$ is higher relative to $d\tau^-(t)$ we see a higher tendency toward positively skewed distributions and *vice versa*. This approach was first used by Carr and Wu (2007) to model the dynamics of currency options.

In terms of a HJM framework we incorporate the stochastic skew components directly in the forward-rate process

$$df(t, T) = \alpha(t, T)dt + \sigma(t, T) [dY^+(t) + dY^-((t))]$$

where $Y^\bullet(t) = L^\bullet(\tau^\bullet(t))$ and the time-change is given as the integrated activity rate

$$\tau^\bullet(t) = \int_0^t v^\bullet(s)ds$$

Furthermore, we assume that $\sigma(t, T)$ is a deterministic integrable function in \mathbb{R} and $\alpha(t, T)$ is an adapted integrable process in \mathbb{R} . Additionally, we

assume that each Lévy process has characteristic exponent given as⁸

$$\varphi^\bullet(u) = iua^\bullet - \frac{1}{2}(u\sigma^\bullet)^2 + \int_{\mathbb{R}_0} (e^{iux} - 1) \nu^\bullet(dx)$$

where a^\bullet is the drift of the process, σ^\bullet is the diffusion coefficient of the process and $\nu^\bullet(dx)$ is the Lévy measure, which dictates the arrival rates of jumps.⁹

One difference compared to Carr and Wu (2007) is that we do not compensate the jump part of the Lévy process. In a HJM framework the compensation would be superfluous, as it is incorporated into the drift term $\alpha(t, T)$.

As argued above, and given the results of Litterman and Scheinkman (1991), multiple factors are needed to match the data. We therefore consider a model consisting of J skewness factors:

$$df(t, T) = \alpha(t, T)dt + \sum_{j=1}^J \sigma_j(t, T) [dY_j^+(t) + dY^-(t)]$$

Using this specification log zero-coupon bond prices have dynamics given by (see Björk, Kabanov, and Runggaldier (1997))

$$d \log p(t, T) = (r(t) + A(t, T)) dt + \sum_{j=1}^J S_j(t, T) [dY^+(t) + dY^-(t)]$$

where

$$A(t, T) = - \int_t^T \alpha(t, u) du \quad \text{and} \quad S_j(t, T) = - \int_t^T \sigma_j(t, u) du$$

To complete the initial description of the model framework, we need to derive a drift condition for $A(t, T)$, which so far is only defined to be an integrable process in \mathbb{R} .

By using Itô's lemma we obtain the ZCB dynamics

$$\begin{aligned} \frac{dp(t, T)}{p(t, T)} &= (r(t) + A(t, T)) dt + \sum_{j=1}^J \sum_{p \in \{+, -\}} \left[S_j(t, T) a_j^p + \frac{1}{2} (S_j(t, T) \sigma_j^p)^2 \right] d\tau_j^p(t) \\ &\quad + \sum_{j=1}^J \sum_{p \in \{+, -\}} \left[S_j(t, T) dW_j^p(\tau_j^p(t)) + \int_{\mathbb{R}_0} (e^{S_j(t, T)x} - 1) \mu_j^p(dx, d\tau_j^p(t)) \right] \end{aligned}$$

⁸We assume that the Lévy processes are 1-dimensional. An extension to multivariate Lévy processes is straightforward, although the interpretation of positive and negative skewness is harder for a multivariate process.

⁹Note that we are only considering finite variation jump processes. The results can be extended to infinite variation processes by using the Lévy-Itô decomposition.

Taking expectations yields

$$\begin{aligned} \mathbb{E}_t^{\mathbb{Q}} \left[\frac{dp(t, T)}{p(t, T)} \right] &= \left(r(t) + A(t, T) + \sum_{j=1}^J \sum_{p \in \{+, -\}} \left[S_j(t, T) a_j^p + \frac{1}{2} (S_j(t, T) \sigma_j^p)^2 \right] v_j^p(t) \right) dt \\ &\quad + \sum_{j=1}^J \sum_{p \in \{+, -\}} \int_{\mathbb{R}_0} (e^{S_j(t, T)x} - 1) \nu_j^p(dx) v_j^p(t) dt \end{aligned}$$

Standard no-arbitrage arguments imply that the expected drift of the ZCB should be equal to the short rate, $r(t)$. Recognizing the Lévy exponent gives us the drift condition:

$$A(t, T) = - \sum_{j=1}^J [v_j^+(t) \varphi_j^+(-iS_j(t, T)) + v_j^-(t) \varphi_j^-(-iS_j(t, T))]$$

The result above yields an extension to the results found in Eberlein and Raible (1999). Furthermore, compared to the drift condition when we allow $\sigma(t, T)$ to be stochastic, our expression is much simpler (see Raible (2000) and Filipovic and Tappe (2008)).¹⁰

Deriving the characteristic function

Ultimately we are interested in calibrating our model to data. One important variable when using Lévy processes and models with stochastic volatility is the characteristic function. The characteristic function allows for calculation of the model implied distribution. This distribution can be used to calculate option prices and moments in the LIBOR distributions.

Using the dynamics stated above we can express the ZCB price as

$$\begin{aligned} p(t, T) &= p(0, T) \exp \left(\int_0^t (r(s) + A(s, T)) ds \right. \\ &\quad \left. + \sum_{p \in \{+, -\}} \sum_{j=1}^J \int_0^t S_j(s, T) dY_j^p(s) \right) \end{aligned}$$

¹⁰Both Raible (2000) and Filipovic and Tappe (2008) find that the drift condition is given by

$$A(t, T, \omega) = -\varphi(-iS(t, T, \omega))$$

where $\omega \in \Omega$. Due to the non-linear form of the Lévy exponent, this approach will be unlikely to be used in practice, except in a few special cases.

which alternatively can be written as

$$p(t, T) = \frac{p(0, T)}{p(0, t)} \exp \left(\int_0^t (A(s, T) - A(s, t)) ds + \sum_{p \in \{+, -\}} \sum_{j=1}^J \int_0^t (S_j(s, T) - S_j(s, t)) dY_j^p(s) \right)$$

Using this expression, we calculate the characteristic function

$$\psi_t(u, T_0, T_1, T_k) = E_t^{T_k} [\exp(iu \log p(T_0, T_1))]$$

As shown in Proposition 1 the expectation involved in calculating the characteristic function can be simplified significantly, leaving us with a relatively simple expression:

Proposition 1. *Suppose $\xi_j(t, u)$ is a continuous, bounded function for all $j = 1, \dots, J$, and let $\xi_j(s, u) \in \vartheta_j$ for all $s : t \leq s \leq T_0$ and fixed u , where ϑ_j is the set of values where the characteristic exponents of the Lévy processes L_j^+ and L_j^- are finite.*

Then the characteristic function of the log zero-coupon bond price under the T_k -forward measure

$$\psi_t(u, T_0, T_1, T_k) = E_t^{T_k} [\exp(iu \log p(T_0, T_1))]$$

is given by the following expectation

$$\psi_t(u, T_0, T_1, T_k) = \left(\frac{p(t, T_1)}{p(t, T_0)} \right)^{iu} E_t^\xi \left[\exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \int_t^{T_0} g_j^p(s, u) v_j^p(s) ds \right) \right]$$

where $E_t^\xi [\bullet]$ denotes expectation with regards to the probability measure defined by

$$\frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \Big|_{\mathcal{F}_t} = \exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \int_0^t \xi_j(s, u) dY_j^p(s) - \int_0^t \varphi_j^p(-i\xi_j(s, u)) v_j^p(s) ds \right)$$

and

$$g_j^p(t, u) = iu[\varphi_j^p(-iS_j(t, T_0)) - \varphi_j^p(-iS_j(t, T_1))] - \varphi_j^p(-iS_j(t, T_k)) + \varphi_j^p(-i\xi_j(t, u))$$

$$\xi_j(t, u) = iu[S_j(t, T_1) - S_j(t, T_0)] + S_j(t, T_k)$$

where $\varphi_j^p(u)$ is the characteristic exponent of the Lévy process L_j^p under the measure \mathbb{Q} .

Proof. See Appendix 1.8. □

The result above does not give an explicit expression for the characteristic function, since we are yet to define the processes v_j^p . The result is encouraging, since for many specifications the expectation will have a tractable solution. The solution may not be analytical, but is typically semi-analytical, i.e. up to solving a set of ordinary differential equations.

We next turn to two applications of the characteristic function. First, we show how to price caplets using Fourier inversion and secondly, we show how to use the characteristic function to get model-implied moments of the LIBOR distribution.

Pricing interest-rate caplets

In this section we briefly describe the results on using Fourier inversion for pricing derivative contracts. For more precise details we refer to Kluge (2005) or Raible (2000).

As mentioned above an interest rate caplet is an option on LIBOR. It can be shown that an interest rate caplet can be related to a put-option on a ZCB:

$$\mathcal{C}(t, T_0, T_1, K) = (1 + (T_1 - T_0)K) \mathbf{PZCB} \left(t, T_0, T_1, \frac{1}{1 + (T_1 - T_0)K} \right)$$

where $\mathbf{PZCB}(t, T_0, T_1, K)$ is the time t price of a European put-option with maturity T_0 on a ZCB maturing at T_0 , where the option has exercise price K . To price this option, thus getting semi-analytical prices for caps, we use Fourier inversion.

In general, Raible (2000) shows that the price of a European option can be obtained through Fourier inversion, by the knowledge of the characteristic function of the (log-)underlying asset and the Fourier transformed contract function. In the case of European call and put options, the transforms are in functional form identical; however, they differ in the range of permitted values for the integration strip of the Fourier integral.

The conclusion from Raible (2000) and Kluge (2005) is that the price can be obtained by calculating the integral given in Proposition 2:

Proposition 2. *Assume that there exists a $\beta > 0$ such that $\psi_t(i\beta, T_0, T_1, T_0) < \infty$. Then the price of a European put option with maturity T_0 and strike K on a zero coupon bond maturing at time T_1 is given by*

$$\mathbf{PZCB}(t, T_0, T_1, K) = \frac{p(t, T_0)}{\pi} \int_0^\infty \Lambda(\beta + iu, K) \psi_t(i\beta - u, T_0, T_1, T_0) du$$

where

$$\Lambda(v, K) = \frac{e^{(1+v) \log K}}{v(v+1)}$$

Moments of the LIBOR distribution

In this section we consider model-based moments of the distribution of LIBOR. We show that the moments of the LIBOR distribution can be obtained up to simple evaluations of the characteristic function.

The expression for LIBOR is

$$L(T_0, T_1) = \frac{1}{T_1 - T_0} \left(\frac{1}{p(T_0, T_1)} - 1 \right)$$

We consider the moments under the T_1 -forward measure. The first moment follows by construction of the T_1 -forward measure

$$E_t^{T_1} [L(T_0, T_1)] = \frac{1}{T_1 - T_0} \left(\frac{p(t, T_0)}{p(t, T_1)} - 1 \right) = L(t, T_0, T_1)$$

The second non-central moment is given by

$$E_t^{T_1} [L(T_0, T_1)^2] = \left(\frac{1}{T_1 - T_0} \right)^2 \left[1 + E_t^{T_1} [e^{-2 \log p(T_0, T_1)}] - 2[1 + (T_1 - T_0)L(t, T_0, T_1)] \right]$$

The third Non-Central moment is given by

$$E_t^{T_1} [L(T_0, T_1)^3] = \left(\frac{1}{T_1 - T_0} \right)^3 \left[E_t^{T_1} [e^{-3 \log p(T_0, T_1)}] - 3E_t^{T_1} [e^{-2 \log p(T_0, T_1)}] + 3[1 + (T_1 - T_0)L(t, T_0, T_1)] - 1 \right]$$

Finally, we can relate the non-central moments by using the standard relations

$$\begin{aligned} \mu &= L(t, T_0, T_1) \\ \sigma^2 &= E_t^{T_1} [L(T_0, T_1)^2] - \mu^2 \\ \text{skewness} &= \frac{E_t^{T_1} [L(T_0, T_1)^3] - 3\mu\sigma^2 - \mu^3}{\sigma^3} \end{aligned}$$

Notice that all expectations of the form $E_t^{T_1} [e^{-u \log p(T_0, T_1)}]$ are simply evaluations of the characteristic function given in Proposition 1, making it easy to get model-based moments.

1.4 Specifying the Model

In this section we focus on specifying the model. We focus on the considered Lévy processes, activity rates and volatility loadings. Finally, we only consider a model with two skewness factors, i.e. $J = 2$.

Specifying the Lévy process

We now turn to the specification of the Lévy process. We emphasize that we only consider one specification; many different specifications are indeed possible.

With respect to the specification of the Lévy process, Carr and Wu (2007) find that a model based on a diffusion component and a compound Poisson process with exponentially distributed jump sizes performs similar to more advanced (infinite activity) models.

Inspired by these results we consider the Lévy processes

$$L_j^+(t) = \sigma_j dW_j^+(t) + \sum_{n=1}^{N_j^+(t)} Q_j^+(t)$$

$$L_j^-(t) = \sigma_j dW_j^-(t) + \sum_{n=1}^{N_j^-(t)} Q_j^-(t)$$

where N_j^+ and N_j^- are Poisson-processes both with intensity λ_j , $Q_j^+(t)$ are positive jumps, such that $Q_j^+(t)$ follows an exponential distribution with parameter $1/\delta_j$ and $Q_j^-(t)$ are negative jumps, such that $-Q_j^-(t)$ follows an exponential distribution with parameter $1/\delta_j$.¹¹ Finally, all Wiener processes, Poisson processes and jumps are independent.

¹¹By using this specification δ_j corresponds to the expected jump size.

The characteristic exponents are then given by

$$\begin{aligned}\varphi_j^+(u) &= -\frac{1}{2}(u\sigma_j)^2 + \lambda_j \left[\left(\frac{1/\delta_j}{1/\delta_j - iu} \right) - 1 \right] \\ \varphi_j^-(u) &= -\frac{1}{2}(u\sigma_j)^2 + \lambda_j \left[\left(\frac{1/\delta_j}{1/\delta_j + iu} \right) - 1 \right]\end{aligned}$$

Using these specifications time-varying skewness is solely captured by the activity rate-processes v_j^p . For instance, if we only considered one activity rate process, the innovation distribution would be symmetric.

Specifying the activity rates

We also need to specify the activity rates of the time-changes, i.e. the stochastic processes which drive the stochastic volatility and skewness.

Here we consider square-root processes. More precisely, each Lévy process L_j^p has its own activity rate v_j^p . The activity rate has dynamics given by

$$dv_j^p(t) = (\theta_j - \kappa_j v_j^p(t)) + \eta_j^p \sqrt{v_j^p(t)} dZ_j^p(t)$$

where Z_j^p is a Wiener process with $[dW_j^p(t), dZ_j^p(t)] = \rho_j^p dt$, but otherwise independent of other Wiener processes in the model. For identification we set η_j^p equal to one.

Using this specification the characteristic function will be given as (see for instance Duffie, Pan, and Singleton (2000) or Duffie, Filipovic, and Schachermayer (2003) for a derivation):

$$\psi_t(u, T_0, T_1, T_k) = \exp \left(A(t, T) + \sum_{j=1}^J [B_j^+(t, T) + B_j^-(t, T)] \right)$$

where the functions $A(t, T)$, $B_j^+(t, T)$ and $B_j^-(t, T)$ solve the ODEs

$$\begin{aligned}\frac{dA(t, T)}{dt} &= \sum_{j=1}^J \theta_j [B_j^+(t, T) + B_j^-(t, T)] \\ \frac{dB_j^+(t, T)}{dt} &= -(\rho_j^+ \sigma_j \xi_j(t, u) - \kappa_j) B_j^+(t, T) - \frac{1}{2} B_j^+(t, T)^2 - g_j^+(t, u) \\ \frac{dB_j^-(t, T)}{dt} &= -(\rho_j^- \sigma_j \xi_j(t, u) - \kappa_j) B_j^-(t, T) - \frac{1}{2} B_j^-(t, T)^2 - g_j^-(t, u)\end{aligned}$$

where $A(T, T) = B_j^+(T, T) = B_j^-(T, T) = 0$ for all j . One noticeable element in the ODEs, are the first terms, $\rho_j^+ \sigma_j \xi_j(t, u)$ and $\rho_j^- \sigma_j \xi_j(t, u)$. These terms reflect the change of measure from \mathbb{Q} to \mathbb{Q}^ξ and correct for the fact that the activity rate processes are correlated with the Lévy processes.

Specifying the volatility loadings

To complete the model, we need to specify the volatility loadings, $\sigma_j(t, T)$. The volatility loadings need to be sufficiently flexible so that they can capture the volatility patterns shown in data. For instance Dai and Singleton (2002) show that unconditional volatilities exhibit a hump-shape, and based on time series estimations using both yields and interest-rate derivatives Trolle and Schwartz (2009) show that a hump-shaped volatility loading is preferable.

Following the results in Trolle and Schwartz (2009) we specify the volatility loadings as

$$\sigma_j(t, T) = (1 + \beta_j(T - t)) e^{-\gamma_j(T-t)}$$

This implies that the impact on the bond price $S_j(t, T)$ is given by

$$S_j(t, T) = \left(\frac{e^{-\gamma_j(T-t)} - 1}{\gamma_j} \right) - \beta_j \left(\frac{1 - e^{-\gamma_j(T-t)}(1 + \gamma_j(T - t))}{\gamma_j^2} \right)$$

This specific shape can generate flat volatility loadings when $\beta_j = 0$ and $\gamma_j \rightarrow 0$, factor loadings as in a Vasicek-model when $\beta_j = 0$, and when all parameters are free we get a flexible hump-shaped volatility loading. Finally, notice that we do not introduce a scaling of the volatility loading as in Trolle and Schwartz (2009), as this scaling is already included into the Lévy process.

1.5 Model calibration

To calibrate the model we use the derived standard deviations and skewness measures directly. Alternatively we could estimate the model using the cap and floor data directly. This latter approach is similar to Carr and Wu (2007), but is computationally expensive compared to directly calibrating to standard deviations and skewness measures. We acknowledge, that we miss out on the finer information in caps and floors, e.g., fat tails, where the advantage of a jump component would be further emphasized, but we consider our calibration method to be sufficient to show the applicability of

the modeling framework.¹² Finally we use a Bayesian Markov Chain Monte Carlo method to handle non-linearities and positivity constraints on the activity rates.

We express the model in non-linear state-space form, i.e. we have the measurement equation

$$D_t = f(v_t|\Theta) + \varepsilon_t$$

where D_t is our data, i.e. an n -vector, where the first half of the elements are standard deviations and the last half of the elements are the skewness measures.¹³ $f(v_t|\Theta)$ is a function that relates the activity rates v_t , given the model parameters Θ , to standard deviations and skewness measures. The entire collection of states $\{D_t\}_{t=1}^T$ is denoted by D . ε_t is an n -vector and is an error term, where the elements of error terms are independent. The error terms are normally distributed with each element having its own standard error:

$$[\varepsilon_t]_j \sim \mathcal{N}(0, \sigma_{\varepsilon,j}^2)$$

Finally, to complete the model we must consider state dynamics and risk premia. We consider the simplest possible risk premia, the completely affine risk premia. We thus change the mean-reversion speed of the activity rates, giving us the state dynamics (after using an Euler approximation)

$$v_{j,t+1}^p = v_{j,t}^p + (\theta_j - \hat{\kappa}_j v_{j,t}^p) \Delta t + \sqrt{v_{j,t}^p \Delta t} \varepsilon_t^v$$

where ε_t^v is standard normal. Finally, the entire collection of activity rates (across j and p) are given by v_t and the entire time series of activity rates, $\{v_t\}_{t=1}^T$, are given by v .

With regard to parameters we denote 1) the parameters that only enter into standard deviations and skewness measures by $\Theta^Q = (\sigma_j, \lambda_j, \delta_j, \kappa_j, \gamma_j, \beta_j)'$, 2) parameters that enter into both state dynamics and observations as $\Theta^{QP} = (\theta_j)'$ and 3) parameters that only enter into state dynamics as $\Theta^P = (\hat{\kappa}_j)'$.

¹²This calibration method could for instance also be used to get good initial parameters and states for an estimation based on caps and floors.

¹³We do not include yields in the calibration. It is possible (see Trolle and Schwartz (2010)), but we prefer to focus on fitting stochastic volatilities and skews. By not using yields, our approach is similar to Carr and Wu (2007) who do not include data on exchange rates.

Estimation using MCMC

When estimating the model we are interested in sampling from the target distribution of parameters and state variables, $p(\Theta, v|D)$. To sample from this distribution the Hammersley-Clifford theorem (Hammersley and Clifford (1974) and Besag (1974)) implies that this can be done by sampling from the complete conditionals

$$\begin{aligned} p(\Theta^Q|\Theta_{\setminus Q}, D, v) \\ \vdots \\ p(v|\Theta, D) \end{aligned}$$

Thus MCMC handles the sampling from the complicated target distribution $p(\Theta, v|D)$, by sampling from the simpler conditional distributions. More specifically, this is handled by sampling in cycles from the conditional distributions. If one can sample directly from the conditional distribution, the resulting algorithm is called a Gibbs sampler (see Geman and Geman (1984)). If it is not possible to sample from this distribution one can sample using the Metropolis-Hastings algorithm (see Metropolis, Rosenbluth, Rosenbluth, Teller, and E. (1953)). In this paper we use a combination of the two (a so-called hybrid MCMC algorithm) since not all the conditional distributions are known. More precisely, we have the following MCMC algorithm:

$$\begin{aligned} p(v|\Theta, D) &\sim \text{Random Block-Size Metropolis-Hastings} \\ p(\Theta^Q|\Theta_{\setminus Q}, v, D) &\sim \text{Metropolis-Hastings} \\ p(\Theta^{QP}|\Theta_{\setminus QP}, v, D) &\sim \text{Metropolis-Hastings} \\ p(\Theta^P|\Theta_{\setminus P}, v, D) &\sim \text{Metropolis-Hastings} \\ p(\sigma_\varepsilon|\Theta_{\setminus \sigma_\varepsilon}, v, D) &\sim \text{Inverse Gamma} \end{aligned}$$

A more precise description of the algorithm and the conditional distributions is found in Appendix 1.9. Furthermore, when sampling activity rates we use a random-block sampler, where the size of the random block is Poisson distributed (with parameter 10). For each draw of the parameters we perform 5 state draws. These two modifications significantly improve the convergence of the Markov Chain compared to univariate single state sampling. We also impose priors on the intensities λ_j , such that they are assumed to be normal with mean 0.5 and variance 1.0 (draws below zero are rejected).

The Markov chain is run for 10 million simulations, where the standard errors of the Random Walk Metropolis-Hasting algorithms are calibrated

	γ_j	β_j	σ_j	λ_j
$j = 1$	0.3167 (0.1224 , 0.5161)	4.5209 (3.8086 , 5.3152)	0.0022 (0.0019 , 0.0024)	0.8486 (0.0096 , 2.4272)
$j = 2$	5.2777 (2.8368 , 8.1037)	51.5802 (36.9804 , 71.9038)	0.013 (0.0095 , 0.0161)	0.8656 (0.0088 , 2.4518)
	δ_j	θ_j	κ_j	$\hat{\kappa}_j$
$j = 1$	1.399 (0.0666 , 5.8175)	0.0082 (0.0002 , 0.0193)	0.1805 (0.1563 , 0.1966)	0.1958 (0.0071 , 0.5316)
$j = 2$	4.8226 (0.1809 , 18.9201)	0.3244 (0.0586 , 0.6256)	55.5035 (55.4882 , 55.5247)	0.4027 (0.0327 , 0.8569)
	ρ_1^+	ρ_1^-	ρ_2^+	ρ_2^-
	0.4188 (0.3432 , 0.5)	-0.8877 (-0.8996 , -0.8546)	0.3095 (0.0414 , 0.5679)	-0.0503 (-0.0946 , -0.0073)

Table 1.1: **Parameters Estimates.** Parameter estimates are based on the mean of the MCMC samples. 95 pct. confidence intervals based on MCMC samples are reported in brackets. δ_1 and δ_2 are multiplied by 10,000 and are thus reported in basis points

to yield acceptance probabilities between 10 and 40 pct. Finally we save each 1000th draw and use an additional 2 million simulations of the chain, leaving 2000 draws for inference.

1.6 Results

In this section we consider the results from the calibration of the model.

Table 1.1 shows the estimated parameters and Figure 1.5 shows root mean squared errors (RMSEs) arising from our calibration procedure. Rather than interpreting directly on the estimated parameters, we use model output such as factor loadings and impact on yields in case of a jump.

Considering the RMSE we see that the RMSEs are highest for short term maturities, indicating that an additional third skewness component that captures these maturities might be beneficial. The RMSE for the skewness is around 1.0, which compared to Figure 1.3 seems a bit large.

However, when considering Figure 1.6, which shows model- and market-based standard deviations and skewness measures, the model captures the overall pattern in the skewness, albeit with some deviations. These deviations happen when the medium-term (1.5 year) skewness slightly decouples

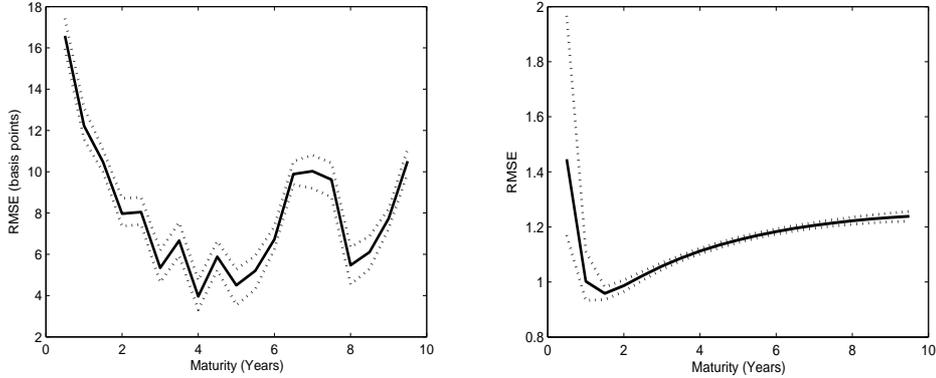


Figure 1.5: **Left:** Root Mean Squared Errors of fit to standard deviations. Root Mean Squared Errors are reported in basis points. Dashed lines indicate 95 percent confidence bands based on MCMC samples. **Right:** Root Mean Squared Errors of fit to skewness measure. Dashed lines indicate 95 percent confidence bands based on MCMC samples.

from the short-term skewness (0.5 year). Again, this figure indicates that an additional skewness factor could relieve some of this tension. With regard to the standard deviations, we see a reasonable description of the market-based standard deviations.

Next we consider the finer aspects of the model structure. The left-hand panel in Figure 1.7 shows the factor loadings implied by the model. We define the factor loadings as

$$FL_j(T-t) = -\frac{S_j(t, T)}{T-t}$$

i.e. the impact of a yield with maturity $T-t$. We see that the factor loadings are similar to the ones found in Figure 1.4, i.e. that the first factor loadings show a level-component for long-term maturities, but drop for short-term maturities. The second factor loading has a significant hump at short-term maturities, which quickly dies out and does not affect long-term yields.

Another interesting issue is to relate the estimated jump sizes to effects on the yield curve. The right-hand panel in Figure 1.7 shows the impact on yields, with different maturities of a jump of size δ_j in the specific factor. Jumps in the first factor appear to be fairly moderate with effects on long-term rates of around 6 basis points. Jumps in the second factor are more pronounced, with effects of around 17 basis points in short-term yields, but due to the shape of the factor loading it has virtually no effect on long-term yields.

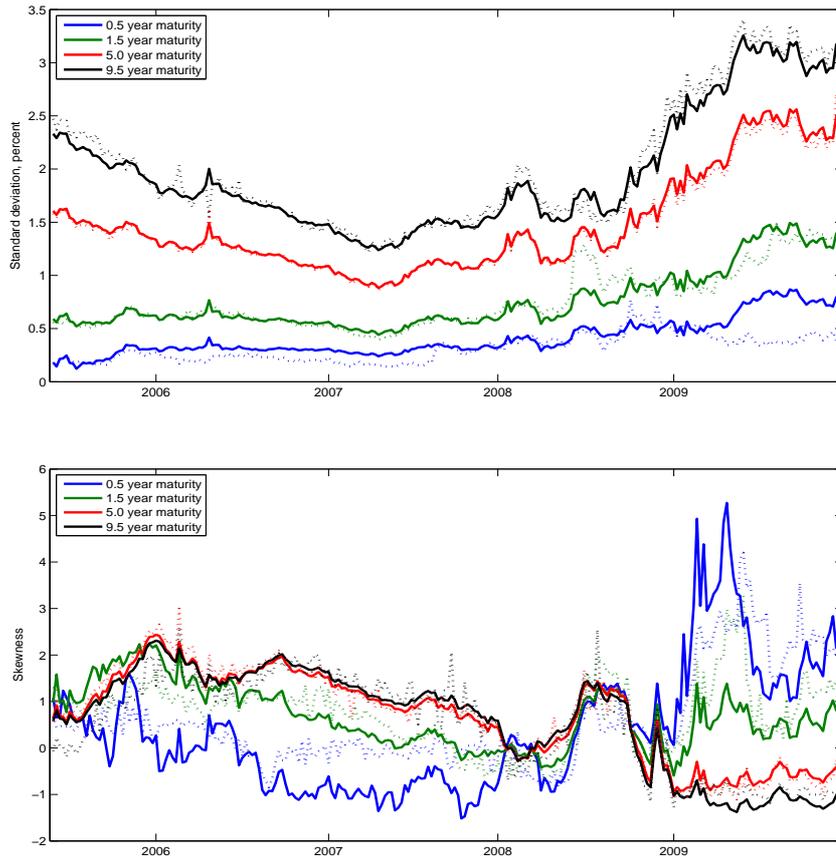


Figure 1.6: **Top:** Model and market standard deviation. Solid lines are model based standard deviations and dashed lines are standard deviations derived from market data. **Bottom:** Model and market skewness. Solid lines are model based skewness measures and dashed lines are skewness measures derived from market data.

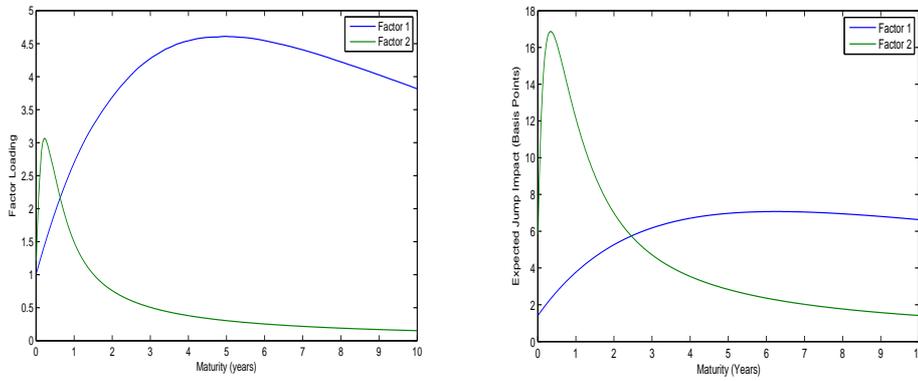


Figure 1.7: **Left:** Factor loadings on the two skewness factors. The factor loadings are based on the mean across MCMC samples. **Right:** Expected impact of a jump in the compound Poisson processes. The expected jump impact is based on the mean across MCMC samples.

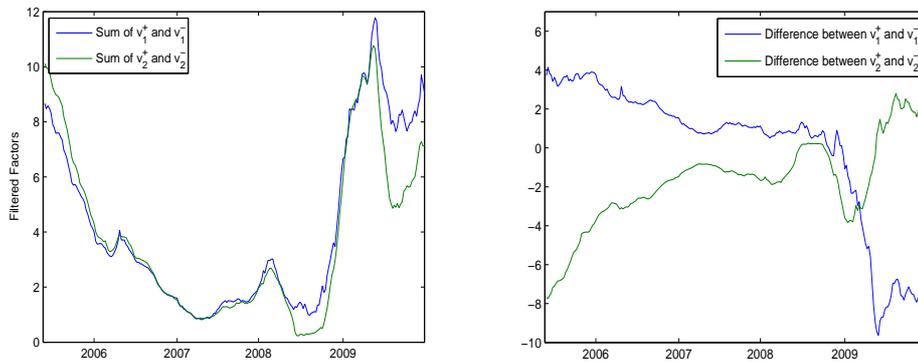


Figure 1.8: **Left:** Sum of the two factors driving each skewness component. **Right:** Difference between the two factors driving each skewness component.

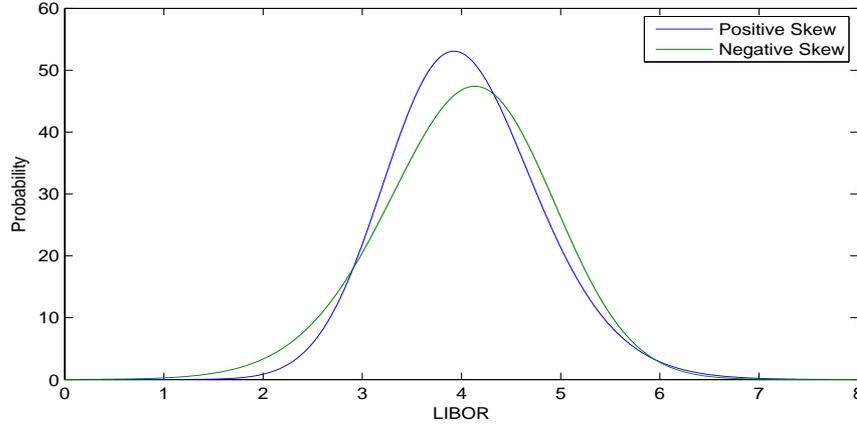


Figure 1.9: **Model implied LIBOR densities for different activity rate values.** Model implied LIBOR densities are derived from model-based caplet prices of a 6-month caplet. The skewness of the positively skewed distribution is 0.31 and -0.22 for the negatively skewed distribution.

In terms of interpreting the dynamics of the activity rates, the size of v_j^+ relative to v_j^- generates the stochastic skewness. Thus the difference, $v_j^+ - v_j^-$ is a model based proxy of the skewness induced by skewness component j . Similarly, since the Lévy processes are independent and variances are additive, the sum, $v_j^+ + v_j^-$, is a proxy for the volatility.

Figure 1.8 shows the sum of the activity rates, $v_j^+ + v_j^-$, and the difference between the activity rates, $v_j^+ - v_j^-$. The sum of the factors show similar patterns for the two stochastic skewness components, and they are very similar to the volatilities in Figure 1.2. The fact that the two time series are so similar, is also consistent with Figure 1.2 and the results in the PCA. The differences in the activity rates are more interesting, and show significantly different patterns for the two skewness components. For the first factor it corresponds (roughly) to the skewness derived from long-term caplets. For the second factor the time series corresponds to short-term skewness. The small values at the start of the data sample reflect that the difference, $v_1^+ - v_1^-$ is high, and to compensate for this $v_2^+ - v_2^-$ is low ¹⁴.

Finally we consider the effect of different choices of activity rates on the model-implied probability densities (PDFs). Figure 1.9 shows the model-implied LIBOR PDFs for different activity rates. We construct the PDF as follows; using Proposition 10 we calculate caplet prices for various strikes

¹⁴Recall that the factor loading for factor 1, is not zero for short term maturities

and then use the results of Breeden and Litzenberger (1978) to obtain the PDF via numerical differentiation. Furthermore, we assume a flat yield curve at 4 percent and when constructing the positively skewed PDF we set the 'positive' activity rates equal to 3 and the 'negative' activity rates equal to 0. When constructing the negatively skewed PDF we do the opposite.

It is evident that the variation in the activity rates does, indeed, produce different PDFs. The difference in skew is clearly visible, with a significantly heavier left tail in the PDF with the negative skews. Such an asymmetric heavy tail could have significant impact on the pricing of out-of-the-money floorlets relative to out-of-the-money caplets.

1.7 Conclusion

In this paper we have shown how to model stochastic skewness present in interest-rate options.

By using interest-rate caps and floors we have derived model-free estimates of the volatility and skewness of LIBOR. We have shown that at least two factors should be used to capture the dynamics of the volatility and skewness, as there is a decoupling between short-term and long-term skews.

We explicitly model the skewness in Heath-Jarrow-Morton framework, where innovations in the forward-rates are driven by time-changed Lévy processes. By using Lévy processes we can easily obtain positively and negatively skewed distributions, and the activity rates offer a simple interpretation as the amount of negative and positive skew at each time-point.

Our calibration shows that the model is able to capture the stylized facts in the data. There is still room for improvement for the skewness measures, but overall our results are encouraging.

Since we only considered one specification of the model, potential future research areas could be the more precise specification of the model along with a more consistent time-series estimation using yields, caps and floors directly.

1.8 Appendix: Proof of proposition 1

Given the log-ZCB expression

$$\begin{aligned} \log p(T_0, T_1) &= \log \left(\frac{p(t, T_0)}{p(t, T_1)} \right) + \sum_{j=1}^J \sum_{p \in \{+, -\}} \int_t^{T_0} [S_j(s, T_1) - S_j(s, T_0)] dY_j^p(s) \\ &\quad \sum_{j=1}^J \sum_{p \in \{+, -\}} \int_t^{T_0} [\varphi_j^p(-iS_j(s, T_0)) - \varphi_j^p(-iS_j(s, T_1))] v_j^p(s) ds \end{aligned}$$

we are interested in the expectation

$$\psi_t(u, T_0, T_1, T_k) = \mathbb{E}_t^{T_k} [e^{iu \log p(T_0, T_1)}]$$

To change the measure from the T_k -forward measure to the risk neutral measure \mathbb{Q} , consider the Radon-Nikodym derivative

$$\begin{aligned} \left. \frac{d\mathbb{Q}^{T_k}}{d\mathbb{Q}} \right|_{\mathcal{F}_t} &= N(t) = \frac{p(t, T_k)}{\beta(t)} \frac{\beta(0)}{p(0, T_k)} \\ &= \exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \left[\int_0^t S_j(s, T_k) dY_j^p(s) - \int_0^t \varphi_j^p(-iS_j(s, T_k)) v_j^p(s) ds \right] \right) \end{aligned}$$

then

$$\begin{aligned} \psi_t(u, T_0, T_1, T_k) &= \mathbb{E}_t^{T_k} [e^{iu \log p(T_0, T_1)}] = \mathbb{E}_t^{\mathbb{Q}} \left[e^{iu \log p(T_0, T_1)} \frac{N(T_0)}{N(t)} \right] \\ &= \left(\frac{p(t, T_1)}{p(t, T_0)} \right)^{iu} \mathbb{E}_t^{\mathbb{Q}} \left[\exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} h_j^p(u, s) v_j^p(s) ds \right) \right. \\ &\quad \left. \times \exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \xi_j(u, s) dY_j^p(s) \right) \right] \end{aligned}$$

where

$$\begin{aligned} h_j^p(u, s) &= iu [\varphi_j^p(-iS_j(s, T_0)) - \varphi_j^p(-iS_j(s, T_1))] + \varphi_j^p(-iS_j(s, T_k)) \\ \xi_j(u, s) &= iu [S_j(s, T_1) - S_j(s, T_0)] + S_j(s, T_k) \end{aligned}$$

Now define a new probability measure by

$$\frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \Big|_{\mathcal{F}_t} = \exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \int_0^t \xi_j(u, s) dY_j^p(s) \right) \times \exp \left(- \sum_{j=1}^J \sum_{p \in \{+, -\}} \int_0^t \varphi_j^p(-i\xi_j(u, s)) v_j^p(s) ds \right)$$

Then the characteristic function will be given by

$$\psi_t(u, T_0, T_1, T_k) = \left(\frac{p(t, T_1)}{p(t, T_0)} \right)^{iu} E_t^\xi \left[\exp \left(\sum_{j=1}^J \sum_{p \in \{+, -\}} \int_t^{T_0} g_j^p(s, u) v_j^p(s) ds \right) \right]$$

where

$$g_j^p(t, u) = iu[\varphi_j^p(-iS_j(s, T_0)) - \varphi_j^p(-iS_j(s, T_1))] - \varphi_j^p(-iS_j(s, T_k)) + \varphi_j^p(-i\xi_j(s, u))$$

which completes the proof.

1.9 Appendix: MCMC details

Conditional distribution

The conditional $p(D|v, \Theta)$

The conditional for the observations can be written as

$$p(D|v, \Theta) = \prod_{t=1}^T p(D_t|v_t, \Theta)$$

where we can write $p(D_t|v_t, \Theta)$ as

$$p(D_t|v_t, \Theta) \propto \prod_{j=1}^n \sigma_{\varepsilon, j}^{-1/2} \exp \left(-\frac{1}{2} \left(\frac{[D_t - f(v_t|\Theta)]_j}{\sigma_{\varepsilon, j}} \right)^2 \right)$$

The conditional $p(v|\Theta)$

Using an Euler scheme we can write the dynamics as

$$v_{t+1} = v_t + [\theta - \hat{\kappa}v_t] \Delta t + \sqrt{v_t \Delta t} \varepsilon_t^v$$

where $\varepsilon_t^v \sim \mathcal{N}(0, I_4)$.

The conditionals for the state transition can be written as

$$p(v|\Theta) \propto \left(\prod_{t=1}^T p(v_t|v_{t-1}, \Theta) \right) p(v_0) \propto \prod_{t=2}^T p(v_t|v_{t-1}, \Theta)$$

where we have assumed independence with v_0 . The conditional $p(v_t|v_{t-1}, \Theta)$ can be written as

$$p(v_{t+1}|v_t, \Theta) = \frac{1}{\det(\Sigma(v_t))} \exp\left(-\frac{1}{2} h_t' \Sigma(v_t)^{-1} h_t\right)$$

where

$$\begin{aligned} h_t &= v_{t+1} - v_t - [\theta - \hat{\kappa}v_t] \Delta t \\ \Sigma(v_t) &= \text{diag}(v_t) \Delta t \end{aligned}$$

Sampling parameters and states

Sampling σ_ε

To sample the elements of σ_ε we use that

$$p(\sigma_\varepsilon|D, v, \Theta_{\setminus\sigma_\varepsilon}) \propto p(D|v)$$

which implies that we can use Gibbs sampling:

$$\sigma_{\varepsilon,j}^2 = \mathcal{IG}\left(\frac{T}{2} + 1, \frac{1}{2} \sum_{t=1}^t ([D_t - f(v_t|\Theta)]_j)^2\right)$$

Sampling Θ^Q

To sample the risk neutral parameters we use a Random-Walk Metropolis-Hastings sampler, i.e. we draw a new parameter value as

$$\Theta^{Q*} = \Theta^Q + \epsilon$$

where ϵ is a zero mean normally distributed variable with a variance that needs calibration.

We accept the draw with probability α

$$\alpha = \min\left(1, \frac{p(D|v, \Theta^*)p(\Theta^*)}{p(D|v, \Theta)p(\Theta)}\right)$$

In all cases, except for λ_j , we furthermore assume uninformative priors.

Sampling Θ^{QP}

To sample the parameters entering in both the risk-neutral moments and the activity rate dynamics, we use a Random-Walk Metropolis-Hastings sampler, i.e. we draw a new parameter value as

$$\Theta^{QP*} = \Theta^{QP} + \epsilon$$

where ϵ is a zero mean normally distributed variable with a variance that needs calibration.

We accept the draw with probability α

$$\alpha = \min \left(1, \frac{p(D|v, \Theta^*)p(v|\Theta^*)}{p(D|v, \Theta)p(v|\Theta)} \right)$$

Sampling Θ^P

To sample the parameters only entering in activity rate dynamics, we use a Random-Walk Metropolis-Hastings sampler, i.e. we draw a new parameter value as

$$\Theta^{P*} = \Theta^P + \epsilon$$

where ϵ is a zero mean normally distributed variable with a variance that needs calibration.

We accept the draw with probability α

$$\alpha = \min \left(1, \frac{p(v|\Theta^*)}{p(v|\Theta)} \right)$$

Sampling v

To sample the activity rates we use a Random-Block size Metropolis-Hastings sampler.

We sample the states as:

- First we sample the initial state v_1 , i.e. draw a new state

$$v_1^* = v_1 + \epsilon$$

and accept it with probability

$$\alpha = \min \left(\frac{p(D_1|v_1^*, \Theta)p(v_2|v_1^*, \Theta)}{p(D_1|v_1, \Theta)p(v_2|v_1, \Theta)} \right)$$

- While $1 < t < T - 1$ then do

$$\hat{t} = t + w, \quad w \sim \text{Poisson}(q)$$

Then sample new parameters

$$v_{t:\hat{t}}^* = v_{t:\hat{t}} + \epsilon$$

and accept the draw with probability

$$\alpha = \min \left(\frac{p(D_{t:\hat{t}}|v_{t:\hat{t}}^*, \Theta)p(v_{t:\hat{t}+1}^*|v_{t-1}, \Theta)}{p(D_{t:\hat{t}}|v_{t:\hat{t}}, \Theta)p(v_{t:\hat{t}+1}|v_{t-1}, \Theta)} \right)$$

- Finally sample the last state v_T , i.e. draw a new state

$$v_T^* = v_T + \epsilon$$

and accept it with probability

$$\alpha = \min \left(\frac{p(D_T|v_T^*, \Theta)p(v_T^*|v_{T-1}, \Theta)}{p(D_T|v_T, \Theta)p(v_T|v_{T-1}, \Theta)} \right)$$

Essay 2

Inflation derivatives modeling using time changed Lévy processes

Abstract

We model inflation derivatives by using the time changed Lévy processes of Carr and Wu (2004), in a Heath-Jarrow-Morton framework. We derive drift conditions for nominal and real forward rates and zero-coupon bonds. Similarly, a drift condition for the consumer price index is found. We show how to price standard inflation derivatives by considering a complex (time dependent) measure. By specifying the subordinator as an affine process, the prices of the considered derivatives can be obtained up to ordinary differential equations and possibly Fourier inversion. Finally, we calibrate our model to market data. Our results show that even though Lévy processes can improve the fit to data, an investigation in the exact specification of the Lévy process and volatility loading is still needed.

Keywords: Inflation derivatives, HJM-framework, Lévy processes, Time Change, Affine processes

JEL Classification: G12, G13, C02, C19

¹I would like to thank Bjarne Astrup Jensen, Thomas Kokholm, Fred Espen Benth, Anne-Sofie Reng Rasmussen, participants at the 2009 Nordic Finance Network Research Workshop and participants at the 2009 Quantitative Methods in Finance Conference for useful comments.

2.1 Introduction

Over the last two decades trading of inflation indexed products has seen a large increase, both in terms of volume and the number of products traded. More specifically, since the 1980s governments have issued inflation linked bonds. Such bonds have their face value as well as their current coupon payments linked to a reference consumer pricer index, henceforth shortened CPI.

In the 1990s markets trading inflation linked derivatives began to develop. The premier examples of such derivatives is inflation indexed swaps. Today, investment banks offer inflation linked derivatives linked to consumer price indices in the United Kingdom, the United States, the Euro area and Japan. Even though the inflation indexed markets are still young, standard products, such as swap style products trade at a reasonable depth. For example, quotes on Euro Area and United States CPI linked zero-coupon inflation swaps have been available from Bloomberg since 2004, and quotes for caps and floors on Euro area CPI have been available since mid-2007.

The rise of inflation indexed markets has, of course, sparked some academic research activity. Among early papers is Jarrow and Yildirim (2003) who consider a three factor HJM model, based on a foreign exchange analogy, that can be calibrated to historical nominal rates and historical United States Treasury Inflation Protected Securities (henceforth TIPS). Using these parameters Jarrow and Yildirim (2003) calculate theoretical prices of options on the United States CPI.²

Using an approach similar to that of the LIBOR-market model, Mercurio (2005) considers modeling forward inflation rates. Using this framework Mercurio (2005) derives prices for Year-on-Year swaps and inflation linked caps and floors.

All of the above mentioned papers operate under the assumption of deterministic volatilities and innovations driven by Wiener processes. Recently, three papers have tried to go beyond these assumptions.

First, Mercurio and Moreni (2006) model forward inflation rates with stochastic volatility in a framework, based on the LIBOR-market-style framework found in Mercurio (2005). They manage to derive prices for Year-on-Year swaps and inflation linked caps and floors, using transform techniques as found in Duffie, Pan, and Singleton (2000).

²Another often referenced paper is Hughston (1998), who also considers modeling real and nominal bonds in a HJM-framework. However, the derivations in Hughston (1998) are not clear, and we prefer to use Jarrow and Yildirim (2003) as the central reference.

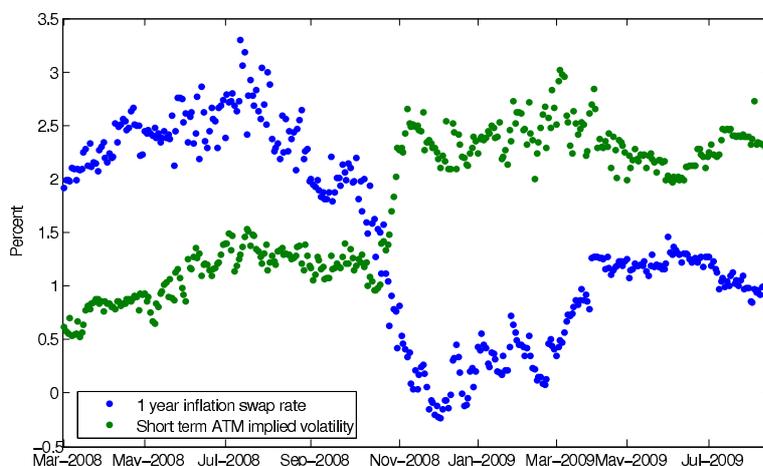


Figure 2.1: *Developments in short term EURO area inflation swap rates and implied volatilities from March 2008 to August 2009. Source: Bloomberg.*

Secondly, Hinnerich (2008) considers a more general framework with stochastic volatility and jumps, extending the HJM framework found in Björk, Kabanov, and Runggaldier (1997). Furthermore, Hinnerich (2008) shows that the foreign exchange analogy and the real bank account are not needed as a priori assumptions in order to derive no-arbitrage conditions; the framework delivers the foreign exchange analogy as a by-product of the no-arbitrage conditions. Even though Hinnerich (2008) considers no-arbitrage conditions for jump processes, a model based on jump processes is only considered for Year-on-Year inflation swaps.

Thirdly, Mercurio and Moreni (2009) describe a forward inflation framework, where the stochastic volatility is given by SABR processes. Their approach requires a SABR process for each maturity of the considered inflation caplets, i.e. for their 15 considered maturities, 45 parameters have to be estimated. Finally, in order to price other derivatives than Year-on-Year inflation caplets - for instance zero-coupon Inflation Caplets - they need to resort to approximate dynamics.

In this paper we seek to close the two gaps left by the above mentioned papers. More precisely, we consider a model based on the HJM framework where the underlying source of uncertainty is driven by a time changed Lévy process, as seen, e.g., in equity modeling in Carr and Wu (2004). We extend this framework to allow for both stochastic volatility and jumps in inflation and interest rates, similar to what was done in Jarrow and Yildirim (2003) and Hinnerich (2008).

Neither jumps nor stochastic volatility in nominal interest rates are novel features. A number of papers have identified these effects (see Johannes (2004) or Andersen, Benzoni, and Lund (2004) for examples). However, this evidence is not the only reason to include jumps and stochastic volatility in an inflation modeling setup. During the second part of 2008 a worsening of the macroeconomic outlook made short term inflation swap rates drop from around 2 percent to somewhere between 0 and 1 percent, as seen in Figure 2.1. At the same time implied inflation cap and floor volatilities more than doubled; for some maturities and degrees of moneyness, the implied volatility even rose to more than four times its previous value. In many cases the volatility smile steepened, pricing in more possible extreme inflation and deflation events. Both facts show evidence of stochastic volatility and jump risk in inflation markets.

The structure of the paper is as follows: In section 2.2 we briefly describe standard inflation linked products and in section 2.3 we describe the Jarrow-Yildirim model and the enhancements that we propose in this paper. Section 2.4 describes time changed Lévy processes, and section 2.5 describes the HJM framework, i.e. drift conditions, different model representations and characteristic functions. In section 2.6 we consider the pricing of the inflation linked derivatives considered in section 2.2. Section 2.7 describes how the time change can be specified, and in section 2.8 we calibrate the model to market data. Finally, section 2.9 concludes the paper.

2.2 Inflation linked products

In this section we describe the typical inflation linked products traded in the market. A good description of inflation derivatives can also be found in Barclays Capital (2008), which is more exhaustive than the description given in this section. In this section we refer to the Consumer Price Index (henceforth CPI), which is the price of a consumer basket measured in Dollars or the representative local currency.

An *Inflation Protected zero-coupon bond* is a bond where the payoff at maturity, T , is compounded by the CPI³

$$p_{IP}(T, t_0, T) = \frac{I_T}{I_{t_0}}$$

³Typically inflation protected bonds are linked to the CPI some months prior to maturity, however we use this simpler specification to enhance the understanding of the product.

where I_T is the value of the CPI at time T . The denominator $I(t_0)$ normalize the dependence of the CPI, such that the inflation indexation is initiated at the issuance of the bond. The price of an inflation protected bond will be given by the expectation

$$p_{IP}(t, t_0, T) = E_t^Q \left[\exp \left(- \int_t^T n_s ds \right) \frac{I_T}{I_{t_0}} \right] = p_n(t, T) E_t^T \left[\frac{I_T}{I_{t_0}} \right]$$

where n_t is the nominal spot rate and $p_n(t, T)$ is the nominal zero-coupon bond with maturity T , observed at time t .

Based on the observed market price of an inflation protected ZCB, we define the real ZCB as

$$p_r(t, T) = \frac{p_{IP}(t, t_0, T) I_{t_0}}{I_t}$$

such that $p_r(T, T) = 1$. Note that the relation above tells us that $p_r(t, T)$ is measured in units of the CPI-basket, and the real bond will give the investor one CPI-basket at time T . Also note that real bonds are derived quantities and thus not directly traded. Finally differences between yields from nominal and real ZCBs are termed *Break Even Inflation Rates*, as it reflects the inflation compensation required by investors.

A *zero-coupon Inflation Indexed Swap (ZCIIS)* is a swap agreement where one party pays the percentage change on the CPI over the period $[t, T]$ and the other party pays a fixed amount K . The payoff, at maturity, for the holder of the ZCIIS is then given by

$$\begin{aligned} ZCIIS_T(t, T, K) &= \left(\frac{I_T}{I_t} - 1 \right) - \left((1 + K)^{T-t} - 1 \right) \\ &= \frac{I_T}{I_t} - (1 + K)^{T-t} \end{aligned}$$

ZCIISs are initiated with a value of zero and are quoted in terms of the fixed payment K , and thus ZCIIS quotes offer a term structure of the expected (risk adjusted) future inflation, also known as *Swap Break Even Inflation Rates*.

Although it may not appear so, the pricing of a ZCIIS is completely model independent. This follows as the value of such a contract is given by

$$\begin{aligned} ZCIIS_t(t, T, K) &= \frac{p_n(t, T)}{I_t} E_t^T [I_T] - p_n(t, T) (1 + K)^{T-t} \\ &= \frac{p_n(t, T)}{I_t} E_t^T \left[I_T \frac{p_r(T, T)}{p_n(T, T)} \right] - p_n(t, T) (1 + K)^{T-t} \end{aligned}$$

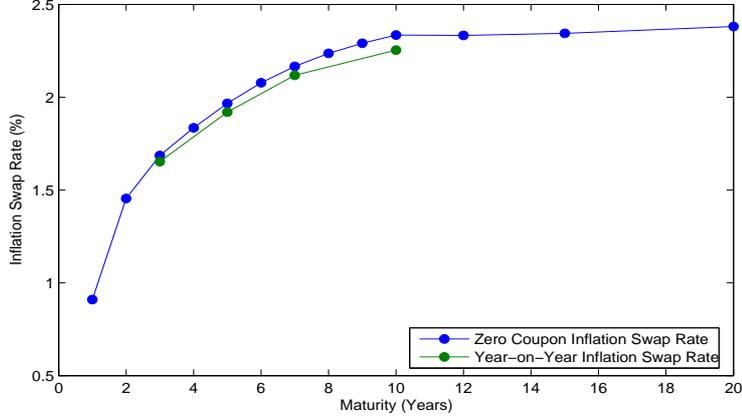


Figure 2.2: Market quotes of Inflation Swap rates. Data is from August 10th 2009. Source: Bloomberg.

Since $I_t p_r(t, T) / p_n(t, T)$ is a martingale under the nominal T -forward measure, we obtain

$$ZCIIS_t(t, T, K) = p_r(t, T) - p_n(t, T) (1 + K)^{T-t}$$

In terms of the fixed payment, K :

$$K = \left(\frac{p_r(t, T)}{p_n(t, T)} \right)^{1/(T-t)} - 1$$

Furthermore, as noted in Hinnerich (2008), this result can also be obtained via a replication argument. Finally quotes of ZCIIS rates and nominal interest rates can be used to derive real yields - in this sense ZCIISs form a central part of the inflation linked market.

A *Year-on-Year Inflation Indexed swap* (henceforth *YYIIS*) is a swap agreement similar to the ZCIIS, however with the difference that the pay-off is linked to forward inflation (i.e. over the period $[T_{i-1}, T_i]$, $t < T_{i-1} < T_i$). At time T_i the YYIIS has pay-off given by

$$YYIIS_{T_i}(t, T_{i-1}, T_i, K) = \Delta_i \left[\left(\frac{I_{T_i}}{I_{T_{i-1}}} - 1 \right) - \left((1 + K)^{T_i - T_{i-1}} - 1 \right) \right]$$

where $\Delta_i = T_i - T_{i-1}$. Note that for $T_{i-1} = t$ and $T_i = T$ the YYIIS collapses to a ZCIIS. Typically YYIIS agreements are traded as portfolios of the single payment swap given above

$$YYIIS_t(t, n, K) = \sum_{j=1}^n YYIIS_t(t, T_{j-1}, T_j, K)$$

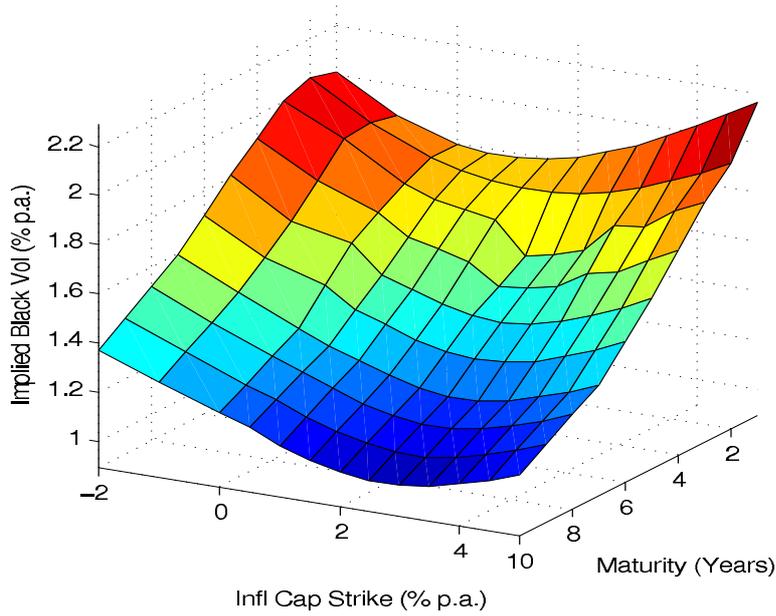


Figure 2.3: *Implied volatility surface derived from EURO area Inflation Year-on-Year caps and floors. Data is from August 10th 2009. Source: Bloomberg.*

Unfortunately, the price of a YYIIS is not model independent, and therefore the exact quote K will depend on model choice.

Figure 2.2 shows market quotes of both ZCIIS and YYIIS rates. The dataset is from August 10th 2009 and shows an upward sloping term structure. The shape of the term structure reflects the state of the EURO area economy, i.e. that a low rate of inflation is expected over the next periods. Over longer horizons the inflation swap rates are around 2.4 percent.⁴

A *zero-coupon Inflation Cap (ZCIC)* is an option on inflation. More precisely (as the ZCIIS) it is linked to the inflation from t to T . Typically the period considered, $[t, T]$, is set such that there exists a ZCIIS which is linked to that same period, as this will ease hedging of the option. The pay-off at time T is then given by:

$$ZCIC_T(t, T, K) = \left(\left(\frac{I_T}{I_t} - 1 \right) - \left((1 + K)^{T-t} - 1 \right) \right)^+$$

⁴Typically surveys on EURO area HICP have long term expected inflation rates below but close to 2 percent. The difference in survey and swap rates reflect an inflation risk premia.

The pricing of an option on inflation will require a model for the CPI evolution, as the price of the option crucially depends on the distribution of the CPI.

Finally a *Year-on-Year Inflation Caplet (YYIC)* is an option on inflation, which links to CPI in the same way as the *YYIIS* links to the CPI, i.e. to forward inflation. Thus the YYIC pays the owner at time T_i

$$YYIC_{T_i}(t, T_{i-1}, T_i, K) = \Delta_i \left[\left(\frac{I_{T_i}}{I_{T_{i-1}}} - 1 \right) - \left((1 + K)^{T_i - T_{i-1}} - 1 \right) \right]^+$$

As with a ZCIC the price of a YYIC is highly model dependent, and a model for the CPI evolution is needed in order to price the option.

Finally, as in the case of an interest rate cap, an inflation indexed cap will be a portfolio of inflation indexed caplets. Inflation caps typically show volatility smiles, as seen in Figure 2.3. This indicates that inflation markets are pricing more extreme movements, than can be described by a log-normal Black's formula (i.e. that inflation rates are normally distributed).

2.3 The Jarrow-Yildirim model

In this section we briefly describe the Jarrow-Yildirim model, and describe the enhancements to their model, which we propose in this paper.

The Jarrow-Yildirim model assumes that nominal forward rates, $f_n(t, T)$, real forward rates, $f_r(t, T)$, and Consumer Price Index, I_t are driven by a three dimensional Wiener process:

$$\begin{aligned} df_n(t, T) &= \alpha_n(t, T)dt + \sigma_n(t, T)dW_t^n \\ df_r(t, T) &= \alpha_r(t, T)dt + \sigma_r(t, T)dW_t^r \\ \frac{dI_t}{I_t} &= \alpha_I(t)dt + \sigma_I(t)dW_t^I \end{aligned}$$

Jarrow and Yildirim (2003) then show that the drift conditions, $\alpha_i(t, T)$ and $\alpha_I(t)$, have the following form:

$$\begin{aligned} \alpha_n(t, T) &= \sigma_n(t, T) \int_t^T \sigma_n(t, u) du \\ \alpha_r(t, T) &= \sigma_r(t, T) \left(\int_t^T \sigma_r(t, u) du - \rho_{Ir} \sigma_I(t) \right) \\ \alpha(t) &= n_t - r_t \end{aligned}$$

where n_t is the nominal spot interest rate and r_t is the real spot interest rate. Furthermore the drift conditions preserve the intuition of the regular HJM-framework, i.e. that the nominal drift condition is a variance correction to the nominal forward rate drift, which ensures no-arbitrage. Similarly, the real drift condition reflects that inflation protected bonds, rather than real bonds, are traded in the market. Finally, the CPI drift condition tells us that the Fisher equation holds under the risk neutral martingale measure.

In terms of analytical tractability the rates in the Jarrow-Yildirim model are normally distributed, which implies that prices of standard inflation derivatives have analytical solutions.

In this paper we propose two modifications to this framework. First we add stochastic volatility to the model:

$$\begin{aligned} df_i(t, T) &= \alpha_i(t, T)dt + \sigma_i(t, T)\sqrt{v_t}dW_t^i, & i = n, r \\ \frac{dI_t}{I_t} &= \alpha_I(t)dt + \sigma_I(t)\sqrt{v_t}dW_t^I \end{aligned}$$

Using the scaling properties of the Wiener process, this can also be expressed through a time change:

$$\begin{aligned} df_i(t, T) &= \alpha_i(t, T)dt + \sigma_i(t, T)dW_{\tau_t}^i, & i = n, r \\ \frac{dI_t}{I_t} &= \alpha_I(t)dt + \sigma_I(t)dW_{\tau_t}^I \end{aligned}$$

The details of this time change are described in the following section.

The other modification, is to replace the Wiener process W with a Lévy process L :

$$\begin{aligned} df_i(t, T) &= \alpha_i(t, T)dt + \sigma_i(t, T)dL_{\tau_t}^i, & i = n, r \\ \frac{dI_t}{I_t} &= \alpha_I(t)dt + \sigma_I(t)dL_{\tau_t}^I \end{aligned}$$

These two modifications allow us to capture two important effects. First, the time changes allows for stochastically varying volatility, including clustering of volatility. Secondly, the use of Lévy processes allows for extreme movements (i.e. jumps), which adds to the flexibility when modeling the underlying distribution of interest rates and inflation⁵.

In the following section we will briefly describe the time changed Lévy processes outlined above. That is their construction and relevant analytical properties which are essential to our modeling of inflation derivatives.

⁵The Jarrow-Yildirim model with stochastic volatility can be recovered when letting the Lévy process be a Wiener process (i.e. no jumps).

2.4 Time changed Lévy processes

As mentioned above we alter the Jarrow-Yildirim model by introducing time inhomogeneous effects, such as effects of stochastic volatility and non-normal innovations, through a time changed Lévy process. More precisely, we consider the approach taken in Carr and Wu (2004). We will describe this framework and present the necessary modifications to the results in Carr and Wu (2004) in order for us to be able to use these processes in interest rate and inflation modeling.

We consider a complete stochastic basis $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{Q})$. On this basis we define a d -dimensional Lévy process L with the characteristic function given by

$$E^{\mathbb{Q}}[\exp(iu'L_t)] = \exp(\varphi(u)t)$$

$$\varphi(u) = iu'\hat{a} - \frac{1}{2}u'\Sigma u + \int_{\mathbb{R}_0^d} \left(e^{iu'x} - 1 - iux1_{|x|<1} \right) \nu(dx)$$

A finite variation Lévy process can alternatively be expressed with the truncation function incorporated into the drift \hat{a} , thus yielding the characteristic exponent

$$\varphi(u) = iu'a - \frac{1}{2}u'\Sigma u + \int_{\mathbb{R}_0^d} \left(e^{iu'x} - 1 \right) \nu(dx)$$

We will use the latter specification of the characteristic exponent throughout this paper⁶.

Lévy processes are time homogeneous processes, and although a wide variety of distributions can be obtained by using Lévy processes, this will only allow us to fit data from shorter periods without having to revise parameter estimates very frequently. Furthermore, stylized facts from interest rate markets indicate that interest rates exhibit stochastic volatility, i.e. (stochastically) time varying quadratic variation. To introduce this variation Carr and Wu (2004) make use of the time change technique and introduce a subordinated Lévy process L . The subordinator τ - being a strictly increasing process - takes the general form

$$\tau_t = \alpha_t + \int_{\mathbb{R}_0^+} x\mu(dx, dt)$$

⁶Even though we use this specification, our proofs hold for a more general Lévy process. The proofs are easily extended using the Lévy-Itô decomposition of the Lévy process.

where α_t represents the locally deterministic part of the process and the last term is a jump process that only exhibits positive jumps.

As in Carr and Wu (2004) we restrict the subordinator to have only a locally deterministic part; i.e. , we assume that τ can be represented as

$$\tau_t = \int_0^t v_{s-} ds$$

The process v is a positive semi-martingale, possibly correlated with the Lévy process L .

The subordinated Lévy process $Y_t = L_{\tau_t}$ will be the driving process in our framework. As shown in Monroe (1978), any semi-martingale can be reached through subordination of a Wiener process, thus in principle the subordination method will give us a large degree of flexibility, when it comes to modeling inflation dynamics.

Carr and Wu (2004) manage to derive the characteristic function of the process Y by using a change of measure. This eliminates the need to derive the joint characteristic function of Y and τ , when L and τ are correlated. This methodology is similar to the by now well known change of numeraire technique, originally introduced in Geman, Karoui, and Rochet (1995).

In an interest rate setting, real as well as nominal, the results in Carr and Wu (2004) cannot be applied directly, because volatility coefficients (volatility loadings) depend on time or time to maturity. Thus we would like to derive the characteristic function of the process X which is a stochastic integral with respect to a - possibly complex - deterministic time dependent function, i.e.

$$X(t, T) = \int_t^T \xi(s)' dY_s$$

To simplify the intuition of obtaining the characteristic function of this stochastic variable, we first consider the case where L and τ are independent.

If we consider the stochastic integral on an n point partition of $[t, T]$ and perform iterated expectations conditional on τ , then heuristically we obtain the following result when appropriate convergence of the integrals is

assumed:

$$\begin{aligned}
 E_t^{\mathbb{Q}} [\exp (iuX(t, T))] &= E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T iu\xi(s)' dY_s \right) \right] \\
 &= E_t^{\mathbb{Q}} \left[E_t^{\mathbb{Q}} \left[\exp \left(\lim_{n \rightarrow \infty} \sum_{l=0}^{n-1} iu\xi(t_l)' (L_{\tau_{t_{l+1}}} - L_{\tau_{t_l}}) \right) \middle| \tau \right] \right] \\
 &= E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T \varphi (u\xi(s)) v_{s-} ds \right) \right]
 \end{aligned}$$

This specific expectation is quite common in the interest rate literature, and for tractable specifications of v the expectation above will yield solutions given by the exponentially affine form

$$E_t^{\mathbb{Q}} [\exp (iuX(t, T))] = \exp (A(t, T, u) + B(t, T, u)'v_t)$$

When L and τ are no longer independent, then conditioning on τ will affect the expectation of the Lévy process L . To overcome this issue, we consider, as Carr and Wu (2004), a change of measure into a complex measure, where we only need to calculate the characteristic function of the time change. The following proposition gives us the solution to the problem of finding the characteristic function.⁷

Proposition 3. *Suppose that $\xi : \mathbb{R}_+ \mapsto \mathbb{C}^d$ is a continuous, bounded function, and let $\xi(s) \in \vartheta$ for all $s : t \leq s \leq T$, where ϑ is the set of values where the characteristic exponent of the Lévy process L is finite.*

Then the transform given by

$$\psi(t, T) = E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T \xi(s)' dY_s \right) \right]$$

can also be expressed as the following expectation

$$E_t^{\xi} \left[\exp \left(\int_t^T \varphi (-i\xi(s)) v_{s-} ds \right) \right] \tag{2.1}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} and $E_t^{\xi} [\bullet]$ denotes expectation with regard to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^{\xi}}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = M_t = \exp \left(- \int_0^t \varphi (-i\xi(s)) v_{s-} ds + \int_0^t \xi(s)' dY_s \right)$$

⁷Note that the term iu can be incorporated into the function ξ , thus leaving us with a simpler expectation.

Proof. See Appendix 2.10 □

The result in proposition 1 is a generalization of the result in Carr and Wu (2004); however, it still allows us to obtain a semi-analytical solution to the expectation given above, provided that the process v is sufficiently tractable.

Also note that - as with a standard change of numeraire - we only need to consider one stochastic variable under this transformed measure, instead of two stochastic variables under the regular measure, although this measure has an explicit dependence on the function ξ .

Furthermore, as mentioned in Carr and Wu (2004), when the subordinator τ is independent of the Lévy process L , the new measure \mathbb{Q}^ξ will coincide with the measure \mathbb{Q} . Intuitively, this is seen as the Radon-Nikodym derivative only contains elements from the Lévy process. This is also consistent with the heuristic calculations earlier in this section.

2.5 An inflation HJM framework based on time changed Lévy processes

Obtaining a drift condition

One of the great advantages of the HJM framework, originally developed in Heath, Jarrow, and Morton (1992), is that it fully describes how to specify forward rates under a risk neutral martingale measure. In a framework driven by Wiener processes, one only needs to specify volatility loadings, and thus the drift of forwards rates - or, equivalently, zero-coupon bond prices - are completely determined from the no-arbitrage principle. When the driving process is a Lévy process, Eberlein and Raible (1999) show that the drift condition is a composite of the volatility loading and the characteristic exponent of the Lévy process.

In our specification of the HJM framework we assume that nominal and real forward rates, respectively, are driven by the processes

$$\begin{aligned} df_n(t, T) &= \alpha_n(t, T)dt + \sigma_n(t, T)'dY_t \\ df_r(t, T) &= \alpha_r(t, T)dt + \sigma_r(t, T)'dY_t \end{aligned}$$

Here we assume that $\sigma_\bullet(t, T)$ is a deterministic integrable function in \mathbb{R}^d and $\alpha_\bullet(t, T)$ is an adapted integrable process in \mathbb{R} .

In a similar fashion we assume the the log consumer price index (henceforth CPI) is given by

$$d \log I_t = \alpha_I(t)dt + \sigma_I(t)'dY_t$$

where $\sigma_I(t)$ is a deterministic integrable function in \mathbb{R}^d and $\alpha_I(t)$ is an adapted integrable process in \mathbb{R} .

Applying the results from Björk, Kabanov, and Runggaldier (1997), nominal and real log zero-coupon bond prices have dynamics given by

$$\begin{aligned} d \log p_n(t, T) &= (n_t + A_n(t, T)) dt + S_n(t, T)'dY_t \\ d \log p_r(t, T) &= (r_t + A_r(t, T)) dt + S_r(t, T)'dY_t \end{aligned}$$

where n_t is the nominal short rate, r_t the real short rate and

$$A_\bullet(t, T) = - \int_t^T \alpha_\bullet(t, u)du \quad \text{and} \quad S_\bullet(t, T) = - \int_t^T \sigma_\bullet(t, u)du$$

This gives us the expression for nominal and real zero-coupon bond prices

$$\begin{aligned} p_n(t, T) &= p_n(0, T) \exp \left(\int_0^t (n_s + A_n(s, T)) ds + \int_0^t S_n(s, T)'dY_s \right) \\ p_r(t, T) &= p_r(0, T) \exp \left(\int_0^t (r_s + A_r(s, T)) ds + \int_0^t S_r(s, T)'dY_s \right) \end{aligned}$$

The usual nominal bank account, β^n , and the derived stochastic process β^r , termed the real bank account, are given by

$$\beta_t^n = \exp \left(\int_0^t n_s ds \right), \quad \beta_t^r = \exp \left(\int_0^t r_s ds \right)$$

Using these expressions to solve for the two bank accounts, expressed in terms of their respective zero-coupon bond prices, we obtain alternative expressions for these two bank accounts:

$$\begin{aligned} \beta_t^n &= \frac{1}{p_n(0, t)} \exp \left(- \int_0^t A_n(s, t) ds - \int_0^t S_n(s, t)'dY_s \right) \\ \beta_t^r &= \frac{1}{p_r(0, t)} \exp \left(- \int_0^t A_r(s, t) ds - \int_0^t S_r(s, t)'dY_s \right) \end{aligned}$$

The remaining part of this section is dedicated to finding the drift conditions, i.e. the conditions on $A_n(t, T)$, $A_r(t, T)$ and $\alpha_I(t)$ that ensure no

arbitrage. To derive such conditions we use the fact that the following processes must be martingales under the risk neutral martingale measure:

$$\frac{p_n(t, T)}{\beta_t^n}, \quad \frac{p_r(t, T)I_t}{\beta_t^n}$$

The first is the condition for a nominal zero-coupon bond and the other is the condition for an inflation protected zero-coupon bond.

As in Hinnerich (2008) we do not assume that the real bank account is a traded asset; the drift conditions can be derived without this assumption. The full derivation is given in Appendix 2.10; the result is given in the proposition below:

Proposition 4. *For every t and T , such that $t \leq T$, $A_n(t, T)$, $A_r(t, T)$, $\alpha_I(t)$, $S_n(t, T)$, $S_r(t, T)$ and $\sigma_I(t)$ must satisfy the following relationships:*

$$\begin{aligned} A_n(t, T) &= -v_{t-} \varphi(-iS_n(t, T)) \\ \alpha_I(t) &= n_t - r_t - v_{t-} \varphi(-i\sigma_I(t)) \\ A_r(t, T) &= -v_{t-} [\varphi(-i(S_r(t, T) + \sigma_I(t))) - \varphi(-i\sigma_I(t))] \end{aligned}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} .

Proof. See Appendix 2.10. □

The first drift condition is equivalent to a purely nominal HJM model (see Andersen (2011) for a verification of this). The second condition makes sure that the instantaneous drift of the CPI is equal to $I_{t-}(n_t - r_t)dt$; this is the Fisher relation, which was also found in Jarrow and Yildirim (2003). The last drift condition is also similar to the one found in Jarrow and Yildirim (2003); i.e. that the real drift condition is a variance and covariance correction of the real zero-coupon bond drift, which reflects that inflation linked bonds, and not real bonds, are traded in the market.

Model dynamics and characteristic functions

Having found the drift conditions under \mathbb{Q} we can now state the dynamics for real and nominal zero-coupon bonds as well as inflation that is needed to price derivatives. First, it follows by using the nominal bank account

that the nominal zero-coupon bond can be expressed as

$$p_n(t, T) = \frac{p_n(0, T)}{p_n(0, t)} \exp \left(\int_0^t (\varphi(-iS_n(s, t)) - \varphi(-iS_n(s, T))) v_{s-} ds \right) \times \exp \left(\int_0^t (S_n(s, T) - S_n(s, t))' dY_s \right)$$

In a similar fashion we can derive the real zero-coupon bonds

$$p_r(t, T) = \frac{p_r(0, T)}{p_r(0, t)} \exp \left(\int_0^t (S_r(s, T) - S_r(s, t))' dY_s \right) \times \exp \left(\int_0^t (\varphi(-i(S_r(s, t) + \sigma_I(s))) - \varphi(-i(S_r(s, T) + \sigma_I(s)))) v_{s-} ds \right)$$

Finally the CPI, can be described as

$$I_t = I_0 \exp \left(\int_0^t (n_s - r_s - \varphi(-i\sigma_I(s)) v_{s-}) ds + \int_0^t \sigma_I(s)' dY_s \right)$$

which has the alternative representation, obtained by using the two bank accounts

$$I_t = I_0 \frac{p_r(0, t)}{p_n(0, t)} \exp \left(\int_0^t [\varphi(-iS_n(s, t)) - \varphi(-i(S_r(s, t) + \sigma_I(s)))] v_{s-} ds \right) \times \exp \left(\int_0^t (\sigma_I(s) + S_r(s, t) - S_n(s, t))' dY_s \right)$$

The term $I_0 \frac{p_r(0, t)}{p_n(0, t)}$ is the forward CPI, similar to the forward price term, $\frac{p_{\bullet}(0, T)}{p_{\bullet}(0, t)}$, for real and nominal bonds. Furthermore, the term $\frac{p_r(0, t)}{p_n(0, t)}$ is observable directly from inflation swap quotes, cf. section 2.2.

Next we consider several characteristic functions, that will be relevant, when pricing different derivatives. To begin with, we consider the characteristic function of the nominal log zero-coupon bond, $\log p_n(T, T_k)$ under the nominal T_l -forward measure. We then get

Proposition 5. *Suppose $\xi(t, u)$ is a continuous, bounded function, and let $\xi(s, u) \in \vartheta$ for all $s : t \leq s \leq T$ and fixed u , where ϑ is the set of values where the characteristic exponent of the Lévy process L is finite.*

Then the characteristic function of the nominal log zero-coupon bond price under the nominal T_l -forward measure

$$\psi_t^n(u, T, T_k, T_l) = E_t^{T_l} [\exp(iu \log p_n(T, T_k))]$$

is given by the following expectation

$$\psi_t^n(u, T, T_k, T_l) = \left(\frac{p_n(t, T_k)}{p_n(t, T)} \right)^{iu} E_t^\xi \left[\exp \left(\int_t^T g(s, u) v_{s-} ds \right) \right]$$

where $E_t^\xi[\bullet]$ denotes expectation with regards to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = \exp \left(- \int_0^t \varphi(-i\xi(s, u)) v_{s-} ds + \int_0^t \xi(s, u)' dY_s \right)$$

and

$$\begin{aligned} g(t, u) &= iu (\varphi(-iS_n(t, T)) - \varphi(-iS_n(t, T_k))) - \\ &\quad \varphi(-iS_n(t, T_l)) + \varphi(-i\xi(t, u)) \\ \xi(t, u) &= iu (S_n(t, T_k) - S_n(t, T)) + S_n(t, T_l) \end{aligned}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} .

Proof. See Appendix 2.10. □

In a similar fashion we can derive the characteristic function of the real log zero-coupon bond, $\log p_r(T, T_k)$, under the nominal T_l -forward measure:

Proposition 6. *Suppose $\xi(t, u)$ is a continuous, bounded function, and let $\xi(s, u) \in \vartheta$ for all $s : t \leq s \leq T$ and fixed u , where ϑ is the set of values where the characteristic exponent of the Lévy process L is finite.*

Then the characteristic function of the real log zero-coupon bond price under the nominal T_l -forward measure

$$\psi_t^r(u, T, T_k, T_l) = E_t^{T_l} [\exp(iu \log p_r(T, T_k))]$$

is given by the following expectation

$$\psi_t^r(u, T, T_k, T_l) = \left(\frac{p_r(t, T_k)}{p_r(t, T)} \right)^{iu} E_t^\xi \left[\exp \left(\int_t^T g(s, u) v_{s-} ds \right) \right]$$

where $E_t^\xi[\bullet]$ denotes expectation with regards to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = \exp \left(- \int_0^t \varphi(-i\xi(s, u)) v_{s-} ds + \int_0^t \xi(s, u)' dY_s \right)$$

and

$$\begin{aligned} g(t, u) &= iu [\varphi(-i(S_r(t, T) + \sigma_I(t))) - \varphi(-i(S_r(t, T_k) + \sigma_I(t)))] - \\ &\quad \varphi(-iS_n(t, T_l)) + \varphi(-i\xi(t, u)) \\ \xi(t, u) &= iu(S_r(t, T_k) - S_r(t, T)) + S_n(t, T_l) \end{aligned}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} .

Proof. Similar to the proof of the nominal log zero-coupon bond. Replace all nominal quantities with real quantities, except when changing from the nominal T_l -forward measure to the regular nominal risk neutral forward measure. \square

Finally, and very important, in order to price inflation derivatives we need the characteristic function of inflation and the inflation prevailing between two time points in the future. We start by considering the time t characteristic function, of the log-change in the inflation index until time T , $\log I_T/I_t$. The variable is considered under the nominal T_l -forward measure:

Proposition 7. *Suppose $\xi(t, u)$ is a continuous, bounded function, and let $\xi(s, u) \in \vartheta$ for all $s : t \leq s \leq T$ and fixed u , where ϑ is the set of values where the characteristic exponent of the Lévy process L is finite.*

Then the characteristic function of the change in the log inflation index, under the nominal T_l -forward measure,

$$\psi_t^I(u, T, T_l) = E_t^{T_l} \left[\exp \left(iu \log \frac{I_T}{I_t} \right) \right]$$

is given by the following expectation

$$\psi_t^I(u, T, T_l) = \left(\frac{p_r(t, T)}{p_n(t, T)} \right)^{iu} E_t^\xi \left[\exp \left(\int_t^T g(s, u) v_{s-} ds \right) \right]$$

where $E_t^\xi [\bullet]$ denotes expectation with regards to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = \exp \left(- \int_0^t \varphi(-i\xi(s, u)) v_{s-} ds + \int_0^t \xi(s, u)' dY_s \right)$$

and

$$\begin{aligned} g(t, u) &= iu [\varphi(-iS_n(t, T)) - \varphi(-i(S_r(t, T) + \sigma_I(t)))] - \\ &\quad \varphi(-iS_n(t, T_l)) + \varphi(-i\xi(t, u)) \\ \xi(t, u) &= iu (\sigma_I(t) + S_r(t, T) - S_n(t, T)) + S_n(t, T_l) \end{aligned}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} .

Proof. See Appendix 2.10. □

In some applications, we are not interested in the inflation prevailing from the present point in time t and until a time in the future T . Instead one is more interested in the conditional distribution at time t of the inflation prevailing between time T_j and T_{j+1} . Assuming that the solution to the characteristic function in proposition 7 is exponentially affine, we can indeed derive the characteristic function of such a quantity:

Proposition 8. *Suppose $\xi(t, u)$ is a continuous, bounded function, and let $\xi(s, u) \in \vartheta$ for all $s : t \leq s \leq T_k$ and fixed u , where ϑ is the set of values where the characteristic exponent of the Lévy process L is finite.*

Furthermore, assume that $\psi_T^I(u, T_k, T_l)$ has an exponentially affine representation:

$$\begin{aligned} \psi_T^I(u, T_k, T_l) &= E_T^{T_l} \left[\exp \left(iu \log \frac{I_{T_k}}{I_T} \right) \right] \\ &= \left(\frac{p_r(T, T_k)}{p_n(T, T_k)} \right)^{iu} \exp (A(T, T_k) + B(T, T_k)v_{T_-}) \end{aligned}$$

Then the characteristic function of the change in the log inflation index, between time T and T_k , under the nominal T_l -forward measure,

$$\psi_t^I(u, T, T_k, T_l) = E_t^{T_l} \left[\exp \left(iu \log \frac{I_{T_k}}{I_T} \right) \right]$$

is given by the following expectation

$$\begin{aligned} \psi_t^I(u, T, T_k, T_l) &= \left(\frac{p_r(t, T_k)}{p_r(t, T)} \frac{p_n(t, T)}{p_n(t, T_k)} \right)^{iu} \times \\ &\quad E_t^\xi \left[\exp \left(A(T, T_k) + B(T, T_k)v_{T_-} + \int_t^T g(s, u)v_{s_-} ds \right) \right] \end{aligned}$$

where $E_t^\xi[\bullet]$ denotes expectation with regards to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^\xi}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = \exp \left(- \int_0^t \varphi(-i\xi(s, u)) v_{s-} ds + \int_0^t \xi(s, u)' dY_s \right)$$

and

$$\begin{aligned} g(t, u) = & iu \left(\varphi(-i(S_r(t, T) + \sigma_I(t))) - \varphi(-iS_n(t, T)) - \right. \\ & \left. \varphi(-i(S_r(t, T_k) + \sigma_I(t))) + \varphi(-iS_n(t, T_k)) \right) - \\ & \varphi(-S_n(t, T_l)) + \varphi(-i\xi(t, u)) \\ \xi(t, u) = & iu \left(S_r(t, T_k) - S_n(t, T_k) - S_r(t, T) + S_n(t, T) \right) + S_n(t, T_l) \end{aligned}$$

where $\varphi(u)$ is the characteristic exponent of the Lévy process L under the measure \mathbb{Q} .

Proof. See Appendix 2.10. □

2.6 Pricing inflation products

This section will describe how to price inflation dependent products. We will consider the products given in section 2.2, which are model dependent; that is the Year-on-Year Inflation linked Swap, the zero-coupon Inflation Cap and the Year-on-Year Inflation Caplet.

Year-on-Year Inflation Indexed swaps

In a (payer) Year-on-Year Inflation Indexed swap (henceforth YYIIS), cash flows are exchanged at several time points. Let the T_n be the start time of the swap, and the $T_j = T_{n+1}, \dots, T_m$ be the payment dates of the swap. For two time points T_j and T_{j-1} , we have that $T_j - T_{j-1} \equiv \Delta_j$. For every $T_j = T_{n+1}, \dots, T_m$ the holder of the payer YYIIS pays

$$\Delta_j K$$

At the same time, the holder receives

$$\Delta_j \left(\frac{I_{T_j}}{I_{T_{j-1}}} - 1 \right)$$

The value of the fixed leg is found as the discounted value

$$V_{\text{Fixed}}(t, n, m) = \sum_{j=n+1}^m p_n(t, T_j) \Delta_j K$$

The value of the inflation indexed leg can be found as

$$\begin{aligned} V_{\text{II}}(t, n, m) &= \sum_{j=n+1}^m \Delta_j E_t^Q \left[e^{-\int_t^{T_j} n_s ds} \left(\frac{I_{T_j}}{I_{T_{j-1}}} - 1 \right) \right] \\ &= \sum_{j=n+1}^m \Delta_j \left(E_t^Q \left[e^{-\int_t^{T_{j-1}} n_s ds} E_{T_{j-1}}^Q \left[e^{-\int_{T_{j-1}}^{T_j} n_s ds} \frac{I_{T_j}}{I_{T_{j-1}}} \right] \right] - p_n(t, T_j) \right) \\ &= \sum_{j=n+1}^m \Delta_j \left(p_n(t, T_{j-1}) E_t^{T_{j-1}} [p_r(T_{j-1}, T_j)] - p_n(t, T_j) \right) \end{aligned}$$

Since

$$E_t^{T_{j-1}} [p_r(T_{j-1}, T_j)] = E_t^{T_{j-1}} [\exp(\log p_r(T_{j-1}, T_j))] = \psi_t^r(-i, T_{j-1}, T_j, T_{j-1})$$

where $\psi_t^r(-i, T_{j-1}, T_j, T_{j-1})$ is given in proposition 6, the value of the inflation indexed leg is given by

$$V_{\text{II}}(t, n, m) = \sum_{j=n+1}^m \Delta_j (p_n(t, T_{j-1}) \psi_t^r(-i, T_{j-1}, T_j, T_{j-1}) - p_n(t, T_j))$$

The value of the payer YIIS will then be

$$\begin{aligned} YIIS_t(n, m) &= V_{\text{II}}(t, n, m) - V_{\text{Fixed}}(t, n, m) \\ &= \sum_{j=n+1}^m \Delta_j (p_n(t, T_{j-1}) \psi_t^r(-i, T_{j-1}, T_j, T_{j-1}) - (1 + K)p_n(t, T_j)) \end{aligned}$$

Market quotes are in terms of the fixed payments K , such that the value of the swap is equal to zero. Thus the market quote will be

$$\begin{aligned} K &= \frac{\sum_{j=n+1}^m \Delta_j (p_n(t, T_{j-1}) \psi_t^r(-i, T_{j-1}, T_j, T_{j-1}) - p_n(t, T_j))}{\sum_{j=n+1}^m \Delta_j p_n(t, T_j)} \\ &= \frac{\left(\sum_{j=n+1}^m \Delta_j p_n(t, T_{j-1}) \psi_t^r(-i, T_{j-1}, T_j, T_{j-1}) \right) - PVBP_t(n, m)}{PVBP_t(n, m)} \end{aligned}$$

where $PVBP_t(n, m)$ is the present value of a basis point, also known from the interest rate swap(tion) literature.

zero-coupon Inflation Caps

We recall the definition from section 2.2, i.e. that a zero-coupon Inflation Cap over the period $[t, T]$ pays at time T

$$\begin{aligned} ZCIC_T(t, T, K) &= \left(\left(\frac{I_T}{I_t} - 1 \right) - \left((1 + K)^{T-t} - 1 \right) \right)^+ \\ &= \left(\frac{I_T}{I_t} - (1 + K)^{T-t} \right)^+ \end{aligned}$$

The price of the option is given by the risk neutral expectation

$$ZCIC_T(t, T, K) = p_n(t, T) E_t^T \left[\left(\frac{I_T}{I_t} - (1 + K)^{T-t} \right)^+ \right]$$

The value of the option crucially depends on the distribution of the CPI. Unfortunately, we do not know the distribution, but we know the characteristic function.

The knowledge of the characteristic function is fortunate, since Raible (2000) shows that the price of European options can be obtained through Fourier inversion, by the knowledge of the characteristic function of the (log-) underlying asset and the Fourier transformed contract function.

By use of proposition 7, we find the characteristic function of the log-Inflation index (under the nominal T -forward measure) as

$$\phi_t^I(u, t, T) = E_t^T [\exp(iu \log I_T/I_t)] = \psi_t^I(u, T, T)$$

where $\psi_t^I(u, T, T_l)$ is given in proposition 7.

Finally the results from Raible (2000) or Kluge (2005) show us how to translate from the characteristic function into actual Inflation zero-coupon Cap prices:

Proposition 9. *Assume that there exists a $\beta < -1$ such that $\phi_t^I(i\beta, T_n) < \infty$. Then the price of an zero-coupon Inflation Cap with maturity T and strike K is given by*

$$IZCC_t(t, T, K) = \frac{p_n(t, T)}{\pi} \int_0^\infty \Lambda(\beta + iu, \kappa) \phi_t^I(i\beta - u, T) du$$

where

$$\begin{aligned} \Lambda(v, \kappa) &= \frac{e^{(1+v) \log \kappa}}{v(v+1)} \\ \kappa &= (1 + K)^{T-t} \end{aligned}$$

In theory it would be most efficient to use a Fast Fourier Transform (FFT) method, cf. Carr and Madan (1999). However, in our model, since the characteristic function is the solution to a set of ordinary differential equations, it is expensive to evaluate compared to a model with a closed form characteristic function, making Gaussian quadrature a good alternative to FFT. Depending on parameter choices, between 50 and 100 integration points are needed to get a reasonable accuracy when using Gaussian quadrature, whereas at least 512 points are needed to achieved the same accuracy when using the FFT method.

Year-on-Year Inflation caplets

A Year-on-Year Inflation caplet is a call option on future inflation, as derived from the CPI. At time T_i the caplet pays⁸

$$\begin{aligned} YYIC_{T_i}(t, T_{i-1}, T_i, K) &= \Delta_i \left[\left(\frac{I_{T_i}}{I_{T_{i-1}}} - 1 \right) - \left((1 + K)^{T_i - T_{i-1}} - 1 \right) \right]^+ \\ &= \Delta_i \left[\frac{I_{T_i}}{I_{T_{i-1}}} - (1 + K)^{T_i - T_{i-1}} \right]^+ \end{aligned}$$

Hence, it can be interpreted as an option on the inflation from time T_{i-1} to time T_i with strike inflation rate K .

For pricing purposes, we consider rewriting the pay off as

$$\Delta_i \left(\frac{I_{T_i}}{I_{T_{i-1}}} - \kappa \right)^+$$

where $\kappa = (1 + K)^{T_i - T_{i-1}}$. The time t price can be found as the risk neutral expectation

$$YYIC_t(t, T_{i-1}, T_i, K) = p_n(t, T_i) \Delta_i E_t^{T_i} \left[\left(\frac{I_{T_i}}{I_{T_{i-1}}} - \kappa \right)^+ \right]$$

As in the case with the zero-coupon Inflation Cap, we can price this option by Fourier methods, if we know the time t conditional characteristic function of $\log I_{T_i}/I_{T_{i-1}}$. This can be found with the help of proposition 8, namely

⁸It is assumed that the option pay off is dependent on the inflation from time T_{i-1} to time T_i . The characteristic functions derived in section 2.5 also allows for the case where caplet dependent on the inflation between time T_{i-1} and T_i , is payed at time T_{i+1} ($T_{i-1} < T_i < T_{i+1}$).

using that

$$E_t^{T_i} \left[\exp \left(iu \log \frac{I_{T_i}}{I_{T_{i-1}}} \right) \right] = \psi_t^I(u, T_{i-1}, T_i, T_i)$$

Obviously, this relies on the assumption that the characteristic function in proposition 7 has an exponentially affine solution. However, as shown in the next section, even with a very general specification of the time change, the solution will indeed be exponentially affine.

As in the previous section, we can rely on the results from Raible (2000) or Kluge (2005), to price the floorlet.

Proposition 10. *Assume that there exists a $\beta < -1$ such that $\psi_t^I(i\beta, T_{i-1}, T_i, T_i) < \infty$. Then the price of a Year-on-Year Inflation Caplet with maturity T_i and strike K on the inflation between T_{i-1} and T_i , is given by*

$$YYIC_t(t, T_{i-1}, T_i) = \frac{p_n(t, T_i)\Delta_i}{\pi} \int_0^\infty \Lambda(\beta + iu, \kappa) \psi_t^I(i\beta - u, T_{i-1}, T_i, T_i) du$$

where

$$\Lambda(v, \kappa) = \frac{e^{(1+v)\log \kappa}}{v(v+1)}$$

$$\kappa = (1 + K)^{T_i - T_{i-1}}$$

2.7 Specification of the time-change

So far we have left the time-change unspecified, except for the fact that it must be a positive semi-martingale. As mentioned above, specifications of v that allow for easy calculation of the expectation in proposition 3 are indeed preferable. We propose using the affine processes described in Duffie, Filipovic, and Schachermayer (2003). These processes are generalizations of the affine jump diffusion models in Duffie, Pan, and Singleton (2000). Furthermore, as shown in Cheng and Scaillet (2007), the linear quadratic model found in Lieppold and Wu (2002), e.g., can be fitted into an affine jump diffusion framework. The affine processes described in Duffie, Filipovic, and Schachermayer (2003) also include non-Gaussian Ornstein-Uhlenbeck processes (see Barndorff-Nielsen and Shephard (2001)) and CBI processes (Conservative Continuous state Branching processes with Immigration, see for instance Filipovic (2001)).

In the following we describe the affine process, which we assume drives the time-change or activity in the economy. We will confine ourselves to the

one-dimensional case, however a generalization to a multivariate process is indeed possible (see Duffie, Filipovic, and Schachermayer (2003)).

We assume that the process v satisfies the stochastic differential equation (subject to parameter constraints that ensures positivity)

$$dv_t = (\theta + \kappa v_{t-}) dt + \sqrt{\alpha + \beta v_{t-}} dZ_t + \int_{\mathbb{R}_0^+} x \mu(dx, dt, v_{t-})$$

where Z is a Wiener process and the last term denotes a jump process possibly dependent of the activity state v . Both processes can be correlated with the Lévy process L . Furthermore the generator of the process is given by

$$\begin{aligned} \mathcal{A}f(v) &= (\hat{\theta} + \hat{\kappa}v_t) f'(v) + \frac{1}{2} (\alpha + \beta v_t) f''(v) \\ &\quad + \int_{\mathbb{R}_0^+} (f(v+x) - f(v) - xf'(v)1_{|x|<1}) (m(dx) + vm(dx)) \end{aligned}$$

where

$$\begin{aligned} \hat{\theta} &= \theta + f'(v) \int_{\mathbb{R}_0^+} x 1_{|x|<1} m(dx) \\ \hat{\kappa} &= \kappa + f'(v) \int_{\mathbb{R}_0^+} x 1_{|x|<1} n(dx) \end{aligned}$$

and the two measures $n(dx)$ and $m(dx)$ must satisfy the following condition

$$\int_{\mathbb{R}_0^+} (1 \wedge x) m(dx) + \int_{\mathbb{R}_0^+} (1 \wedge x^2) n(dx) < \infty$$

This implies that the jumps dictated by the measure $m(dx)$ must show finite variation, where the jumps related to the measure $n(dx)$ only have to show finite quadratic variation. Thereby the affine jump diffusions found in Duffie, Pan, and Singleton (2000) is a special case of this framework, namely where both $n(dx)$ and $m(dx)$ show finite activity.

As shown in Duffie, Filipovic, and Schachermayer (2003), the expectation needed to perform pricing has an exponential affine solution

$$E_t \left[\exp \left(r + qv_T + \int_t^T f(s)v_s ds \right) \right] = \exp (A(t, T) + B(t, T)v_t)$$

where $A(t, T)$ and $B(t, T)$ solve the generalized Riccati equations

$$\begin{aligned} \frac{dA(t, T)}{dt} &= -\hat{\theta}B(t, T) - \frac{1}{2}\alpha B(t, T)^2 \\ &\quad - \int_{\mathbb{R}_0^+} (e^{B(t, T)x} - 1 - xB(t, T)1_{|x|<1}) m(dx) \\ \frac{dB(t, T)}{dt} &= -\hat{\kappa}B(t, T) - \frac{1}{2}\beta B(t, T)^2 \\ &\quad - \int_{\mathbb{R}_0^+} (e^{B(t, T)x} - 1 - xB(t, T)1_{|x|<1}) n(dx) - f(t) \end{aligned}$$

subject to $A(T, T) = r$ and $B(T, T) = q$.

This direct application of the result in Duffie, Filipovic, and Schachermayer (2003) will, however, only be applicable when the activity rate process is independent of the driving Lévy process. As shown above, when this is no longer the case a change of measure is needed to preserve the analytical tractability. However, in most cases this will only result in changing the parameters of the activity rate process with time dependent parameters - thus the functional form of the solution remains the same. The specific change of parameters is highly model dependent. This depends on the choice of Lévy process, the choice of activity rate and the dependence between the two.

With regard to the dependence between the activity rate process and the driving Lévy process, recall that any purely continuous process is independent from any pure jump process. Furthermore, any finite activity jump process is also independent from any infinite activity jump process. Hence, if dependence between the activity rate process and the driving Lévy process is needed, then both must contain either a Wiener process, a finite activity jump process or an infinite activity process. The processes could possibly contain all three. However, as described in Geman, Madan, and Yor (2001), an infinite activity jump process could replace both a Wiener process and a finite activity jump process.

2.8 Calibration

In this section we consider the calibration of the model given in this paper. We consider data on inflation swaps and caps/floors linked to the EURO area HICP ex. tobacco. Furthermore, to identify parameters related to the nominal term structure, we include At-The-Money caps (which are not

Maturity	1	2	3	4	5
ZCIIS	0.91 %	1.45 %	1.69 %	1.84 %	1.97 %
YYIIS	N/A	N/A	1.65 %	N/A	1.92 %
NomZCB	0.986	0.962	0.933	0.9	0.865
RealZCB	0.995	0.99	0.981	0.968	0.954
Maturity	6	7	8	9	10
ZCIIS	2.08 %	2.17 %	2.24 %	2.29 %	2.34 %
YYIIS	N/A	2.12 %	N/A	N/A	2.25 %
NomZCB	0.83	0.794	0.759	0.725	0.691
RealZCB	0.939	0.923	0.906	0.888	0.871

Table 2.1: *Inflation zero-coupon Inflation Indexed Swap rates, Year-on-Year Inflation Indexed Swap rates, Nominal zero-coupon Bonds and Real zero-coupon Bonds from Inflation zero-coupon Inflation. Indexed Swaps Inflation swap rates are linked to EURO area HICP ex. Tobacco. Data is from August 10th 2009. Source: Bloomberg.*

Type	Floor	Floor	Floor	Cap	Cap	Cap	Cap
Mat. / Strike	-1 %	0 %	1%	2%	3%	4%	5%
1	15	35	75	31	14	6	3
2	33	66	129	125	69	39	23
3	60	102	187	214	127	81	54
5	97	175	289	443	267	169	114
7	122	195	332	651	390	243	162
10	151	241	409	931	558	388	227

Table 2.2: *Year-on-Year Inflation Cap and Floor prices (in basis points) linked to EURO area HICP ex. Tobacco. Data is from August 10th 2009. Source: Bloomberg.*

reported here). Our data is obtained from Bloomberg and is from August 10th 2009.

First, Table 2.1 shows market quotes of zero-coupon Inflation Indexed Swaps and Year-on-Year Inflation Indexed Swap rates. The table shows that due to expectations of low short term inflation rates, the term structure of ZCIIS break even rates is upward sloping, reflecting a gradual recovery of the EURO area economy.

Secondly, Table 2.2 shows market quotes (mid-prices) of Year-on-Year Inflation caps and floors. In the actual calibration a finer strike grid, than

presented in Figure 2.2, has been used. Furthermore, we extract prices for intermediate maturities by extracting flat volatilities from the caps and floors by using Black's formula (as for instance done in Mercurio and Moreni (2009)); we then interpolate flat volatilities for different maturities (and same strike rate) in order to get prices for all maturities. The observed and interpolated flat volatilities are shown in Figure 2.3. Finally, by taking differences between cap prices, we can extract caplet prices and implied volatilities. More precisely we minimize the squared percentage errors measured by implied volatilities (nominal caps and inflation caplets) and Year-on-Year swap rates.

Choice of volatility loading and driving process

To calibrate our model we need to make assumptions on the shape of the volatility loading $\sigma_i(t, T)$, the driving Lévy process L and the time change v .

We start by considering the volatility loadings. As a benchmark, we consider the Jarrow-Yildirim model, hence we choose a Vasicek-style volatility loading, i.e. :

$$\sigma_n(t, T) = \begin{pmatrix} \sigma_n e^{-\alpha_n(T-t)} \\ 0 \\ 0 \end{pmatrix}, \quad \sigma_r(t, T) = \begin{pmatrix} 0 \\ \sigma_r e^{-\alpha_r(T-t)} \\ 0 \end{pmatrix}, \quad \sigma_I(t) = \begin{pmatrix} 0 \\ 0 \\ \sigma_I \end{pmatrix}$$

In terms of the Lévy process, we consider a (correlated) Wiener process, i.e. a process with characteristic exponent:

$$\varphi(u) = -\frac{1}{2}u'\Sigma u$$

where Σ is the correlation matrix⁹.

We also consider a Lévy process based on a Variance Gamma (VG) process¹⁰. To include correlation between the elements in the Lévy process, we consider the following specification of a multivariate Lévy process, as introduced in Ballotta and Bonfiglioli (2010):

$$\begin{pmatrix} L_{1,t} \\ L_{2,t} \\ L_{3,t} \end{pmatrix} = \begin{pmatrix} X_{1,t} + a_1 X_{Z,t} \\ X_{2,t} + a_2 X_{Z,t} \\ X_{3,t} + a_3 X_{Z,t} \end{pmatrix}$$

⁹In the calibration routine, we fix the correlations to historical estimates obtained from time series of nominal rates and inflation swaps.

¹⁰The VG process is obtained by subordinating an arithmetic Brownian motion with a Gamma process. Appendix 2.11 provides a brief description of the VG process.

where X_i is a VG process with parameters $(\beta_i, \gamma_i, \nu_i)$. This process has a characteristic exponent given by

$$\begin{aligned} \varphi(u) = & -\frac{1}{\nu_Z} \log \left[1 - i\beta_Z \nu_Z \left(\sum_{j=1}^3 a_j u_j \right) + \frac{\gamma_Z^2}{2} \nu_Z \left(\sum_{j=1}^3 a_j u_j \right)^2 \right] \\ & - \sum_{j=1}^3 \left(\frac{1}{\nu_i} \log \left[1 - iu_j a_j \beta_j \nu_j + u_j^2 \frac{a_j^2 \gamma_j^2}{2} \nu_j \right] \right) \end{aligned}$$

Finally, we consider the time change. For the Wiener process and Variance Gamma based model we use a CIR-process, i.e. that the rate of the time change v solves the SDE

$$dv_t = \kappa(\theta - v_t)dt + \eta\sqrt{v_t}dZ_t$$

For identification purposes, and to preserve the intuition of v being a time change, we fix θ to be equal to one. When we calibrate the Wiener process based model, we assume a common correlation coefficient γ , between all three Wiener processes in the driving process W .

Calibration results

With the above model specification, we calibrate four different models to the market data; 1) a standard Jarrow-Yildirim model, 2) a Jarrow-Yildirim model with time change, 3) a Variance Gamma based model without time change and finally 4) a Variance Gamma model with time change. We show the calibration results in Figure 2.4, where the different models are compared to the extracted inflation caplet implied volatilities (3, 5 and 10 year maturities).

The first observation is that none of the models obtain a perfect fit to the data. Worst is the Jarrow-Yildirim model; not surprisingly, the assumption of a Gaussian model is too restrictive. In terms of adding stochastic volatility (time change) to the Jarrow-Yildirim model, the main improvement arises from fitting longer term data better, since the time change adds a minor 'smile' at longer maturities. In addition the time change seems to help in fitting the overall level of volatilities, i.e. removing some of the restrictions implied by the Vasicek-type volatility loading.

In terms of the Variance Gamma based models, we see that there is little difference between including the time change or not. However, due to the

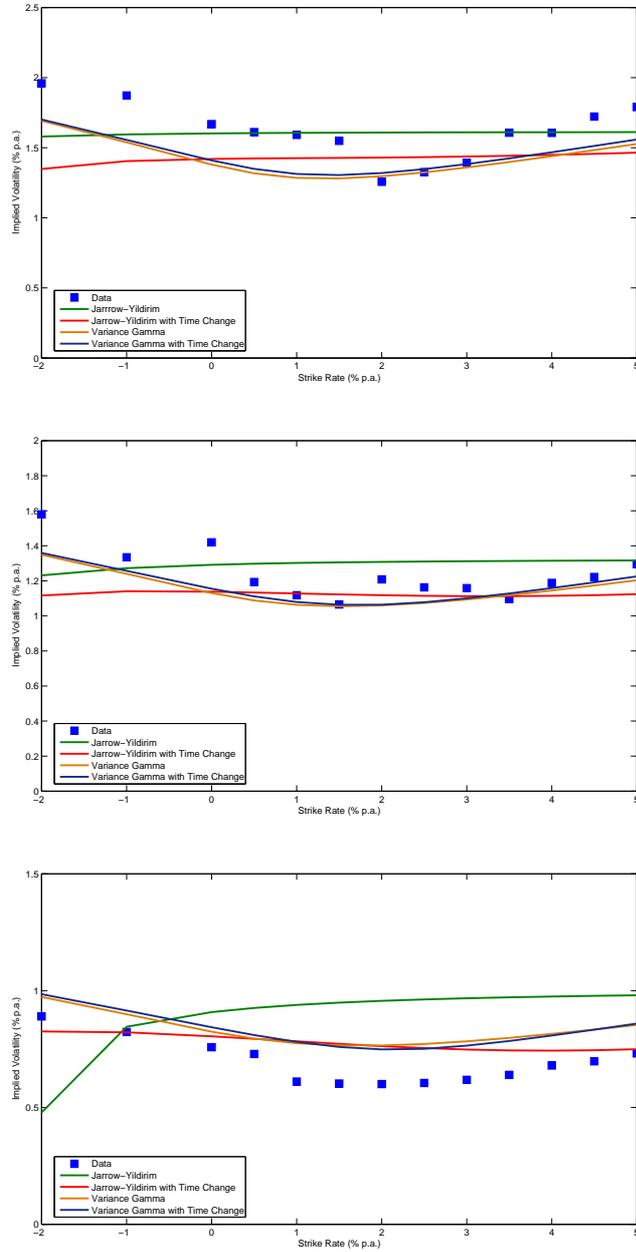


Figure 2.4: Calibration of different models to market data (inflation caplets). The upper graph is the 3 year inflation caplet, the middle graph is the 5 year inflation caplet and the lower graph is the 10 year inflation caplet.

flexibility of the VG process, a time series study is probably needed in order to assess the importance of the time change.¹¹

Interestingly, the Variance Gamma based models seem to capture the shape of the volatility smile. However, volatility tends to be underestimated for shorter maturities, but overestimated for longer maturities. Again, we believe that this can be attributed to the restrictive Vasicek volatility loading.

These findings indicate that further work has to be done in different areas. First, the exact specification of the driving Lévy process and time change could possibly lead to more flexible volatility smiles. For instance, in terms of interest rate modeling, Trolle and Schwartz (2009) show that a three factor model with stochastic volatility, as opposed to a one factor model in this paper, is needed in order to capture the dynamics of nominal caps and swaptions.

Secondly, easing the restriction on the volatility loading should also add flexibility to the modeling. Trolle and Schwartz (2009) show that a hump shaped volatility loading is indeed preferable compared to a Vasicek specification.

2.9 Conclusion

In this paper, we have proposed an inflation modeling framework based on the Heath-Jarrow-Morton approach found in Jarrow and Yildirim (2003). The novelty of our framework arises from the use of a time changed Lévy process, which allows more realistic modeling of real and nominal forward rate and inflation dynamics.

By using a Heath-Jarrow-Morton framework based on time changed Lévy processes, we show that the drift conditions can be expressed as a compound of the volatility loading, the characteristic exponent of the Lévy process and the current state of the subordinator. As described in Hinnerich (2008), we do not need to assume the foreign exchange analogy as in Jarrow and Yildirim (2003). This comes as a by-product of the general derivations.

Due to the analytical tractability of the time changed Lévy processes, we manage to derive a range of characteristic functions. Namely the charac-

¹¹This could possibly be due to the fact, that a time change of a Lévy process corresponds to a scaling of the Lévy measure, i.e. controlling how frequent jumps occur. For the Variance Gamma model, the ν parameters also control how frequent jumps arrive, which implies that the effect of the subordination using a CIR process would mostly be beneficial in the time series dimension.

teristic functions for nominal and real zero-coupon bond, and for inflation and forward inflation.

These characteristic functions allow us to price standard inflation products, such as zero-coupon Inflation Indexed Swaps, Year-on-Year Inflation Indexed swap and finally inflation indexed caplets and floorlets.

By assuming that the subordinator is driven by the generalized affine processes (see Duffie, Filipovic, and Schachermayer (2003)), the prices of inflation indexed derivatives are known up to ordinary differential equations and possibly a Fourier inversion.

Finally, we calibrate the model to market data. We find that a model based on a Variance Gamma model captures the shape of the volatility smile, whereas Gaussian models more seem to capture the overall level of the volatility smile. We also find that the volatility loadings implied by the original Jarrow-Yildirim model, may be too restrictive in order to fit the volatility surface properly.

2.10 Appendix: Proofs

Proof of proposition 3

Proof. We first prove that M is a \mathbb{Q} -martingale. The dynamics of $\log M$ are given by

$$d \log M_t = -\varphi(-i\xi(s)) v_{t-} dt + \xi(t)' dY_t$$

This implies the dynamics of M

$$\frac{dM_t}{M_{t-}} = \xi(t)' \Sigma^{1/2} dW_{\tau_t} + \int_{\mathbb{R}_0^d} \left(e^{\xi(t)'x} - 1 \right) (\mu(dx, dt, v_{t-}) - \nu(dx) v_{t-} dt)$$

Which is a martingale if the following condition is satisfied

$$E^{\mathbb{Q}} [|M_t|] = E^{\mathbb{Q}} \left[\exp \left(- \int_0^t \varphi(-i\xi(s)) v_{s-} ds + \int_0^t \xi(s)' dY_s \right) \right] < \infty$$

The fact that M is a martingale implies

$$\begin{aligned} \psi(t, T) &= E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T \xi(s)' dY_s \right) \right] \\ &= E_t^{\mathbb{Q}} \left[\exp \left(\pm \int_t^T \varphi(-i\xi(s)) v_{s-} ds + \int_t^T \xi(s)' dY_s \right) \right] \\ &= E_t^{\mathbb{Q}} \left[\exp \left(\int_t^T \varphi(-i\xi(s)) v_{s-} ds \right) \frac{M_T}{M_t} \right] \\ &= E_t^{\xi} \left[\exp \left(\int_t^T \varphi(-i\xi(s)) v_{s-} ds \right) \right] \end{aligned}$$

Where $E_t^{\xi}[\bullet]$ denotes expectation with regard to the probability measure defined by

$$\left. \frac{d\mathbb{Q}^{\xi}}{d\mathbb{Q}} \right|_{\mathcal{F}_t} = M_t = \exp \left(- \int_0^t \varphi(-i\xi(s)) v_{s-} ds + \int_0^t \xi(s)' dY_s \right)$$

□

Proof of proposition 4

We start by defining the discounted nominal bond and discounted inflation protected bond:

$$F_t \equiv \frac{p_n(t, T)}{\beta_t^n}, \quad H_t \equiv \frac{p_r(t, T)I_t}{\beta_t^n}$$

We need to find conditions such that these two processes are martingales. First for nominal bonds:

The dynamics of $\log F_t$ follow easily

$$d \log F_t = (A_n(t, T) + S_n(t, T)' av_{t-}) dt + S_n(t, T)' \Sigma^{1/2} dW_{\tau_t} + \int_{\mathbb{R}_0^d} S_n(t, T)' x \mu(dx, d\tau_t)$$

Which gives us the dynamics of F_t

$$\begin{aligned} \frac{dF_t}{F_{t-}} &= (A_n(t, T) + S_n(t, T)' av_{t-} + S_n(t, T)' \Sigma S_n(t, T) v_{t-}) dt + \\ &\int_{\mathbb{R}_0^d} \left(e^{S_n(t, T)' x} - 1 \right) \nu(dx) v_{t-} dt + S_n(t, T)' \Sigma^{1/2} dW_{\tau_t} \\ &\int_{\mathbb{R}_0^d} \left(e^{S_n(t, T)' x} - 1 \right) (\mu(dx, d\tau_t) - \nu(dx) v_{t-} dt) \end{aligned}$$

The last two terms are martingales, thus the nominal drift condition is

$$A_n(t, T) = -v_{t-} \varphi(-i S_n(t, T))$$

Finally we seek to derive a drift condition for the inflation protected zero-coupon bonds. We consider the dynamics of $\log H_t$

$$d \log H_t = (r_t + A_r(t, T) + \alpha_I(t) - n_t + (S_r(t, T) + \sigma_I(t))' av_{t-}) dt + (S_r(t, T) + \sigma_I(t))' \Sigma^{1/2} dW_{\tau_t} + \int_{\mathbb{R}_0^d} (S_r(t, T) + \sigma_I(t))' x \mu(dx, d\tau_t)$$

The dynamics of H_t are then

$$\begin{aligned} \frac{dH_t}{H_{t-}} &= (A_r(t, T) + \alpha_I(t) + r_t - n_t + (S_r(t, T) + \sigma_I(t))' av_{t-}) dt + \\ &\left(\frac{1}{2} (S_r(t, T) + \sigma_I(t))' \Sigma (S_r(t, T) + \sigma_I(t)) v_{t-} \right) dt + \\ &\int_{\mathbb{R}_0^d} \left(e^{(S_r(t, T) + \sigma_I(t))' x} - 1 \right) \nu(dx) v_{t-} dt + (S_r(t, T) + \sigma_I(t))' \Sigma^{1/2} dW_{\tau_t} + \\ &\int_{\mathbb{R}_0^d} \left(e^{(S_r(t, T) + \sigma_I(t))' x} - 1 \right) (\mu(dx, d\tau_t) - \nu(dx) v_{t-} dt) \end{aligned}$$

The last two terms are martingales, giving us the condition:

$$\begin{aligned} A_r(t, T) + \alpha_I(t) + r_t - n_t + ((S_r(t, T) + \sigma_I(t))' a) v_{t-} + \\ \frac{1}{2} (S_r(t, T) + \sigma_I(t))' \Sigma (S_r(t, T) + \sigma_I(t)) v_{t-} + \\ \int_{\mathbb{R}_0^d} \left(e^{(S_r(t, T) + \sigma_I(t))' x} - 1 \right) \nu(dx) v_{t-} dt = 0 \end{aligned}$$

Or equivalently

$$\begin{aligned} A_r(t, T) + (\varphi(-i(S_r(t, T) + \sigma_I(t))) - \varphi(-i\sigma_I(t))) v_{t-} = \\ -\alpha_I(t) + (n_t - r_t - v_{t-} \varphi(-i\sigma_I(t))) \end{aligned}$$

Since this must be satisfied for all $t < T$ we must have

$$\begin{aligned} \alpha_I(t) &= n_t - r_t - v_{t-} \varphi(-i\sigma_I(t)) \\ A_r(t, T) &= -v_{t-} (\varphi(-i(S_r(t, T) + \sigma_I(t))) - \varphi(-i\sigma_I(t))) \end{aligned}$$

Which completes the proof.

Proof of proposition 5

We are interested in

$$\begin{aligned} \psi_t^n(u, T, T_k, T_l) &= E_t^{T_l} [\exp(iu \log p_n(T, T_k))] \\ &= E_t^Q \left[\exp(iu \log p_n(T, T_k)) \frac{N_T}{N_t} \right] \end{aligned}$$

Where

$$\begin{aligned} N_t &= \frac{dQ^{T_l}}{dQ} \Big|_{\mathcal{F}_t} = \frac{p_n(t, T_l)}{\beta_t^n} \frac{\beta_0^n}{p_n(0, T_l)} \\ &= \exp \left(- \int_0^t \varphi(-iS_n(s, T_l)) v_{s-} ds + \int_0^t S_n(s, T_l)' dY_s \right) \end{aligned}$$

Thus we are interested in the expectation

$$\psi_t^n(u, T, T_k, T_l) = \left(\frac{p_n(t, T_k)}{p_n(t, T)} \right)^{iu} E_t^Q \left[\exp \left(\int_t^T h(s, u) v_{s-} ds + \int_t^T \xi(s, u)' dY_s \right) \right]$$

where

$$\begin{aligned} h(s, u) &= iu [\varphi(-iS_n(s, T)) - \varphi(-iS_n(s, T_k))] - \varphi(-iS_n(s, T_l)) \\ \xi(s, u) &= iu (S_n(s, T_k) - S_n(s, T)) + S_n(s, T_l) \end{aligned}$$

Using the result in proposition 3 gives us the result.

Proof of proposition 7

We are interested in the expectation

$$\begin{aligned}\psi_t^I(u, T, T_l) &= E_t^{T_l} [\exp(iu(\log I_T - \log I_t))] \\ &= \left(\frac{p_r(t, T)}{p_n(t, T)}\right)^{iu} E_t^{T_l} \left[\exp\left(\int_t^T h(s, u)v_{s-} ds + \int_t^T \xi(s, u)' dY_s\right) \right]\end{aligned}$$

where

$$\begin{aligned}h(s, u) &= iu [\varphi(-iS_n(s, T)) - \varphi(-i(S_r(s, T) + \sigma_I(s)))] \\ \xi(s, u) &= iu (\sigma_I(s) + S_r(s, T) - S_n(s, T))\end{aligned}$$

Changing to the nominal risk neutral measure, yields

$$\psi_t^I(u, T, T_l) = \left(\frac{p_r(t, T)}{p_n(t, T)}\right)^{iu} E_t^Q \left[\exp\left(\int_t^T \hat{h}(s, u)v_{s-} ds + \int_t^T \hat{\xi}(s, u)' dY_s\right) \right]$$

where

$$\begin{aligned}\hat{h}(s, u) &= iu [\varphi(-iS_n(s, T)) - \varphi(-i(S_r(s, T) + \sigma_I(s)))] - \varphi(-iS_n(s, T_l)) \\ \hat{\xi}(s, u) &= iu (\sigma_I(s) + S_r(s, T) - S_n(s, T)) + S_n(s, T_l)\end{aligned}$$

Using the result in proposition 3 gives us the result.

Proof of proposition 8

We start by considering

$$\psi_t^I(u, T, T_k, T_l) = E_t^{T_l} \left[\exp\left(iu \log \frac{I_{T_k}}{I_T}\right) \right] = E_t^{T_l} \left[E_T^{T_l} \left[\exp\left(iu \log \frac{I_{T_k}}{I_T}\right) \right] \right]$$

Using the assumption of the exponentially affine solution, gives us

$$\begin{aligned}\psi_t^I(u, T, T_k, T_l) &= E_t^{T_l} \left[\left(\frac{p_r(T, T_k)}{p_n(T, T_k)}\right)^{iu} \exp(A(T, T_k) + B(T, T_k)v_{T-}) \right] \\ &= \left(\frac{p_r(t, T_k)}{p_r(t, T)} \frac{p_n(t, T)}{p_n(t, T_k)}\right)^{iu} \times \\ &\quad E_t^{T_l} \left[\exp\left(\int_t^T g(s, u)v_{s-} ds + \int_t^T \xi(s, u)' dY_s\right) \times \right. \\ &\quad \left. \exp(A(T, T_k) + B(T, T_k)v_{T-}) \right]\end{aligned}$$

where

$$g(s, u) = iu \left(\varphi(-i(S_r(s, T) + \sigma_I(s))) - \varphi(-iS_n(s, T)) - \right. \\ \left. \varphi(-i(S_r(s, T_k) + \sigma_I(s))) + \varphi(-iS_n(s, T_k)) \right) \\ \xi(s, u) = iu \left(S_r(s, T_k) - S_n(s, T_k) - S_r(s, T) + S_n(s, T) \right)$$

Changing to the nominal martingale measure \mathbb{Q} gives us

$$\psi_t^I(u, T, T_k, T_l) = \left(\frac{p_r(t, T_k) p_n(t, T)}{p_r(t, T) p_n(t, T_k)} \right)^{iu} \times \\ E_t^{T_l} \left[\exp \left(\int_t^T h(s, u) v_{s-} ds + \int_t^T k(s, u)' dY_s \right) \times \right. \\ \left. \exp(A(T, T_k) + B(T, T_k) v_{T-}) \right]$$

where

$$h(s, u) = iu \left(\varphi(-i(S_r(s, T) + \sigma_I(s))) - \varphi(-iS_n(s, T)) - \right. \\ \left. \varphi(-i(S_r(s, T_k) + \sigma_I(s))) + \varphi(-iS_n(s, T_k)) \right) - \varphi(-iS_n(s, T_l)) \\ k(s, u) = iu \left(S_r(s, T_k) - S_n(s, T_k) - S_r(s, T) + S_n(s, T) \right) + S_n(s, T_l)$$

Applying proposition 3 gives us the result.

2.11 Appendix: The Variance Gamma process

The Variance Gamma process (as introduced in Madan and Seneta (1990)) arises from subordinating an arithmetic Brownian motion with a Gamma process.

More precisely consider the arithmetic Brownian Motion

$$Z(t) = \beta t + \gamma W(t)$$

Next, the Variance Gamma process arises from subordinating this process with a Gamma Process, $G(t) \sim \Gamma(a, b)$

$$X(t) = Z(G(t)) = \beta G(t) + \gamma W(G(t))$$

With respect to the Gamma process, Madan, Carr, and Chang (1998) show that it is sufficient to consider a gamma process where $a = b = 1/\nu$. In this case the Gamma process reflects an unbiased clock (i.e. $E[G(t)] = t$). Using this specification, we say the process X is $VG(\beta, \gamma, \nu)$.

The VG process is a Lévy process (with infinite activity and finite variation) and has characteristic exponent given by

$$\varphi(u) = -\frac{1}{\nu} \log \left(1 - iu\beta\nu + u^2 \frac{\beta^2}{2} \nu \right)$$

Finally the VG process has mean, variance, skewness and kurtosis given as

$$\begin{aligned} E[X(t)] &= \beta t \\ \text{Var}(X(t)) &= (\gamma^2 + \beta^2 \nu) t \\ \text{Skew}(X(t)) &= \frac{(3\gamma^2 + 2\beta^2 \nu) \beta \nu}{(\gamma^2 + \beta^2 \nu)^{3/2} \sqrt{t}} \\ \text{Kurt}(X(t)) &= \frac{(3\gamma^4 + 12\gamma^2 \beta^2 \nu + 6\beta^4 \nu^2) \nu}{(\gamma^2 + \beta^2 \nu)^2 t} \end{aligned}$$

Essay 3

Inflation risk premia in the term structure of interest rates: Evidence from Euro area inflation swaps

Abstract

We estimate inflation risk premia in the Euro area using inflation swaps. By proposing a no-arbitrage model for econometric analysis, and estimating it using Markov Chain Monte Carlo, we find estimates of inflation risk premia, that on average show an upward sloping term structure, with 1 year risk premia of 18 bps and 10 year risk premia of 43 bps, however with fluctuation in risk premia over time. Our estimates suggest that surveys are important in identifying inflation expectations and thus inflation risk premia. We relate estimates of inflation risk premia to agents beliefs, and find that skews in short term inflation perceptions drive short term inflation risk premia, where beliefs on GDP growth drive longer term risk premia.

Keywords: Inflation risk premia, Inflation expectations, Inflation swaps, Surveys, Affine Term Structure Models, Markov Chain Monte Carlo

JEL Classification: C11, C58, E31, E43, G12

¹I would like to thank Thomas Werner, Jacob Ejsing, Kasper Lorenzen, Mads Stenbo Nielsen, Peter Feldhütter, Jesper Lund, Anne-Sofie Reng Rasmussen and participants at presentations at the Capital Markets/Financial Structure Division (ECB), Danmarks Nationalbank, Copenhagen Business School and the 2010 WHU Campus for Finance Conference for useful suggestions.

3.1 Introduction

The ability to correctly estimate inflation risks are vital to investors, as well as central banks. One such measure is the Break Even Inflation Rate (BEIR), which is the difference in yield between a nominal and real bond.

Another measure is provided by inflation swaps. More precisely zero-coupon inflation indexed swaps, are swap agreements who at maturity pay the change in the reference index (the Consumer Price Index) as the floating leg and a pre-specified fixed payment as the fixed leg. The fixed leg is set, so that the contract has a value of zero at initiation. Hence the quotes of inflation swaps gives an additional measure of the BEIR. Typically inflation swaps require less capital to hold, than inflation linked bonds, making these contracts less prone to market distortions. In fact around the collapse of Lehman Brothers (end-2008), the spread between inflation swap rates and BEIRs from inflation indexed bonds, widened due to the financial crises and liquidity effects (see for instance Campbell, Schiller, and Viceira (2009) for an elaboration on this issue).

Recently a number of papers have tried to estimate inflation risk premia using various methodologies. On US-data the analysis have mainly been focused on using CPI data, surveys and/or US treasury inflation protected securities (TIPS) to estimate the inflation risk premia (see Ang, Bekaert, and Wei (2008), D'Amico, Kim, and Wei (2008), Chernov and Mueller (2008) and Christensen, Lopez, and Rudebusch (2008)). The only paper to use inflation swap data is Haubrich, Pennacchi, and Ritchken (2008), who use US inflation swap data.

With regard to Euro Area data, we are aware of three papers, namely Tristani and Hördahl (2007), Garcia and Werner (2010) and Tristani and Hördahl (2010). All papers extract real yields from inflation indexed bonds, and then estimate inflation expectations and inflation risk premia.

Overall only a few of these studies agree on the size of the inflation risk premia. Some papers have inflation risk premia of up to 300 basis points (Chernov and Mueller (2008)), where others show more moderate fluctuations (-50 to 50 basis points, see for instance Christensen, Lopez, and Rudebusch (2008)). These differences seem to arise from small differences in data periods and the data included, e.g. for instance the inclusion of surveys or not. Finally, only Tristani and Hördahl (2007) present confidence bands on their of estimates inflation risk premia. They find that their estimate of inflation risk premia is statistically insignificant for most of the considered maturities.

In this paper we focus on Euro area inflation risk premia. However, instead

of using inflation indexed bonds to identify real yields, we use inflation swaps. We choose to use inflation swaps, since inflation swaps linked to the Euro area HICP have developed into a fairly liquid market.² As mentioned above, swaps require less capital to hold, and swap rates are less likely to be distorted by market related issues compared to cash products. Furthermore inflation swap rates have the advantage, that they can be included directly into an estimation, making use of the data less prone to errors from interpolation. Finally, to our knowledge, we are the first to conduct an analysis on inflation risk premia using inflation swaps.

Rather than trying to relate inflation risk premia to a large framework including agents, GDP, etc., we use a reduced form approach. The choice of a reduced form model is motivated by the large degree of disagreement on the inflation risk premia. We rely on the existing literature on continuous time term structure models (see Duffie and Kan (1996) and Dai and Singleton (2002)), extended with an inflation process similar to D'Amico, Kim, and Wei (2008), although with slight differences. Thereby real zero-coupon bonds (and inflation swaps) can be priced through no-arbitrage methods.

To more easily identify inflation risk premia we follow Garcia and Werner (2010) and include the ECB survey of professional forecasters. Since we in this paper use a fairly short time series (data from 1999), we are likely to face a small-sample bias. The use of surveys may help to reduce such a bias and help identify the model. Furthermore, to derive the inflation risk premia, we want to model the inflation expectations of market participants. With all likelihood one can construct a model, which fits realised inflation better than surveys, however such a model may not be representative of the actual inflation expectation, thus leading to wrong estimates of inflation risk premia.

We estimate our model using a Bayesian approach, namely Markov Chain Monte Carlo. This allows us to draw precise inference on derived variables such as inflation expectations and risk premia. By using draws from the Markov Chain Monte Carlo estimation, we examine the effect of including surveys, and find that surveys improve the identification of inflation expectations and thus inflation risk premia.

In terms of our estimate of inflation risk premia, we obtain estimates of average inflation risk premia that are increasing in time to maturity, with 1 year risk premia of 18 basis points and 10 year risk premia of 43 basis points. These show significant fluctuations with 1 year inflation risk premia being

²In terms of US inflation linked markets, TIPS are still by far the most actively traded product, thus having a significant negative effect on the US inflation swap markets.

between -146 and 68 basis points, with the lowest value being in the time after the collapse of Lehman Brothers. Longer term inflation risk premia (10 year) show less variation, with inflation risk premia between -30 and 81 basis points.

Finally we relate the estimated risk premia to agents beliefs on the outcome of the economy. We find that short term inflation risk premia are mainly driven by the skewness of the distribution of inflation as measured by the ECB survey of professional forecasters, where longer term risk premia are driven by GDP expectations.

The paper is structured as follows: Section 3.2 describes the data and provides an ad-hoc measure of inflation risk premia. Section 3.3 and 3.4 introduce the no-arbitrage model which we use to estimate inflation risk premia, and section 3.5 describes our estimation methodology. Section 3.6 describes the empirical results and finally section 3.7 concludes the paper.

3.2 Data: Inflation swap rates and the nominal term structure

Initial description

In this section we will describe the data on inflation swap rates, and its connection to the nominal term structure. A zero-coupon inflation swap is a swap agreement where the floating leg pays the percentage change on the reference consumer price index over some reference period $[t, T]$, which for the Euro area is the HICP ex. tobacco index:

$$ZCIIIS_T(t, T, K) = \left(\frac{I(T)}{I(t)} - 1 \right) - ((1 + K)^{T-t} - 1)$$

Zero-coupon inflation swap rates are quoted in terms of the fixed rates K , and the quotes will therefore reflect a market based inflation expectation over the considered period. It can be shown that inflation swap rates can be derived through nominal and real interest-rates. Here we term real interest-rates, as the *ex-ante* real rates, as for instance can be derived from normal inflation linked bonds. On the other hand, due to this relationship, real rates can also be derived from inflation swaps and nominal interest-rates.³

³We ignore the fact that inflation swaps have an indexation lag, i.e. that the swaps fix to the CPI released 3 month prior to the maturity of the swaps. This approximation will mainly affect 1 year inflation risk premia.

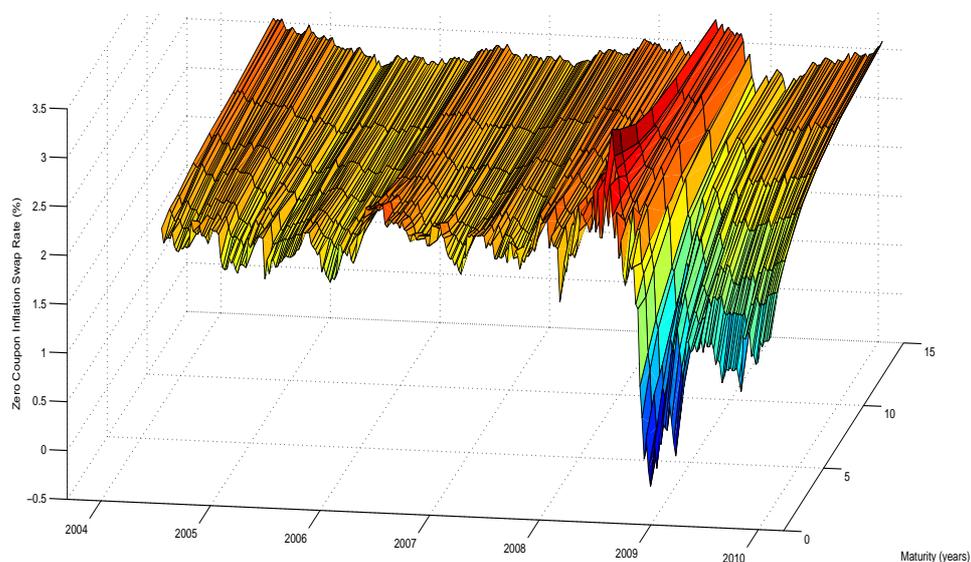


Figure 3.1: Time series of zero-coupon inflation swap rates. The data sample is June 2004 to January 2010. Source: Bloomberg.

Next we turn to our data. From Bloomberg we collect weekly data on zero-coupon inflation swaps on Euro area HICP ex. tobacco from June 2004 to January 2010. Similarly we collect swap rates (also from Bloomberg) which range from January 1999 to January 2010. Figure 3.1 shows time series of inflation swap rates and Figure 3.2 shows the times series on nominal swap rates.⁴

As seen from Figure 3.1, inflation swap rates saw large variability through 2008. First inflation swap rates rose in the first half of 2008 due to rising commodity prices, and in the latter part of 2008 the fact that the financial crisis spread to the real economy triggered strong downward revisions of inflation swap rates.⁵ Apart from this period, inflation swap rates have been fairly stable with long term rates around 2.5 percent and short term rates being more affected by short term fluctuations in inflation.

⁴We perform weekly sampling of the data on Wednesdays to avoid weekday effects, see Lund (1997).

⁵Part of this drop in inflation rates can also be related to liquidity reasons, although inflation swaps have been less affected than inflation linked bonds, as a consequence of the swap structure (vs. the cash structure of inflation linked bonds).

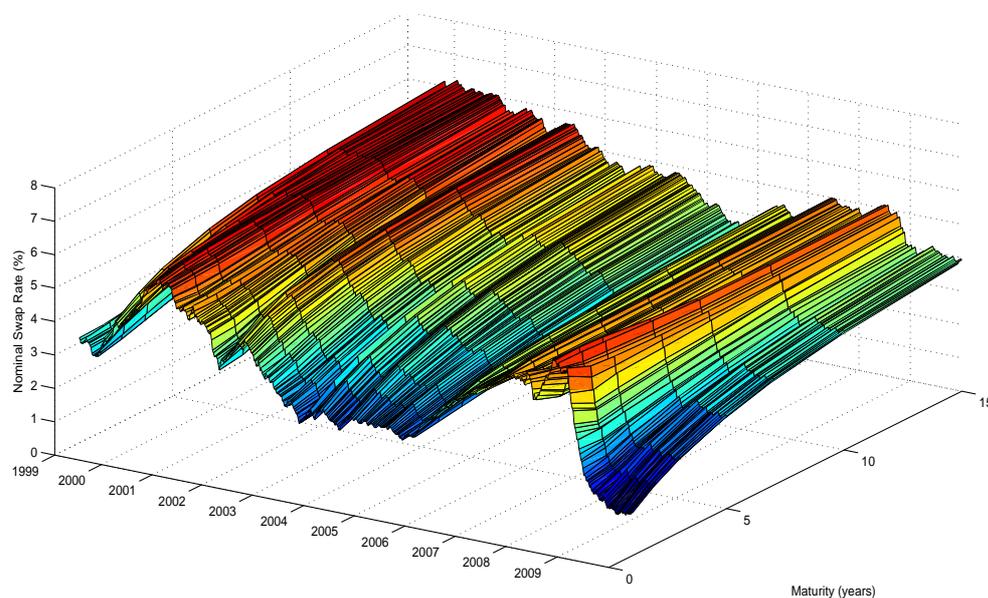


Figure 3.2: Time series of nominal swap rates. The data sample is January 1999 to January 2010. Source: Bloomberg.

Linking the nominal term structure and inflation swaps

As first shown in Litterman and Scheinkman (1991), the nominal term structure can be described by a number of principal components, typically three. From Figure 3.1 and 3.2 there is visual evidence that at least some of the variation of inflation swap rates is captured by the nominal term structure, and hence its principle components. Thus to capture the structure between the data, we find the principal components of the changes in nominal swap rates, and perform a regression where changes in inflation swap rates are explained by the principal components. The top panel in Table 3.1 shows the result from the principal components analysis (PCA) of the nominal interest-rate data.⁶

First of all, our PCA on the nominal term structure confirms the usual findings, i.e. that three principal components are sufficient to describe the nominal term structure. Also, our three principal components have the usual interpretation of level, slope and curvature, although the first two factors also could be described as flat and steep slope factors.

⁶A PCA performed on the swap rate levels gives a similar result, albeit significantly higher R^2 's are obtained when regressing the principal components on inflation swap levels.

PC	% Explained by PC	Nominal Yield Maturity						
		1Y	2Y	3Y	5Y	7Y	10Y	15Y
1st PC	87.34 %	0.2718	0.3977	0.4232	0.4067	0.3971	0.3703	0.3586
2nd PC	10.25 %	-0.5067	-0.4217	-0.2792	0.0026	0.2198	0.3787	0.5438
3rd PC	1.48 %	-0.7490	0.1055	0.3898	0.3316	0.1072	-0.1286	-0.3713

	Zero Coupon Inflation Swap Rate Maturity						
	1Y	2Y	3Y	5Y	7Y	10Y	15Y
Constant	0.0000 (-0.3847)	0.0000 (-0.2858)	0.0000 (-0.2076)	0.0000 (-0.1241)	0.0000 (-0.0484)	0.0000 (0.0782)	0.0000 (0.0540)
1st PC	0.1304 (2.5699)	0.1449 (2.9958)	0.1194 (2.8087)	0.0997 (2.6789)	0.0735 (3.4606)	0.0718 (3.6847)	0.0447 (2.3013)
2nd PC	0.0639 (0.5099)	0.0281 (0.18720)	0.0189 (0.1541)	-0.0163 (-0.1676)	-0.0232 (-0.2674)	0.004 (0.0702)	-0.0533 (-0.7956)
3rd PC	-0.2538 (-0.7031)	-0.4495 (-1.4749)	-0.3223 (-1.2015)	-0.0997 (-0.5528)	-0.0269 (-0.1402)	0.0477 (0.3135)	0.1351 (0.7657)
R^2	0.0498	0.1134	0.1045	0.1026	0.0764	0.1010	0.0545

Table 3.1: **Top Panel:** Results from Principal Components Analysis on changes in nominal swap rates. **Bottom Panel:** Regression of Zero Coupon Inflation Swap Rates on Principal Components from the nominal swap rates. Newey-West t -statistics are given in brackets.

Table 3.2: Regression of changes in Zero Coupon Inflation Swap Rates on Principal Components from the nominal term structure and the first principal component from the residuals of the regression in the bottom panel in figure 3.1. Newey-West t -statistics are given in brackets.

Zero Coupon Inflation Swap Rate Maturity		1Y	2Y	3Y	5Y	7Y	10Y	15Y
Constant	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1st PC	(-1.0516)	0.1304	0.1449	0.1194	(-0.3681)	(-0.1228)	(0.1558)	(0.0864)
2nd PC	(4.9332)	0.0639	0.0281	0.0189	(4.8650)	(7.7004)	(7.4474)	(3.0978)
3rd PC	(0.9020)	(0.5751)	(0.6135)	(-0.4764)	(-0.6919)	(-0.1644)	0.0477	(-1.0386)
ZCIS PC	(-2.0367)	0.6229	0.4788	(-4.0473)	(-0.8844)	(-0.2737)	0.1927	(0.8579)
R^2	(12.8493)	0.7897	0.8673	(17.5182)	0.7593	0.7129	0.6149	0.4416

Next we regress the change in each inflation swap rate on the principal components to see how much of the variation in inflation swap rates is explained by the nominal principal components. The bottom panel in Table 3.1 shows the results from these regressions. Our first observation is that the R^2 's from the different regressions are between 4 and 11 percent. This is in contrast to the explanation percentage of about 99 percent in the PCA on the nominal term structure. This implies that part of the variation in inflation swap rates is not captured by the nominal term structure.

In practice this implies that we would model the nominal term structure with three factors, but we would need (at least) one more factor to model the inflation swap rates. To address this issue, we perform another PCA on the residuals from the regressions mentioned above. We then repeat our regressions from before, but we also include the first principal component from the PCA on the residuals. The results are given in Table 3.2.

The inclusion of the additional principal component increases the R^2 's in all the regressions. Thus the additional principal component seems to capture fluctuations in shorter term inflation swap rates. Typically one would expect these inflation swap rates to be more influenced by news on inflation and macro economic fundamentals (since the pay off is directly linked to the CPI), than short term interest-rates which to a larger extent are driven by central bank policies.

Inferring inflation risk premia

Ultimately we would like to infer inflation risk premia. One ad-hoc way of doing it, would be to take a measure of inflation expectations (i.e. a real world expectation) and extract it from the inflation swap rates (i.e. a risk neutral expectation). One such measure could be the European Central Bank Survey of Professional Forecasters (ECB SPF).

In the SPF a number of financial and non-financial professionals submit their point estimate for inflation and probabilities that inflation will fall in prespecified intervals.⁷ More specifically they submit such a forecast of the year-on-year inflation for a horizon of 1,2 and 5 years ahead (a 1 year forecast, a 1 year forward forecast of the 1 year inflation and finally a 4 year forward forecast of the 1 year inflation).⁸

⁷The survey is also conducted for real GDP and unemployment, see Garcia (2003) for further details.

⁸To be specific the survey is done for the present and following calendar years, as well as rolling horizons of 1 and 2 years, i.e. year-on-year forecast of a horizon of 1 and 2 year. The 5 year forecast is a forecast of the calendar year 5 years ahead, hence it will

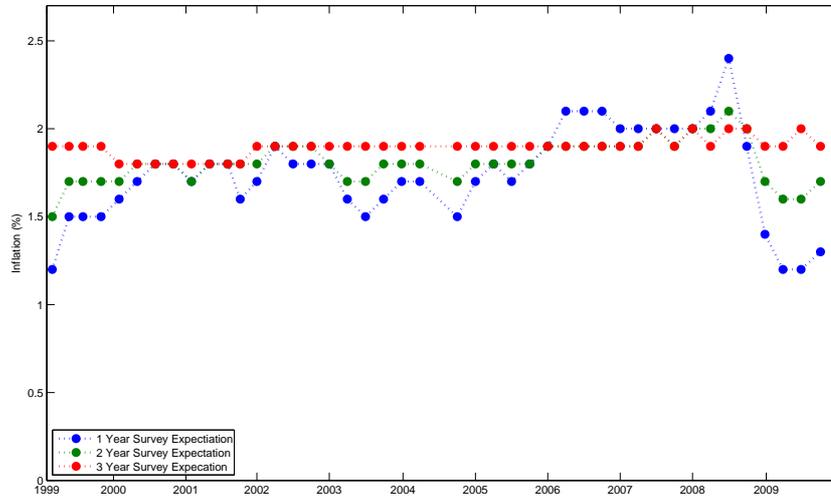


Figure 3.3: The ECB Survey of professional forecasters (SPF). Source: ECB Website.

However, using the ECB SPF has one problem. We only have the survey expectation on a quarterly basis, and hence we are only able to extract the inflation risk premia at these quarterly points. To get an ad-hoc measure of inflation expectations, we use linear interpolation between each quarter. Since the ECB SPF considers expectations as one-year annual inflation rates (and forward inflation rates) and the inflation swap rates are average inflation rates over a longer horizon, we need to convert ECB SPF expectations into an average expectation. We propose using simple compounding of inflation rates (and thus ignoring 'Jensen/convexity' terms):

$$E_t[\Pi(t, t + n)] = \left[\prod_{k=t}^{t+n-1} (1 + E_t[\Pi(k, k + 1)]) \right]^{1/n} - 1$$

When a specific expectation is not available (for instance the 3 year expectation) we use linear interpolation and for expectations with maturities longer than 5 years we keep the expectation fixed at the 5 year level.

Figure 3.4 shows the estimated inflation risk premia. One obvious observation is the big drop in short term inflation risk premia in late-2008. This

be of varying horizon, i.e. between 4.5 and 5.5 years - for simplicity we implement this as a constant maturity 5 year forecast. The very low variation of the 5 year forecast (see Figure 3.3), implies that this approximation is of minor importance.

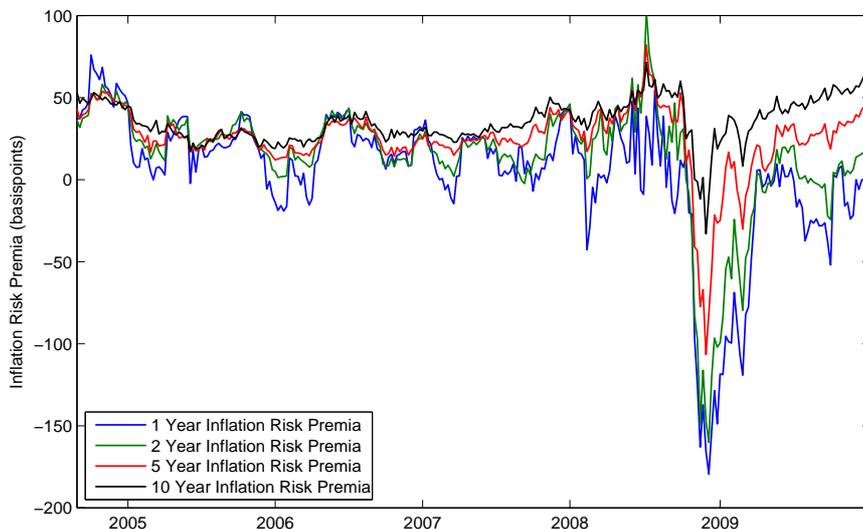


Figure 3.4: Estimate of inflation risk premia obtained by using linear interpolation on survey data. The risk premia is given in basis points.

corresponds to the large drop in inflation swap rates. However cf. Figure 3.3 the drop in survey expectations was smaller in magnitude. Since these results are based on rather ad-hoc means, we prefer estimating a more coherent model to data, which is done in the following sections.

3.3 Inflation risk premia: What theory predicts

Before we describe our model, we consider identification of the risk premia in a theoretical framework. To do so, we consider the no-arbitrage relationship between nominal and real pricing kernels

$$M^R(t) = M^N(t)I(t)$$

This implies that

$$E_t^P \left[\frac{M^R(T)}{M^R(t)} \right] = E_t^P \left[\frac{M^N(T)}{M^N(t)} \right] E_t^P \left[\frac{I(T)}{I(t)} \right] + \text{Cov}^P \left[\frac{M^N(T)}{M^N(t)}, \frac{I(T)}{I(t)} \right]$$

Or equivalently in terms of ZCB prices

$$p_r(t, T) = p_n(t, T) \times E_t^P \left[\frac{I(T)}{I(t)} \right] \times \left(1 + \frac{\text{Cov}^P \left[\frac{M^N(T)}{M^N(t)}, \frac{I(T)}{I(t)} \right]}{E_t^P \left[\frac{M^N(T)}{M^N(t)} \right] E_t^P \left[\frac{I(T)}{I(t)} \right]} \right)$$

In terms of yields this can be written as⁹

$$y_n(t, T) - y_r(t, T) = E_t^P [\Pi(t, T)] + RP(t, T)$$

where

$$E_t^P [\Pi(t, T)] = \frac{1}{T-t} \log E_t^P \left[\frac{I(T)}{I(t)} \right]$$

and

$$RP(t, T) = \frac{1}{T-t} \log \left(1 + \frac{\text{Cov}^P \left[\frac{M^N(T)}{M^N(t)}, \frac{I(T)}{I(t)} \right]}{E_t^P \left[\frac{M^N(T)}{M^N(t)} \right] E_t^P \left[\frac{I(T)}{I(t)} \right]} \right)$$

Hence the BEIR can be decomposed into an inflation expectation and a risk premia. The risk premia is related to the covariance between the nominal stochastic discount factor and inflation. To gain some more intuition on this result we recall that under suitable assumptions in a C-CAPM framework (CRRA utility and log-normality), the inflation risk premia can be described as a function of risk aversion and the covariance between consumption growth and inflation:

$$RP(t, t + \Delta t) \approx -\gamma \text{Cov}^P \left(\frac{C(t + \Delta t)}{C(t)}, \Pi(t, t + \Delta t) \right)$$

All things being equal, a rise in inflation will decrease real consumption, leading to a negative covariance term - thus we would expect inflation risk premia to be positive. Obviously short term fluctuations can turn inflation risk premia negative. Consider the case where the economy is in a recession, here we would expect inflation to be low, or even negative. At the same time due to the recession we could also see a negative growth in real consumption, thus leading to a positive correlation and a negative inflation risk premia. This also provides us with a simple sanity check - we should have somewhat similar dynamics of GDP growth and inflation risk premia.

⁹As shown in section 3.6, the risk premia component, $RP(t, T)$, also includes a convexity term.

3.4 A no-arbitrage model of nominal and inflation swap rates

As mentioned above we prefer a more robust method to derive the inflation risk premia. Thus to estimate the inflation risk premia we consider a continuous time model. A variant similar to the one found here, can be found in D'Amico, Kim, and Wei (2008). More precisely this relies on the affine models proposed in Duffie and Kan (1996).

To begin with, we consider the no-arbitrage relationship between pricing kernels

$$M^R(t) = M^N(t)I(t)$$

This implies, that in a no-arbitrage setting we can model nominal rates and inflation, and then infer real rates. We thus model observable quantities, and by using this approach we follow D'Amico, Kim, and Wei (2008). We therefore assume that the model is driven by four latent factors, and satisfies the following relationships:

$$\begin{aligned} n(t) &= \delta_0 + \delta'_X X(t) \\ \pi(t) &= \gamma_0 + \gamma'_X X(t) + \gamma_Y Y(t) \\ dX(t) &= -K_X X(t)dt + dW_X^Q(t) \\ dY(t) &= -K_Y Y(t)dt + dW_Y^Q(t) \\ \frac{dI(t)}{I(t)} &= \pi(t)dt + \eta dZ^Q(t) \end{aligned}$$

where $n(t)$ is the instantaneous nominal rate, $\pi(t)$ is the instantaneous expected inflation, X is a 3-vector of latent factors driving the yield curve and inflation and Y is a scalar latent factor which only enters into inflation, cf. the regressions above. Finally I is the CPI, which accumulates equal to the expected inflation and a component which is independent of the yield curve, the Wiener process Z . The latter noise term is motivated by the fact that inflation is affected by factors not spanned by the nominal yield curve (and possibly inflation swaps), as argued in Kim (2007). To satisfy the identification constraints in Dai and Singleton (2002), K_X has zeros above the diagonal and δ_0 , δ_X and γ_Y have to be positive.¹⁰

¹⁰We have also estimated a model where the CPI has explicit dependence on the Wiener process W . Our results are not affected by not including this dependence.

Nominal yields and inflation swap rates

As the model draws on the existing literature on affine term structure models, the nominal zero-coupon bond (henceforth nominal ZCB) price can be found using results from Duffie and Kan (1996), i.e. that ZCB prices are an exponentially affine function of the states:

$$p_n(t, T) = \exp(A_n(t, T) + B_n(t, T)'X(t))$$

where $A_n(t, T)$ and $B_n(t, T)$ solve ordinary differential equations (henceforth ODEs).¹¹

Swap rates can then be calculated by using the standard expression for an interest swap rate:

$$S_n(t, T) = \frac{1 - p_n(t, T)}{\sum_{T_k=t+1}^T \Delta_i p_n(t, T_k)}$$

where Δ_i is the tenor between two payment dates and is approximately equal to one (depending on day count convention).

Here we use the methodology of Jarrow and Yildirim (2003) to derive the price of a real ZCB in our model. The result is that the real ZCB also is exponentially affine in the state variables (see Appendix 3.8 for the full derivation):

$$p_r(t, T) = \exp(A_r(t, T) + B_r(t, T)'X(t) + C_r(t, T)Y(t))$$

where $A_r(t, T)$ and $B_r(t, T)$ solve ODEs.

However, real bonds are not traded in the market, making a direct application of the pricing formula impossible. One way around this problem is to estimate the real curve using inflation protected bonds as in Ejsing, Garcia, and Werner (2007). Here we use inflation swap rates. One advantage is that inflation swap rates are quoted in the market, and no estimation methodology has to be used to estimate the real yields. The application of inflation swap rates, follows from Brigo and Mercurio (2006), who show that the zero-coupon Inflation Indexed Swap rate (henceforth ZCIIS rate) can be expressed through nominal and real bonds:

$$\begin{aligned} ZCIIS(t, t + \tau) &= \left(\frac{p_r(t, t + \tau)}{p_n(t, t + \tau)} \right)^{1/\tau} - 1 \\ &= e^{([A_r(t, t+\tau) - A_n(t, t+\tau)] + [B_r(t, t+\tau) - B_n(t, t+\tau)]'X(t) + C_r(t, t+\tau)Y(t))\frac{1}{\tau}} - 1 \end{aligned}$$

¹¹We have omitted the actual ODEs in the main text, they can however be found in Appendix 3.8.

Risk premia: Including surveys

The ultimate purpose of this paper is to estimate inflation risk premia. To identify the premia we need to establish a link between the risk neutral and the real world probability measure. This is established by the (nominal) stochastic discount factor:

$$\frac{dM^N(t)}{M^N(t)} = -n(t)dt - \Lambda_X(t)'dW_X^P(t) - \Lambda_Y(t)dW_Y^P(t)$$

The first two terms of the SDE are the same as in a regular nominal yields models, however the last term relates to inflation risk premia. Obviously the term does not affect the nominal term premia; however, it will affect inflation indexed yields, thus inducing an inflation risk premia.¹²

Furthermore, the no-arbitrage relationship and the model dynamics imply that the real stochastic discount factor evolves according to the SDE

$$\frac{dM^R(t)}{M^N(t)} = -r(t)dt - \Lambda_X(t)'dW_X^P(t) - \Lambda_Y(t)'dW_Y^P(t) + \eta dZ^P(t)$$

where

$$r(t) = n(t) - \pi(t)$$

One thing is evident - we can identify the pure term premia $\Lambda_X(t)$ from nominal yields, however the inflation risk premia, $\Lambda_Y(t)$, requires inflation linked products. Furthermore, to make the model applicable in practice, we need to assume a form for the market price of risk processes, $\Lambda_X(t)$ and $\Lambda_Y(t)$. Here we apply the essentially affine risk premia proposed in Duffee (2002). One obvious advantage is that the state variables stay affine under both the real world measure P and the risk neutral measure Q . Thus the risk premia will be given by

$$\begin{aligned}\Lambda_X(t) &= \lambda_{0X} + \lambda_{XX}X(t) + \lambda_{XY}Y(t) \\ \Lambda_Y(t) &= \lambda_{0Y} + \lambda'_{YX}X(t) + \lambda_{YY}Y(t)\end{aligned}$$

where λ_{0X} is a 3-vector, λ_{XX} is a 3×3 matrix, λ_{XY} is a 3-vector, λ_{0Y} is a scalar, λ_{YX} is a 3-vector and λ_{YY} is a scalar.

With this specification the factors evolve as

$$\begin{aligned}dX(t) &= (\lambda_{0X} + (\lambda_{XX} - K_X)X(t) + \lambda_{XY}Y(t)) dt + dW_X^P(t) \\ dY(t) &= (\lambda_{0Y} + \lambda'_{YX}X(t) + (\lambda_{YY} - K_Y)Y(t)) dt + dW_Y^P(t)\end{aligned}$$

¹²We have tried to include an additional risk factor related to the last Wiener process Z . The results do not change when using this specification, thus we prefer the more parsimonious specification given above.

To identify risk premia, we could just use time series of CPI and inflation swaps, however this might lead to a weak identification of the dynamics. As a consequence, a number of papers (see Ang, Bekaert, and Wei (2007), Ang, Bekaert, and Wei (2008), D’Amico, Kim, and Wei (2008) and Garcia and Werner (2010)) have identified that using surveys improves inflation forecasts and model performance.

One advantage of maintaining the affine structure of the CPI under the real world measure (as specified above), is that the expectation of the CPI is exponentially affine (here $\tau = 0, 1, 4$):

$$E_t^P \left[\frac{I(t + \tau + 1)}{I(t + \tau)} \right] = \exp (A_s(\bullet) + B_s(\bullet)'X(t) + C_s(\bullet)'Y(t))$$

where $A_s(\bullet) = A_s(t, t + \tau, t + \tau + 1)$, $B_s(\bullet) = B_s(t, t + \tau, t + \tau + 1)$ and $C_s(\bullet) = C_s(t, t + \tau, t + \tau + 1)$ solve ODEs.

Using this result, the time t survey expectation of the year-on-year inflation with maturity τ can then be expressed as

$$\begin{aligned} S(t, t + \tau) &= E_t^P \left[\frac{I(t + \tau + 1)}{I(t + \tau)} \right] - 1 \\ &= \exp (A_s(\bullet) + B_s(\bullet)'X(t) + C_s(\bullet)'Y(t)) - 1 \end{aligned}$$

Finally surveys may be biased estimators of inflation. We resolve this issue by incorporating a constant in the measurement of each survey maturity:

$$S(t, t + \tau) = \alpha_\tau + (\exp (A_s(\bullet) + B_s(\bullet)'X(t) + C_s(\bullet)'Y(t)) - 1)$$

3.5 Model Estimation

In this paper we adopt a Bayesian approach. Admittedly Bayesian methods are more computationally cumbersome than for instance Quasi Maximum Likelihood methods, however an estimation based on Bayesian methods allows for direct draws of the posterior distribution, which will indeed be useful in terms of interpreting the inflation risk premia. In this section we will describe the notation used in the estimation, the specification of conditional distributions and the implemented hybrid MCMC algorithm used. A survey article on MCMC is Johannes and Polson (2003), where textbook treatments can be found in Gamerman and Lopes (2006) and Robert and Casella (2004). Our approach is also inspired by Feldhütter (2008).

Notation

In this paper we observe nominal swap rates, inflation swap rates, surveys and the CPI. Let us denote the observed nominal swap rates at time t by $R_t^n = (R_{1,t}^n, \dots, R_{N,t}^n)'$ and let the observed inflation swap rates at time t be given by $R_t^k = (R_{1,t}^k, \dots, R_{K,t}^k)'$. Similarly we denote the observed survey forecast at time t by $Y_t^s = (R_{1,t}^s, \dots, R_{S,t}^s)'$. Finally the log-CPI at time t is denoted by $\log I_t$, and the change between two publications of the CPI at time $t-k$ and t is given by $\Delta \log I_t$. Since not all observations occur at each time point we let \mathcal{T}_N be the set of times where nominal yields are observed, we let \mathcal{T}_K be the set of times where inflation swap rates are observed, we let \mathcal{T}_S be the set of times where surveys are observed and finally we let \mathcal{T}_I be the set of times where the CPI is observed. The entire collection of observations is denoted by R .

With regard to parameters, we denote the risk neutral parameters of the nominal interest-rate model and the risk neutral factor dynamics $(\delta_0, \delta_X, K_X, K_Y)$ by Θ^Q and the risk premia parameters (all λ 's) are denoted by Θ^P . The Risk neutral inflation process and the inflation variance parameters $(\gamma, \gamma_X, \gamma_X, \eta)$ are denoted by Θ^π . Finally measurement bias and errors are given by $\sigma_n, \sigma_k, \sigma_s$ and α . The entire collection of parameters is given as $\Theta = (\Theta^Q, \Theta^P, \Theta^\pi, \sigma_n, \sigma_k, \sigma_s, \alpha)'$.

Estimation using MCMC

At time $t \in \mathcal{T}_N$ we observe N nominal swap rates, which are stacked in the N -Vector R_t^n . We assume that the yields are observed with measurement errors:

$$R_t^n = S_n(t, T) + \varepsilon_{n,t}$$

where $S_n(t, T)$ is the nominal swap rate. Furthermore, we assume that the measurement errors are normally distributed with common variance

$$\varepsilon_{n,t} \sim \mathcal{N}(0, \sigma_n^2 I_N)$$

Similarly at time $t \in \mathcal{T}_K$ we observe K inflation swap rates, which also are stacked in a K -Vector. Again these observations are also observed with errors:

$$R_t^k = e^{([A_r(t,t+\tau) - A_n(t,t+\tau)] + [B_r(t,t+\tau) - B_n(t,t+\tau)]' X_t + C_r(t,t+\tau) Y_t) \frac{1}{\tau}} - 1 + \varepsilon_{k,t}$$

As above we assume that the measurement errors are normally distributed with common variance

$$\varepsilon_{k,t} \sim \mathcal{N}(0, \sigma_k^2 I_K)$$

At times $t \in \mathcal{T}_S$ we observe S survey expectations, which are stacked in the S -vector R_t^s . The surveys are also observed with a measurement error which have a common variance:

$$R_t^s = \alpha_\tau + \left(e^{A_s(t,t+\tau,t+\tau+1)+B_s(t,t+\tau,t+\tau+1)'X_t+C_s(t,t+\tau,t+\tau+1)'Y_t} - 1 \right) + \varepsilon_{s,t}$$

$$\varepsilon_{s,t} \sim \mathcal{N}(0, \sigma_s^2 I_S)$$

Finally the CPI is observed as log-CPI and is assumed to be observed without error.

When estimating the model we are interested in sampling from the target distribution of parameters and state variables, $p(\Theta, X, Y|R)$. To sample from this distribution the Hammersley-Clifford theorem (Hammersley and Clifford (1974) and Besag (1974)) implies that this can be done by sampling from the complete conditionals

$$p(\Theta^Q | \Theta_{\setminus Q}, X, Y, R)$$

$$\vdots$$

$$p(X, Y | \Theta, R)$$

Thus MCMC handles the sampling from the complicated target distribution $p(\Theta, X|Y)$, by sampling from the simpler conditional distributions. More specifically this is handled by sampling in cycles from the conditional distributions. If one can sample directly from the conditional distribution, the resulting algorithm is called a Gibbs sampler (see Geman and Geman (1984)). If it is not possible to sample from this distribution one can sample using the Metropolis-Hastings algorithm (see Metropolis, Rosenbluth, Rosenbluth, Teller, and E. (1953)). In this paper we use a combination of the two (a so-called hybrid MCMC algorithm) since not all the conditional distributions are known. More precisely we have the following MCMC algorithm:

$$p(X, Y | \Theta, R) \sim \text{Metropolis-Hastings}$$

$$p(\Theta^Q | \Theta_{\setminus Q}, X, Y, R) \sim \text{Metropolis-Hastings}$$

$$p(\Theta^P | \Theta_{\setminus P}, X, Y, R) \sim \text{Metropolis-Hastings}$$

$$p(\Theta^\pi | \Theta_{\setminus \pi}, X, Y, R) \sim \text{Metropolis-Hastings}$$

$$p(\sigma_n, \sigma_k, \sigma_s | \Theta_{\setminus \sigma}, X, Y) \sim \text{Inverse Gamma}$$

$$p(\alpha | \Theta_{\setminus \alpha}, X, Y) \sim \text{Normal}$$

It should be noted that nominal swap rates and inflation swaps depend on Q -parameters and surveys depend on P -parameters. This makes the estimation slightly harder, as both P and Q -parameters depend non-linearly on

the states, through the pricing functions A , B and C . Thus the estimation of P -parameters has to be done by Metropolis-Hastings sampling, rather than Gibbs sampling, which would normally be the case when estimating the P -parameters in a model of the nominal term structure.

A more precise description of the algorithm and the conditional distributions is found in Appendix 3.9. Furthermore, we block sample the entire history of each state, and for each draw of parameters, we perform 10 state draws. This improves the convergence of the Markov Chain compared to univariate single state sampling.

The Markov chain is run for 10 million simulations¹³, where the standard errors of the Random Walk Metropolis-Hasting algorithms are calibrated to yield acceptance probabilities between 10 and 40 pct. We successively remove insignificant parameters, such that the reported model is the minimal model required to fit the data. Finally, we save each 1000th draw and use an additional 1 million simulations of the chain, leaving 1000 draws for inference.

3.6 Empirical results

Parameter estimates and model fit

In this section we consider the parameter estimates and model fit.

Table 3.3 shows the model fit, as measured by root mean squared errors (RMSEs). We see that the fit to data is good - nominal yields have RMSEs of around 2 basis points, and surveys and inflation swaps are around 5-9 basis points, with the 1 year inflation swap rate, however, having a RMSE of 12 basis points. Given our data we find the model fit to be satisfactory (e.g. the ECB SPF is reported with a precision of 0.1 percent).

Table 3.4 presents the parameter estimates from the MCMC estimation. Parameter estimates are based on the mean of the MCMC samples, where confidence bands present the 2.5 % and 97.5 % quantiles of the MCMC samples. One interesting finding is that the vector λ_{XY} is significant, which implies that the factor specific to inflation swaps can help in explaining the dynamics of nominal yields.¹⁴

¹³The choice of 10 million simulations is somewhat arbitrary. It is sufficiently high to ensure convergence of the Markov chain without having to run more simulations.

¹⁴This is also found in Christensen, Lopez, and Rudebusch (2008). It would be interesting to explore if this additional factor can improve forecasts of nominal yields, this however is outside the scope of this paper.

	Swap rates	Inflation swaps	Surveys
1 year	2.3726 (2.2381 , 2.5162)	12.3711 (11.764 , 12.9555)	8.7507 (7.8122 , 9.8146)
2 years	2.7567 (2.6583 , 2.8582)	6.4908 (6.0717 , 6.9253)	6.8518 (6.2207 , 7.6378)
3 years	1.8743 (1.774 , 1.9846)	6.5815 (6.2003 , 6.9597)	-
5 years	2.2179 (2.1306 , 2.3113)	7.5487 (7.2932 , 7.8174)	4.9447 (4.6308 , 5.716)
7 years	2.0004 (1.9112 , 2.0904)	8.375 (8.18 , 8.5885)	-
10 years	1.7131 (1.6191 , 1.8038)	9.0494 (8.8674 , 9.234)	-
15 years	2.5488 (2.4427 , 2.657)	8.4802 (8.3158 , 8.6563)	-

Table 3.3: **Root Mean Squared Errors.** The RMSEs are measured in basis points and are based on the mean of the MCMC samples. 95 pct. confidence intervals based on MCMC samples are reported in brackets.

Rather than directly interpreting on all the parameters, we consider the estimated factor loadings and filtered factors. Factor loadings based on the estimated parameters are given in Figure 3.5 and the filtered states are given in Figure 3.6.

The factor loadings for the nominal yields imply that factor 2 and 3 can be interpreted as steep and flat slope factors, respectively. The first factor has the interpretation of a curvature factor. Overall this seems consistent with the principal component analysis performed above.

Our inflation specific factor affects the slope of the real yield curve. The first three factors preserve the same interpretation for real yields, although with a smaller absolute effect for the slope factor. This also implies that the curvature of the yield curve has little effect on the BEIRs.

We also plot factor loadings for inflation expectations and inflation swaps, cf. Figure 3.5. The inflation expectation factor loadings are based on the expected growth rate of the CPI index.¹⁵, and the factor loadings for inflation swaps are based on a first order Taylor expansion of the inflation

¹⁵The Inflation growth rate is given by

$$\frac{1}{\tau} \log E_t^P \left[\frac{I(t+\tau)}{I(t)} \right] = \frac{A_s(\bullet)}{\tau} + \frac{B_s(\bullet)'}{\tau} X(t) + \frac{C_s(\bullet)}{\tau} Y(t)$$

	$k = 1$	$k = 2$	$k = 3$
δ_0	0.0568 (0.0564 , 0.0571)	-	-
δ_X	0.0006 (0.0000 , 0.0014)	0.0369 (0.0355 , 0.0381)	0.0283 (0.0277 , 0.0288)
γ_0	0.0234 (0.0233 , 0.0235)	-	-
γ_X	0.0000 -	-0.0034 (-0.0037 , -0.0032)	0.0000 -
γ_Y	0.0125 (0.0113 , 0.0141)	-	-
$K_X(1, k)$	1.1391 (1.0833 , 1.1974)	-	-
$K_X(2, k)$	0.0903 (0.0716 , 0.1193)	0.0162 (0.0135 , 0.0189)	-
$K_X(3, k)$	0.4618 (0.4218 , 0.5231)	0.5463 (0.5143 , 0.568)	0.3279 (0.3217 , 0.3349)
K_Y	0.6609 (0.6297 , 0.6922)	-	-
η	0.0014 (0.0013 , 0.0016)	-	-
$\lambda_{0X}(k)$	2.2328 (1.5514 , 3.1729)	0.0000 -	0.0000 -
λ_{0Y}	0.0000 -	-	-
$\lambda_{XX}(1, k)$	-1.9732 (-2.9321 , -1.2027)	0.0000 -	0.5434 (-2.9321 , -1.2027)
$\lambda_{XX}(2, k)$	0.0000 -	-0.9169 (-1.4661 , -0.4603)	-0.462 (-0.7109 , -0.2387)
$\lambda_{XX}(3, k)$	0.0000 -	0.0000 -	0.0000 -
$\lambda_{XY}(k)$	1.2213 (0.2395 , 2.2709)	0.1296 (0.1296 , 0.1296)	0.727 (0.1337 , 1.3399)
$\lambda_{YX}(k)$	-1.0454 (-1.5168 , -0.649)	3.9591 (3.1637 , 4.9191)	1.8344 (1.4377 , 2.2741)
λ_{YY}	-3.4897 (-4.3312 , -2.8321)	-	-
$\sigma(k)$	0.0002 (0.0002 , 0.0002)	0.0009 (0.0008 , 0.0009)	0.0007 (0.0006 , 0.0008)
$\alpha(k)$	0.0000 (-0.0002 , 0.0002)	0.0000 (-0.0002 , 0.0002)	0.0000 (-0.0002 , 0.0002)

Table 3.4: **Parameter Estimates in no-arbitrage model.** Parameter estimates are based on the means of the MCMC samples. 95 pct. confidence intervals based on MCMC samples are reported in brackets. $\sigma(1)$ is the measurement error of nominal yields, $\sigma(2)$ is the measurement error of surveys and $\sigma(3)$ is the measurement error of inflation swaps. Parameters with no confidence intervals are fixed at the reported value.

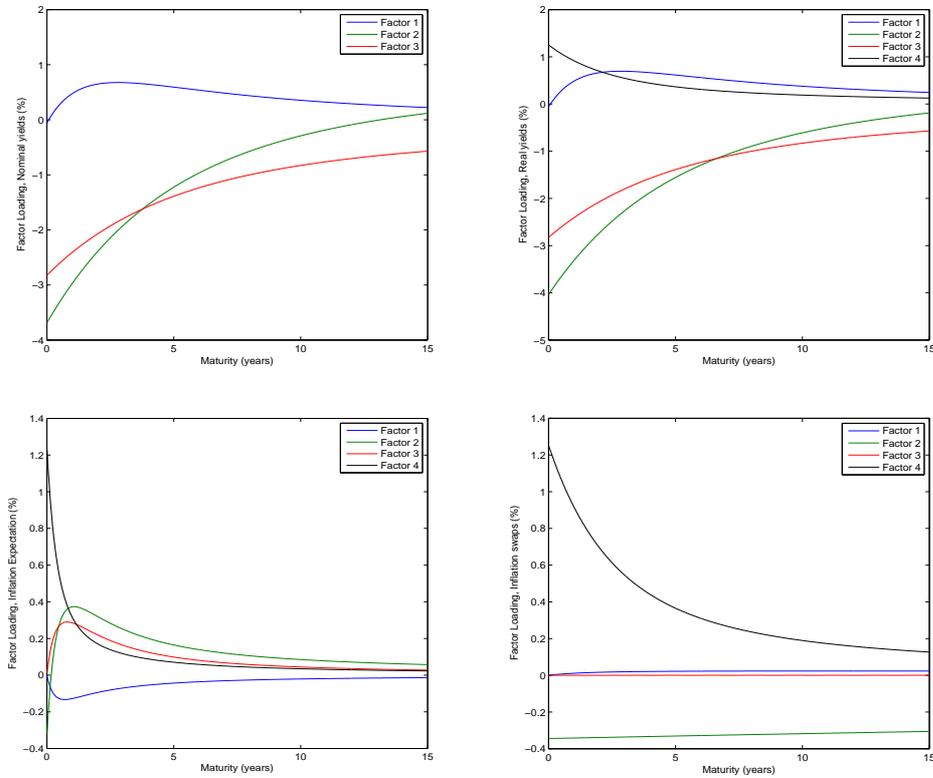


Figure 3.5: **Upper left:** Factor Loadings for nominal yields. **Upper right:** Factor Loadings for real yields. **Lower left:** Factor Loadings for inflation expectation. **Lower right:** Factor Loadings for inflation swaps.

swap quote¹⁶

One interesting finding when comparing the factor loadings for inflation expectations and swaps, is that the factor loading related to the inflation factor decay much faster for inflation expectations, than for inflation swaps. This implies that shocks to the inflation factor have a greater effect on long term inflation swaps than on long term inflation expectations. Thus this factor is instrumental in modeling inflation risk premia. Another interesting finding is that the slope factor has a constant effect on all inflation swaps,

¹⁶The Taylor expansion gives us

$$ZCIS(t, t + \tau) \approx \frac{A_r(\bullet) - A_n(\bullet)}{\tau} + \frac{B_r(\bullet) - B_n(\bullet)'}{\tau} X(t) + \frac{C_r(\bullet) - C_n(\bullet)}{\tau} Y(t)$$

which is equivalent to a continuous time Break Even Inflation Rate.

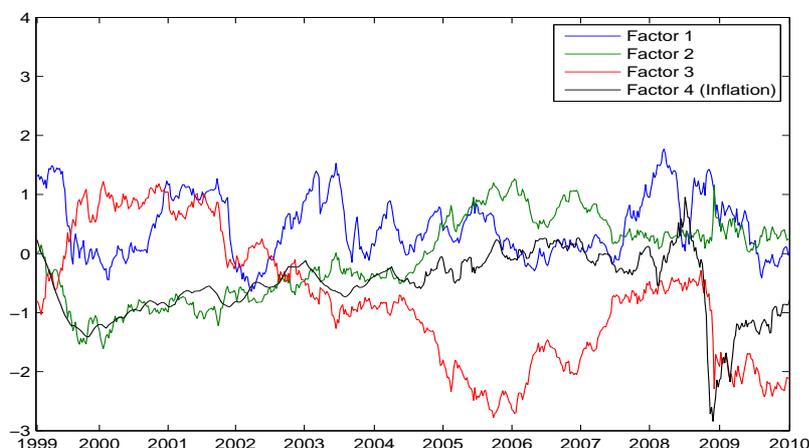


Figure 3.6: **Time series of filtered factors.** The filtered estimate is based on the mean of the MCMC samples.

where the long term effects of the slope factor on inflation expectations are close to zero.

When considering the filtered factors (Figure 3.6), we see that when comparing the factors to Figure 3.2, the flat slope factor (factor 3) is the main driver of the level of interest-rates, where the steep slope factor (factor 2) drives slope of the yield curve. When considering the inflation specific factor it only shows minor variation in the period until 2008, but shows a spike in the summer of 2008 and again a drop around end-2008. This pattern is similar to Figure 3.1, and describes the rise in commodity prices during the summer of 2008 and worries regarding the macro economy post the Lehman Brothers collapse.

Decomposing nominal yields and inflation compensation

In this section we consider the estimated inflation risk premia and how nominal yields and inflation compensation can be decomposed into real yields, inflation expectations, inflation risk premia and convexity terms. Explaining the inflation risk premia as a function of macro economic and financial factors is postponed until section 3.6.

First we show in Appendix 3.8, that the break even inflation rate ($BEIR =$

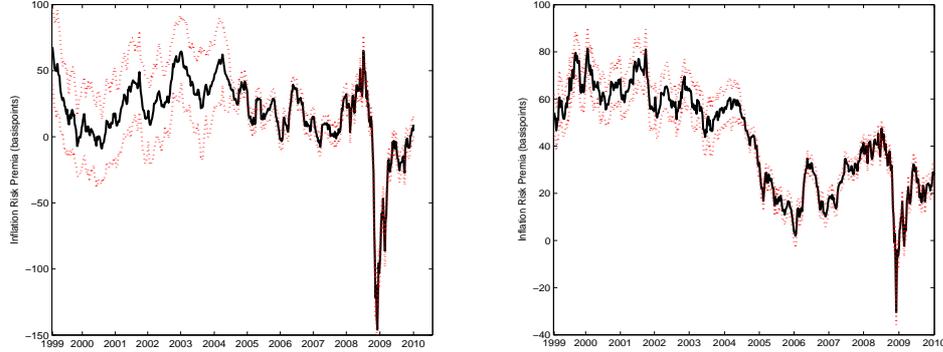


Figure 3.7: **Inflation risk premia.** The solid line represents the 1 year inflation risk premia and the 10 year inflation risk premia. 95 pct. confidence intervals based on MCMC samples are reported as dashed lines.

$y_n(t, T) - y_r(t, T)$ can be decomposed as

$$\begin{aligned}
 BEIR = & \underbrace{\frac{1}{T-t} \int_t^T (\gamma_0 + \gamma_X' E_t^P [X(s)]) ds}_{\text{Inflation expectations}} \\
 & + \underbrace{\frac{1}{2} \frac{1}{T-t} \int_t^T (B_r(s, T)' B_r(s, T) - B_n(s, T)' B_n(s, T)) ds}_{\text{Convexity correction}} \\
 & + \underbrace{\frac{1}{T-t} \int_t^T [(B_n(s, T) - B_r(s, T))' (\lambda_0 + \lambda_X E_t^P [X(s)])] ds}_{\text{Inflation risk premia}}
 \end{aligned}$$

The expression for the inflation expectation is simply the average integrated (spot-)inflation over the considered period, and the inflation risk premia can be described as

$$IRP(t, T) = \frac{1}{T-t} \int_t^T \left[\underbrace{(B_n(s, T) - B_r(s, T))'}_{\text{Amount of risk}} \underbrace{(\lambda_0 + \lambda_X E_t^P [X(s)])}_{\text{Market price of risk}} \right] ds$$

which implies that the risk premia is equal to the market price of risk taken in a specific maturity segment times the amount of risk taken in the specific maturity segment.

Figure 3.7 shows estimated inflation risk premia along with 95 percent confidence bands based on MCMC samples. Considering the 1 year inflation

risk premia, we see some degree of variation, with risk premia fluctuating between -146 and 68 basis points. The smallest risk premia is in end-2008, indicating that the market was pricing very severe scenarios.¹⁷ The highest inflation risk premia is measured when commodity prices spiked, i.e. during the summer of 2008. During the remainder of the period the risk premia shows fluctuations between -5 and 65 basis points, with the 95 percent confidence band being between 15 and 50 basis points wide.¹⁸

The 10 year inflation risk premia, also shows a higher level of inflation risk premia until 2005. The risk premia in this period is between 5 and 81 basis points. After 2005 the risk premia show more fluctuation but is still between -30 and 45 basis points. The higher risk premia until 2005 reflects that the nominal term structure was steeper in this period, implying that part of the nominal term premia was driven by inflation risk premia.

With regard to similarity to other studies our estimated risk premia is very similar to, if slightly higher than, the ones found in Garcia and Werner (2010). With respect to the 10 year inflation risk premium, our estimates are similar to Tristani and Hördahl (2010). The slightly higher inflation risk premia that we estimate can probably be related to the inflation linked data used. We use inflation swaps where Garcia and Werner (2010) and Tristani and Hördahl (2010) use inflation linked bonds. Inflation swaps provide an easier hedge than inflation linked bonds given the simpler nature of the swaps. This implies a convenience premia that could explain the slight differences between our estimates and the ones found in Garcia and Werner (2010) and Tristani and Hördahl (2010).

Figure 3.8 shows the decomposition of the nominal yield into real yield, inflation expectation, inflation risk premia and convexity. It is evident that the main components in the variation of nominal yields are variations in real yields and inflation risk premia. Real yields account for the majority of the variation. When considering inflation expectations we see that they are fairly constant.

Table 3.5 reports average levels for the decomposition of nominal yields, along with a variance decomposition. Table 3.6 shows a decomposition of the inflation compensation.

Table 3.5 shows that on average there is an upward sloping term structure in both nominal and real yields, as well as inflation expectations and risk

¹⁷Part of this drop could also be related to liquidity reasons, however as mentioned in the introduction, inflation swaps were less affected than linkers in this period.

¹⁸When considering the period from 1999 to mid-2004, where inflation swaps are not available the typical width of the confidence bands are 50 basis points, whereas from mid-2004 and ahead the width is around 15 basis points.

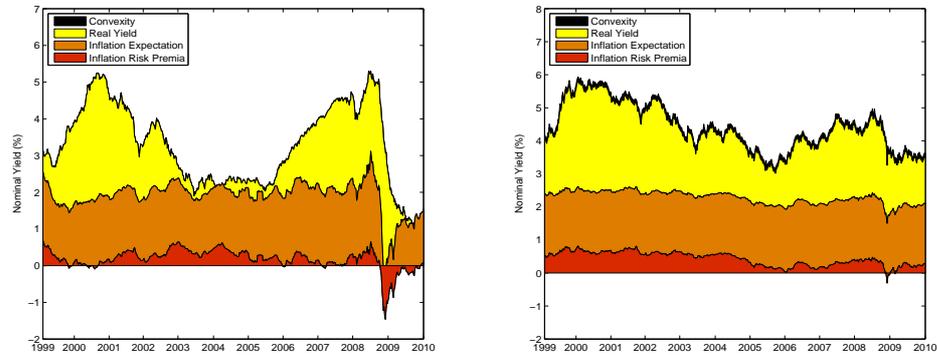


Figure 3.8: **Decomposition of nominal yields.** The figure decomposes the 1 year nominal yield (left) and 10 year nominal yield (right) into real yield, inflation expectation, inflation risk premia and convexity.

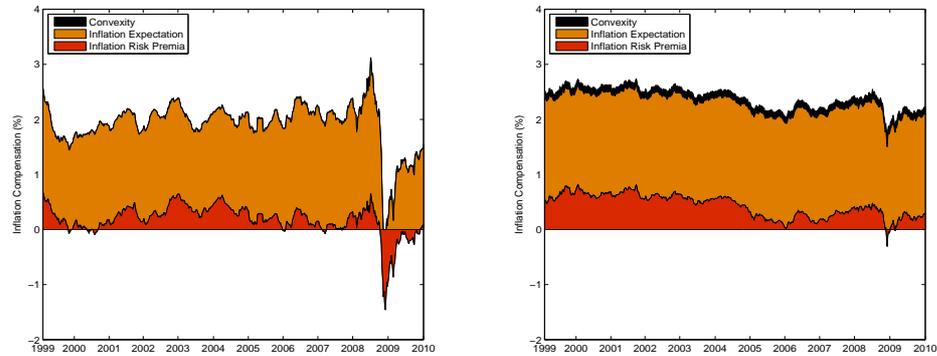


Figure 3.9: **Decomposition of inflation compensation (break even inflation rate).** The figure decomposes the 1 year inflation compensation (left) and 10 year inflation compensation (right) into inflation expectation, inflation risk premia and convexity.

Maturity		Nominal Yield	Real Yield	Inflation Expectations	Inflation Risk Premia	Convexity
1 Year	Mean	3.2773 (3.2756 , 3.279)	1.3735 (1.3173 , 1.4296)	1.7232 (1.6993 , 1.7486)	0.1753 (0.1218 , 0.2267)	0.0054 (0.005 , 0.0059)
	Std. Dev.	1.0938 (1.0921 , 1.0954)	1.0308 (0.9924 , 1.0743)	0.2455 (0.2265 , 0.2644)	0.303 (0.2722 , 0.3374)	- -
	Var. Decomp.	100 -	80.921 (0.7206 , 0.8994)	19.8106 (0.1618 , 0.2347)	-0.7334 (-0.0714 , 0.0559)	- -
2 Year	Mean	3.4585 (3.4571 , 3.4598)	1.4247 (1.3817 , 1.4685)	1.7606 (1.7424 , 1.7801)	0.2558 (0.2115 , 0.2986)	0.0174 (0.016 , 0.0186)
	Std. Dev.	0.9786 (0.9775 , 0.9796)	0.904 (0.8726 , 0.9406)	0.1598 (0.1435 , 0.1754)	0.2803 (0.2537 , 0.3105)	- -
	Var. Decomp.	100 -	82.2518 (0.7588 , 0.8839)	12.4116 (0.093 , 0.1559)	5.3328 (-0.002 , 0.1111)	- -
5 Year	Mean	3.9551 (3.9539 , 3.9562)	1.7022 (1.6776 , 1.7272)	1.818 (1.7999 , 1.8374)	0.3724 (0.3424 , 0.4004)	0.0625 (0.0575 , 0.0666)
	Std. Dev.	0.8043 (0.8033 , 0.8053)	0.6942 (0.6753 , 0.7167)	0.0747 (0.0641 , 0.086)	0.2266 (0.2055 , 0.2504)	- -
	Var. Decomp.	100 -	80.6528 (0.7728 , 0.8397)	3.1634 (0.0127 , 0.0531)	16.1842 (0.1276 , 0.1978)	- -
10 Year	Mean	4.4732 (4.4719 , 4.4746)	2.0771 (2.0598 , 2.0925)	1.8521 (1.8307 , 1.875)	0.4296 (0.4032 , 0.4551)	0.1144 (0.1049 , 0.1231)
	Std. Dev.	0.7157 (0.7144 , 0.7169)	0.5623 (0.5484 , 0.5769)	0.0396 (0.0329 , 0.0473)	0.2123 (0.194 , 0.2316)	- -
	Var. Decomp.	100 -	74.3415 (0.7195 , 0.766)	-0.3539 (-0.0232 , 0.0144)	26.0126 (0.2344 , 0.2886)	- -

Table 3.5: **Decomposition of nominal yields.** The table reports mean and standard deviation for each of the variables. Reported numbers are measured in percentages. 95 pct. confidence intervals based on MCMC samples are reported in brackets.

Maturity		Inflation Compensation	Inflation Expectations	Inflation Risk Premia	Convexity
1 Year	Mean	1.9039 (1.848 , 1.9604)	1.7232 (1.6993 , 1.7486)	0.1753 (0.1218 , 0.2267)	0.0054 (0.005 , 0.0059)
	Std. Dev.	0.4753 (0.4561 , 0.4983)	0.2455 (0.2265 , 0.2644)	0.303 (0.2722 , 0.3374)	- -
	Var. Decomp.	100 -	34.1906 (0.2905 , 0.3961)	65.8289 (0.6041 , 0.7098)	- -
2 Year	Mean	2.0338 (1.9905 , 2.0774)	1.7606 (1.7424 , 1.7801)	0.2558 (0.2115 , 0.2986)	0.0174 (0.016 , 0.0186)
	Std. Dev.	0.366 (0.3496 , 0.386)	0.1598 (0.1435 , 0.1754)	0.2803 (0.2537 , 0.3105)	- -
	Var. Decomp.	100 -	23.0752 (0.1859 , 0.2798)	76.94 (0.7203 , 0.8142)	- -
5 Year	Mean	2.2529 (2.2282 , 2.2774)	1.818 (1.7999 , 1.8374)	0.3724 (0.3424 , 0.4004)	0.0625 (0.0575 , 0.0666)
	Std. Dev.	0.2353 (0.2202 , 0.252)	0.0747 (0.0641 , 0.086)	0.2266 (0.2055 , 0.2504)	- -
	Var. Decomp.	100 -	12.8364 (0.0897 , 0.1696)	87.1743 (0.8305 , 0.9104)	- -
10 Year	Mean	2.3961 (2.3805 , 2.413)	1.8521 (1.8307 , 1.875)	0.4296 (0.4032 , 0.4551)	0.1144 (0.1049 , 0.1231)
	Std. Dev.	0.1998 (0.186 , 0.2146)	0.0396 (0.0329 , 0.0473)	0.2123 (0.194 , 0.2316)	- -
	Var. Decomp.	100 -	4.4215 (-0.0006 , 0.0879)	95.5874 (0.9122 , 1.0006)	- -

Table 3.6: **Decomposition of inflation compensation.** The table reports mean and standard deviation for each of the variables. Reported numbers are measured in percentages. 95 pct. confidence intervals based on MCMC samples are reported in brackets.

premia. The inflation expectation shows the least slope with a one year inflation expectation of 1.72 percent and a 10 year expectation of 1.85 percent. We find that average inflation risk premia are moderate - between 17 and 43 basis points, however the mean might not be representative of the inflation risk premia in a normal scenario, due to a large drop in risk premia in end-2008, cf. Figure 3.7 and 3.8.

To assess the drivers of the variation of nominal yields we consider the unconditional variance decompositions used in for instance Ang, Bekaert, and Wei (2008) and Garcia and Werner (2010). The variance decomposition is given by¹⁹

$$1 = \frac{\text{Cov}(\Delta y_n, \Delta y_r)}{\text{Var}(\Delta y_n)} + \frac{\text{Cov}(\Delta y_n, \Delta IE)}{\text{Var}(\Delta y_n)} + \frac{\text{Cov}(\Delta y_n, \Delta IRP)}{\text{Var}(\Delta y_n)}$$

where Δy_n is the change in the nominal yield, Δy_r is the change in the real yield, ΔIE is the change in the inflation expectation and ΔIRP is the change in the inflation risk premia.

The variance decomposition of nominal yields shows that short term variation is mainly driven by variation in real yields (81 percent) and to a lesser degree inflation expectations (20 percent). Changes in inflation risk premia in the short run appear to be more or less uncorrelated to changes in nominal yields. That changes in inflation risk premia are uncorrelated with changes in nominal interest-rates, could be explained by the fact that nominal swap markets and inflation swap markets are not fully integrated, i.e. that inflation traders react to all news on inflation, where swap traders do not. For nominal yields with a longer time to maturity (e.g. 10 years), inflation expectations are very anchored and do not add to the variation of nominal yields. Instead the variation is driven by real yields (74 percent) and inflation risk premia (26 percent).

In terms of inflation risk premia and variance decompositions, we are not only interested in nominal yields. Another interesting variable is the inflation compensation, i.e. the sum of the inflation expectation and risk premia, which is equivalent to a BEIR or an inflation swap rate. We see that the main driver of variation of inflation compensation is the inflation risk premia. For short term inflation compensation (1 year) the variation in risk premia corresponds to 66 percent, where for the long term inflation compensation (10 years) it corresponds to almost all variation in the inflation compensation (95 percent), again confirming an anchoring of inflation expectations in the Euro area.

¹⁹The decomposition for the inflation compensation is performed in an analogous fashion.

Comparing to Garcia and Werner (2010), who also report variance decompositions for Euro area inflation compensation, we see that the degree of variation generated by inflation risk premia is higher in our analysis. We consider two factors to be the main drivers. First, we include a longer sample period. The sample period in Garcia and Werner (2010) includes data until end-2006, and the period from 2007 is especially volatile compared to the period before 2007. Secondly, Garcia and Werner (2010) estimate their model using real yields extracted from inflation linked bonds, whereas we use inflation swap quotes directly, which could induce more noise in our measure of inflation compensation.

Surveys and model output

Although a number of papers (see Ang, Bekaert, and Wei (2007), Ang, Bekaert, and Wei (2008), D’Amico, Kim, and Wei (2008) and Garcia and Werner (2010)) have identified that using surveys improves inflation forecasts and model performance, to our knowledge no papers have assessed the effect on the identification of inflation risk premia.

To assess this issue, we re-estimate the model without using surveys. We do not report parameter estimates, but as expected more parameters are insignificant and overall the filtered states are similar to ones given in Figure 3.6.²⁰

As our main purpose is to estimate inflation risk premia, the exclusion of surveys could have a profound effect on the estimated risk premia. To see this consider the inflation risk premia

$$IRP(t, T) = \frac{1}{T-t} \int_t^T \left[\underbrace{(B_n(s, T) - B_r(s, T))'}_{\text{Amount of risk}} \underbrace{(\lambda_0 + \lambda_X E_t^P [X(s)])}_{\text{Market price of risk}} \right] ds$$

The first term (the amount of risk taken) is identified directly from nominal swap rates and inflation swaps and is thus identified very precisely. The second term (the market price of risk) is determined from the dynamics of the factors, i.e. the P -dynamics. Since surveys are based on a P -expectation, including surveys would improve the identification of the market price of risk.

Figure 3.10 shows the estimated risk premia arising from the two estimations. The left figure, shows the 1 year inflation risk premia. On average the inflation risk premia is of the same size, albeit the confidence bands are

²⁰Parameter estimates and derived variables are available upon request.

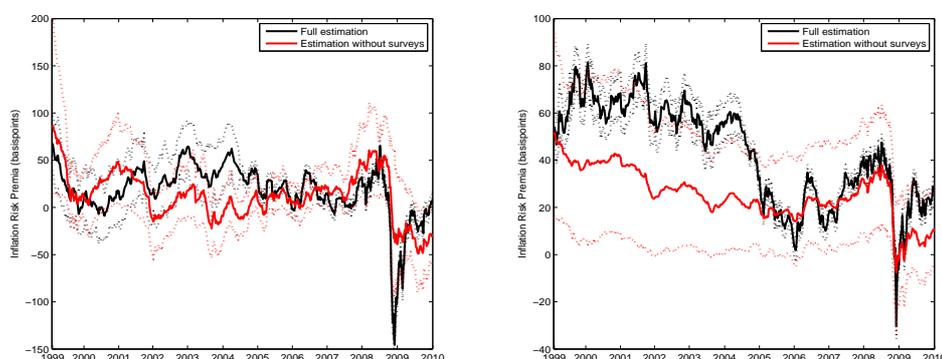


Figure 3.10: **Comparison of inflation risk premia from different estimations.** The figures show the 1 year inflation risk premia (left) and the 10 year inflation risk premia (right) from an estimation with and without surveys. 95 pct. confidence intervals based on MCMC samples are reported as dashed lines.

wider in the estimation with out surveys. The difference in the width of the confidence bands is around 25 basis points.

The right figure shows the estimated 10 year inflation risk premia. The inflation risk premia from the estimation without surveys is approximately at the same level from 2005 and onwards (i.e. where inflation swaps are available). Before 2005 the inflation risk premia from the estimation without surveys is approximately 20-30 basis points lower than in the estimation including surveys. Most interesting is the difference in the width of the confidence bands, which is around 35 basis points.

One conclusion from Figure 3.10 is that studies that report risk premia should ideally report confidence bands as well. In the case of inflation risk premia for the Euro area, the inclusion of surveys massively improve the identification of long term risk premia.²¹

²¹For instance Tristani and Hördahl (2007) estimate their model without surveys and report time series for inflation risk premia which do not look like the one found in our study and the study by Garcia and Werner (2010). Furthermore, they conclude that the inflation risk premia are insignificant. In a more recent estimation of their model, Tristani and Hördahl (2010) include surveys and find inflation risk premia similar to our paper.

Explaining inflation risk premia

Although the size and dynamics of inflation risk premia is an interesting issue on its own, relating the inflation risk premia to fundamentals or agents beliefs are important for understanding the behavior of inflation risk premia.

Here we focus on the beliefs of the agents in the economy, as measured by the ECB SPF. Beside asking the participants about the expected outcome of inflation, GDP and unemployment, the ECB SPF also asks the participants to put probabilities on given outcomes of inflation, GDP and unemployment, thus giving more detailed information on the belief of the participants.

Using these probabilities for inflation and GDP, we construct the mean, standard deviation and skewness of the distributions as perceived by the participants of the survey. This approach is similar to Trolle and Schwartz (2010), who relate moments of swap distributions to survey moments. By using the survey moments, we restrict ourselves to quarterly observations, giving 44 observations for the 1 year inflation risk premia, and 38 for the 5 year inflation risk premia.²²

To account for other factors, which could potentially be important drivers of inflation risk premia, we consider controls for overall market sentiment and volatility as well as overall liquidity in the market.

As a proxy for market sentiment and volatility we use the VSTOXX volatility index, which gives a model free volatility estimate, based on options written on the Eurostoxx 50 index.

To account for liquidity we use the spread between the 3 month Overnight Index Swap (OIS) rate and the 3 month German Treasury (Bubill) yield. Since OIS rate is an expected rate over the life of the contract, and is fully collateralized, it is virtually free of counterparty risk.²³ This makes the OIS-Bubill spread a very good liquidity proxy, see also Krishnamurthy (2010). Typically the spread is quite low, around 20 basis points, but around the collapse of Lehman Brothers, it spiked to around 150 basis points.

Table 3.7 reports the regressions of the survey measures and controls on the estimated inflation risk premia. For the sake of brevity, we do not show regressions relating GDP standard deviation, GDP skew, inflation

²²The difference in the number of observations is due to the fact, that the ECB SPF did only ask participants for 5 year forecasts on an annual basis in the first two years of the survey.

²³The OIS contract is based on an unsecured overnight interest-rate, which implies that the OIS rate still has a credit risk element. The credit element is however smaller than in a 3 month LIBOR contract.

	1 year Inflation Risk Premia					
Constant	4.486 (0.3195)	-5.4823 (-0.2669)	19.4416 (3.6939)	7.6386 (0.8277)	14.2806 (1.4838)	4.3771 (0.2706)
GDP Expectation	8.5782 (1.1714)	8.0259 (0.95)			2.934 (0.5581)	1.2789 (0.2124)
Inflation Skewness			142.8496 (2.464)	152.6385 (2.7196)	124.8086 (2.735)	146.2607 (2.7253)
VSTOXX		0.6009 (1.7738)		0.6765 (2.2805)		0.7058 (2.2372)
Liquidity		-0.2868 (-1.7515)		-0.3547 (-3.8313)		-0.3403 (-3.1035)
R^2	0.1179	0.242	0.2405	0.4185	0.2505	0.42
# Observations	44	44	44	44	44	44

	5 year Inflation Risk Premia					
Constant	-35.1787 (-3.0597)	-38.2436 (-3.861)	7.1797 (4.6748)	1.8175 (0.5763)	-36.326 (-2.957)	-42.2722 (-3.4237)
GDP Expectation	19.0485 (3.5077)	18.7042 (5.0865)			19.2796 (3.5499)	20.2611 (4.3958)
Inflation Skewness			-33.2488 (-0.8941)	-30.4797 (-0.9013)	-34.6455 (-1.0074)	-37.8574 (-1.1608)
VSTOXX		0.1327 (0.9636)		0.2221 (1.5352)		0.1124 (0.9294)
Liquidity		0.02 (0.9722)		-0.024 (-1.2357)		0.0409 (1.7158)
R^2	0.2773	0.3265	0.0778	0.1796	0.3617	0.4221
# Observations	38	38	38	38	38	38

Table 3.7: **Regression of survey measures on inflation risk premia.** Each table reports an OLS regression of surveys measures (GDP expectation and Inflation skewness) along with controls for investor sentiment (VSTOXX) and overall market liquidity. Newey-West t -statistics are given in brackets.

expectation and inflation standard deviation to inflation risk premia, since these variables were all insignificant in our regressions.²⁴

First we consider the 1 year inflation risk premia. Here we see very little evidence that GDP expectations drive inflation risk premia. On the other hand our inflation skewness measure is significant in our regressions, even when controlling for overall market sentiment and liquidity ($t = 2.72$). Our liquidity measure is also very significant ($t = -3.83$), and when included it improves the R^2 significantly. This implies that the changes in short term inflation risk premia in end-2008 have a significant liquidity component.

²⁴Results from these regressions are available upon request.

That inflation skewness is significant is similar to Garcia and Werner (2010), who use the skewness measure from Garcia and Manzanares (2007). Garcia and Werner (2010) find that the inflation skewness from the 5 year survey is significant when regressed on their estimate of the 5 year inflation risk premia.²⁵

Considering the 5 year inflation risk premia we see that the measure of inflation skewness is insignificant. This is not consistent with the finding in Garcia and Werner (2010). However, we believe that part of the explanation could be due to the way the inflation skewness measure is constructed. We use the survey data directly, whereas the measure constructed by Garcia and Manzanares (2007) assume a skew-normal distribution for the inflation. The GDP expectation on the other hand is very significant. In the joint regression with inflation skewness and controls, the t -value is 4.40. For this longer term inflation risk premia, both liquidity ($t = 1.72$) and the VSTOXX index ($t = 0.93$) are still insignificant.

Overall these results indicate that agents form their decisions on inflation risk premia based on both inflation skewness and GDP expectations. Short term inflation risk premia are mostly based on the perceived inflation skewness, as this is the most direct measure of inflation risk. In the short term, fluctuations in GDP do not materialize into changes in salaries and prices. Longer periods with high economic growth do on the other hand most likely translate into price pressures, which would affect longer term inflation risk premia.

3.7 Conclusion

In this paper we have estimated a joint model for nominal interest-rates, inflation swaps and inflation expectations. The model is estimated using Markov Chain Monte Carlo on Euro Area data, i.e. nominal swap rates, inflation swap rates, Euro area HICP inflation and surveys.

In terms of our estimated inflation risk premia, we obtain estimates of average inflation risk premia that are increasing in time to maturity, with 1 year risk premia of 18 basis points and 10 year risk premia of 43 basis points. The risk premia show significant fluctuations with 1 year inflation risk premia being between -156 and 68 basis points, with the lowest value being in the time after the collapse of Lehman Brothers. Longer term

²⁵Garcia and Werner (2010) do not include information on GDP forecasts and market liquidity in their regressions. In terms of GDP information, they only include the current output gap, i.e. no forward looking information.

inflation risk premia (10 year) show less variation, with inflation risk premia between -30 and 81 basis points.

By using draws from Markov Chain Monte Carlo we examined the effect of including surveys, and found that it improves the identification of inflation expectations and thus inflation risk premia. In fact inflation risk premia are mostly statistically insignificant when surveys are excluded, where they are statistically significant when surveys are included to enhance the identification of parameters.

Finally, we relate the estimated risk premia to agents beliefs on the outcome of the economy. We found that short term inflation risk premia are mainly driven by the skewness of the distribution of inflation (as measured by the ECB survey of professional forecasters), where longer term risk premia are driven by GDP expectations. This link between GDP and inflation risk premia is interesting, as inflation swaps could possibly give better real time information on GDP expectations - this is however outside the scope of this paper.

3.8 Appendix: Derivation of nominal ZCB prices, real ZCB prices and Inflation expectations

In this section we derive the different ODEs for pricing and inflation expectation. For simplicity we assume that the both the nominal factors and the inflation factor are collected in one factor $X(t)$. The parameter restrictions to obtain the model structure given in the main text are easily derived.

Furthermore the factor dynamics under Q and P can be written as

$$\begin{aligned}dX(t) &= -KX(t)dt + dW^Q(t) \\dX(t) &= (\lambda_0 + (\lambda_X - K)X(t))dt + dW^P(t)\end{aligned}$$

Nominal ZCB prices

The prices of nominal ZCB prices follow directly from Duffie and Kan (1996), i.e. that ZCB prices are an exponentially affine function of the states:

$$p_n(t, T) = \exp(A_n(t, T) + B_n(t, T)'X(t))$$

where $A_n(t, T)$ and $B_n(t, T)$ solve ordinary differential equations (henceforth ODEs)

$$\begin{aligned}\frac{dA_n(t, T)}{dt} &= -\frac{1}{2}B_n(t, T)'B_n(t, T) + \delta_0 \\ \frac{dB_n(t, T)}{dt} &= K'B_n(t, T) + \delta_X\end{aligned}$$

where $A_n(T, T) = 0$ and $B_n(T, T) = 0$.

Finally the system of ODEs has an explicit solution, where $B_n(t, T)$ is given by

$$B_n(t, T) = (I - \exp(-K'(T - t)))(K')^{-1}\delta_X$$

and where $A_n(t, T)$ is a rather lengthy expression containing integrals of matrix exponentials.

Real ZCB prices

To derive the real ZCB price we use that the expected rate of return from an inflation protected ZCB will be the nominal risk free rate, when the asset

is considered under the risk neutral martingale measure:

$$E_t^Q [d(p_r(t, T)I(t))] = n(t)dt$$

The dynamics of the inflation protected bond is

$$d(p_r(t, T)I(t)) = I(t)dp_r(t, T) + p_r(t, T)dI(t) + d[p_r(t, T), I(t)]$$

Assuming that the real ZCB price is exponentially affine:

$$p_r(t, T) = \exp(A_r(t, T) + B_r(t, T)'X(t))$$

Then

$$E_t^Q \left[\frac{d(p_r(t, T)I(t))}{p_r(t, T)I(t)} \right] = \left(\frac{dA_r(t, T)}{dt} + \frac{dB_r(t, T)}{dt}X(t) \right) dt - B_r(t, T)'KX(t)dt + \frac{1}{2}B_r(t, T)'B_r(t, T)dt + \pi(t)dt = n(t)dt$$

Such that $A_r(t, T)$ and $B_r(t, T)$ solve ODEs

$$\begin{aligned} \frac{dA_r(t, T)}{dt} &= -\frac{1}{2}B_r(t, T)'B_r(t, T) + \delta_0 - \gamma_0 \\ \frac{dB_r(t, T)}{dt} &= K'B_r(t, T) + \delta_X - \gamma_X \end{aligned}$$

where $A_r(T, T) = 0$ and $B_r(T, T) = 0$.

Again the explicit solution for $B_r(t, T)$ is given by

$$B_r(t, T) = (I - \exp(-K'(T-t)))(K')^{-1}(\delta_X - \gamma_X)$$

Inflation dynamics

Assume that CPI can be described as

$$\frac{dI(t)}{I(t)} = (\gamma_0 + \gamma_X'X(t))dt + \eta\tilde{\Sigma}(t)dZ^P(t)$$

Thus the dynamics of log CPI is given by

$$d \log I(t) = \left(\gamma_0 + \gamma_X'X(t) - \frac{1}{2}\eta^2 \right) dt + \eta dZ(t)$$

Inflation Expectations

In this section we will derive the inflation expectation:

$$\begin{aligned} E_t^P \left[\frac{I(T)}{I(t)} \right] &= E_t^P \left[\exp \left(\int_t^T d \log I(s) \right) \right] \\ &= E_t^P \left[\exp \left(\int_t^T \left(\gamma_0 + \gamma'_X X(s) - \frac{1}{2} \eta^2 \right) ds \right) \right] E_t^P \left[\exp \left(\int_t^T \eta dZ(s) \right) \right] \\ &= E_t^P \left[\exp \left(\int_t^T (\gamma_0 + \gamma'_X X(s)) ds \right) \right] \end{aligned}$$

Using the results from Duffie, Pan, and Singleton (2000) we obtain

$$E_t \left[\frac{I(T)}{I(t)} \right] = \exp (A_s(t, T) + B_s(t, T)' X(t))$$

where $A_I(t, T)$ and $B_I(t, T)$ solve ODEs

$$\begin{aligned} \frac{dA_s(t, T)}{dt} &= -\lambda'_0 B_s(t, T) - \frac{1}{2} B_s(t, T)' B_s(t, T) - \gamma_0 \\ \frac{dB_s(t, T)}{dt} &= -(\lambda_X - K)' B_s(t, T) - \gamma_X \end{aligned}$$

subject to $A_s(T, T) = 0$ and $B_s(T, T) = 0$.

The explicit solution for $B_s(t, T)$ is given by

$$B_s(t, T) = (I - \exp((K - \lambda_X)'(T - t))) ((K - \lambda_X)')^{-1} \gamma_X$$

Forward Inflation Expectations

In this section we will derive the forward inflation expectation:

$$\begin{aligned} E_t^P \left[\frac{I(T_1)}{I(T_0)} \right] &= E_t^P \left[E_{T_0} \left[\frac{I(T_1)}{I(T_0)} \right] \right] \\ &= E_t^P [\exp (A_s(T_0, T_1) + B_s(T_0, T_1)' X(t))] \end{aligned}$$

Using the results from Duffie, Pan, and Singleton (2000) we obtain

$$E_t^P \left[\frac{I(T_1)}{I(T_0)} \right] = \exp (A_s(t, T_0, T_1) + B_s(t, T_0, T_1)' X(t))$$

where $A_s(t, T_0, T_1)$ and $B_s(t, T_0, T_1)$ solve ODEs

$$\begin{aligned} \frac{dA_s(t, T_0, T_1)}{dt} &= -\lambda'_0 B_s(t, T_0, T_1) - \frac{1}{2} B_s(t, T_0, T_1)' B_s(t, T_0, T_1) \\ \frac{dB_s(t, T_0, T_1)}{dt} &= -(\lambda_X - K)' B_s(t, T) \end{aligned}$$

subject to $A_s(T_0, T_0, T_1) = A_s(T_0, T_1)$ and $B_s(T_0, T_0, T_1) = B_s(T_0, T_1)$.

Decomposition of the Break Even Inflation Rate

We now consider decomposing the Break Even Inflation Rate (BEIR) into inflation expectations, inflation risk premia and convexity terms.

First the P dynamics of the log-ZCB price can be written as

$$d \log p_k(t, T) = \left[\frac{dA_k(t, T)}{dt} + \frac{dB_k(t, T)'}{dt} X(t) + B_k(t, T)' (\lambda_0 + (\lambda_X - K)X(t)) \right] dt + B_k(t, T)' dW^P(t), \quad k = n, r$$

Since $\log(p_k(T, T)) = \log(p_k(t, T)) + \int_t^T d \log p(s, T)$ and $p_k(T, T) = 1$, we obtain a decomposition for the nominal yield

$$y_n(t, T) = \frac{1}{T-t} \int_t^T (\delta_0 + \delta'_X E_t^P [X(s)]) ds - \frac{1}{2} \frac{1}{T-t} \int_t^T B_n(s, T)' B_n(s, T) ds + \frac{1}{T-t} \int_t^T B_n(s, T)' (\lambda_0 + \lambda_X E_t^P [X(s)]) ds$$

and similarly for the real yield

$$y_r(t, T) = \frac{1}{T-t} \int_t^T ((\delta_0 - \gamma_0) + (\delta_X - \gamma_X)' E_t^P [X(s)]) ds - \frac{1}{2} \frac{1}{T-t} \int_t^T B_r(s, T)' B_r(s, T) ds + \frac{1}{T-t} \int_t^T B_r(s, T)' (\lambda_0 + \lambda_X E_t^P [X(s)]) ds$$

Thus the break even inflation rate can be written as $BEIR = y_n(t, T) - y_r(t, T)$

$$BEIR = \underbrace{\frac{1}{T-t} \int_t^T (\gamma_0 + \gamma_X' E_t^P [X(s)]) ds}_{\text{Inflation expectations}} + \underbrace{\frac{1}{2} \frac{1}{T-t} \int_t^T (B_r(s, T)' B_r(s, T) - B_n(s, T)' B_n(s, T)) ds}_{\text{Convexity correction}} + \underbrace{\frac{1}{T-t} \int_t^T [(B_n(s, T) - B_r(s, T))' (\lambda_0 + \lambda_X E_t^P [X(s)])] ds}_{\text{Inflation risk premia}}$$

where

$$E_t^P [X(T)] = \exp(-(K - \lambda_X)(T - t)) X(t) + (I - \exp(-(K - \lambda_X)(T - t))) (K - \lambda_X)^{-1} \lambda_0$$

3.9 Appendix: MCMC estimation of the model

The conditionals $p(R_t^\bullet | \Theta, X, Y)$

The conditional of the yield observations can be found to be

$$\begin{aligned} p(R^n | \Theta, X, Y) &= \prod_{t \in \mathcal{T}_N} p(R_t^n | \Theta, X, Y) \\ &\propto \prod_{t \in \mathcal{T}_N} (\sigma_n^2)^{-N/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_n^2} (e_t^n)' (e_t^n)\right) \\ &\propto (\sigma_n^2)^{-\sum_{t \in \mathcal{T}_N} N/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_n^2} \sum_{t \in \mathcal{T}_N} (e_t^n)' (e_t^n)\right) \end{aligned}$$

where

$$e_t^n = R_t^n - S_n(t, T)$$

Similarly the conditional of the inflation swap rate observations can be found to be

$$\begin{aligned} p(R^k | \Theta, X, Y) &= \prod_{t \in \mathcal{T}_K} p(R_t^k | \Theta, X, Y) \\ &\propto \prod_{t \in \mathcal{T}_K} (\sigma_k^2)^{-K/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_k^2} (e_t^k)' (e_t^k)\right) \\ &\propto (\sigma_k^2)^{-\sum_{t \in \mathcal{T}_K} K/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_k^2} \sum_{t \in \mathcal{T}_K} (e_t^k)' (e_t^k)\right) \end{aligned}$$

where

$$e_t^k = R_t^k - \left(e^{([A_r(t, t+\tau) - A_n(t, t+\tau)] + [B_r(t, t+\tau) - B_n(t, t+\tau)]' X_t + C_r(t, t+\tau) Y_t) \frac{1}{\tau}} - 1 \right)$$

Finally the conditional for the survey observations is given by

$$\begin{aligned} p(R^s | \Theta, X, Y) &= \prod_{t \in \mathcal{T}_S} p(R_t^s | \Theta, X, Y) \\ &\propto \prod_{t \in \mathcal{T}_S} (\sigma_s^2)^{-S/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_s^2} (e_t^s)' (e_t^s)\right) \\ &\propto (\sigma_s^2)^{-\sum_{t \in \mathcal{T}_S} S/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_s^2} \sum_{t \in \mathcal{T}_S} (e_t^s)' (e_t^s)\right) \end{aligned}$$

where

$$e_t^s = R_t^s - \alpha_\tau - \left(e^{A_s(t,t+\tau,t+\tau+1)+B_s(t,t+\tau,t+\tau+1)'X_t+C_s(t,t+\tau,t+\tau+1)Y_t} - 1 \right)$$

The conditional $p(X, Y|\Theta, R)$

First we collect X and Y in one vector \hat{X} which has dynamics

$$d\hat{X}(t) = \left(\theta^P + K^P \hat{X}(t) \right) dt + IdW^P(t)$$

Which by an Euler approximation gives us

$$\hat{X}_{t+\Delta} - \hat{X}_t = \left(\theta^P + K^P \hat{X}_t \right) \Delta + \sqrt{\Delta} \varepsilon_t$$

where $\varepsilon_t \sim \mathcal{N}(0, I)$. This implies

$$p(X_t, Y_t|\Theta) \propto \exp \left(-\frac{1}{2} \frac{1}{\Delta} v_t' v_t \right)$$

where $v_t = \hat{X}_{t+1} - \hat{X}_t - \left(\theta^P + K^P \hat{X}_t \right) \Delta$.

The conditional for all the state observations are then given by:

$$\begin{aligned} p(X, Y|\Theta) &= \prod_{t=0}^{T-1} p(X_t, Y_t|\Theta) p(X_0) p(Y_0) \\ &\propto \exp \left(-\frac{1}{2} \frac{1}{\Delta} \sum_{t=0}^{T-1} v_t' v_t \right) \end{aligned}$$

where we have assumed independent prior for X_0 and Y_0 .

The conditional $p(\log I|Y, X, \Theta)$

We assume that (log-) inflation is observed without error. To obtain the density we use an Euler approximation

$$\Delta \log I_{t_k} = \sum_{t=t_{k-h}}^{t_k} \left(\gamma_0 + \gamma_X' X_t + \gamma_Y' Y_t - \frac{1}{2} \eta^2 \right) \Delta + \sum_{t=t_{k-h}}^{t_k} \eta \varepsilon_{t,\pi}$$

Which implies that

$$\Delta \log I_{t_k} \sim \mathcal{N}(M_t, V_t)$$

$$M_t = \sum_{t=t_k-h}^{t_k} \left(\gamma_0^P + \gamma_X^{P'} X_t + \gamma_Y^P Y_t - \frac{1}{2} (R_X' R_X + R_Y^2 + \eta^2) \right) \Delta + [R_X \quad R_Y] v_t$$

$$V_t = \Delta \sum_{t=t_k-h}^{t_k} \eta^2$$

Thus the conditional density will be

$$\begin{aligned} p(\Delta \log I|X, \Theta) &= \prod_{t \in \mathcal{T}_I} p(\Delta \log I_t|X, \Theta) \\ &\propto \prod_{t \in \mathcal{T}_I} p(\Delta \log I_t|X, \Theta) \\ &\propto \prod_{t \in \mathcal{T}_I} \frac{1}{\sqrt{V_t}} \exp \left(-\frac{1}{2} \frac{1}{V_t} (\Delta \log I_t - M_t)^2 \right) \end{aligned}$$

Sampling σ_n, σ_k and σ_s

To sample the measurement errors we use Bayes theorem to obtain

$$p(\sigma_n|R, X, Y, \Theta_{\setminus \sigma_n}) \propto p(R|\Theta, X, Y)$$

Thus we can sample the measurement errors through Gibbs sampling and obtain Inverse Gamma draws

$$\begin{aligned} \sigma_n^2 &\sim \mathcal{IG} \left(\frac{\sum_{t \in \mathcal{T}_N} N}{2} + 1, \frac{1}{2} \sum_{t \in \mathcal{T}_N} (e_t^n)' (e_t^n) \right) \\ \sigma_k^2 &\sim \mathcal{IG} \left(\frac{\sum_{t \in \mathcal{T}_K} K}{2} + 1, \frac{1}{2} \sum_{t \in \mathcal{T}_K} (e_t^k)' (e_t^k) \right) \\ \sigma_s^2 &\sim \mathcal{IG} \left(\frac{\sum_{t \in \mathcal{T}_S} S}{2} + 1, \frac{1}{2} \sum_{t \in \mathcal{T}_S} (e_t^s)' (e_t^s) \right) \end{aligned}$$

Sampling α

To sample the measurement bias we use Bayes theorem to obtain

$$p(\alpha_\tau|R, X, Y, \Theta_{\setminus \alpha_\tau}) \propto p(R_{s,\tau}|\Theta, X, Y)$$

Thus we can sample the measurement bias through Gibbs sampling and obtain Normal draws

$$\begin{aligned}\alpha_\tau &\sim \mathcal{N}(\mu_\alpha, \sigma_\alpha^2) \\ \mu_\alpha &= \frac{\sum_{t \in \mathcal{T}_S} R_t^s - (e^{A_s(t, t+\tau, t+\tau+1)} + B_s(t, t+\tau, t+\tau+1))' X_t + C_s(t, t+\tau, t+\tau+1) Y_t}{N} \\ \sigma_\alpha^2 &= \frac{\sigma_s^2}{N} \\ N &= \sum_{t \in \mathcal{T}_S} 1\end{aligned}$$

Sampling X and Y

To sample the latent states we use a Random-Walk Metropolis Hastings (RW-MH) algorithm. In this case new states are sampled as

$$\begin{aligned}X^{m+1} &= X^m + \epsilon_{x, m+1} \\ Y^{m+1} &= Y^m + \epsilon_{y, m+1}\end{aligned}$$

where $\epsilon_{\bullet, m+1}$ is a zero mean random variable with a variance that needs to be calibrated. The conditional for the states can be written as

$$\begin{aligned}p(X, Y | \Theta, R) &\propto p(R | X, Y, \Theta) p(X, Y | \Theta) \\ &\propto p(R^n | X, Y, \Theta) p(R^k | X, Y, \Theta) p(R^s | X, Y, \Theta) p(\Delta \log I | X, Y, \Theta) p(X, Y | \Theta)\end{aligned}$$

And the draw will be accepted with probability

$$\alpha = \max\left(\frac{p(R | X^{m+1}, Y^{m+1}, \Theta) p(X^{m+1}, Y^{m+1} | \Theta)}{p(R | X^m, Y^m, \Theta) p(X^m, Y^m | \Theta)}, 1\right)$$

Sampling Θ^Q

To sample the risk neutral parameters in the nominal interest-rate process and the latent factors, we use a RW-MH algorithm. The conditional is given by

$$\begin{aligned}p(\Theta^Q | R, X, Y, \Theta_{\setminus Q}) &\propto p(R | X, Y, \Theta) p(X, Y | \Theta) \\ &\propto p(R^n | X, Y, \Theta) p(R^k | X, Y, \Theta) p(R^s | X, Y, \Theta) p(X, Y | \Theta)\end{aligned}$$

Where the surveys and states enter since that P -parameters are the sum of risk neutral parameters and risk premia, cf. above.

The draws will be accepted with probability

$$\alpha = \max\left(\frac{p(R^n | X, Y, \Theta^{m+1}) p(R^k | X, Y, \Theta^{m+1}) p(R^s | X, Y, \Theta^{m+1}) p(X, Y | \Theta^{m+1})}{p(R^n | X, Y, \Theta^m) p(R^k | X, Y, \Theta^m) p(R^s | X, Y, \Theta^m) p(X, Y | \Theta^m)}, 1\right)$$

Sampling Θ^P

To sample the risk premia parameters we use a RW-MH algorithm. The conditional is given by

$$p(\Theta^P | R, X, Y, \Theta_{\setminus P}) \propto p(R^s | X, Y, \Theta) p(\Delta \log I | X, Y, \Theta) p(X, Y | \Theta)$$

The draws will be accepted with probability

$$\alpha = \max \left(\frac{p(R^s | X, Y, \Theta^{m+1}) p(\Delta \log I | X, Y, \Theta^{m+1}) p(X, Y | \Theta^{m+1})}{p(R^s | X, Y, \Theta^m) p(\Delta \log I | X, Y, \Theta^m) p(X, Y | \Theta^m)}, 1 \right)$$

Sampling Θ^π

To sample the risk premia parameters we use a RW-MH algorithm. The conditional is given by

$$p(\Theta^\pi | R, X, Y, \Theta_{\setminus \pi}) \propto p(R^k | X, Y, \Theta) p(R^s | X, Y, \Theta) p(\Delta \log I | X, Y, \Theta) p(X, Y | \Theta)$$

The draws will be accepted with probability

$$\alpha = \max \left(\frac{p(R^k | X, Y, \Theta^{m+1}) p(R^s | X, Y, \Theta^{m+1}) p(\Delta \log I | X, Y, \Theta^{m+1}) p(X, Y | \Theta^{m+1})}{p(R^k | X, Y, \Theta^m) p(R^s | X, Y, \Theta^m) p(\Delta \log I | X, Y, \Theta^m) p(X, Y | \Theta^m)}, 1 \right)$$

Essay 4

Affine Nelson-Siegel Models and Risk Management Performance

Abstract

In this paper we assess the ability of the Affine Nelson-Siegel model-class with stochastic volatility to match observed distributions of Danish Government bond yields. Based on data from 1987 to 2010 and using a Markov Chain Monte Carlo estimation approach we estimate 7 different model specifications and test their ability to forecast yields (both means and variances) out of sample. We find that models with 3 CIR-factors perform the best in short term predictions, while models with a combination of CIR and Gaussian factors perform well on 1 and 5-year horizons. Overall our results indicate that no single model should be used for risk management, but rather a suite of models.

Keywords: Affine Term Structure Models, Nelson-Siegel, Markov Chain Monte Carlo, Variance Forecasts

JEL Classification: G12, G17, C11, C58

¹I would like to thank Jacob Ejsing, Jesper Lund, Bjarne Astrup Jensen, Anne-Sofie Reng Rasmussen and the Government Debt Management section at Danmarks Nationalbank for useful suggestions.

4.1 Introduction

The ability of a risk management model to match observed data is of great importance. In this setting Value-at-Risk (VaR) is an often used tool to measure the riskiness of a given portfolio. VaR measures portfolio losses at a given quantile, e.g. at a 99 percent level. A risk measure used by debt issuers is Cost-at-Risk (CaR), which measures the cost of issuing bonds at a given quantile level, see Danmarks Nationalbank (2005).

In this paper we examine the ability of a class of interest-rate models to capture the time varying distribution of Danish government bond yields, here measured by means and variances.² We consider both short-term forecasts, i.e. 1 month, and long-term forecasts, i.e. up to 5 years. The long horizons are motivated by the desire to evaluate the considered models in the setting of a debt issuer, where debt is issued over long horizons.

We consider the Affine Nelson-Siegel model-class with stochastic volatility, as introduced in Christensen, Lopez, and Rudebusch (2010).³ In these models stochastic volatility can be generated by the level, slope or curvature factors (or a combination of these). This could possibly imply very different results for the predicted means and variances for each of these models. In Christensen, Lopez, and Rudebusch (2010) the preferred model is a model where all three factors, i.e. level, slope and curvature, drive stochastic volatility. This model performs well in describing the monthly volatility patterns of US treasuries, UK gilts and US swap rates.

On the issue of describing interest-rate volatility Collin-Dufresne, Goldstein, and Jones (2008) argue that spanned yield curve factors cannot describe the volatility of yield changes. Instead they argue that at least one additional factor which does not describe yields should be included to model interest-rate volatility, so-called *Unspanned Stochastic Volatility*. On the other hand Jacobs and Karoui (2009) argue that the results from Collin-Dufresne, Goldstein, and Jones (2008) does to some extent depend on the considered data and sampling period.

In the context of forecasting densities Egorov, Hong, and Li (2006) consider density forecasts for affine term structure models. They consider a non-parametric test of the realization of the entire yield curve. They find that affine models outperform a random walk in describing the joint probability

²In an appendix we also consider density forecasts. The results are overall consistent with the mean and variance forecasts.

³We also include a 2-factor Cox-Ingersoll-Ross model as this model has similarities to the preferred affine Nelson-Siegel model, and has previously been used in risk management, see for instance Danmarks Nationalbank (2005).

of yields, however they still fail in making satisfactory density forecasts overall.

As mentioned above, we consider both mean and variance forecasts, as forecasting means and variances provide more simple and transparent inference on model performance, compared to density forecasts such as in Egorov, Hong, and Li (2006).

In terms of short term forecasts (one month), we find results similar to Christensen, Lopez, and Rudebusch (2010), i.e. that a model only based on CIR-factors performs well both in terms of forecasting means and variances.

For long-term forecasts we find a mixed picture, which is to some extent driven by the considered time period, which shows a downward trend in interest-rates; and for an extended period in the data sample, the yield curve was inverted. The preferred model of Christensen, Lopez, and Rudebusch (2010) tends to produce point forecasts that are too high and volatility forecasts that are too low, implying a limitation from only using CIR-factors. The models based on a combination of CIR and Gaussian factors, and where the level factor drives stochastic volatility, perform reasonably well in forecasting the mean of the distributions, but for very long-horizon forecasts (5 years), the distribution is wide compared to actual data.

The structure of the paper is as follows; section 4.2 describes our data, section 4.3 describes affine term structure models in general and sections 4.4 and 4.5 describe the multifactor CIR and affine Nelson-Siegel models. Section 4.6 describes the model estimation, section 4.7 describes the in-sample behavior of the models and section 4.8 describes the results on forecasting out-of-sample. Finally, section 4.9 concludes the paper.

4.2 The Danish Government Bond Term Structure

As mentioned above we are interested in assessing the performance of the model candidates with a view towards a risk management setting (Value-at-Risk). We focus on the Danish government bond term structure.

Our data consist of monthly zero-coupon yields, sampled at the first trading day of the month. The yields are extracted from the government bonds traded in the market, using an extended Nelson-Siegel approach. We have yields for the 3 month, 2 year, 5 year, 10 year and 15 year maturity, and time series of the yields are presented in Figure 4.1.⁴

⁴Recently, a 30 year bond has been issued. however, we do not use yields of this

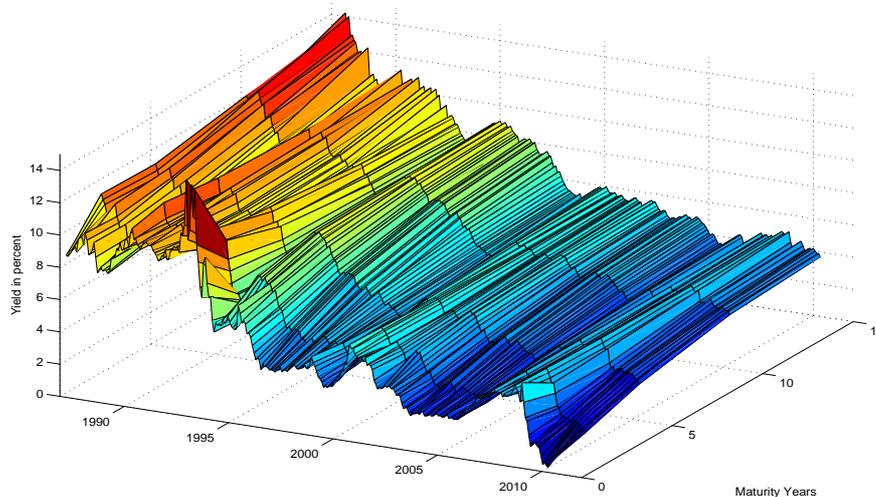


Figure 4.1: **Time series of Danish government zero-coupon bond yields.** The data sample is January 1987 to May 2010. Yields are extracted from Danish Government bonds using an extended Nelson-Siegel approach.

Typically the term structure is upward sloping, with a general downward trend in the interest-rate level over the period. We also see significant downward sloping term structures during the ERM crisis in 1992.

Table 4.1 provides descriptive statistics for the yield curve data. Again we see that the average term structure is upward sloping, with an unconditional standard deviation between 2.5 and 3.1 percent. With regard to monthly changes in yields, we see that the standard deviation is between 28 and 57 basis points and that all changes in yields show excess kurtosis, with the 3 month yield being the most significant.

Researchers have typically found that three factors are sufficient to model the term structure of interest-rates (see Litterman and Scheinkman (1991)). The three factors are determined through use of a Principal Component Analysis (henceforth PCA), which we also perform on our data. We perform the PCA on both levels and changes in yields, to control for the fact that yields typically are nearly-integrated processes.

The results from the PCA are found in Table 4.1 and Figure 4.2. We confirm the typical pattern, i.e. that the first factor accounts for the majority of the variation in the data. When the PCA is performed on the levels the first maturity in our analysis due to a very short time series of 30 year yields.

Maturity	# obs	Levels				Changes			
		Mean (%)	Std. Dev. (%)	Skew.	Ex.Kurt.	Mean (%)	Std. Dev. (%)	Skew.	Ex.Kurt.
3 Months	281	5.416	3.072	0.632	-0.64	-0.029	0.571	-0.759	14.399
2 Years	281	5.565	2.751	0.509	-1.083	-0.033	0.325	-0.254	1.957
5 Years	281	5.952	2.514	0.485	-1.15	-0.031	0.306	0.061	1.255
10 Years	281	6.481	2.514	0.577	-0.908	-0.03	0.282	0.629	2.067
15 Years	281	6.834	2.673	0.741	-0.531	-0.031	0.382	-0.79	3.983

Maturity	Levels			Changes		
	1st P.C.	2nd P.C.	3rd P.C.	1st P.C.	2nd P.C.	3rd P.C.
3 Months	0.4283	0.7276	-0.5068	0.2841	0.7883	-0.479
2 Years	0.4543	0.2503	0.5103	0.5066	0.1924	0.1996
5 Years	0.4579	-0.0634	0.5011	0.5118	0.026	0.4577
10 Years	0.4516	-0.3711	-0.074	0.5063	-0.2743	0.1384
15 Years	0.4433	-0.5159	-0.4756	0.38	-0.5154	-0.7086
Pct Explained	94.63	4.76	0.46	66.95	19.46	10.03

Table 4.1: **Top panel:** Descriptive statistics of the time series of Danish government bond yields. The data sample is January 1987 to May 2010. Yields are extracted from Danish Government bonds using an extended Nelson-Siegel approach. **Bottom panel:** Factor loadings on the first three principal components. The final row shows the proportion of the variation which is captured by each principal component.

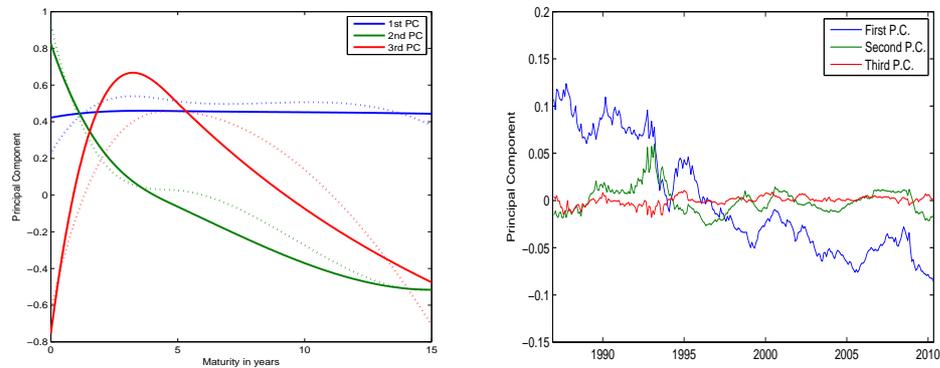


Figure 4.2: **Left:** Factor loadings on yields on the first three principal components. Solid lines are factor loadings from the PCA performed on levels and dashed lines are factor loadings from the PCA performed on changes. **Right:** Time series of the first three principal components (levels).

factor accounts for almost 95 percent of the variation in data, where the second and third factor account for a little less than 5 and 1 percent of the variation. When the PCA is performed on changes in yields, the factors account for 67, 19 and 10 percent, respectively, which emphasizes the value of performing the PCA on both levels and changes. Regardless of whether the PCA is performed on levels or changes, the factor loadings are also similar to the typical pattern, i.e. the first factor can be interpreted as a level factor, the second factor as a slope factor, and the third factor as a curvature factor.

The right panel in Figure 4.2 shows the time series of the extracted principal components. The first factor confirms the downward trend in the interest-rate level. During the ERM crisis in 1992, the slope factor also shows a significant spike, confirming the interpretation as a slope factor.

4.3 Affine Term Structure Models

In this section we consider the class of affine term structure models, i.e. we consider the results of Duffie and Kan (1996).

Consider a complete stochastic basis $(\Omega, \mathcal{F}, (\mathcal{F}_t), Q)$ where Q is a risk neutral martingale measure and the filtration $(\mathcal{F}_t) = \{\mathcal{F}_t, t \geq 0\}$ satisfies the usual conditions (see Protter (2005)). The term structure is driven by a number of factors, n , such that the short rate is defined as $r(t) = \delta_0 + \delta_1' X(t)$ (where δ_0

is a scalar and δ_1 is an n -vector). The factors are adapted to the filtration, \mathcal{F}_t , and the factors are assumed to be a Markov process on \mathbb{R}^n which solves the following stochastic differential equation:

$$dX(t) = K(\theta - X(t))dt + SD(X(t))dW^Q(t)$$

where W^Q is a standard Wiener process in \mathbb{R}^n , θ is an n -vector and K is an $n \times n$ -matrix and S is an $n \times n$ -matrix. Finally, the matrix $D(X(t))$ is an $n \times n$ matrix taking the form:

$$D(X(t)) = \begin{pmatrix} \sqrt{\alpha_1 + \beta_1 X(t)} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sqrt{\alpha_k + \beta_k X(t)} \end{pmatrix}$$

where $\alpha_j, j = 1, \dots, k$ are scalars and $\beta_j, j = 1, \dots, k$ are k -vectors ⁵.

Duffie and Kan (1996) show that zero-coupon bond prices are exponentially affine functions of the states

$$p(t, T) = E_t^Q \left[\exp \left(- \int_t^T r(s) ds \right) \right] = \exp(A(t, T) + B(t, T)'X(t))$$

where $A(t, T)$ and $B(t, T)$ are solutions to ordinary differential equations (henceforth ODEs)

$$\begin{aligned} \frac{dA(t, T)}{dt} &= \delta_0 - B(t, T)'K\theta - \frac{1}{2} \sum_{j=1}^n [S'B(t, T)B(t, T)'S]_{jj} \alpha_j \\ \frac{dB(t, T)}{dt} &= \delta_1 + K'B(t, T) - \frac{1}{2} \sum_{j=1}^n [S'B(t, T)B(t, T)'S]_{jj} \beta_j' \end{aligned}$$

subject to the initial conditions $A(T, T) = 0$ and $B(T, T) = 0$.

The implication of the above expression is that zero-coupon yields are affine functions of the factors:

$$y(t, T) = -\frac{1}{T-t} \log p(t, T) = -\frac{A(t, T)}{T-t} - \frac{B(t, T)'}{T-t} X(t)$$

To complete the model, we need to consider P -dynamics, i.e. risk premia. Here there are three obvious choices; the *completely affine* considered by

⁵In this general specification of an affine term structure model, not all parameters are identifiable. Dai and Singleton (2000) show when various specifications of the affine term structure models are identifiable. These constraints hold in all the considered models.

Dai and Singleton (2000), the *essentially affine* considered by Duffee (2002) and the *extended affine* by Cheredito, Filipovic, and Kimmel (2007). These three specifications have different flexibility, and for the extended affine specification the Feller condition must also hold for the square root factor.⁶ For each model, we explicitly state the model dynamics and risk premia, see Appendix 4.10 for further details.

4.4 Multi-factor Cox-Ingersoll-Ross models

In this section we consider the multi-factor Cox-Ingersoll-Ross (henceforth CIR) models⁷. The classical CIR model (see Cox, Ingersoll, and Ross (1985)) is defined by the stochastic differential equation for the short rate, r

$$dr(t) = (\theta - \kappa r(t)) dt + \sigma \sqrt{r(t)} dW^Q(t)$$

This model belongs to the affine class of interest-rate models, and therefore zero-coupon bond prices are exponentially affine:

$$p(t, T) = \exp(A(t, T) + B(t, T)r(t))$$

where $A(t, T)$ and $B(t, T)$ solve Ricatti ODEs, which in this case has analytical solutions given by

$$A(t, T) = \left(\frac{\theta}{\sigma^2}\right) \left[2 \log \left(\frac{2\gamma}{(\kappa + \gamma)(e^{\gamma(T-t)} - 1) + 2\gamma} \right) + (\kappa + \gamma)(T - t) \right]$$

$$B(t, T) = \frac{2(1 - e^{\gamma(T-t)})}{(\kappa + \gamma)(e^{\gamma(T-t)} - 1) + 2\gamma}$$

$$\gamma = \sqrt{\kappa^2 + 2\sigma^2}$$

In the multi-factor CIR models, the short rate is defined to be the sum of the factors $r(t) = \sum_{k=1}^n X_k(t)$. The factors are defined as independent

⁶The Feller condition ensures that a square root process does not reach its boundary value, i.e. zero for a CIR-process. For the CIR process the Feller condition is $\theta > \frac{1}{2}\sigma^2$.

⁷For a risk management application of multi-factor CIR models see Danmarks Nationalbank (2005).

processes which evolve as the short rate in the classical CIR model:

$$\begin{pmatrix} dX_1(t) \\ \vdots \\ dX_n(t) \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix} - \begin{pmatrix} \kappa_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \kappa_n \end{pmatrix} \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sigma_n \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sqrt{X_n(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ \vdots \\ dW_n^Q(t) \end{pmatrix}$$

The independence of the factors in the multi-factor CIR model imply that zero-coupon bond prices take the form:

$$p(t, T) = \exp \left(A(t, T) + \sum_{k=1}^n B_k(t, T) X_k(t) \right)$$

where

$$A(t, T) = \sum_{k=1}^n \left(\frac{\theta_k}{\sigma_k^2} \right) \left[2 \log \left(\frac{2\gamma_k}{(\kappa_k + \gamma_k)(e^{\gamma_k(T-t)} - 1) + 2\gamma_k} \right) + (\kappa_k + \gamma_k)(T - t) \right] \\ B_k(t, T) = \frac{2(1 - e^{\gamma_k(T-t)})}{(\kappa_k + \gamma_k)(e^{\gamma_k(T-t)} - 1) + 2\gamma_k} \\ \gamma_k = \sqrt{\kappa_k^2 + 2\sigma_k^2}$$

To complete the model, we need to specify the risk premia. The market price of risk is chosen as the Completely Affine specification, which implies the P -dynamics.

$$\begin{pmatrix} dX_1(t) \\ \vdots \\ dX_n(t) \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \vdots \\ \theta_n \end{pmatrix} - \begin{pmatrix} \kappa_1^P & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \kappa_n^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_1 & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sigma_n \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ \vdots & \ddots & \vdots \\ 0 & 0 & \sqrt{X_n(t)} \end{pmatrix} \begin{pmatrix} dW_1^P(t) \\ \vdots \\ dW_n^P(t) \end{pmatrix}$$

This implies that we have a single risk premium parameter for each factor. The risk premia specification implies that both mean reversion speed and level are different under the real world measure, compared to the risk neutral measure.

The factor loadings in a two-factor model typically look similar to level and slope, as in the PCA above. However, since the factors are independent and with the same structure, we cannot say in advance which factor will act as level and which will act as slope. One solution to this problem is to impose a structure on a general affine term-structure model, such that level, slope and curvature are forced upon the factor loadings - the Affine Nelson-Siegel Models of Christensen, Diebold, and Rudebusch (2011) offer this kind of identification.

4.5 Affine Nelson-Siegel models

In this section we consider the affine Nelson-Siegel models introduced by Christensen, Diebold, and Rudebusch (2011).

Recall the classical Nelson-Siegel model for fitting the cross section of yields:

$$y(t, t + \tau) = \beta_0 + \beta_1 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + \beta_2 \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

where $y(t, t + \tau)$ is the time t zero-coupon yield with maturity τ and $\beta_0, \beta_1, \beta_2$ and λ are parameters.

The Nelson-Siegel approach offers a static fit of the term structure, and parameters will vary to fit the term structure across different trading days. To remedy this fact, Diebold and Li (2006) suggest allowing the β coefficients to be interpreted as time varying level, slope and curvature, creating the *Dynamic Nelson-Siegel model*:

$$y(t, t + \tau) = B_1(t, t + \tau)L(t) + B_2(t, t + \tau)S(t) + B_3(t, t + \tau)C(t)$$

where

$$\begin{aligned} B_1(t, t + \tau) &= 1 \\ B_2(t, t + \tau) &= \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) \\ B_3(t, t + \tau) &= \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right) \end{aligned}$$

Obviously the notion of level, slope and curvature is justified by the factor loadings, $B_\bullet(t, T)$, which resemble the level, slope and curvature factor loadings from the PCA. Figure 4.3 shows the factor loadings with λ equal to 0.5.

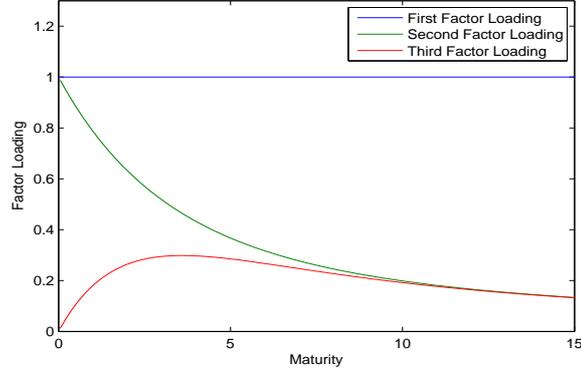


Figure 4.3: **Nelson-Siegel factor loadings.** Factor loadings of yields on the level, slope and curvature in the Nelson-Siegel specification. λ is equal to 0.5.

As argued in Christensen, Diebold, and Rudebusch (2011), the Dynamic Nelson-Siegel model is not arbitrage-free. By casting the model into the affine term structure framework, Christensen, Diebold, and Rudebusch (2011) are able to consider an arbitrage free version of the Dynamic Nelson-Siegel model, coined the *Affine Nelson-Siegel model*. This model obviously preserves the three factor loadings from the Dynamic Nelson-Siegel model. It is obtained by assuming that the short rate is the sum of the level and slope, $r(t) = X_1(t) + X_2(t)$ and that the factors solve the stochastic differential equation:

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

where X_1 is the level factor, X_2 is the slope factor and X_3 is the curvature factor. Since the model is part of the affine term structure framework, zero-coupon bond prices are exponentially affine:

$$p(t, T) = \exp(A(t, T) + B_1(t, T)X_1(t) + B_2(t, T)X_2(t) + B_3(t, T)X_3(t))$$

where

$$\begin{aligned} B_1(t, T) &= -(T - t) \\ B_2(t, T) &= - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right) \\ B_3(t, T) &= - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} - (T - t)e^{-\lambda(T-t)} \right) \end{aligned}$$

and where $A(t, T)$ is a convexity adjustment term due to Jensen's inequality. Furthermore $A(t, T)$ has an explicit solution. The factor loadings given above yield an expression for zero-coupon yields that is consistent with the classical Nelson-Siegel representation, except from the convexity adjustment term which ensures that the model is indeed arbitrage free.

One implication of matching the factor loadings in the affine Nelson-Siegel model is that yields are normally distributed with no time variation in the volatility. This is not a desirable property of a model used for risk management, where capturing volatility correctly is of great importance.⁸

One solution to this problem is offered by Christensen, Lopez, and Rudebusch (2010), who let level, slope and/or curvature be the drivers of stochastic volatility. By letting the risk-neutral mean-reversion matrix be equal (or very similar) to the original affine Nelson-Siegel model, the model factors keep their interpretation as level, slope and curvature (see also Figure 4.4 for our estimated factor loadings).

Christensen, Lopez, and Rudebusch (2010) derive 5 different specifications:

- $AFNS_1 - L$: A model where the volatility is driven by the level factor.
- $AFNS_1 - C$: A model where the volatility is driven by the curvature factor.
- $AFNS_2 - LC$: A model where the volatility is driven by the level and curvature factors.
- $AFNS_2 - SC$: A model where the volatility is driven by the slope and curvature factors.
- $AFNS_3$: A model where the volatility is driven by all three factors.

In Appendix 4.10 we describe all of the above mentioned models, along with the risk premia specification considered by Christensen, Lopez, and Rudebusch (2010) .

⁸Compare also to Table 4.1 which shows the non-normality of the data.

For instance, the $AFNS_1 - L$ will, under the risk neutral measure, follow the SDE:

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & 1 & 0 \\ \sigma_{31} & \sigma_{32} & 1 \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21}X_1(t)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31}X_1(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

We notice that the level factor now is a square root process, therefore the unit root behavior, is replaced by a near unit root behavior, since in the first diagonal element, the zero is replaced by a small number ε . Furthermore, we see that in the volatility loading matrix, $D(X(t))$ there is now a dependence on the level factor. This representation therefore offers a natural combination of the $\mathbb{A}_1(3)$ of Dai and Singleton (2000) and the Gaussian Affine Nelson-Siegel model. The other Affine Nelson-Siegel models with stochastic volatility are defined in an analogous fashion.

For the $AFNS_3$ we also estimate a version with a Completely Affine risk premia specification. For the $AFNS_3$ model, the Feller condition has to hold for the slope and curvature factors, due to the extended affine risk premia specification. The Feller condition is the reason to consider a different specification; our estimates show that the Feller condition is binding for the slope factor, implying a higher mean-reversion level, than what otherwise would have been estimated. To assess the implication of the Feller condition in our analysis, we consider the completely affine specification.⁹ We label this model $AFNS_3 - CA$.

4.6 Model Estimation

We adopt a Bayesian approach. Admittedly, Bayesian methods are more computationally cumbersome than for instance Quasi Maximum Likelihood methods, however Bayesian methods allow for more precise filtering of volatility, compared to the Kalman filter or assuming that some yields are observed without error.

Furthermore, a Bayesian estimation allows for precise inference of each estimation, i.e. that convergence is indeed obtained. Furthermore a Bayesian

⁹Feldhütter (2008) also finds that the Feller condition is limiting the behavior of some affine term structure models. He solves the problem by considering the semi-affine risk premia by Duarte (2004).

method allows for easier implementation of the parameter restrictions imposed by various Feller conditions.

In this section we will describe the notation used in the estimation, the specification of conditional distributions and the implemented hybrid MCMC algorithm used. A survey article on MCMC is Johannes and Polson (2003), where text book treatments can be found in Gamerman and Lopes (2006) and Robert and Casella (2004).

Notation

We observe N different zero-coupon yields each observation day. For each observation date, t , the yields are collected in the vector Y_t . The full collection of yields are given by $Y = \{Y_t\}_{t=1}^T$. We assume that the yields are observed with measurement errors:

$$Y_t = -\frac{(A(t, t + \tau) + B(t, t + \tau)'X_t)}{\tau} + \varepsilon_{n,t}$$

which are independent and normally distributed with common variance

$$\varepsilon_{n,t} \sim \mathcal{N}(0, \sigma_\varepsilon I_N)$$

With respect to parameters, we denote those that only enter in the risk neutral (Q) drift by Θ^Q and those related to the real-world (P) drift Θ^P . Parameters entering both under the risk-neutral and real-world probability measures, such as diffusion parameters, are denoted Θ^D . The entire collection of parameters, including the measurement error, are given by $\Theta = (\Theta^Q, \Theta^P, \Theta^D, \sigma_\varepsilon)'$.

Estimation using MCMC

When estimating the model we are interested in sampling from the target distribution of parameters and state variables, $p(\Theta, X|Y)$. To sample from this distribution, the Hammersley-Clifford theorem (Hammersley and Clifford (1974) and Besag (1974)) implies that this can be done by sampling from the complete conditionals

$$\begin{aligned} & p(\Theta^Q | \Theta_{\setminus Q}, X, Y) \\ & p(\Theta^P | \Theta_{\setminus P}, X, Y) \\ & p(\Theta^D | \Theta_{\setminus D}, X, Y) \\ & p(X | \Theta, Y) \\ & p(\sigma_\varepsilon | \Theta_{\setminus \sigma_\varepsilon}, X, Y) \end{aligned}$$

Thus MCMC handles the sampling from the complicated target distribution $p(\Theta, X|Y)$ by sampling from the simpler conditional distributions. More specifically this is handled by sampling in cycles from the conditional distributions. If one can sample directly from the conditional distribution, the resulting algorithm is called a Gibbs sampler (see Geman and Geman (1984)). If it is not possible to sample from this distribution one can sample using the Metropolis-Hastings algorithm (see Metropolis, Rosenbluth, Rosenbluth, Teller, and E. (1953)). In this paper we use a combination of the two (a so-called hybrid MCMC algorithm), since not all the conditional distributions are known. More precisely, we have the following MCMC algorithm:

$$\begin{aligned} p(X|\Theta, Y) &\sim \text{Metropolis-Hastings} \\ p(\Theta^Q|\Theta_{\setminus Q}, X, Y) &\sim \text{Metropolis-Hastings} \\ p(\Theta^P|\Theta_{\setminus P}, X, Y) &\sim \text{Metropolis-Hastings} \\ p(\Theta^D|\Theta_{\setminus D}, X, Y) &\sim \text{Metropolis-Hastings} \\ p(\sigma_\varepsilon|\Theta_{\setminus \sigma_\varepsilon}, X, Y) &\sim \text{Inverse Gamma} \end{aligned}$$

A more precise description of the algorithm and the conditional distributions are found in Appendix 4.11.

We first run the estimation on the full data sample, and run the Markov chain until convergence - typically 2-3 million simulations - and we save the last 1 million draws of the chain with a thin factor of 1000, leaving 1000 draws for inference. Secondly, since we are interested in out-of-sample performance, we successively remove the most insignificant parameter, defined by an absolute t statistic below 1.65.¹⁰

Approach for recursive model estimation

The main objective of this paper is to examine the out-of-sample performance of the different Affine Nelson-Siegel term structure models. To assess this issue we estimate each model on 10 year data samples. Thus we first estimate a model on data from January 1987 to December 1997, then move the estimation window one month forward, i.e. estimate the model from February 1987 to January 1998. In total this gives us 162 estimations on 10 year samples.

For each model we start the MCMC estimation using the parameters and states estimated on the full data sample. For each data sample we use a

¹⁰We define the t -statistic as the mean of the 1000 draws divided by the standard deviation of the 1000 draws: $t = E[\Theta_i]/\sqrt{\text{Var}[\Theta_i]}$

burn-in size of 1 million draws, and we run an additional 1 million draws of the chain with a thin factor of 1000, leaving 1000 draws for inference. Finally, for each data sample the reduced models are used, cf. above.

4.7 Empirical Results: In-Sample

This section discusses the in-sample results generated by the different models. This includes generated factor loadings and volatilities, and how the models fit the data.

Tables 4.2 and 4.3 show the parameter estimates from the different models. Parameter estimates are based on the full data sample, i.e. from 1987 to 2010. As mentioned above, the models are successively reduced until all parameters are significant.¹¹

One interesting finding is that the parameters in the 2 factor CIR model are estimated close to the ones from the $AFNS_3$ model. Even though κ^Q is estimated freely, our estimates imply that κ_{22}^Q is close to zero, and κ_{11}^Q is estimated in the range of the λ 's from the Affine Nelson-Siegel models. This has two consequences; first, the parameter restrictions in the Affine Nelson-Siegel models are not unreasonable if we obtain similar estimates when the parameters are estimated freely. Secondly, the two factor CIR model can be interpreted as a Nelson-Siegel model reduced to only level and slope.

The resemblance of the factor loadings can also be seen from Figure 4.4, which shows the factor loadings generated by the different models. The $CIR - 2$ model has level and slope loading which is in line with the level and slope loadings generated by the affine Nelson-Siegel models. One other interesting fact is that the $AFNS_3$ and $AFNS_3 - CA$ models have a slightly different curvature factor loading compared to the other models; the hump on the factor loading is placed on shorter maturities compared to the other models. This is likely to be due to the positivity constraint of all three factors.

Another interesting finding is that the Feller condition is binding for the slope factor in the $AFNS_3$ model, implying a limited flexibility for this model class. The $AFNS_3$ model with a completely affine risk premia specification implies a lower mean reversion level for the slope factor. As we are interested in long-term forecasts, the Feller condition could impact the results.

¹¹Parameters that ensure stationarity and admissibility are included, even if they are insignificant.

	<i>CIR - 2</i>	<i>AFNS₁ - L</i>	<i>AFNS₁ - C</i>	<i>AFNS₂ - LC</i>	<i>AFNS₂ - SC</i>	<i>AFNS₃</i>	<i>AFNS₃ - CA</i>
σ_{11}	0.0976 (0.0058)	0.0359 (0.0024)	1.0000 (-)	0.0301 (0.0017)	1.0000 (-)	0.0301 (0.0014)	0.0480 (0.0028)
σ_{12}	- (-)	- (-)	-0.2506 (0.0332)	- (-)	-0.0150 (0.0021)	- (-)	- (-)
σ_{13}	- (-)	- (-)	-0.0231 (0.0042)	- (-)	-0.0155 (0.0036)	- (-)	- (-)
σ_{21}	- (-)	-0.0286 (0.0070)	- (-)	-0.0366 (0.0060)	- (-)	- (-)	- (-)
σ_{22}	0.0358 (0.0016)	1.0000 (-)	1.0000 (-)	1.0000 (-)	0.0628 (0.0036)	0.0824 (0.0033)	0.0896 (0.0047)
σ_{23}	- (-)	- (-)	0.0000 (-)	-0.0149 (0.0041)	- (-)	- (-)	- (-)
σ_{31}	- (-)	-0.0421 (0.0147)	- (-)	- (-)	- (-)	- (-)	- (-)
σ_{32}	- (-)	-0.7979 (0.1473)	- (-)	- (-)	- (-)	- (-)	- (-)
σ_{33}	- (-)	1.0000 (-)	0.1546 (0.0144)	0.0822 (0.0065)	0.1289 (0.0111)	0.1482 (0.0099)	0.1409 (0.0111)
α_1	- (-)	- (-)	0.0001 (0.0000)	- (-)	0.0000 (0.0000)	- (-)	- (-)
α_2	- (-)	0.0000 (-)	0.0003 (0.0000)	0.0000 (-)	- (-)	- (-)	- (-)
α_3	- (-)	0.0000 (-)	- (-)	- (-)	- (-)	- (-)	- (-)
β_{11}	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
β_{12}	- (-)	- (-)	- (-)	- (-)	0.0000 (-)	- (-)	- (-)
β_{13}	- (-)	- (-)	0.0000 (-)	- (-)	0.0000 (-)	- (-)	- (-)
β_{21}	- (-)	0.0047 (0.0005)	- (-)	0.0041 (0.0004)	- (-)	- (-)	- (-)
β_{22}	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
β_{23}	- (-)	- (-)	0.0000 (-)	0.0000 (-)	- (-)	- (-)	- (-)
β_{31}	- (-)	0.0139 (0.0018)	- (-)	- (-)	- (-)	- (-)	- (-)
β_{32}	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
β_{33}	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)

Table 4.2: **Parameter Estimates (1)**: Estimates of the diffusion parameters in the 2 factor CIR and Affine Nelson-Siegel models. Standard errors based on MCMC draws are reported in brackets. Estimates are based on the full data sample. Parameters without standard errors are fixed at the reported value.

	<i>CIR</i> – 2	<i>AFNS</i> ₁ – <i>L</i>	<i>AFNS</i> ₁ – <i>C</i>	<i>AFNS</i> ₂ – <i>LC</i>	<i>AFNS</i> ₂ – <i>SC</i>	<i>AFNS</i> ₃	<i>AFNS</i> ₃ – <i>CA</i>
λ	- (-)	0.4569 (0.0230)	0.3756 (0.0118)	0.5108 (0.0231)	0.3969 (0.0138)	0.6312 (0.0307)	0.6260 (0.0399)
κ_{11}^Q	0.5622 (0.0250)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
κ_{22}^Q	0.0001 (0.0001)	- (-)	- (-)	- (-)	- (-)	- (-)	- (-)
θ/θ_1	0.0140 (0.0007)	0.0017 (0.0001)	0.0454 (0.0018)	0.0014 (0.0001)	0.0031 (0.0010)	0.0015 (0.0001)	0.0019 (0.0002)
θ_2	0.0022 (0.0001)	- (-)	- (-)	0.0748 (0.0040)	0.0398 (0.0017)	0.0036 (0.0003)	0.0005 (0.0004)
θ_3	- (-)	- (-)	- (-)	- (-)	- (-)	0.0171 (0.0010)	0.0186 (0.0011)
θ_1^P	- (-)	- (-)	-0.0441 (0.0100)	- (-)	-0.0359 (0.0061)	- (-)	- (-)
θ_2^P	- (-)	0.0000 (-)	0.0000 (-)	-0.1086 (0.0269)	0.0241 (0.0107)	0.0045 (0.001)	- (-)
θ_3^P	- (-)	0.0000 (-)	0.1202 (0.0348)	0.1314 (0.0402)	0.0720 (0.021)	0.0152 (0.0032)	- (-)
κ_{11}^P	0.5729 (0.1311)	0.0894 (0.0272)	0.0634 (0.0425)	0.0743 (0.0254)	0.0548 (0.0397)	0.1814 (0.0365)	0.1923 (0.0590)
κ_{12}^P	- (-)	- (-)	0.0000 (-)	- (-)	0.0000 (-)	- (-)	- (-)
κ_{13}^P	- (-)	- (-)	-0.5128 (0.1017)	- (-)	-0.4477 (0.0847)	- (-)	- (-)
κ_{21}^P	- (-)	0.0000 (-)	0.0000 (-)	0.0000 (-)	- (-)	0.0000 (-)	0.0000 (-)
κ_{22}^P	0.1995 (0.0431)	0.3531 (0.1536)	0.1575 (0.0958)	0.4070 (0.1629)	0.2920 (0.1346)	0.1888 (0.0991)	0.1046 (0.0747)
κ_{23}^P	- (-)	-0.6171 (0.1668)	0.0000 (-)	-0.7673 (0.1911)	0.0000 (-)	0.0000 (-)	0.0000 (-)
κ_{31}^P	- (-)	- (-)	- (-)	0.0000 (-)	- (-)	0.0000 (-)	0.0000 (-)
κ_{32}^P	- (-)	- (-)	- (-)	- (-)	0.0000 (-)	0.0000 (-)	0.0000 (-)
κ_{33}^P	- (-)	0.8009 (0.3124)	1.5029 (0.4352)	0.984 (0.3033)	1.0614 (0.3193)	0.9027 (0.2825)	0.9216 (0.2130)
σ_ε	0.0032 (0.0001)	0.0020 (0.0001)	0.0021 (0.0001)	0.0020 (0.0000)	0.0021 (0.0001)	0.0025 (0.0001)	0.0027 (0.0001)

Table 4.3: **Parameter Estimates (2)**: Estimates of the drift parameters and the measurement error standard deviation in the 2 factor CIR and Affine Nelson-Siegel models. Standard errors based on MCMC draws are reported in brackets. Estimates are based on the full data sample. Parameters without standard errors are fixed at the reported value.

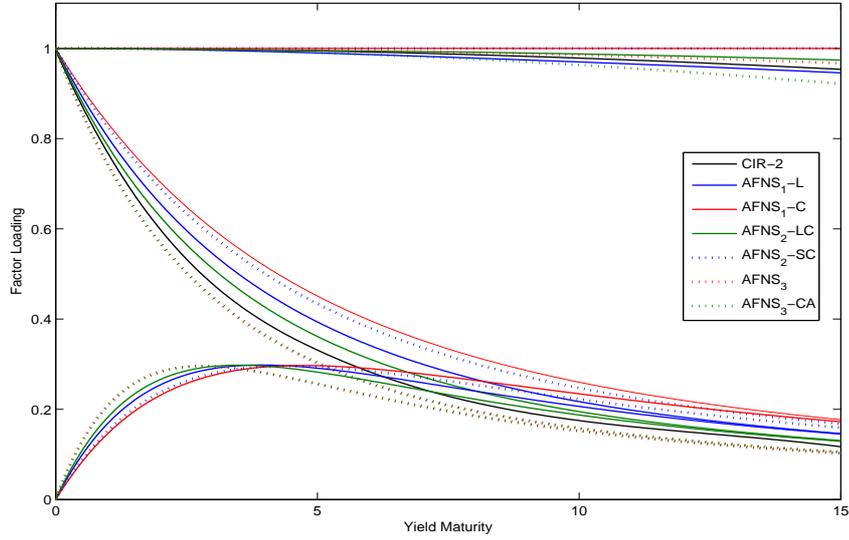


Figure 4.4: Factor loadings generated by the different models.

	3M	2Y	5Y	10Y	15Y
<i>CIR - 2</i>	34.49	32.15	29.90	23.15	38.07
<i>AFNS₁ - L</i>	20.33	20.77	17.33	16.36	22.29
<i>AFNS₁ - C</i>	23.49	22.41	17.68	17.09	22.93
<i>AFNS₂ - LC</i>	20.58	20.20	17.77	16.05	22.88
<i>AFNS₂ - SC</i>	23.48	23.16	17.80	17.80	23.75
<i>AFNS₃</i>	26.36	22.86	23.47	19.95	32.28
<i>AFNS₃ - CA</i>	27.45	22.89	24.32	20.10	35.10

Table 4.4: In-sample Root Mean Squared Errors (RMSE) based on the different models. RMSEs are reported in basis points.

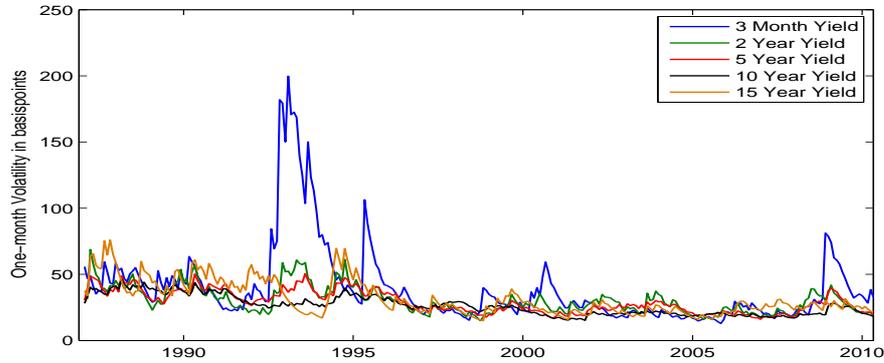


Figure 4.5: **E-GARCH estimates of one-month volatilities.** Volatilities are measured in basis points

Table 4.4 shows Root Mean Squared Errors (RMSEs) for the considered models. Obviously, since these figures are in-sample figures, the model with the least number of factors (the 2 factor CIR) performs the worst. Interestingly, the $AFNS_3$ models perform the worst after the CIR model. One explanation could be that the positivity of the factors limit the in-sample flexibility of the model. Finally, the models where the level enters in the volatility matrix, perform the best.

Next we examine the model’s ability to generate the correct volatility pattern. To assess this issue we estimate an E-GARCH model to yield changes, for each maturity (h):

$$x_t^h \equiv Y(t + 1, t + 1 + h) - Y(t, t + h) = \epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, \sigma_{t,h}^2)$$

$$\log \sigma_{t,h}^2 = \omega + \theta z_{t-1} + \gamma \left(|z_t| - \sqrt{2/\pi} \right) + \beta \log \sigma_{t-1,h}^2$$

where $z_t = \epsilon_t / \sigma_{t-1,h}$ is the standardized innovation.

Figure 4.5 shows the estimated volatilities. It is evident that the short term (especially the 3-month yield) shows the highest degree of variation. Especially, the ERM-crisis shows up in the data, as well as the impact of the financial crisis in 2008. This downward trend in volatility, as a function of maturity, is consistent with the findings in Christensen, Lopez, and Rudebusch (2010) and Jacobs and Karoui (2009).

Figure 4.6 shows the E-GARCH volatilities along with model generated volatilities. Table 4.5 shows the correlations between E-GARCH volatilities and the correlations between changes in E-GARCH volatilities and changes

Levels	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	0.5168	0.4322	0.6728	0.8378	0.7706
<i>AFNS</i> ₁ – <i>L</i>	0.301	0.5527	0.7555	0.8683	0.7524
<i>AFNS</i> ₁ – <i>C</i>	-0.4556	-0.0968	-0.2408	-0.2079	0.0004
<i>AFNS</i> ₂ – <i>LC</i>	0.2813	0.5556	0.7189	0.8707	0.7828
<i>AFNS</i> ₂ – <i>SC</i>	0.4406	0.1405	-0.2258	-0.216	-0.0079
<i>AFNS</i> ₃	0.5118	0.4575	0.5404	0.7965	0.7879
<i>AFNS</i> ₃ – <i>CA</i>	0.5568	0.5793	0.7	0.851	0.7728
Changes	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	0.1678	0.0726	0.0965	0.0909	0.0011
<i>AFNS</i> ₁ – <i>L</i>	0.0589	-0.0765	0.0379	0.087	-0.0278
<i>AFNS</i> ₁ – <i>C</i>	-0.0273	-0.103	-0.0085	-0.0751	-0.0782
<i>AFNS</i> ₂ – <i>LC</i>	0.0489	-0.1202	0.0122	0.016	-0.0831
<i>AFNS</i> ₂ – <i>SC</i>	0.1131	0.107	0.0401	-0.046	-0.0599
<i>AFNS</i> ₃	0.1699	0.0588	0.0899	0.0522	0.0021
<i>AFNS</i> ₃ – <i>CA</i>	0.1654	0.0555	0.082	0.0625	-0.043

Table 4.5: **Correlations between model volatilities and E-GARCH estimates for the different maturities.**

in model volatilities.¹² In Christensen, Lopez, and Rudebusch (2010) the preferred model is the *AFNS*₃ model. This model, together with the *CIR*–2 and the *AFNS*₃–*CA* models, also performs well in our data sample, both by visually inspecting Figure 4.6 and by showing the highest correlations both for levels and changes. The correlations for changes are moderate. However, compared to the results from Collin-Dufresne, Goldstein, and Jones (2008), where the three factor affine model with spanned stochastic volatility shows negative correlations with GARCH estimates, our results are slightly better. The fact that the 2 factor CIR model also performs well, should not come as a great surprise as it is very similar to the *AFNS*₃ and *AFNS*₃–*CA* models.

Another model that performs well is the *AFNS*₁–*L*. It cannot generate the volatility of the 3-month yield, but performs similar to the *AFNS*₃ and *CIR*–2 models for longer maturities. Finally, the remaining models show little, or even negative correlations with the E-GARCH estimate. Finally,

¹²The model generated volatility is calculated as:

$$\text{Vol}_t(y(t+1, t+h+1)) = \sqrt{\frac{1}{h^2} B(h)' SD(X_t) D(X_t) S' B(h) \Delta}$$

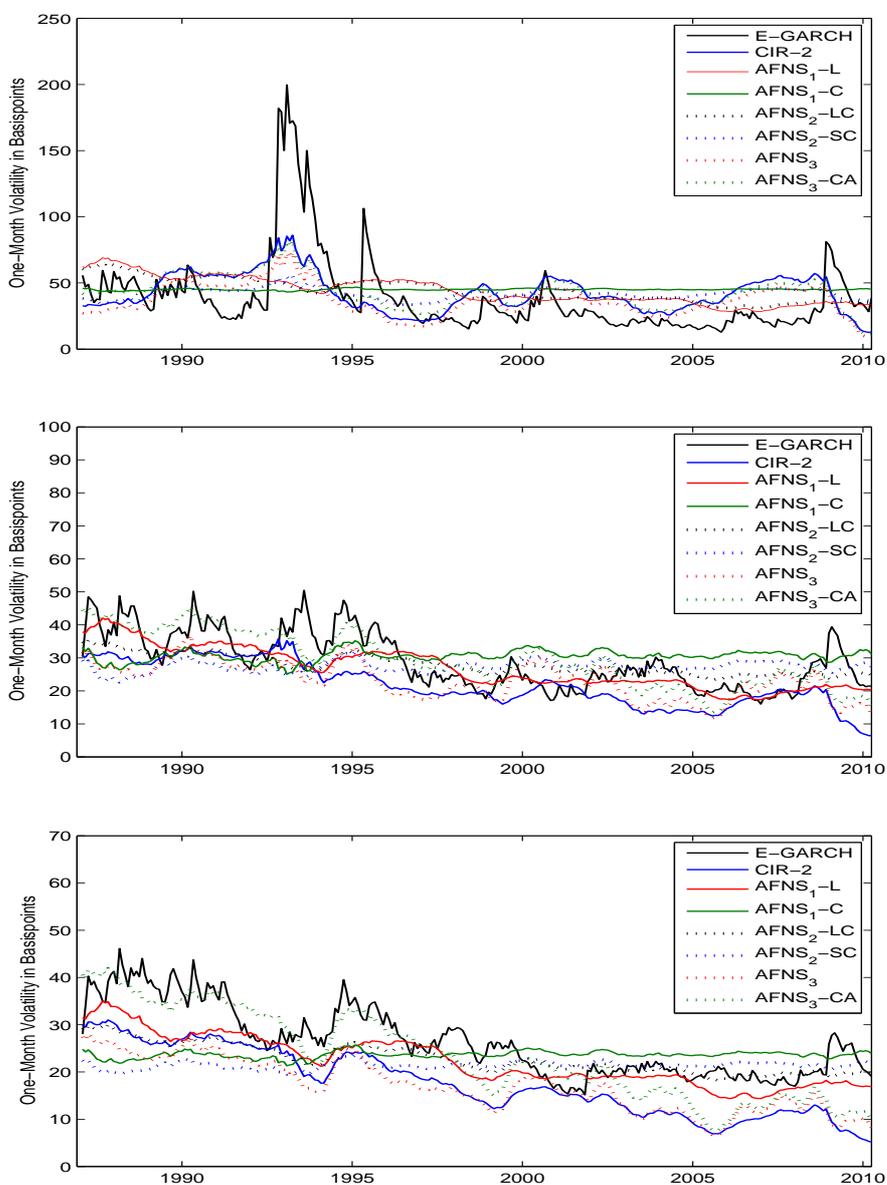


Figure 4.6: **Top:** Comparison of model volatility to E-GARCH volatility for the 3 Month maturity. **Middle:** Comparison of model volatility to E-GARCH volatility for the 5 Year maturity. **Bottom:** Comparison of model volatility to E-GARCH volatility for the 10 Year maturity.

when considering changes in the volatility related to the 15-year yield, the results show little correlation with the E-GARCH estimate.

Overall, these findings are indications that the level factor offers a good description of the overall level of volatility. However, to capture the changes in interest-rate volatility, unspanned stochastic volatility might be needed.

4.8 Empirical Results: Out-of-Sample

This section considers the performance of each model with regard to out-of-sample testing. We first consider the evaluation of point forecasts, and then turn to evaluating forecasts of volatilities. For the sake of brevity tables are placed at the end of the paper.

Point Forecasts

Tables 4.6 to 4.8 present mean errors for the yield forecast 1 month, 1 year and 5 years ahead.

For forecasts 1 month ahead, most models perform well, as measured by the mean errors. The exception is the $AFNS_2 - SC$ -model which has mean errors significantly different from zero for all maturities except for the 3 month yield. Another observation is that for the 15 year yield, the models which have all factors in the variance matrix, i.e. $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$, are also biased, implying that the limitations in structure induced by these models limit the forecast of long-term yields.

The same pattern is visible for the 1-year forecast. The $AFNS_2 - SC$ -model has a highly significant bias for all yields, and the $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$ models show a bias for the 15-year yield. For the $AFNS_2 - SC$ -model the errors are over 100 basis points, indicating a very bad model. The level of the term structure has mostly been downward trending, cf. Figure 4.2, and in the $AFNS_2 - SC$ -model the level factor is mostly negative due to the positivity restrictions on slope and curvature. All in all this implies that the level factor mean-reverts to a negative value, giving yield curves with negative yields! For the $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$ models, errors are more moderate and in the range of 20-40 basis points, indicating that the positivity constraint of the factors, in combination with the downward trending interest-rate level, may induce forecasts that are too high.

For the 5-year forecasts we see the same pattern as above. The $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$ models show a bias of around 80-150 basis points

on the 5-year forecast horizon, with all but the 3-month yield being significantly different from zero. Again this seems consistent with the downward trending term structure and the model restrictions. On the other hand, the models that performed well for the shorter maturities, still has mean errors that are not statistically different from zero. However, mean errors are now around 10-50 basis points.

So far we have only considered mean errors. This measure only gives information on the overall mean errors, not the variation and size at each point in time. Therefore, we also consider absolute errors.¹³

Tables 4.9 to 4.11 show mean absolute errors for each maturity and model, along with pairwise comparisons of the models using the Diebold and Mariano (1995) test. This test allows more precise evidence on the model differences, than just comparing the mean absolute errors.

For the one-month forecasts the mean absolute errors are for all the considered models between 15 and 30 basis points. Still the $AFNS_2 - SC$ model is significantly outperformed by most of the other models, and the $AFNS_3$ and $AFNS_3 - CA$ models are outperformed for long-term yields. Furthermore, the $CIR - 2$ model is outperformed for all but the 3-month yield.

For the one-year forecasts the mean absolute errors are very similar for all models, except the $AFNS_2 - SC$ model. The mean absolute errors are between 170 and 280 basis points for the $AFNS_2 - SC$ model, whereas for the other models the mean absolute errors are between 55 and 105 basis points. In terms of statistically significant out-performance only the $AFNS_2 - SC$ model is outperformed.

In terms of the 5-year forecast, the $AFNS_2 - SC$ model again performs very poorly, as the absolute mean errors are around 700 basis points. For the other models the $AFNS_1 - L$ and $AFNS_2 - LC$ model perform the best, with absolute mean errors between 50 and 130 basis points. For the $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$ models the mean absolute errors are between 105 and 150, although with a slightly higher mean absolute error for the $CIR - 2$ model. In terms of statistical significance, the $AFNS_1 - L$ and $AFNS_2 - LC$ models outperform the other models. For the shorter maturities, the $AFNS_3$ and $AFNS_3 - CA$ models perform comparable to these two models.

All in all, based on point forecasts we would rank the $AFNS_1 - L$ and $AFNS_2 - LC$ models as the best, followed by the $AFNS_3$ and $AFNS_3 - CA$

¹³The results presented below are robust to other choices of forecast errors, such as Root Mean Squared Errors.

models. *CIR-2* and *AFNS₁-C* are joint third and finally *AFNS₂-SC* is the worst model.

Variance forecasts

We also consider how the models forecast variances or rather volatilities over 1 month, 1 year and 5 year time horizons.¹⁴

The existing literature on variance forecasting for affine models is mainly focused on short-term forecasts, i.e. one week or one month (see Collin-Dufresne, Goldstein, and Jones (2008), Jacobs and Karoui (2009) and Christensen, Lopez, and Rudebusch (2010)). Here, in addition to one month forecasts, we also perform 1 year and 5 year forecasts.

To construct a measure of realized volatility, we consider taking the square root of squared (de-meanned) yield changes:¹⁵

$$RV(t, \tau, n) = \sqrt{\sum_{t_j=t}^{t+n} (\Delta y(t_j, t_j + \tau) - \mu_y)^2}$$

where t is the point in time, τ is the maturity of the yield, n is the number of months in the forecast horizon and μ_y is the mean of the yield changes (see Table 4.1). We construct the realized volatility for each forecast horizon (i.e. $n = 1$, $n = 12$ and $n = 60$) and each yield.

To generate the model-implied volatilities, we use an Euler approximation of the state dynamics, with each month being split into 25 steps. We generate 25,000 samples and use the standard deviation of the 25,000 samples as our volatility estimate. We construct the errors as the difference between the realized volatilities and the model based volatility:

$$\hat{\varepsilon}_t = RV(t, \tau, n) - MV(t, \tau, n)$$

where MV is the model based volatility.

Figure 4.7 shows the realized volatility along with the model based volatilities. In general, there is some indication that the model capture the level of

¹⁴In Appendix 4.12 we also describe density and quantile forecasts for all the considered models. We have chosen to place these results in an appendix as they are slightly harder to interpret compared to mean and variance forecasts. The results in Appendix 4.12 are consistent with the results from the mean and variance forecasts.

¹⁵We have also considered an E-GARCH measure. The results are consistent with the results in this section, albeit with a slightly better performance for the *AFNS₁-L*-model in the 1-month forecasts.

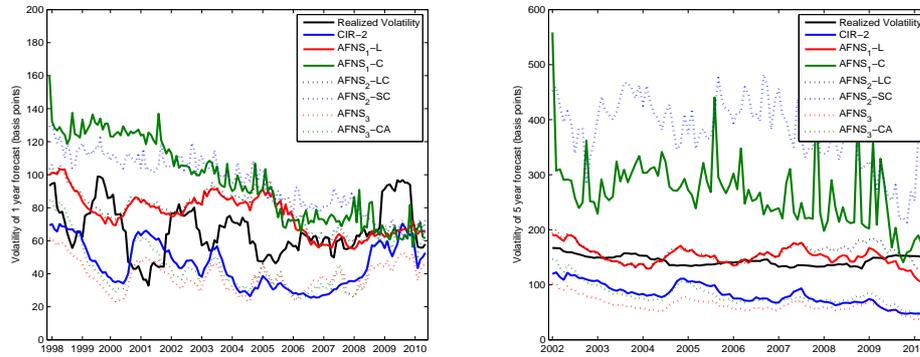


Figure 4.7: **Left:** Realized volatility and model volatility for the 1-year forecast of the 10-year yield. **Right:** Realized volatility and model volatility for the 5-year forecast of the 10-year yield.

volatilities rather than the dynamics. However, for 1 and 5-year forecasts, it is highly unlikely that most models would generate the correct volatility dynamics.

Tables 4.12 to 4.14 present mean errors from the volatility forecasting for each forecast horizon. Here, positive mean errors correspond to underestimating volatility and *vice versa*.

For the 1-month forecasts the mean errors are quite small, between 1 and 4 basis points. The exceptions are mainly found in the 15-year yield volatility, where the $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$ models underestimate the volatility. As with the mean forecasts, the $AFNS_2 - SC$ model performs badly.

In terms of the 1-year forecast it is only the $AFNS_1 - L$ which can show unbiased forecasts. Though the forecasts are unbiased, the volatilities are on average overestimated by around 10 basis points. For some maturities the $AFNS_2 - LC$ model can also perform unbiased forecasts. The bias for this model is between 6 and 18 basis points. The CIR-based models ($CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$) typically underestimate the volatility. The bias is between 10 and 30 basis points. The $AFNS_2 - SC$ overestimates volatilities by 7 to 29 basis points.

The $AFNS_1 - L$ model also performs the best for the 5-year volatility forecast. Forecasts are statistically unbiased, with mean errors between 6 and 21 basis points. Similar to the other forecast-horizons, the $AFNS_1 - L$ also perform well. The CIR based models typically underestimate volatility by 60 to 120 basis points, with the $CIR - 2$ model performing the best.

Again the $AFNS_2 - SC$ model is the worst, with an overestimation of volatility of about 200 basis points!

To directly compare the performance of the model, we perform Diebold and Mariano (1995) tests on mean absolute errors. Tables 4.15 to 4.17 present the results from the tests, along with mean absolute errors for the different models. The mean absolute errors show the same pattern as Tables 4.12 to 4.14, with small modifications for the 1-month forecasts (see below).

In terms of the 1-month forecast the models that perform the best are the $AFNS_3$ and $AFNS_3 - CA$ models. They show the smallest mean absolute errors, and for most maturities outperform the other models. With respect to the remaining model they are mostly at par; however, with a large mean absolute error in the 3 month yield for the $CIR - 2$ model.

Considering the 1-year forecast we see that the $AFNS_1 - L$ model performs the best, although it does not statistically outperform the CIR-based models and the $AFNS_1 - LC$ models, except for a few maturities. As previously the $AFNS_2 - SC$ performs the worst.

The differences in the models are very visible when considering the 5-year forecast. The $AFNS_1 - L$ and $AFNS_2 - LC$ models perform the best, generally outperform all the other models, except each other. The $CIR - 2$ model performs the third-best, followed by the $AFNS_3$ and $AFNS_3 - CA$ models. Interestingly, the $AFNS_3 - CA$ outperform the $AFNS_3$ indicating that the Feller condition is limiting the model. The $AFNS_2 - SC$ again shows a very bad performance with mean absolute errors over 200 basis points!

To shed some more light on the differences between the models, we plot the densities of the 10-year yield, when forecasting 1 month, 1 year and 5 years ahead. The states and parameters are based on the last data in our sample, i.e. May 2005.

It is evident that the CIR-based models perform alike; however, with more probability toward higher yields in the $CIR - 2$ model. The $AFNS_1 - L$ and $AFNS_2 - SC$ model show densities that are closer to normal distributions, although with significant skew in the 5-year forecast distribution.

In general for the 1-month forecast the densities look fairly similar, with a slightly wider distribution for the $AFNS_1 - L$ model. With respect to the $AFNS_2 - SC$ model we re-confirm that both mean and variance are biased, which is even more pronounced for the 1 and 5-year forecasts.

In terms of the 1 and 5-year forecasts we see that the CIR-based model could be limited by the fact that the factors need to be positive. The lowest possible value for the yield appears to be the same for the 1 and

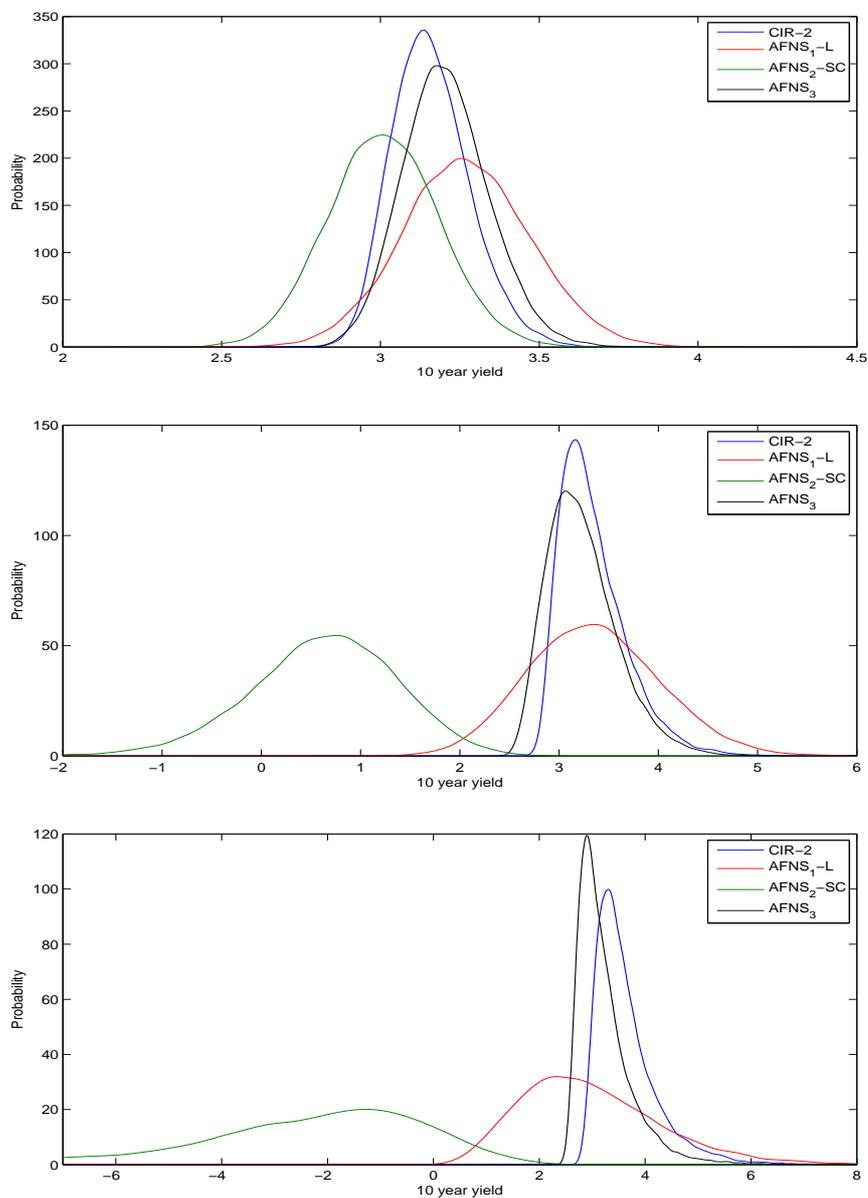


Figure 4.8: **Top:** 1-month forecast of the 10-year yield performed in May 2010. **Middle:** 1-year forecast of the 10-year yield performed in May 2010. **Bottom:** 5-year forecast of the 10-year yield performed in May 2010.

5-year forecast, whereas there is more flexibility with respect to the upper tail. The $AFNS_1 - L$ model on the other hand shows more flexibility in both tails of the distribution, and significantly wider distributions.

Overall Figure 4.8 re-confirms the results from the volatility forecasts, i.e. that the $AFNS_1 - L$ slightly overestimates the volatility, where the CIR-based model underestimates the volatility. Results based on probability density forecasts, given in Appendix 4.12, also indicate that the CIR-based models have troubles capturing the lower tails of the distributions, where the $AFNS_1 - L$ has a tendency to overestimate the risks in the tails of the distribution. These results emphasize that rather than using a single model, a suite of models is more appropriate. For instance the $AFNS_1 - L$ could be used as a slightly pessimistic risk estimate in VaR-calculations.

4.9 Conclusion

In this paper we have examined the stochastic volatility Affine Nelson-Siegel models of Christensen, Lopez, and Rudebusch (2010), and their ability to forecast interest-rates and quantiles of the distribution of interest-rates on Danish government bond term structure data.

We have found that the tested models cannot capture all the desired features of an interest-rate model.

In terms of forecasting the realized interest-rates (i.e. point estimates), models with a combination of CIR-factors and Gaussian factors perform well, especially if the level-factor generates the stochastic volatility.

Models based on CIR-factors ($AFNS_3$ and $CIR - 2$) perform well in forecasting the mean and variance in the short-term, i.e. one month forecasts. These results are in line with the results in Christensen, Lopez, and Rudebusch (2010).

For longer forecasts we find a mixed picture, which is to some extent driven by the considered data. The preferred model of Christensen, Lopez, and Rudebusch (2010) tends to produce forecasts that are too high and variances that are too low, implying a bias in model quantiles. The one- and two-factor models, where the level factor drives stochastic volatility, perform reasonably well in forecasting the mean and variances of the distributions, but for very long horizon forecasts, 5 years, the distributions tend to be wide compared to actual data. Overall, this indicates that a suite of models, rather than a single model, is preferable.

For future research it would be interesting to assess the long-term forecasting properties of models with unspanned stochastic volatility, as this type

of models might provide a better description of the volatility over the longer horizons.

4.10 Appendix: Affine Nelson-Siegel models with stochastic volatility

Models with one stochastic volatility factor

AFSN₁ – L

The first affine Nelson-Siegel model, has the level factor, X_1 , as the driving factor of the interest-rate volatility. The model evolves according to the following system of stochastic differential equations under the risk neutral martingale measure Q :

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & 1 & 0 \\ \sigma_{31} & \sigma_{32} & 1 \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21}X_1(t)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31}X_1(t)} \end{pmatrix}$$

X_1 acts as the level factor, and in this model class, it also drives the stochastic volatility. β_{21} and β_{31} measure the volatility sensitivity of the slope and curvature factors. Finally $\varepsilon = 10^{-6}$ to approximate the unit root behaviour in the Gaussian affine Nelson-Siegel model.

In this model class the factor loadings on slope and curvature are given by

$$B_2(t, T) = - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right) \\ B_3(t, T) = (T - t)e^{-\lambda(T-t)} - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right)$$

and the factor loading on the level factor solves the ODE

$$\frac{dB_1(t, T)}{dt} = 1 + \varepsilon B_1(t, T) - \frac{1}{2} [H_1^2 + H_2^2\beta_{21} + H_3^2\beta_{31}]$$

where

$$H_1 = \sigma_{11}B_1(t, T) + \sigma_{21}B_2(t, T) + \sigma_{31}B_3(t, T) \\ H_2 = B_2(t, T) + \sigma_{32}B_3(t, T) \\ H_3 = B_3(t, T)$$

The yield adjustment term, $A(t, T)$ solves the ODE

$$\frac{dA(t, T)}{dt} = -B_1(t, T)\theta - \frac{1}{2} [H_2^2\alpha_2 + H_3^2\alpha_3]$$

Finally to complete the model, we need to specify the risk premia. For this model class we consider the essentially affine risk premia introduced by Duffee (2002).¹⁶ The P dynamics are then given by

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & 1 & 0 \\ \sigma_{31} & \sigma_{32} & 1 \end{pmatrix} \begin{pmatrix} dW_1^P(t) \\ dW_2^P(t) \\ dW_3^P(t) \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21}X_1(t)} & 0 \\ 0 & 0 & \sqrt{\alpha_3 + \beta_{31}X_1(t)} \end{pmatrix}$$

AFSN₁ – C

The second affine Nelson-Siegel model, has the curvature factor, X_3 , as the driving factor of the interest-rate volatility.¹⁷ The model evolves according to the following system of stochastic differential equations under the risk neutral martingale measure Q :

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} 0 \\ -\theta \\ \theta \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ 0 & 1 & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{\alpha_1 + \beta_{13}X_3(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{23}X_3(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix}$$

X_3 acts as the curvature factor, and in this model class, it also drives the stochastic volatility. β_{13} and β_{23} measure the volatility sensitivity of the level and slope factors.

In this model class the factor loadings on level and slope are given by

$$B_1(t, T) = - (T - t) \\ B_2(t, T) = - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right)$$

and the factor loading on the level factor solves the ODE

$$\frac{dB_3(t, T)}{dt} = \lambda [B_3(t, T) - B_2(t, T)] - \frac{1}{2} [H_1^2\beta_{13} + H_2^2\beta_{23} + H_3^2]$$

¹⁶Following Christensen, Lopez, and Rudebusch (2010) we do not consider the extended affine specification for the *AFSN₁ – L* model class, as we cannot expect the feller condition to hold under the risk neutral measure.

¹⁷The slope factor cannot drive stochastic volatility in the models with one stochastic volatility factor due to model admissibility, see Christensen, Lopez, and Rudebusch (2010) for further details.

where

$$\begin{aligned} H_1 &= B_1(t, T) \\ H_2 &= \sigma_{12} B_1(t, T) + B_2(t, T) \\ H_3 &= \sigma_{13} B_1(t, T) + \sigma_{23} B_2(t, T) + \sigma_{33} B_3(t, T) \end{aligned}$$

The yield adjustment term, $A(t, T)$ solves the ODE

$$\frac{dA(t, T)}{dt} = \theta [B_2(t, T) - B_3(t, T)] - \frac{1}{2} [H_1^2 \alpha_1 + H_2^2 \alpha_2]$$

Finally we specify risk premia as an extended affine risk premia. This implies the P dynamics

$$\begin{aligned} \begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} &= \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ 0 & 0 & \kappa_{33}^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ 0 & 1 & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{\alpha_1 + \beta_{13} X_3(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{23} X_3(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix} \end{aligned}$$

For the Feller condition to hold, the following conditions must hold

$$\theta > \frac{1}{2} \sigma_{33}^2, \quad \theta_3^P > \frac{1}{2} \sigma_{33}^2$$

Models with two stochastic volatility factors

AFSN₂ – LC

The third Nelson-Siegel model has two stochastic volatility factors, namely that volatility is driven by level and curvature. The model class follows the stochastic differential equation under the risk neutral measure Q :

$$\begin{aligned} \begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} &= \left[\begin{pmatrix} \theta_1 \\ -\theta_2 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & 1 & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} X_1(t) + \beta_{23} X_3(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix} \end{aligned}$$

In this model class volatility is driven by level and curvature, and β_{21} and β_{23} measure the effect of these two factors on the volatility of the slope factor.

In this model class the factor loading on the slope is given by

$$B_2(t, T) = - \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda} \right)$$

and the loadings on level and curvature solve the ODEs

$$\begin{aligned} \frac{dB_1(t, T)}{dt} &= 1 + \varepsilon B_1(t, T) - \frac{1}{2} [H_1^2 + H_2^2 \beta_{21}] \\ \frac{dB_3(t, T)}{dt} &= \lambda [B_3(t, T) - B_2(t, T)] - \frac{1}{2} [H_2^2 \beta_{23} + H_3^2] \end{aligned}$$

where

$$\begin{aligned} H_1 &= \sigma_{11} B_1(t, T) + \sigma_{21} B_2(t, T) \\ H_2 &= B_2(t, T) \\ H_3 &= \sigma_{23} B_2(t, T) + \sigma_{33} B_3(t, T) \end{aligned}$$

The yield adjustment term, $A(t, T)$, solves the ODE

$$\frac{dA(t, T)}{dt} = -\theta_1 B_1(t, T) + \theta_2 [B_2(t, T) - B_3(t, T)] - \frac{1}{2} H_2^2 \alpha_2$$

Finally the risk premia is a combination of an essentially and extended affine risk premia specification. Thus we have the following P -dynamics.

$$\begin{aligned} \begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} &= \left[\begin{pmatrix} \theta_1 \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & 0 & \kappa_{33}^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & 1 & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ &\times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{\alpha_2 + \beta_{21} X_1(t) + \beta_{23} X_3(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix} \end{aligned}$$

where the Feller condition must hold for X_3 , i.e.

$$\theta > \frac{1}{2} \sigma_{33}^2, \quad \theta_3^P > \frac{1}{2} \sigma_{33}^2$$

and for model admissibility we must have that $\kappa_{31}^P \leq 0$.

AFSN₂ – SC

The third Nelson-Siegel model has two stochastic volatility factors, namely that volatility is driven by slope and curvature. The model class follows the

stochastic differential equation under the risk neutral measure Q :

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} 0 \\ \theta_1 \\ \theta_2 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{\alpha_1 + \beta_{12}X_2(t) + \beta_{13}X_3(t)} & 0 & 0 \\ 0 & \sqrt{X_2(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

In this model class volatility is driven by level and curvature, and β_{12} and β_{13} measure the effect of these two factors on the volatility of the level factor.

In this model class the factor loading on the slope is given by

$$B_1(t, T) = -(T - t)$$

and the loadings on level and curvature solve the ODEs

$$\begin{aligned} \frac{dB_2(t, T)}{dt} &= 1 + \lambda B_2(t, T) - \frac{1}{2} [H_1^2 \beta_{12} + H_2^2] \\ \frac{dB_3(t, T)}{dt} &= \lambda [B_3(t, T) - B_2(t, T)] - \frac{1}{2} [H_1^2 \beta_{13} + H_3^2] \end{aligned}$$

where

$$\begin{aligned} H_1 &= B_1(t, T) \\ H_2 &= \sigma_{12} B_1(t, T) + \sigma_{22} B_2(t, T) \\ H_3 &= \sigma_{13} B_1(t, T) + \sigma_{33} B_3(t, T) \end{aligned}$$

The yield adjustment term, $A(t, T)$, solve the ODE

$$\frac{dA(t, T)}{dt} = -\theta_1 B_2(t, T) - \theta_2 B_3(t, T) - \frac{1}{2} H_1^2 \alpha_1$$

Finally we use an extended risk premia specification, giving us the P dynamics:

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ 0 & \kappa_{22}^P & \kappa_{23}^P \\ 0 & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} 1 & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{\alpha_1 + \beta_{12}X_2(t) + \beta_{13}X_3(t)} & 0 & 0 \\ 0 & \sqrt{X_2(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

For the Feller condition to hold, the following conditions must be satisfied

$$\theta_1 > \frac{1}{2}\sigma_{22}^2, \quad \theta_2 > \frac{1}{2}\sigma_{33}^2, \quad \theta_2^P > \frac{1}{2}\sigma_{22}^2, \quad \theta_3^P > \frac{1}{2}\sigma_{33}^2$$

For admissibility the following conditions must hold

$$\kappa_{23}^P \leq 0, \quad \kappa_{32}^P \leq 0$$

Models with three stochastic volatility factors

In the fifth Nelson-Siegel model the volatility is driven by all three factors. The model class follows the stochastic differential equation under the risk neutral measure Q :

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{pmatrix} - \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{X_2(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

In this model class the factor loadings are given as solutions to the ODEs

$$\begin{aligned} \frac{dB_1(t, T)}{dt} &= 1 + \varepsilon B_1(t, T) - \frac{1}{2}\sigma_{11}^2 B_1(t, T) \\ \frac{dB_2(t, T)}{dt} &= 1 + \lambda B_2(t, T) - \frac{1}{2}\sigma_{22}^2 B_2(t, T) \\ \frac{dB_3(t, T)}{dt} &= \lambda [B_3(t, T) - B_2(t, T)] - \frac{1}{2}\sigma_{33}^2 B_3(t, T) \end{aligned}$$

and the yield adjustment term is given as the solution to the ODE

$$\frac{dA(t, T)}{dt} = -\theta_1 B_1(t, T) - \theta_2 B_2(t, T) - \theta_3 B_3(t, T)$$

Finally the risk premia is a combination of an essentially and extended affine risk premia. This implies the P dynamics:

$$\begin{pmatrix} dX_1(t) \\ dX_2(t) \\ dX_3(t) \end{pmatrix} = \left[\begin{pmatrix} \theta_1 \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \begin{pmatrix} X_1(t) \\ X_2(t) \\ X_3(t) \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \\ \times \begin{pmatrix} \sqrt{X_1(t)} & 0 & 0 \\ 0 & \sqrt{X_2(t)} & 0 \\ 0 & 0 & \sqrt{X_3(t)} \end{pmatrix} \begin{pmatrix} dW_1^Q(t) \\ dW_2^Q(t) \\ dW_3^Q(t) \end{pmatrix}$$

The Feller condition holds when

$$\theta_2 > \frac{1}{2}\sigma_2^2, \quad \theta_3 > \frac{1}{2}\sigma_3^2, \quad \theta_2^P > \frac{1}{2}\sigma_2^2, \quad \theta_3^P > \frac{1}{2}\sigma_3^2$$

and for admissibility we must have that $\kappa_{23}^P \leq 0$ and $\kappa_{32}^P \leq 0$.

For the $AFNS_3$ we also estimate a version with a Completely Affine risk premia specification. This implies that $\theta_i^Q = \theta_i^P$ for all i . Furthermore, the Feller condition does not apply any more, as it is not needed to rule out arbitrage opportunities.

4.11 Appendix: MCMC details

Conditional distributions

The conditional $p(Y|X, \Theta)$

The conditional for the yield observations can be decomposed as

$$p(Y|X, \Theta) = \prod_{t=1}^T p(Y_t|X_t, \Theta)$$

where we can write $p(Y_t, X_t, \Theta)$ as

$$p(Y_t|X_t, \Theta) \propto (\sigma_\varepsilon^2)^{-K/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_\varepsilon^2} (e_t)'(e_t)\right)$$

where

$$e_t = Y_t + \frac{(A(t, t + \tau) + B(t, t + \tau)'X_t)}{\tau}$$

Thus we have that

$$p(Y|X, \Theta) \propto (\sigma_\varepsilon^2)^{-TK/2} \exp\left(-\frac{1}{2} \frac{1}{\sigma_\varepsilon^2} \sum_{t=1}^T (e_t)'(e_t)\right)$$

The conditional $p(X|\Theta)$

Using an Euler scheme we can write the dynamics as

$$X_{t+\Delta} = X_t + [\Theta^P - \kappa^P X_t] \Delta + SD(X_t) \sqrt{\Delta} u_t$$

where $u_t \sim \mathcal{N}(0, I_3)$.

The conditional for the state transitions can be decomposed as

$$p(X|\Theta) = \left(\prod_{t=1}^T p(X_t|X_{t-1}, \Theta) \right) p(X_0) \propto \prod_{t=1}^T p(X_t|X_{t-1}, \Theta)$$

where the last 'proportional to' follows from independent priors. The conditional $p(X_t|X_{t-1}, \Theta)$ can be written as

$$p(X_t|X_{t-1}, \Theta) \propto \frac{1}{\det(\Sigma(X_t))} \exp\left(-\frac{1}{2}v_t'\Sigma(X_t)^{-1}v_t\right)$$

where

$$\begin{aligned} \Sigma(X_t) &= \Delta S D(X_t) D(X_t) S' \\ v_t &= X_{t+\Delta} - X_t - [\Theta^P - \kappa^P X_t] \Delta \end{aligned}$$

Sampling parameters and states

Sampling σ_ε

To sample σ_ε we use that

$$p(\sigma_\varepsilon|\Theta_{\setminus\sigma_\varepsilon}, X, Y) \propto p(Y|\Theta, X)p(\sigma_\varepsilon) \propto p(Y|\Theta, X)$$

where the last proportional to follows from independent priors.

We can then use Gibbs sampling, to sample σ_ε

$$\sigma_\varepsilon^2 \sim \mathcal{IG}\left(\frac{TK}{2} + 1, \frac{1}{2} \sum_{t=1}^T (e_t)'(e_t)\right)$$

Sampling Θ^Q

To sample Θ^Q we use a random walk Metropolis-Hastings algorithm. Hence we propose a new value as

$$\Theta^{Q*} = \Theta^Q + \epsilon$$

where ϵ is a zero mean Normally distributed variable with a variance that needs calibration.

The acceptance probability follows from

$$p(\Theta^{Q*}|\Theta_{\setminus Q}, X, Y) \propto p(Y|\Theta^*, X)p(\Theta^{Q*}) \propto p(Y|\Theta, X)$$

which gives us the acceptance probability

$$\alpha = \min\left(\frac{p(Y|\Theta^*, X)}{p(Y|\Theta, X)}, 1\right)$$

Sampling Θ^P

To sample Θ^P we use a random walk Metropolis-Hastings algorithm. Hence we propose a new value as

$$\Theta^{P*} = \Theta^P + \epsilon$$

where ϵ is a zero mean Normally distributed variable with a variance that needs calibration.

The acceptance probability follows from

$$p(\Theta^{P*} | \Theta_{\setminus P}, X, Y) \propto p(X | \Theta^{P*}) p(\Theta^{P*}) \propto p(X | \Theta)$$

which gives us the acceptance probability

$$\alpha = \min \left(\frac{p(X | \Theta^{P*})}{p(X | \Theta)}, 1 \right)$$

Sampling Θ^D

To sample Θ^D we use a random walk Metropolis-Hastings algorithm. Hence we propose a new value as

$$\Theta^{D*} = \Theta^D + \epsilon$$

where ϵ is a zero mean Normally distributed variable with a variance that needs calibration.

The acceptance probability follows from

$$p(\Theta^{D*} | \Theta_{\setminus D}, X, Y) \propto p(Y | \Theta^{D*}, X) p(X | \Theta^{D*}) p(\Theta^{D*}) \propto p(Y | \Theta, X) p(X | \Theta)$$

which gives us the acceptance probability

$$\alpha = \min \left(\frac{p(Y | \Theta^{D*}, X) p(X | \Theta^{D*})}{p(Y | \Theta, X) p(X | \Theta)}, 1 \right)$$

Sampling X

To sample X we use a random walk Metropolis-Hastings algorithm. Hence we propose a new value as

$$X^* = X + \epsilon$$

where ϵ is a zero mean Normally distributed variable with a variance that needs calibration.

The acceptance probability follows from

$$p(X|\Theta, Y) \propto p(Y|\Theta, X)p(X|\Theta)$$

which gives us the acceptance probability

$$\alpha = \min \left(\frac{p(Y|\Theta, X^*)p(X^*|\Theta)}{p(Y|\Theta, X)p(X|\Theta)}, 1 \right)$$

Efficient calculation of Inverses and Determinants

In the affine Nelson-Siegel models the variance of the transition equation is given by

$$\text{Var}_t(X_{t+\Delta}) \equiv \Sigma(X_t) = \Delta S D(X_t) D(X_t) S'$$

Then the inverse will be given by

$$\begin{aligned} \Sigma(X_t)^{-1} &= \frac{1}{\Delta} [S D(X_t) D(X_t) S']^{-1} \\ &= \frac{1}{\Delta} [D(X_t) S']^{-1} [S D(X_t)]^{-1} \\ &= \frac{1}{\Delta} (S')^{-1} D(X_t)^{-1} D(X_t)^{-1} S^{-1} \end{aligned}$$

Next in the conditional for X we need to calculate

$$\begin{aligned} v_t' \Sigma(X_t)^{-1} v_t &= \frac{1}{\Delta} [v_t' (S')^{-1} D(X_t)^{-1}] [D(X_t)^{-1} S^{-1} v_t] \\ &= \sum_{j=1}^3 \frac{[S^{-1} v_t]_j^2}{[D(X_t)]_{jj}^2 \Delta} \end{aligned}$$

Since $(S')^{-1} = (S^{-1})'$ and $D(X_t)$ is diagonal. In terms of implementation this is far more efficient than calculating the entire quadratic form. Furthermore typically S has a simple form, such that the inverse is simple as well, speeding up the computation.

We also need to calculate the determinant:

$$\begin{aligned} \det(\Sigma(X_t)) &= \Delta \det(S) \det(D(X_t)) \det(D(X_t)) \det(S') \\ &= \Delta \det(S)^2 \det(D(X_t))^2 \end{aligned}$$

where we have used that $\det(S) = \det(S')$. Since $D(X_t)$ is diagonal, we get that

$$\det(D(X_t)) \det(D(X_t)) = \prod_{j=1}^3 [D(X_t)]_{jj}^2$$

Again the determinant of S depends on the structure of S . As with the inverse, S has a simple structure which implies a simple determinant as well.

The following subsection presents the inverses and determinant for the model with one and two stochastic volatility factors. For the two factor CIR and the model with three stochastic volatility factors, these are not presented as they are trivial.

$AFNS_1 - L$

In the $AFNS_1 - L$ the inverse of S is given by

$$S^{-1} = \begin{pmatrix} \frac{1}{\sigma_{11}} & 0 & 0 \\ -\frac{\sigma_{21}}{\sigma_{11}} & 1 & 0 \\ \frac{\sigma_{21}\sigma_{32}-\sigma_{31}}{\sigma_{11}} & -\sigma_{32} & 1 \end{pmatrix}$$

and the determinant is given by

$$\det(S) = \sigma_{11}$$

$AFNS_1 - C$

In the $AFNS_1 - C$ the inverse of S is given by

$$S^{-1} = \begin{pmatrix} 1 & -\sigma_{12} & \frac{\sigma_{12}\sigma_{23}-\sigma_{13}}{\sigma_{33}} \\ 0 & 1 & -\frac{\sigma_{23}}{\sigma_{11}} \\ 0 & 0 & \frac{1}{\sigma_{33}} \end{pmatrix}$$

and the determinant is given by

$$\det(S) = \sigma_{33}$$

$AFNS_2 - L, C$

In the $AFNS_2 - L, C$ the inverse of S is given by

$$S^{-1} = \begin{pmatrix} \frac{1}{\sigma_{11}} & 0 & 0 \\ -\frac{\sigma_{21}}{\sigma_{11}} & 1 & -\frac{\sigma_{23}}{\sigma_{33}} \\ 0 & 0 & \frac{1}{\sigma_{33}} \end{pmatrix}$$

and the determinant is given by

$$\det(S) = \sigma_{11}\sigma_{33}$$

$AFNS_2 - S, C$

In the $AFNS_2 - S, C$ the inverse of S is given by

$$S^{-1} = \begin{pmatrix} 1 & -\frac{\sigma_{12}}{\sigma_{22}} & -\frac{\sigma_{13}}{\sigma_{33}} \\ 0 & \frac{1}{\sigma_{22}} & 0 \\ 0 & 0 & \frac{1}{\sigma_{33}} \end{pmatrix}$$

and the determinant is given by

$$\det(S) = \sigma_{22}\sigma_{33}$$

4.12 Appendix: Density Forecasts

In this appendix we focus on forecasting densities in affine models, rather than just means and variances.

Here, we consider the discrete form of an affine model with observations at time $t = 1, 2, \dots, T$. Using the state-space form of the model implies that

$$p(y_{t+n}(T+n)|\Omega_t) = p(X_{t+1}|X_t) \times \dots \times p(X_{t+n}|X_{t+n-1}) \times p(y_{t+n}(T+n)|X_{t+n})$$

It is evident that if one-step-ahead state predictions, $p(X_{t+1}|X_t)$, are correctly specified, then the model will perform well in forecasting yields over the long horizon.

To generate the densities we use an Euler approximation of the state dynamics, with each month being split into 25 steps. We generate 25,000 samples and use a kernel estimator to smooth the simulated density.

Evaluating density forecasts 1 month ahead

Since we want to evaluate the ability of the models in forecasting 1 month-ahead predictions of densities, we need a formal test. We consider the *Probability Integral Transform* (henceforth PIT), as suggested by Diebold, Gunther, and Tay (1998).

Consider a yield, y_t , with density $f(y_t|\Omega_{t-s})$, i.e. the density of the yield at time t , conditional on the information at time $t - s$. Define the probability integral transform, u_t , as

$$u_t = \int_{-\infty}^{y(t,T)} f(y|\Omega_{t-s}) dy$$

where $y(t, T)$ is the observed yield, with maturity $T - t$.

Using the probability integral transform, we obtain the realized quantiles in the conditional distribution. Consider a sequence of probability integral transforms, $\{u_t\}_{t=1}^T$, then two conditions must hold for the model to be well specified.

First, the probability integral transforms must follow a *Uniform* distribution over the unit interval, $[0, 1]$. Second, the sequence must be independent - the past values of the sequence cannot be used to forecast the next value.

To conduct a formal test, we consider the method of Berkowitz (2001). Consider the sequence $\{z_t\}_{t=1}^T$, where

$$z_t = \Phi^{-1}(u_t)$$

and $\Phi^{-1}(u)$ is the inverse of the standard normal distribution function. If $\{u_t\}_{t=1}^T$ form a uniform and independent sequence, then $\{z_t\}_{t=1}^T$ form a standard Normal and independent sequence.

Next step in the test of Berkowitz (2001) is to consider the *AR(1)* model:¹⁸

$$z_{t+1} - \mu = \rho(z_t - \mu) + \epsilon_{t+1}, \quad \epsilon_{t+1} \sim \mathcal{N}(0, \sigma^2)$$

so that under the hypothesis of independent and standard Normal data, we have that $\mu = 0, \rho = 0$ and $\sigma = 1$. The Likelihood-ratio test size is given by

$$LR_{\text{PIT}} = -2(l(z|0, 0, 1) - l(z|\mu, \rho, \sigma))$$

where (ignoring constant terms)

$$l(z|\mu, \rho, \sigma) = -\frac{1}{2} \left[\log \frac{\sigma^2}{1 - \rho^2} + \frac{(z_1 - \mu/(1 - \rho))^2}{\sigma^2/(1 - \rho^2)} + (T - 1) \log(\sigma^2) + \sum_{t=2}^T \frac{(z_t - (1 - \mu)\rho - \rho z_{t-1})^2}{\sigma^2} \right]$$

Likelihood ratio tests for independence only ($\rho = 0$) and for standard Normality only ($\mu = 0$ and $\sigma = 1$) can be defined in an analogous fashion.

¹⁸More lags and possibly time variation could be included in this specification. We however, choose this simple specification as it is intuitive and should be sufficient to test against a standard normal distribution.

Results from Probability Integral Transform tests

Tables 4.18 to 4.22 present the results from the PIT-tests.

First, the tables show the likelihood ratio-test for zero mean, unit variance and no autocorrelation. It is evident that most of the models fail this test; in fact it is only for the 3-month maturity and for the $AFNS_2 - LC$ -model that the hypothesis of zero mean, unit variance and no autocorrelation is not rejected.

Second to assess to which extent, the results are driven by autocorrelation in the PITs, we consider the test, where the autocorrelation can be estimated freely, but still with zero mean and unit variance.

Interestingly, the $AFNS_3$ and $AFNS_3 - CA$, perform quite well for most maturities, when ignoring autocorrelation in the PIT time series. This implies that, when just considering matching the distribution, i.e. mainly mean and variance, the $AFNS_3$ and $AFNS_3 - CA$ perform well in short term forecasts. Recall that for most models, the one-month forecasts are unbiased, as measured by the t -statistics, and coupled with the high correlations with the E-GARCH volatility, these results are not a big surprise.

Evaluating density forecasts 1 and 5 years ahead

When considering density forecasts 1 and 5 years ahead, we consider forecasting quantiles, rather than the PIT-test above.

Consider a quantile on level α , and let $y(t, t + \tau)$ be the yield at time t with maturity τ , and let $Q(\alpha|\Omega_{t-s})$ be the yield at the quantile level α conditional on the information up to time $t - s$, in one of the considered models. Then define a *hit* function as

$$I(t, \tau, \alpha) = \begin{cases} 1 & \text{if } y(t, t + \tau) \leq Q(\alpha|\Omega_{t-s}) \\ 0 & \text{if } y(t, t + \tau) > Q(\alpha|\Omega_{t-s}) \end{cases}$$

in this manner the hit function forms a sequence of zeros and ones, e.g. $(0, 0, 1, 0, \dots, 1)$, which gives a history of whether the quantile has been exceeded or not. Christoffersen (1998) points out that the problem of determining the accuracy of a VaR model (or in this case quantiles), is to test if the sequence $\{I(t, \tau, \alpha)\}_{t=1}^T$ satisfies two properties.

First, the *Unconditional Coverage Property*, which implies that the probability of exceeding $Q(\alpha|\Omega_{t-s})$ should be equal to α , or more in the previous notation $\Pr(I(t, \tau, \alpha) = 1) = \alpha$. Obviously if this is not the case, the model would under- or overestimate the actual risk.

Second, the *Independence Property*, implies that there is a restriction on how often violations can occur. For instance $I(t_1, \tau, \alpha)$ must be independent of $I(t_2, \tau, \alpha)$, such that $I(t_1, \tau, \alpha)$ cannot be used for forecasting $I(t_2, \tau, \alpha)$ (for $t_1 < t_2$). If a model cannot adapt to changing market conditions, then the model could suffer from violations for multiple periods in a row, making it less useful for risk management.

In the present setting with a downward trend in the interest-rate level, we do not expect the independence property to hold. This is also indicated by the PIT-tests, where all model fail the full test, which includes autocorrelation. Visual inspection of our time series also confirms that the independence property does not hold. Instead we only consider the unconditional coverage property.

Results from the quantile forecasts

Table 4.23 to 4.32 present the results from the quantile forecasts. For each model, the tables show the estimated frequency of quantile exceedences (here termed quantile prediction probabilities (*QPD*)):

$$QPD \equiv \hat{\alpha} = \frac{1}{T} \sum_{t=1}^T I(t, \tau, \alpha)$$

where $I(t, \tau, \alpha)$ is the hit function defined above, and T is the sample size.

Each table presents the QPDs, along with t -statistics based standard errors from the estimation of the QPDs. Standard errors are calculated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated QPD. Thus we do not take autocorrelation in the exceedences into account, however the results in this section are quite clear so the biased standard deviations are not likely to affect the results greatly.

For the one year forecasts (Tables 4.23 to 4.27) it is mainly the $AFNS_1 - L$ and $AFNS_1 - C$ -models that perform well, by matching the quantiles for the majority of the yields. On the other hand, the $AFNS_2 - SC$ model performs the worst, forecasting quantiles that are far too low, compared to the data. For the CIR-based models, $CIR - 2$, $AFNS_3$ and $AFNS_3 - CA$, the results are quite interesting. For the lower quantiles, the models forecast quantiles that are too high compared to the data, i.e. putting to little risk in the lower tail of the distribution. For the higher quantiles the picture is a bit mixed; for the shorter maturities, the Nelson-Siegel-based models forecast quantiles that are too low, i.e. understating the upward risk to interest-rate changes. For longer-term interest-rates all the CIR-based models forecast

the upper quantiles quite well, i.e. there is no statistical difference between the observed and model-based quantiles.

For the five-year forecast (Tables 4.28 to 4.32), there is an interesting finding for the non-CIR-based models ($AFNS_1$ and $AFNS_2$ -models). They all (except for the 3-month yield) forecast lower quantiles that are far too low. For instance for the 10 year yield, there are no realized observations below the 10 percent quantile. In terms of the upper quantiles, the picture is almost the same, there are no or only a few observations above the 90 percent quantile. All in all, this implies that these models may provide a good forecast of the mean, but distributions are generally too wide.

In terms of the CIR-based models (the $CIR-2$, $AFNS_3$ and $AFNS_3-CA$), they do not perform very well either. In term of the lower quantiles, the models typically have a large portion of the observed yields below the 5 percent quantile (i.e. around 40-70 percent). In terms of the upper quantiles these models perform slightly worse than the non-CIR-based models. We believe the poor performance of the CIR-based models to be an artifact of the downward trend in the level of the interest-rates, coupled with a model structure that ensures strict positivity of interest-rates. This is especially a weakness when one is mostly exposed to risk of falling interest-rates.

4.13 Appendix: Tables with out-of-sample forecast results

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> - 2	-0.87 (-0.1279)	-4.51 (-0.6379)	0.64 (0.1189)	-3.84 (-0.7915)	-21.6 (-2.928)
<i>AFNS</i> ₁ - <i>L</i>	0.94 (0.2681)	-2.26 (-0.7225)	-0.57 (-0.2398)	1.07 (0.5529)	-2.65 (-1.0008)
<i>AFNS</i> ₁ - <i>C</i>	-3.62 (-1.2035)	3.06 (1.0062)	-1.42 (-0.9715)	-5.47 (-2.4147)	-0.32 (-0.1242)
<i>AFNS</i> ₂ - <i>LC</i>	0.99 (0.2853)	1.17 (0.3447)	1.72 (0.6691)	2.12 (1.049)	-1.65 (-0.6)
<i>AFNS</i> ₂ - <i>SC</i>	3.16 (1.2152)	31.49 (6.784)	27.74 (14.227)	14.23 (5.4185)	14.93 (5.1105)
<i>AFNS</i> ₃	-2.46 (-0.6899)	-1.02 (-0.2491)	-0.01 (-0.0046)	-3.33 (-1.0534)	-12.99 (-2.4615)
<i>AFNS</i> ₃ - <i>CA</i>	-2.37 (-0.6538)	-0.23 (-0.0557)	2.3 (0.7961)	-3.09 (-0.9874)	-16.13 (-2.9508)

Table 4.6: Mean errors for 1-month point forecasts measured in basis points. *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	-28.87 (-0.83)	-29.12 (-1.0409)	-24.06 (-1.2122)	-30.87 (-1.8414)	-49.75 (-2.7214)
<i>AFNS</i> ₁ – <i>L</i>	3.28 (0.0812)	-5.29 (-0.1645)	-3.56 (-0.1536)	-0.82 (-0.043)	-4.23 (-0.2191)
<i>AFNS</i> ₁ – <i>C</i>	-14.66 (-0.3929)	3.34 (0.1157)	-1.3 (-0.0685)	-11.38 (-0.7252)	-9.09 (-0.5434)
<i>AFNS</i> ₂ – <i>LC</i>	7.98 (0.2005)	15.06 (0.4632)	13.7 (0.5756)	9.48 (0.4745)	3.43 (0.1703)
<i>AFNS</i> ₂ – <i>SC</i>	143.44 (3.9084)	271.6 (8.8893)	279.93 (12.5794)	222.21 (10.8869)	194.04 (9.1715)
<i>AFNS</i> ₃	-24.27 (-0.6472)	-22.84 (-0.7699)	-24.73 (-1.1962)	-30.91 (-1.7502)	-41.28 (-2.1872)
<i>AFNS</i> ₃ – <i>CA</i>	-25.94 (-0.6945)	-20.25 (-0.6743)	-20.28 (-0.9469)	-29.54 (-1.5999)	-44.07 (-2.2393)

Table 4.7: Mean errors for 1-year point forecasts measured in basis points. *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	-134.86 (-2.0061)	-128.96 (-2.8772)	-115.21 (-4.4861)	-119.96 (-6.3831)	-142.21 (-6.9637)
<i>AFNS</i> ₁ – <i>L</i>	57.72 (0.8941)	25.75 (0.5462)	20.5 (0.6497)	20.28 (0.7685)	11.72 (0.4297)
<i>AFNS</i> ₁ – <i>C</i>	33.53 (0.4312)	46.07 (0.7388)	33.44 (0.7156)	14.38 (0.3764)	11.75 (0.3253)
<i>AFNS</i> ₂ – <i>LC</i>	55.79 (1.0067)	49.59 (1.3041)	39.46 (1.7359)	28.35 (1.5804)	15.91 (0.8399)
<i>AFNS</i> ₂ – <i>SC</i>	722.57 (12.546)	838.56 (16.0571)	829.8 (17.4618)	763.45 (14.5777)	734.6 (12.7829)
<i>AFNS</i> ₃	-75.26 (-1.3116)	-93.43 (-2.3579)	-102.15 (-4.3854)	-109.96 (-6.4404)	-123.24 (-6.6778)
<i>AFNS</i> ₃ – <i>CA</i>	-87.42 (-1.5287)	-100.16 (-2.5662)	-107.89 (-4.8115)	-119.99 (-7.3512)	-137.54 (-7.6402)

Table 4.8: Mean errors for 5-year point forecasts measured in basis points. *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	23.17	25.78	23.42	22.24	30.59
<i>AFNS</i> ₁ – <i>L</i>	18.9	19.69	18.5	15.67	16.39
<i>AFNS</i> ₁ – <i>C</i>	18.36	19.18	18.24	16.94	16.5
<i>AFNS</i> ₂ – <i>LC</i>	19.05	19.61	18.34	15.75	16.34
<i>AFNS</i> ₂ – <i>SC</i>	18.15	33.69	29.8	19.72	20.71
<i>AFNS</i> ₃	18.58	20.73	19.56	18.44	23.02
<i>AFNS</i> ₃ – <i>CA</i>	18.89	20.58	19.36	18.57	24.27

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>L</i>	1.685	4.2635	2.567	2.7433	3.1385
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>C</i>	1.6644	4.498	2.1909	2.0006	3.1673
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>LC</i>	1.6003	4.314	2.8971	2.7139	3.134
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>SC</i>	2.0004	-1.8768	-2.8497	0.7577	1.7759
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃	1.5742	2.8889	2.2936	2.1369	2.6163
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃ – <i>CA</i>	1.4351	3.1321	2.6369	2.5295	2.0726
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₁ – <i>C</i>	0.9689	0.9366	0.5128	-1.1807	-0.3881
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	-0.7973	0.1977	0.5121	-0.5027	0.2569
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	1.3896	-3.3667	-6.9106	-2.9107	-1.898
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃	0.5051	-1.5248	-2.2122	-2.0464	-2.674
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.0191	-1.3851	-1.3001	-2.1394	-2.8574
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	-1.112	-3.7136	-0.1302	1.0217	0.4785
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	0.2535	-3.7658	-5.43	-1.2263	-1.9692
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃	-0.474	-1.689	-1.6519	-1.1731	-2.5276
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-1.0286	-1.6187	-1.0486	-1.1269	-2.7335
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	1.3459	-3.6354	-8.2253	-2.9473	-2.0198
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃	0.6593	-1.2768	-3.0961	-1.9299	-2.6085
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.2085	-1.1975	-2.4129	-2.0235	-2.7891
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃	-0.4419	2.8948	6.16	0.4866	-0.5232
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.7529	2.9653	7.364	0.4446	-0.7452
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ – <i>CA</i>	-1.8102	1.4119	0.5083	-0.3649	-2.0927

Table 4.9: **Top Panel:** Mean absolute errors for 1-month point forecasts. **Bottom Panel:** Pairwise comparison of 1-month point forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	93.84	80.31	65.03	54.43	62.38
<i>AFNS</i> ₁ – <i>L</i>	105.29	79.58	66.21	57.98	56.36
<i>AFNS</i> ₁ – <i>C</i>	103.22	83.24	67.73	59.48	59.01
<i>AFNS</i> ₂ – <i>LC</i>	104.8	80.58	67.54	59.67	57.54
<i>AFNS</i> ₂ – <i>SC</i>	168.28	271.64	279.93	222.21	194.04
<i>AFNS</i> ₃	100.32	80.08	66.55	55.5	57.07
<i>AFNS</i> ₃ – <i>CA</i>	101.35	80.92	66.21	56.87	58.83

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>L</i>	-1.1945	0.0786	-0.1567	-0.4634	0.6365
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>C</i>	-1.404	-0.3995	-0.3784	-0.8251	0.424
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>LC</i>	-1.0808	-0.0209	-0.2275	-0.539	0.4436
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>SC</i>	-2.4958	-5.3415	-9.0112	-5.9977	-4.0191
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃	-0.7496	0.0331	-0.678	-0.4212	1.2258
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.8615	-0.0943	-0.3489	-0.7462	0.8458
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₁ – <i>C</i>	0.2791	-0.6019	-0.3082	-0.2363	-0.428
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	0.2173	-0.1833	-0.2917	-0.5328	-0.4609
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-2.3782	-5.4216	-10.7485	-7.2183	-5.0933
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃	0.54	-0.0827	-0.0518	0.3146	-0.0884
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.4078	-0.2445	0.0002	0.1543	-0.266
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	-0.1795	0.4036	0.0341	-0.0247	0.2058
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-2.387	-6.1462	-11.8557	-6.926	-5.145
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃	0.4354	0.394	0.179	0.5185	0.2128
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.2859	0.3556	0.3096	0.352	0.0188
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-2.4903	-5.9434	-13.2706	-7.8794	-5.3009
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃	0.4412	0.048	0.0955	0.402	0.0465
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.3257	-0.0349	0.1497	0.2899	-0.1165
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃	2.1517	5.3065	8.9411	5.7529	4.0876
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	2.1286	5.4251	9.3616	5.7002	3.9135
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ – <i>CA</i>	-1.2831	-0.5823	0.1637	-0.7016	-1.2269

Table 4.10: **Top Panel:** Mean absolute errors for 1-year point forecasts. **Bottom Panel:** Pairwise comparison of 1-year point forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	184.13	146.45	118.48	119.96	142.21
<i>AFNS</i> ₁ – <i>L</i>	132.86	93.87	63.87	55.94	59.61
<i>AFNS</i> ₁ – <i>C</i>	160.13	138.04	111.71	94.51	89.08
<i>AFNS</i> ₂ – <i>LC</i>	120.81	83.07	61.44	50.29	46.74
<i>AFNS</i> ₂ – <i>SC</i>	722.57	838.56	829.8	763.45	734.6
<i>AFNS</i> ₃	142.83	116.54	105.09	109.96	123.24
<i>AFNS</i> ₃ – <i>CA</i>	147.17	120.25	109.82	119.99	137.54

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>L</i>	0.8124	1.0908	1.6418	2.3192	3.0589
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>C</i>	0.4422	0.1584	0.1881	0.9441	1.8479
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>LC</i>	1.065	1.1808	1.7737	3.0022	3.9496
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>SC</i>	-7.6147	-10.9371	-15.6876	-13.9772	-11.4666
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃	1.4283	1.7701	1.3742	1.8582	3.8886
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃ – <i>CA</i>	1.367	1.7071	0.9911	-0.0064	0.9711
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₁ – <i>C</i>	-2.3985	-3.7605	-4.3816	-6.4203	-8.7866
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	1.7462	0.9935	0.3033	0.6774	1.8278
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-10.6331	-15.0459	-14.0469	-11.7654	-10.5389
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃	-0.2028	-0.5501	-1.3959	-2.1464	-2.5987
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.2665	-0.611	-1.5525	-2.5987	-3.2188
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	3.2288	5.0872	4.5721	4.8703	4.6388
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-10.7758	-12.8706	-12.7885	-11.2831	-10.0628
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃	0.398	0.441	0.1899	-0.6195	-1.2952
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.271	0.3523	0.0544	-1.0508	-1.8616
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-11.042	-14.5229	-14.5737	-12.9612	-11.7523
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃	-0.4725	-0.697	-1.4743	-2.7439	-3.392
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.5173	-0.7473	-1.6387	-3.3244	-4.106
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃	8.7173	11.9422	15.8371	13.1773	10.9435
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	8.3437	11.6454	15.6421	13.0456	10.7617
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.9146	-1.5049	-3.1973	-7.6749	-11.1326

Table 4.11: **Top Panel:** Mean absolute errors for 5-year point forecasts. **Bottom Panel:** Pairwise comparison of 5-year point forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	-0.87 (-0.1279)	-4.51 (-0.6379)	0.64 (0.1189)	-3.84 (-0.7915)	-21.6 (-2.928)
<i>AFNS</i> ₁ – <i>L</i>	0.94 (0.2681)	-2.26 (-0.7225)	-0.57 (-0.2398)	1.07 (0.5529)	-2.65 (-1.0008)
<i>AFNS</i> ₁ – <i>C</i>	-3.62 (-1.2035)	3.06 (1.0062)	-1.42 (-0.9715)	-5.47 (-2.4147)	-0.32 (-0.1242)
<i>AFNS</i> ₂ – <i>LC</i>	0.99 (0.2853)	1.17 (0.3447)	1.72 (0.6691)	2.12 (1.049)	-1.65 (-0.6)
<i>AFNS</i> ₂ – <i>SC</i>	3.16 (1.2152)	31.49 (6.784)	27.74 (14.227)	14.23 (5.4185)	14.93 (5.1105)
<i>AFNS</i> ₃	-2.46 (-0.6899)	-1.02 (-0.2491)	-0.01 (-0.0046)	-3.33 (-1.0534)	-12.99 (-2.4615)
<i>AFNS</i> ₃ – <i>CA</i>	-2.37 (-0.6538)	-0.23 (-0.0557)	2.3 (0.7961)	-3.09 (-0.9874)	-16.13 (-2.9508)

Table 4.12: **Mean errors for 1-month volatility forecasts measured in basis points.** *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	-49.33 (-2.9917)	-13.48 (-2.2343)	12.81 (2.4998)	20.65 (5.2513)	24.56 (5.538)
<i>AFNS</i> ₁ – <i>L</i>	6.34 (0.3651)	-11.16 (-1.3397)	-10.76 (-1.4366)	-9.93 (-1.7233)	-8.36 (-1.2653)
<i>AFNS</i> ₁ – <i>C</i>	-1.73 (-0.0877)	-29.38 (-2.2549)	-35.52 (-2.8621)	-30.27 (-2.9855)	-25.85 (-2.5267)
<i>AFNS</i> ₂ – <i>LC</i>	6.89 (0.3934)	-18.38 (-2.0908)	-16.49 (-2.1388)	-12.64 (-2.1425)	-10.01 (-1.4892)
<i>AFNS</i> ₂ – <i>SC</i>	-7.98 (-0.4317)	-20.97 (-2.0598)	-24.69 (-2.5982)	-28.78 (-3.6915)	-29.01 (-3.5362)
<i>AFNS</i> ₃	51.66 (3.9029)	20.6 (3.281)	21.61 (4.2437)	27.54 (8.4627)	31.7 (9.728)
<i>AFNS</i> ₃ – <i>CA</i>	41.85 (3.3766)	11.55 (1.7963)	13.79 (2.4998)	20.61 (5.6274)	25.43 (7.3623)

Table 4.13: Mean errors for 1-year volatility forecasts measured in basis points. *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	-36.69 (-1.0176)	31.92 (2.5187)	66.82 (7.8245)	62.77 (8.2337)	63.45 (7.1063)
<i>AFNS</i> ₁ – <i>L</i>	21.17 (0.9248)	11.77 (1.3716)	8.14 (1.0616)	-6.42 (-1.0238)	-8.33 (-1.1227)
<i>AFNS</i> ₁ – <i>C</i>	-76.63 (-2.6755)	-101.32 (-7.3619)	-110.53 (-7.5727)	-121.05 (-8.0712)	-122.04 (-7.3627)
<i>AFNS</i> ₂ – <i>LC</i>	18.59 (0.8811)	0.32 (0.0351)	-1.4 (-0.1781)	-12.9 (-2.0432)	-14.16 (-2.1166)
<i>AFNS</i> ₂ – <i>SC</i>	-199.96 (-5.4566)	-204.57 (-9.5532)	-211.85 (-9.6411)	-233.06 (-10.3565)	-239.72 (-9.9328)
<i>AFNS</i> ₃	120.92 (5.3999)	77.13 (8.3685)	80.19 (12.3702)	81.29 (19.3075)	83.78 (13.8682)
<i>AFNS</i> ₃ – <i>CA</i>	96.88 (4.32)	63.27 (7.367)	65.32 (8.184)	63.72 (10.1167)	66.83 (9.1737)

Table 4.14: Mean errors for 5-year volatility forecasts measured in basis points. *t*-statistics for zero mean errors are reported in brackets. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	32.13	15.97	11.83	9.79	9.72
<i>AFNS</i> ₁ – <i>L</i>	16.58	15.17	14.97	12.6	12.5
<i>AFNS</i> ₁ – <i>C</i>	16.74	15.76	17.35	13.12	12.79
<i>AFNS</i> ₂ – <i>LC</i>	16.66	17.84	16.94	13.49	13.02
<i>AFNS</i> ₂ – <i>SC</i>	15.94	15.17	16.64	13.05	12.63
<i>AFNS</i> ₃	12.59	11.94	12.03	9.31	8.87
<i>AFNS</i> ₃ – <i>CA</i>	13.12	13.33	13.2	9.71	9.18

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>L</i>	4.0665	0.5943	-3.2091	-3.9874	-3.2214
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>C</i>	4.1781	0.1846	-3.4143	-3.8747	-3.1792
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>LC</i>	4.0408	-1.5194	-4.7053	-5.1325	-3.8466
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>SC</i>	4.3657	0.8523	-3.9321	-4.4958	-3.3305
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃	4.5985	4.0707	-0.7097	1.023	1.5551
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃ – <i>CA</i>	4.8291	2.7861	-3.4276	0.283	1.5573
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₁ – <i>C</i>	-0.1962	-0.5588	-1.954	-0.7133	-0.4144
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	-0.9356	-5.1479	-7.4712	-11.3452	-9.3946
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	0.8368	0.0065	-1.7313	-0.6341	-0.1806
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃	3.0971	2.9776	2.7257	3.6382	3.0744
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	3.7276	2.0566	1.9461	3.4028	3.0683
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	0.0867	-2.6831	0.3715	-0.5055	-0.3173
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	2.0983	1.8715	1.6669	0.3663	1.2069
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃	3.431	2.8813	3.1057	3.1476	2.7091
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	5.3936	2.2072	2.8053	3.077	2.7876
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	0.9574	3.8397	0.357	0.6143	0.5397
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃	3.2775	4.6524	4.1105	4.5248	3.5609
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	3.9263	4.3823	3.6655	4.3609	3.579
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃	3.9003	3.0314	3.451	3.3774	2.7516
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	8.5896	2.1535	3.0598	3.3768	2.8532
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.9161	-3.2542	-2.7995	-1.5879	-1.4881

Table 4.15: **Top Panel:** Mean absolute errors for 1-month volatility forecasts. **Bottom Panel:** Pairwise comparison of 1-month volatility forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> - 2	64.4	21.66	20.38	24.61	27.87
<i>AFNS</i> ₁ - <i>L</i>	34.73	22.77	22.95	17.75	19.2
<i>AFNS</i> ₁ - <i>C</i>	38.24	41.24	43.49	35.21	32.7
<i>AFNS</i> ₂ - <i>LC</i>	34.47	28.58	26.06	19.1	19.71
<i>AFNS</i> ₂ - <i>SC</i>	40.98	32.98	34.24	33.29	33.78
<i>AFNS</i> ₃	52.48	25.79	25.57	28.41	32.01
<i>AFNS</i> ₃ - <i>CA</i>	45.38	23.68	24.23	24.11	27.12

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> - 2 vs. <i>AFNS</i> ₁ - <i>L</i>	2.0112	-0.2429	-0.4941	1.7744	1.5944
<i>CIR</i> - 2 vs. <i>AFNS</i> ₁ - <i>C</i>	1.741	-3.7517	-2.5862	-1.3619	-0.5679
<i>CIR</i> - 2 vs. <i>AFNS</i> ₂ - <i>LC</i>	1.9874	-1.9139	-1.066	1.348	1.3985
<i>CIR</i> - 2 vs. <i>AFNS</i> ₂ - <i>SC</i>	1.7567	-3.7972	-2.5431	-1.6959	-0.8121
<i>CIR</i> - 2 vs. <i>AFNS</i> ₃	0.793	-0.6214	-1.7632	-1.4905	-1.6016
<i>CIR</i> - 2 vs. <i>AFNS</i> ₃ - <i>CA</i>	1.3871	-0.4764	-1.6376	0.4531	0.5093
<i>AFNS</i> ₁ - <i>L</i> vs. <i>AFNS</i> ₁ - <i>C</i>	-0.8994	-2.6361	-2.9884	-3.0377	-2.489
<i>AFNS</i> ₁ - <i>L</i> vs. <i>AFNS</i> ₂ - <i>LC</i>	0.5774	-3.7835	-2.1288	-1.7948	-0.8203
<i>AFNS</i> ₁ - <i>L</i> vs. <i>AFNS</i> ₂ - <i>SC</i>	-2.2944	-3.7053	-4.5132	-4.8371	-4.335
<i>AFNS</i> ₁ - <i>L</i> vs. <i>AFNS</i> ₃	-2.3169	-0.5003	-0.4329	-2.3546	-2.3455
<i>AFNS</i> ₁ - <i>L</i> vs. <i>AFNS</i> ₃ - <i>CA</i>	-1.4405	-0.2201	-0.281	-1.4633	-1.5171
<i>AFNS</i> ₁ - <i>C</i> vs. <i>AFNS</i> ₂ - <i>LC</i>	0.9106	1.9526	2.4222	2.6249	2.2832
<i>AFNS</i> ₁ - <i>C</i> vs. <i>AFNS</i> ₂ - <i>SC</i>	-0.8579	1.7044	1.8094	0.5253	-0.3456
<i>AFNS</i> ₁ - <i>C</i> vs. <i>AFNS</i> ₃	-1.7682	1.544	1.7384	0.8263	0.0845
<i>AFNS</i> ₁ - <i>C</i> vs. <i>AFNS</i> ₃ - <i>CA</i>	-0.9266	2.1744	2.2887	1.4398	0.6874
<i>AFNS</i> ₂ - <i>LC</i> vs. <i>AFNS</i> ₂ - <i>SC</i>	-2.0964	-1.8671	-3.3142	-4.1284	-3.9
<i>AFNS</i> ₂ - <i>LC</i> vs. <i>AFNS</i> ₃	-2.3536	0.4625	0.0789	-1.8681	-2.08
<i>AFNS</i> ₂ - <i>LC</i> vs. <i>AFNS</i> ₃ - <i>CA</i>	-1.4603	1.3422	0.405	-1.077	-1.3123
<i>AFNS</i> ₂ - <i>SC</i> vs. <i>AFNS</i> ₃	-1.3979	1.1247	1.3427	0.8029	0.2451
<i>AFNS</i> ₂ - <i>SC</i> vs. <i>AFNS</i> ₃ - <i>CA</i>	-0.5883	2.3146	2.2973	1.7598	0.9813
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ - <i>CA</i>	3.0646	0.6986	0.5446	2.004	2.455

Table 4.16: **Top Panel:** Mean absolute errors for 1-year volatility forecasts. **Bottom Panel:** Pairwise comparison of 1-year volatility forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2	80.82	36.07	66.82	62.77	63.45
<i>AFNS</i> ₁ – <i>L</i>	44.3	21.46	19.18	17.72	19.05
<i>AFNS</i> ₁ – <i>C</i>	102.02	103.76	111.93	121.61	122.74
<i>AFNS</i> ₂ – <i>LC</i>	42.77	19.94	18.85	19.74	20.56
<i>AFNS</i> ₂ – <i>SC</i>	202.79	204.57	211.85	233.06	239.72
<i>AFNS</i> ₃	120.92	77.13	80.19	81.29	83.78
<i>AFNS</i> ₃ – <i>CA</i>	96.88	63.28	65.37	63.72	66.83

	3M	2Y	5Y	10Y	15Y
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>L</i>	2.2802	1.3761	5.2256	0.0045	5.0882
<i>CIR</i> – 2 vs. <i>AFNS</i> ₁ – <i>C</i>	-1.4412	-3.2208	-2.0978	-0.0059	-2.3946
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>LC</i>	2.4066	1.675	5.3097	0.0043	4.5126
<i>CIR</i> – 2 vs. <i>AFNS</i> ₂ – <i>SC</i>	-3.1699	-5.4425	-5.0027	-0.017	-5.455
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃	-1.9924	-3.5793	-1.6747	-0.0019	-6.3537
<i>CIR</i> – 2 vs. <i>AFNS</i> ₃ – <i>CA</i>	-0.8184	-2.8777	0.285	-0.0001	-1.4049
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₁ – <i>C</i>	-2.7069	-5.7338	-6.0203	-6.5377	-5.9846
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	0.5108	0.3376	0.0849	-0.6247	-0.4988
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-3.3074	-7.8362	-8.1353	-9.0825	-8.7645
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃	-6.4115	-9.4228	-13.0863	-16.1611	-10.1323
<i>AFNS</i> ₁ – <i>L</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-4.6646	-6.6731	-5.9603	-8.0308	-6.9329
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>LC</i>	2.9888	6.6	6.6818	6.3298	5.9367
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-3.5245	-8.3813	-8.1859	-9.3567	-9.9273
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃	-0.6478	1.4735	1.8851	2.2589	1.7947
<i>AFNS</i> ₁ – <i>C</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	0.1774	2.2026	2.3882	2.8751	2.4281
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₂ – <i>SC</i>	-3.4835	-8.3355	-8.5493	-8.7841	-8.6153
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃	-5.8932	-5.9519	-8.8425	-10.5259	-8.4354
<i>AFNS</i> ₂ – <i>LC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	-4.198	-4.8453	-5.2455	-6.2874	-5.787
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃	5.3172	4.5908	5.2843	6.0118	5.3172
<i>AFNS</i> ₂ – <i>SC</i> vs. <i>AFNS</i> ₃ – <i>CA</i>	5.7217	5.2012	5.5662	6.3276	5.7217
<i>AFNS</i> ₃ vs. <i>AFNS</i> ₃ – <i>CA</i>	12.8575	2.9561	3.2687	5.6332	7.7661

Table 4.17: **Top Panel:** Mean absolute errors for 5-year volatility forecasts. **Bottom Panel:** Pairwise comparison of 5-year volatility forecasts using the Diebold and Mariano (1995) test. The table reports *t*-tests for differences between models. *t*-statistics are based on Newey and West (1987) standard errors.

	μ	ρ	σ	$\Pr(H_1 = H_0)$	ρ_2	$\Pr(H_2 = H_0)$
<i>CIR</i> – 2	-0.0331 (-0.2654)	0.6292 (17.6246)	0.6404 (-5.9048)	0.0000	0.6238 (6.6221)	0.0000
<i>AFNS</i> ₁ – <i>L</i>	-0.0503 (-0.3311)	0.2272 (2.7242)	1.4972 (6.4993)	0.0000	0.2288 (4.4775)	0.0000
<i>AFNS</i> ₁ – <i>C</i>	-0.1315 (-1.3501)	0.129 (2.1464)	1.0783 (1.0038)	4.7490	0.1421 (1.9791)	13.4383
<i>AFNS</i> ₂ – <i>LC</i>	0.0329 (0.3274)	0.1706 (2.8818)	1.0616 (0.7959)	7.6944	0.1715 (2.3525)	51.7550
<i>AFNS</i> ₂ – <i>SC</i>	0.1118 (1.1026)	0.1522 (2.4951)	1.0942 (1.2108)	2.5141	0.1604 (2.2623)	12.1432
<i>AFNS</i> ₃	-0.1686 (-1.3499)	0.214 (3.0703)	1.2514 (3.2777)	0.0000	0.2288 (3.7632)	0.0019
<i>AFNS</i> ₃ – <i>CA</i>	-0.1497 (-1.3224)	0.2117 (3.3391)	1.1369 (1.7826)	0.0230	0.2260 (3.3731)	1.8315

Table 4.18: **Probability Integral transform tests for the 3-month yield, 1 month ahead.** The test is based on Berkowitz (2001). In hypothesis H_0 all parameters are estimated freely. In the hypothesis H_1 , $\mu = \rho = 0$ and $\sigma = 1$. In the hypothesis H_2 , ρ is estimated freely, and $\mu = 0$ and $\sigma = 1$. Probabilities from the Likelihoodratio-tests (in percent) and Parameter estimates under H_0 and H_2 along with t -statistics are reported. The t -statistics are based on tests against H_1 . t -statistics are based on Newey and West (1987) standard errors.

	μ	ρ	σ	$\Pr(H_1 = H_0)$	ρ_2	$\Pr(H_2 = H_0)$
<i>CIR - 2</i>	-0.2099 (-1.6023)	0.6791 (19.9735)	0.6095 (-6.6866)	0.0000	0.6837 (7.2657)	0.0000
<i>AFNS₁ - L</i>	-0.0707 (-0.8043)	0.347 (8.4634)	0.7366 (-3.5643)	0.0000	0.3489 (3.4995)	0.0003
<i>AFNS₁ - C</i>	0.0584 (0.6206)	0.3274 (7.2434)	0.8112 (-2.5342)	0.0011	0.3305 (3.6199)	0.1862
<i>AFNS₂ - LC</i>	0.0228 (0.2731)	0.3597 (9.3916)	0.6871 (-4.2514)	0.0000	0.3584 (3.3684)	0.0000
<i>AFNS₂ - SC</i>	0.9467 (10.4262)	0.3145 (7.0993)	0.7946 (-2.8176)	0.0000	0.7072 (11.5745)	0.0000
<i>AFNS₃</i>	-0.0644 (-0.466)	0.4472 (8.1014)	0.9908 (-0.1301)	0.0000	0.4478 (6.2805)	88.6506
<i>AFNS₃ - CA</i>	-0.0051 (-0.0415)	0.4497 (9.1963)	0.8773 (-1.7355)	0.0000	0.4488 (5.5821)	7.9479

Table 4.19: **Probability Integral transform tests for the 2-year yield, 1 month ahead.** The test is based on Berkowitz (2001). In hypothesis H_0 all parameters are estimated freely. In the hypothesis H_1 , $\mu = \rho = 0$ and $\sigma = 1$. In the hypothesis H_2 , ρ is estimated freely, and $\mu = 0$ and $\sigma = 1$. Probabilities from the Likelihoodratio-tests (in percent) and Parameter estimates under H_0 and H_2 along with t -statistics are reported. The t -statistics are based on tests against H_1 . t -statistics are based on Newey and West (1987) standard errors.

	μ	ρ	σ	$\Pr(H_1 = H_0)$	ρ_2	$\Pr(H_2 = H_0)$
<i>CIR</i> - 2	-0.0314 (-0.2336)	0.6181 (15.7277)	0.7053 (-4.7763)	0.0000	0.6141 (7.0749)	0.0000
<i>AFNS</i> ₁ - <i>L</i>	-0.0120 (-0.1491)	0.2675 (6.3842)	0.7528 (-3.2484)	0.0003	0.2663 (2.6419)	0.0020
<i>AFNS</i> ₁ - <i>C</i>	-0.0327 (-0.4529)	0.2019 (4.9364)	0.7339 (-3.4424)	0.0002	0.2023 (1.9285)	0.0003
<i>AFNS</i> ₂ - <i>LC</i>	0.0532 (0.6982)	0.2771 (7.0689)	0.7038 (-3.9025)	0.0000	0.2798 (2.6125)	0.0000
<i>AFNS</i> ₂ - <i>SC</i>	0.8712 (11.4331)	0.2379 (5.7743)	0.7399 (-3.4634)	0.0000	0.6718 (10.0119)	0.0000
<i>AFNS</i> ₃	0.0044 (0.0371)	0.3555 (6.4991)	0.9822 (-0.2418)	0.0049	0.3555 (4.7463)	94.9428
<i>AFNS</i> ₃ - <i>CA</i>	0.0983 (0.8746)	0.3856 (7.7742)	0.8898 (-1.5158)	0.0004	0.3925 (4.8457)	9.6955

Table 4.20: **Probability Integral transform tests for the 5-year yield, 1 month ahead.** The test is based on Berkowitz (2001). In hypothesis H_0 all parameters are estimated freely. In the hypothesis H_1 , $\mu = \rho = 0$ and $\sigma = 1$. In the hypothesis H_2 , ρ is estimated freely, and $\mu = 0$ and $\sigma = 1$. Probabilities from the Likelihoodratio-tests (in percent) and Parameter estimates under H_0 and H_2 along with t -statistics are reported. The t -statistics are based on tests against H_1 . t -statistics are based on Newey and West (1987) standard errors.

	μ	ρ	σ	$\Pr(H_1 = H_0)$	ρ_2	$\Pr(H_2 = H_0)$
<i>CIR</i> – 2	-0.0641 (-0.4334)	0.6611 (16.7792)	0.7075 (-4.9576)	0.0000	0.6617 (8.119)	0.0000
<i>AFNS</i> ₁ – <i>L</i>	0.0578 (0.8463)	0.1253 (2.9622)	0.7592 (-3.0793)	0.0051	0.1297 (1.2654)	0.0029
<i>AFNS</i> ₁ – <i>C</i>	-0.1638 (-2.3982)	0.1373 (3.2847)	0.7495 (-3.2157)	0.0003	0.1759 (1.733)	0.0003
<i>AFNS</i> ₂ – <i>LC</i>	0.0917 (1.3915)	0.1368 (3.3945)	0.7240 (-3.5339)	0.0001	0.1495 (1.4024)	0.0001
<i>AFNS</i> ₂ – <i>SC</i>	0.5635 (6.9056)	0.257 (5.9629)	0.7741 (-2.996)	0.0000	0.5032 (6.3058)	0.0000
<i>AFNS</i> ₃	-0.1637 (-1.0965)	0.4614 (7.9142)	1.0454 (0.6532)	0.0000	0.4741 (7.2052)	38.0497
<i>AFNS</i> ₃ – <i>CA</i>	-0.1335 (-0.9182)	0.5165 (10.0291)	0.9249 (-1.1209)	0.0000	0.5256 (7.3305)	27.6644

Table 4.21: **Probability Integral transform tests for the 10-year yield, 1 month ahead.** The test is based on Berkowitz (2001). In hypothesis H_0 all parameters are estimated freely. In the hypothesis H_1 , $\mu = \rho = 0$ and $\sigma = 1$. In the hypothesis H_2 , ρ is estimated freely, and $\mu = 0$ and $\sigma = 1$. Probabilities from the Likelihoodratio-tests (in percent) and Parameter estimates under H_0 and H_2 along with t -statistics are reported. The t -statistics are based on tests against H_1 . t -statistics are based on Newey and West (1987) standard errors.

	μ	ρ	σ	$\Pr(H_1 = H_0)$	ρ_2	$\Pr(H_2 = H_0)$
<i>CIR</i> – 2	-0.3742 (-1.6955)	0.7580 (18.716)	0.7271 (-4.6412)	0.0000	0.8151 (13.6992)	0.0001
<i>AFNS</i> ₁ – <i>L</i>	-0.0658 (-0.8613)	0.1750 (3.9150)	0.8024 (-2.5398)	0.0569	0.1790 (1.8530)	0.0900
<i>AFNS</i> ₁ – <i>C</i>	0.0134 (0.1703)	0.2032 (4.5663)	0.7995 (-2.5871)	0.0365	0.2029 (2.0982)	0.0916
<i>AFNS</i> ₂ – <i>LC</i>	-0.0271 (-0.3566)	0.1962 (4.5312)	0.778 (-2.8645)	0.0093	0.1960 (1.9718)	0.0170
<i>AFNS</i> ₂ – <i>SC</i>	0.5992 (6.4569)	0.3139 (6.9141)	0.8148 (-2.5027)	0.0000	0.5436 (7.3065)	0.0000
<i>AFNS</i> ₃	-0.5441 (-2.5425)	0.6262 (10.3333)	1.0869 (1.4364)	0.0000	0.6920 (13.5686)	1.3538
<i>AFNS</i> ₃ – <i>CA</i>	-0.5966 (-2.8329)	0.6653 (12.3892)	0.9633 (-0.6295)	0.0000	0.7486 (14.3136)	4.8292

Table 4.22: **Probability Integral transform tests for the 15-year yield, 1 month ahead.** The test is based on Berkowitz (2001). In hypothesis H_0 all parameters are estimated freely. In the hypothesis H_1 , $\mu = \rho = 0$ and $\sigma = 1$. In the hypothesis H_2 , ρ is estimated freely, and $\mu = 0$ and $\sigma = 1$. Probabilities from the Likelihoodratio-tests (in percent) and Parameter estimates under H_0 and H_2 along with t -statistics are reported. The t -statistics are based on tests against H_1 . t -statistics are based on Newey and West (1987) standard errors.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.0933 (1.8244)	0.1 (0)	0.2 (-1.5309)	0.5733 (1.8159)	0.7933 (1.3107)	0.94 (2.0628)	0.9867 (0.0094)
<i>AFNS</i> ₁ – <i>L</i>	0.1533 (3.5125)	0.1667 (2.1909)	0.32 (1.8379)	0.4 (-2.5)	0.6067 (-3.5937)	0.78 (-3.5479)	0.82 (-4.1442)
<i>AFNS</i> ₁ – <i>C</i>	0.1333 (3.0024)	0.2 (3.0619)	0.3533 (2.6476)	0.46 (-0.9829)	0.6733 (-2.0021)	0.82 (-2.5503)	0.8867 (-2.4469)
<i>AFNS</i> ₂ – <i>LC</i>	0.1267 (2.8231)	0.16 (2.0045)	0.3133 (1.6723)	0.4067 (-2.3271)	0.5933 (-3.9062)	0.7467 (-4.3179)	0.8 (-4.5928)
<i>AFNS</i> ₂ – <i>SC</i>	0.06 (0.5157)	0.0667 (-1.6366)	0.0733 (-8.3002)	0.1067 (-15.6058)	0.1933 (-17.264)	0.3267 (-14.9722)	0.4733 (-11.6925)
<i>AFNS</i> ₃	0.38 (8.3267)	0.4067 (7.6462)	0.44 (4.6879)	0.5133 (0.3267)	0.6 (-3.75)	0.7133 (-5.0557)	0.7533 (-5.5876)
<i>AFNS</i> ₃ – <i>CA</i>	0.3467 (7.6347)	0.38 (7.0651)	0.4467 (4.845)	0.5267 (0.6541)	0.6333 (-2.9651)	0.76 (-4.0148)	0.7933 (-4.7387)

Table 4.23: **Quantile prediction probabilities for the 3-month yield 1 year ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and *T* is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.1267 (2.8231)	0.1867 (2.7241)	0.36 (2.8067)	0.6267 (3.2073)	0.78 (0.887)	0.9267 (1.2529)	0.9867 (3.9153)
<i>AFNS</i> ₁ – <i>L</i>	0.0933 (1.8244)	0.1333 (1.201)	0.2267 (-0.6826)	0.5067 (0.1633)	0.72 (-0.8183)	0.8133 (-2.7241)	0.9 (-2.0412)
<i>AFNS</i> ₁ – <i>C</i>	0.08 (1.3543)	0.1 (0)	0.2133 (-1.0962)	0.4867 (-0.3267)	0.7467 (-0.0939)	0.88 (-0.7538)	0.9267 (-1.0963)
<i>AFNS</i> ₂ – <i>LC</i>	0.0667 (0.8183)	0.1 (0)	0.1667 (-2.7386)	0.44 (-1.4804)	0.6933 (-1.5051)	0.7933 (-3.2263)	0.88 (-2.6382)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0.0067 (-74.2479)	0.0267 (-54.9881)	0.06 (-43.3197)	0.1267 (-30.318)
<i>AFNS</i> ₃	0.24 (5.4486)	0.2933 (5.2007)	0.4333 (4.5312)	0.5667 (1.6477)	0.72 (-0.8183)	0.7933 (-3.2263)	0.82 (-4.1442)
<i>AFNS</i> ₃ – <i>CA</i>	0.2 (4.5928)	0.24 (4.0148)	0.4067 (3.9062)	0.56 (1.4804)	0.7333 (-0.4616)	0.8067 (-2.8946)	0.8333 (-3.8341)

Table 4.24: **Quantile prediction probabilities for the 2-year yield 1 year ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.12 (2.6382)	0.2533 (4.3179)	0.42 (4.2185)	0.6067 (2.6744)	0.7533 (0.0947)	0.9267 (1.2529)	0.96 (0.625)
<i>AFNS</i> ₁ – <i>L</i>	0.06 (0.5157)	0.0867 (-0.5804)	0.26 (0.2792)	0.5133 (0.3267)	0.72 (-0.8183)	0.8533 (-1.6156)	0.9467 (-0.1817)
<i>AFNS</i> ₁ – <i>C</i>	0.04 (-0.625)	0.06 (-2.0628)	0.22 (-0.887)	0.48 (-0.4903)	0.8 (1.5309)	0.9467 (2.5436)	0.98 (2.6245)
<i>AFNS</i> ₂ – <i>LC</i>	0.02 (-2.6245)	0.06 (-2.0628)	0.1733 (-2.4805)	0.4533 (-1.1481)	0.7 (-1.3363)	0.8333 (-2.1909)	0.9267 (-1.0963)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0.02 (-76.984)	0.1067 (-33.4598)
<i>AFNS</i> ₃	0.2333 (5.3088)	0.3333 (6.0622)	0.48 (5.6383)	0.6133 (2.8503)	0.72 (-0.8183)	0.8533 (-1.6156)	0.9 (-2.0412)
<i>AFNS</i> ₃ – <i>CA</i>	0.2 (4.5928)	0.2733 (4.7634)	0.4067 (3.9062)	0.6 (2.5)	0.7333 (-0.4616)	0.84 (-2.0045)	0.8733 (-2.8231)

Table 4.25: **Quantile prediction probabilities for the 5-year yield 1 year ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.2267 (5.168)	0.26 (4.4675)	0.4067 (3.9062)	0.6467 (3.7579)	0.8267 (2.4805)	0.9467 (2.5436)	0.9733 (1.7738)
<i>AFNS</i> ₁ – <i>L</i>	0.0067 (-6.5218)	0.0733 (-1.2529)	0.2267 (-0.6826)	0.4533 (-1.1481)	0.7267 (-0.6412)	0.88 (-0.7538)	0.9667 (1.1371)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0.0667 (-1.6366)	0.2533 (0.0939)	0.5333 (0.8183)	0.8133 (1.9907)	0.9667 (4.5486)	1 (-)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0.0133 (-9.2543)	0.2067 (-1.3107)	0.4333 (-1.6477)	0.7067 (-1.1657)	0.86 (-1.4119)	0.9467 (-0.1817)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0.0267 (-54.9881)	0.1733 (-23.5112)	0.26 (-19.266)
<i>AFNS</i> ₃	0.2867 (6.4098)	0.36 (6.634)	0.4333 (4.5312)	0.6533 (3.946)	0.7867 (1.0962)	0.9133 (0.5804)	0.9533 (0.1936)
<i>AFNS</i> ₃ – <i>CA</i>	0.24 (5.4486)	0.3 (5.3452)	0.4267 (4.3747)	0.6067 (2.6744)	0.76 (0.2868)	0.92 (0.9029)	0.9467 (-0.1817)

Table 4.26: **Quantile prediction probabilities for the 10-year yield 1 year ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.2267 (5.168)	0.26 (4.4675)	0.4067 (3.9062)	0.6467 (3.7579)	0.8267 (2.4805)	0.9467 (2.5436)	0.9733 (1.7738)
<i>AFNS</i> ₁ – <i>L</i>	0.0067 (-6.5218)	0.0733 (-1.2529)	0.2267 (-0.6826)	0.4533 (-1.1481)	0.7267 (-0.6412)	0.88 (-0.7538)	0.9667 (1.1371)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0.0667 (-1.6366)	0.2533 (0.0939)	0.5333 (0.8183)	0.8133 (1.9907)	0.9667 (4.5486)	1 (-)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0.0133 (-9.2543)	0.2067 (-1.3107)	0.4333 (-1.6477)	0.7067 (-1.1657)	0.86 (-1.4119)	0.9467 (-0.1817)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0.0267 (-54.9881)	0.1733 (-23.5112)	0.26 (-19.266)
<i>AFNS</i> ₃	0.2867 (6.4098)	0.36 (6.634)	0.4333 (4.5312)	0.6533 (3.946)	0.7867 (1.0962)	0.9133 (0.5804)	0.9533 (0.1936)
<i>AFNS</i> ₃ – <i>CA</i>	0.24 (5.4486)	0.3 (5.3452)	0.4267 (4.3747)	0.6067 (2.6744)	0.76 (0.2868)	0.92 (0.9029)	0.9467 (-0.1817)

Table 4.27: **Quantile prediction probabilities for the 15-year yield 1 year ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.0784 (1.068)	0.1863 (2.238)	0.6176 (7.6406)	0.7255 (5.1031)	0.8431 (2.5865)	0.9608 (3.1626)	0.9902 (4.1202)
<i>AFNS</i> ₁ – <i>L</i>	0.0294 (-1.2307)	0.049 (-2.3847)	0.1569 (-2.5865)	0.3824 (-2.445)	0.5686 (-3.6985)	0.8333 (-1.8067)	0.9118 (-1.3615)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0.0294 (-4.2194)	0.1765 (-1.948)	0.5294 (0.5951)	0.7255 (-0.5547)	0.8922 (-0.2554)	0.9412 (-0.3787)
<i>AFNS</i> ₂ – <i>LC</i>	0.0098 (-4.1202)	0.0294 (-4.2194)	0.1275 (-3.7114)	0.3725 (-2.6623)	0.6078 (-2.9406)	0.8824 (-0.5532)	0.9706 (1.2307)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0.0098 (-75.8728)	0.049 (-39.806)	0.1569 (-22.0262)
<i>AFNS</i> ₃	0.5196 (9.4929)	0.5686 (9.5563)	0.6078 (7.4023)	0.6373 (2.8832)	0.6961 (-1.184)	0.8333 (-1.8067)	0.8922 (-1.8834)
<i>AFNS</i> ₃ – <i>CA</i>	0.5098 (9.2894)	0.5392 (8.8991)	0.598 (7.1692)	0.6373 (2.8832)	0.7451 (-0.1136)	0.8922 (-0.2554)	0.9216 (-1.068)

Table 4.28: **Quantile prediction probabilities for the 3-month yield 5 years ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.3627 (6.5695)	0.5196 (8.4822)	0.6471 (8.3914)	0.8039 (7.7311)	0.9412 (8.2058)	0.9902 (9.2454)	1 (-)
<i>AFNS</i> ₁ – <i>L</i>	0 (-)	0.0294 (-4.2194)	0.0784 (-6.4451)	0.4118 (-1.8107)	0.7255 (-0.5547)	0.951 (2.3847)	0.9804 (2.2138)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0 (-)	0.0686 (-7.2454)	0.4608 (-0.7946)	0.7255 (-0.5547)	0.9608 (3.1626)	0.9902 (4.1202)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0 (-)	0.0392 (-10.9672)	0.2549 (-5.68)	0.6765 (-1.5874)	0.9804 (5.856)	1 (-)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0.0294 (-55.0285)
<i>AFNS</i> ₃	0.5392 (9.9122)	0.5784 (9.785)	0.6373 (8.1347)	0.6961 (4.3055)	0.9216 (6.4451)	0.9706 (4.2194)	1 (-)
<i>AFNS</i> ₃ – <i>CA</i>	0.4608 (8.3231)	0.5392 (8.8991)	0.6373 (8.1347)	0.6863 (4.0544)	0.9314 (7.2454)	0.9902 (9.2454)	1 (-)

Table 4.29: **Quantile prediction probabilities for the 2-year yield 5 years ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.4216 (7.5994)	0.5196 (8.4822)	0.7647 (12.2548)	0.9216 (15.8365)	0.9804 (16.7824)	1 (-)	1 (-)
<i>AFNS</i> ₁ – <i>L</i>	0 (-)	0 (-)	0.049 (-9.4012)	0.2843 (-4.8291)	0.7843 (0.8426)	0.9804 (5.856)	1 (-)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0 (-)	0.0686 (-7.2454)	0.4314 (-1.3995)	0.8235 (1.948)	0.9902 (9.2454)	1 (-)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0 (-)	0.0098 (-24.621)	0.1765 (-8.5711)	0.8431 (2.5865)	1 (-)	1 (-)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
<i>AFNS</i> ₃	0.549 (10.1285)	0.6373 (11.2856)	0.7549 (11.8547)	0.8922 (12.7686)	0.9804 (16.7824)	1 (-)	1 (-)
<i>AFNS</i> ₃ – <i>CA</i>	0.4804 (8.7002)	0.5784 (9.785)	0.7353 (11.1095)	0.8824 (11.9854)	0.9902 (24.621)	1 (-)	1 (-)

Table 4.30: **Quantile prediction probabilities for the 5-year yield 5 years ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.5588 (10.3496)	0.6961 (13.0886)	0.9412 (29.6673)	0.9902 (50.2469)	1 (-)	1 (-)	1 (-)
<i>AFNS</i> ₁ – <i>L</i>	0 (-)	0 (-)	0.0196 (-16.7824)	0.2451 (-5.9849)	0.7843 (0.8426)	1 (-)	1 (-)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0 (-)	0.0784 (-6.4451)	0.5098 (0.1981)	0.8824 (4.1488)	1 (-)	1 (-)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0 (-)	0 (-)	0.2059 (-7.3463)	0.8824 (4.1488)	1 (-)	1 (-)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)
<i>AFNS</i> ₃	0.7059 (14.5378)	0.7941 (17.3373)	0.9412 (29.6673)	0.9902 (50.2469)	1 (-)	1 (-)	1 (-)
<i>AFNS</i> ₃ – <i>CA</i>	0.6569 (12.9098)	0.7745 (16.3009)	0.9314 (27.2191)	0.9902 (50.2469)	1 (-)	1 (-)	1 (-)

Table 4.31: **Quantile prediction probabilities for the 10-year yield 5 years ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Quantile	0.05	0.1	0.25	0.5	0.75	0.9	0.95
<i>CIR</i> – 2	0.8431 (22.0262)	0.9412 (36.1057)	0.9706 (43.0734)	1 (-)	1 (-)	1 (-)	1 (-)
<i>AFNS</i> ₁ – <i>L</i>	0 (-)	0 (-)	0.098 (-5.161)	0.5098 (0.1981)	0.902 (5.161)	1 (-)	1 (-)
<i>AFNS</i> ₁ – <i>C</i>	0 (-)	0 (-)	0.0686 (-7.2454)	0.5392 (0.7946)	0.8824 (4.1488)	1 (-)	1 (-)
<i>AFNS</i> ₂ – <i>LC</i>	0 (-)	0 (-)	0.0098 (-24.621)	0.2647 (-5.3864)	0.8922 (4.6286)	1 (-)	1 (-)
<i>AFNS</i> ₂ – <i>SC</i>	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0 (-)	0.0196 (-67.7722)
<i>AFNS</i> ₃	0.8922 (27.4206)	0.951 (39.806)	0.9804 (53.2036)	1 (-)	1 (-)	1 (-)	1 (-)
<i>AFNS</i> ₃ – <i>CA</i>	0.8824 (26.0913)	0.951 (39.806)	0.9902 (75.8728)	1 (-)	1 (-)	1 (-)	1 (-)

Table 4.32: **Quantile prediction probabilities for the 15-year yield 5 years ahead.** *t*-statistics in brackets. *t*-statistics are based on standard errors estimated as $SD = \sqrt{\hat{\alpha}(1 - \hat{\alpha})/T}$, where $\hat{\alpha}$ is the estimated quantile prediction probability and T is the sample size.

Conclusion

In this thesis we have considered the modeling of risks in interest-rate and inflation markets. Motivated by the size and importance of these two markets, both to investors, policy makers and individuals, we have discussed both interest-rates and inflation from an empirical as well as a modeling perspective.

In the first essay we considered the modeling of the stochastic skewness implicit in interest-rate options. We showed evidence of stochastic skewness in caps and floors linked to Euro area LIBOR. By using a Heath-Jarrow-Morton framework based on time-changed Lévy processes we showed how to capture the stochastic skewness. Our calibration to time-series of volatilities and skewness measures suggests that the model provides a reasonable fit to the skewness data and that the jump components in the time-changed Lévy processes mainly affect short-term maturities.

Inspired by the rise of inflation derivatives and more liquid markets, essay 2 presented a framework for modeling inflation derivatives based on the forward rate approach of Jarrow and Yildirim (2003). By using time-changed Lévy processes, we were able to capture both stochastic volatility and jumps in real and nominal rates as well as inflation. We calibrated our model to market data. Our results showed that even though Lévy processes can improve the fit to data, an investigation into the exact specification of the Lévy process and volatility loading is still needed.

In the third essay, inspired by the demands of for instance central banks, we derived inflation risk premia from inflation swaps, nominal interest-rate swaps, CPI data and surveys. By using a reduced-form no-arbitrage term structure model, we estimated the inflation risk premia using a Markov Chain Monte Carlo estimation methodology. Our estimates of inflation risk premia on average showed an upward sloping term structure, with 1 year risk premia of 18 bps and 10 year risk premia of 43 bps, however with fluctuation in risk premia over time. Furthermore, our estimates suggested that surveys are important in identifying inflation expectations and thus inflation risk premia. Finally, we related estimates of inflation risk premia

to agents beliefs, and found that skews in short term inflation perceptions drive short term inflation risk premia, where beliefs on GDP growth drive longer term risk premia.

In the fourth and final essay, we assessed the performance of a class of term structure models, namely the Affine Nelson-Siegel models with stochastic volatility. We assessed both short- and long-term forecasting ability of the models. More precisely, we forecasted both means and variances of the considered yields. We found that models with 3 CIR-factors perform the best in short term predictions, while models with a combination of CIR and Gaussian factors perform well on 1 and 5-year horizons. Overall, our results indicated that no single model should be used for risk management, but rather a suite of models.

Bibliography

- ANDERSEN, A. S. T. (2011): “Modeling stochastic skewness in a Heath-Jarrow-Morton-framework,” *Working Paper, Copenhagen Business School*.
- ANDERSEN, L., AND J. ANDREASEN (2000): “Volatility Skews and Extensions for the LIBOR market model,” *Applied Mathematical Finance*, 7, 1 – 32.
- ANDERSEN, L., AND R. BROTHERTON-RATCLIFFE (2001): “Extended LIBOR market model with stochastic volatility,” *Working paper*.
- ANDERSEN, T. G., L. BENZONI, AND J. LUND (2004): “Stochastic Volatility, Mean Drift and Jumps in the Short-Term Interest Rate,” *Working paper, Kellogg School of Management, Northwestern University*.
- ANDERSEN, T. G., AND J. LUND (1997): “Estimating continuous-time stochastic volatility models of the short-term interest rate,” *Journal of Econometrics*, 77, 343–377.
- ANG, A., G. BEKAERT, AND M. WEI (2007): “Do Macro Variables, Asset Markets or Survey Forecast Inflation Better,” *Journal of Monetary Economics*, 54, 1163–1212.
- (2008): “The term structure of real rates and inflation expectations,” *Journal of Finance*, 63,2, 797–849.
- BAKSHI, G., N. KAPADIA, AND D. MADAN (2003): “Stock return characteristics, skew laws, and the differential pricing of individual equity options,” *Review of Financial Studies*, 16, 101 – 143.
- BAKSHI, G., AND D. MADAN (2000): “Spanning and derivative-security valuation,” *Journal of Financial Economics*, 55, 205 – 238.
- BALLOTTA, L., AND E. BONFIGLIOLI (2010): “Multivariate Asset Models Using Levy Processes and Applications,” *Working Paper*.

- BARCLAYS CAPITAL (2008): “Global Inflation-Linked Products, a User Guide,” *Barclays Capital, Research*.
- BARNDORFF-NIELSEN, O. E., AND N. SHEPHARD (2001): “Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics,” *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 63(2), 167–241.
- BERKOWITZ, J. (2001): “Testing Density Forecasts with Applications to Risk Management,” *Journal of Business and Economic Statistics*, 19, 465–474.
- BESAG, J. (1974): “Spatial Integration and the Statistical Analysis of Lattice Systems,” *Journal of the Royal Statistical Association, Series B* 36, 192–236.
- BIS (2010): “OTC derivatives market activity in the second half of 2010,” *Bank for International Settlements*.
- BJÖRK, T. (2004): *Arbitrage Theory in Continuous Time*, Finance. Oxford University Press, second edn.
- BJÖRK, T., Y. KABANOV, AND W. RUNGALDIER (1997): “Bond Market Structure in the Presence of Market Point Processes,” *Mathematical Finance*, 7, 211–239.
- BRACE, A., D. GATAREK, AND M. MUSIELA (1997): “The market model of interest rate dynamics,” *Mathematical Finance*, 7, 127–155.
- BREEDEN, D. T., AND R. H. LITZENBERGER (1978): “Prices of State-Contingent Claims Implicit in Option Prices,” *The Journal of Business*, 13(1), 15 – 22.
- BRIGO, D., AND F. MERCURIO (2006): *Interest Rate Models - Theory and Practice*, Springer Finance. Springer Verlag, second edn.
- CAMPBELL, J. Y., R. J. SCHILLER, AND L. M. VICEIRA (2009): “Understanding Inflation-Indexed Bond Markets,” *NBER Working Paper 15014*.
- CARR, P., AND D. MADAN (1999): “Option Valuation Using the Fast Fourier Transform,” *Journal of Computational Finance*, 2(4), 61–73.
- CARR, P., AND D. MADAN (2001): “Optimal positioning in derivatives securities,” *Quantitative Finance*, 1, 19 – 37.

- CARR, P., AND L. WU (2004): “Time-changed Lévy processes and option pricing,” *Journal of Financial Economics*, 71, 113–141.
- (2007): “Stochastic Skew in Currency Options,” *Journal of Financial Economics*, 86, 213–247.
- CASASSUS, J., P. COLLIN-DUFRESNE, AND R. GOLDSTEIN (2005): “Unspanned stochastic volatility and fixed income derivatives pricing,” *Journal of Banking and Finance*, 29, 2733–2749.
- CHENG, P., AND O. SCAILLET (2007): “Linear-quadratic jump-diffusion modelling,” *Mathematical Finance*, 17.
- CHEREDITO, P., D. FILIPOVIC, AND P. KIMMEL (2007): “Market price of risk specifications for affine models: Theory and evidence,” *Journal of Financial Economics*, 83, 123–170.
- CHERNOV, M., AND P. MUELLER (2008): “The term structure of inflation expectations,” *Working paper*.
- CHRISTENSEN, J. E. H., F. X. DIEBOLD, AND G. D. RUDEBUSCH (2011): “The Affine Arbitrage-free Class of Nelson-Siegel Term Structure models,” *Journal of Econometrics*, Forthcoming.
- CHRISTENSEN, J. E. H., AND J. GILLAN (2011): “A Model-Independent Maximum Range for the Liquidity Correction of TIPS Yields,” *Working paper, Federal Reserve Bank of San Francisco*.
- CHRISTENSEN, J. E. H., J. A. LOPEZ, AND G. D. RUDEBUSCH (2008): “Inflation Expectations and Risk Premiums in an Arbitrage-Free model of Nominal and Real Yields,” *Journal of Money, Credit, and Banking*, 42, 143–178.
- (2010): “Can Spanned Term Structure Factors Drive Stochastic Volatility?,” *Working paper, Federal Reserve Bank of San Francisco*.
- CHRISTOFFERSEN, P. (1998): “Evaluating Interval Forecasts,” *International Economic Review*, 39, 841–862.
- COLLIN-DUFRESNE, P., AND R. GOLDSTEIN (2002): “Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility,” *Journal of Finance*, 57, 1685–1730.
- COLLIN-DUFRESNE, P., R. S. GOLDSTEIN, AND C. JONES (2008): “Can Interest Rate Volatility be Extracted from the Cross-Section of Bond Yields?,” *Journal of Financial Economics*, 94, 47–66.

- CONT, R., AND P. TANKOV (2004): *Financial Modelling With Jump Processes*, Financial Mathematics Series. Chapman & Hall / CRC.
- COX, J. C., J. E. INGERSOLL, AND S. A. ROSS (1985): “A Theory of the Term Structure of Interest Rates,” *Econometrica*, 53, 385–407.
- DAI, Q., AND K. SINGLETON (2000): “Specification Analysis of the Term Structure models,” *The Journal of Finance*, 5, 1943–1978.
- (2002): “Expectations puzzles, time-varying risk premia, and affine models of the term structure,” *Journal of Financial Economics*, 63, 415–442.
- (2003): “Term Structure Dynamics in Theory and Reality,” *The Review of Financial Studies*, 16, 3, 631–678.
- D’AMICO, S., D. H. KIM, AND M. WEI (2008): “Tips from TIPS: The information content of Treasury Inflation-Protected Security prices,” *BIS Working Papers*, 248.
- DANMARKS NATIONALBANK (2005): “Danish Government Borrowing and Debt,” .
- DAS, S. (2002): “The Surprise Element: Jumps in Interest Rates,” *Journal of Econometrics*, 106, 27–65.
- DEACON, M., A. DERRY, AND D. MIRFENDERESKI (2004): *Inflation-indexed Securities: Bonds, Swaps and Other Derivatives*, Finance. Wiley, second edn.
- DIEBOLD, F., AND R. MARIANO (1995): “Comparing predictive accuracy,” *Journal of Business and Economics Statistics*, 13, 253–253.
- DIEBOLD, F. X., T. GUNTHER, AND A. TAY (1998): “Evaluating Density Forecasts with Applications to Financial Risk Management,” *International Economic Review*, 39, 863–883.
- DIEBOLD, F. X., AND C. LI (2006): “Forecasting the Term Structure of Government Bond Yields,” *Journal of Econometrics*, 130, 337–364.
- DUARTE, J. (2004): “Evaluating an Alternative Risk Preference in Affine Term Structure Models,” *Review of Financial Studies*, 17(2), 379–404.
- DUFFEE, G. R. (2002): “Term premia and Interest Rate Forecasts in Affine Models,” *Journal of Finance*, 57, 405–443.

- DUFFIE, D. (2001): *Dynamic Asset Pricing Theory*. Princeton University Press, third edn.
- DUFFIE, D., D. FILIPOVIC, AND W. SCHACHERMAYER (2003): “Affine Processes and Applications in Finance,” *The Annals of Applied Probability*, 13, 984–1053.
- DUFFIE, D., AND R. KAN (1996): “A Yield-Factor Model of Interest Rates,” *Mathematical Finance*, 6(4), 379–406.
- DUFFIE, D., J. PAN, AND K. SINGLETON (2000): “Transform analysis and option pricing for affine jump-diffusions,” *Econometrica*, 68, 1343–1376.
- EBERLEIN, E., AND S. RAIBLE (1999): “Term structure models driven by general Lévy processes,” *Mathematical Finance*, 9 (1), 31–53.
- EBERLEIN, E., AND F. ÖZKAN (2005): “The Lévy Libor Model,” *Finance and Stochastics*, 9, 327–348.
- EGOROV, A. V., Y. HONG, AND H. LI (2006): “Validating forecasts of the joint probability density of bond yields: Can affine models beat random walk?,” *Journal of Econometrics*, 135(1-2), 255 – 284.
- EJSING, J., J. GARCIA, AND T. WERNER (2007): “The term structure of euro area break even inflation rates: The impact of seasonality,” *ECB working paper No. 830*.
- FELDHÜTTER, P. (2008): “Can Affine Models Match the Moments in Bond Yields?,” *Working Papers, Copenhagen Business School*.
- FILIPOVIC, D. (2001): “A General Characterization of One Factor Affine Term Structure Models,” *Finance and Stochastics*, 5, 389–412.
- FILIPOVIC, D., AND S. TAPPE (2008): “Existence of Lévy Term Structure Models,” *Finance and Stochastics*, 12, 83–115.
- GAMERMAN, D., AND H. LOPES (2006): *Markov Chain Monte Carlo, Stochastic Simulation for Bayesian Inference*, Texts in Statistical Science. Chapman & Hall / CRC, second edn.
- GARCIA, J. (2003): “An Introduction to the ECB’s Survey of Professional Forecasters,” *ECB Occasional paper No. 8*.
- GARCIA, J., AND T. WERNER (2010): “Inflation risks and inflation risk premia,” *ECB Working paper No. 1162*.

- GARCIA, J. A., AND A. MANZANARES (2007): “Reporting Biases and Survey Results. Evidence from European Professional Forecasters,” *ECB working paper No. 836*.
- GEMAN, H., N. E. KAROUI, AND J. C. ROCHET (1995): “Changes of numeraire, changes of probability measure and pricing of options,” *Journal of Applied Probability*, 32, 442–458.
- GEMAN, H., D. MADAN, AND M. YOR (2001): “Time Changes for Lévy processes,” *Mathematical Finance*, 11, 79–96.
- GEMAN, S., AND D. GEMAN (1984): “Stochastic Relaxation, Gibbs Distributions and the Bayesian Restoration of Images,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 6, 721–741.
- HAMMERSLEY, J., AND P. CLIFFORD (1974): “Markov Fields on Finite Graphs and Lattices,” *Unpublished Manuscript*.
- HAUBRICH, J., G. PENNACCHI, AND P. RITCHKEN (2008): “Estimating Real and Nominal Term Structures using Treasury Yields, Inflation, Inflation Forecasts and Inflation Swap Rates,” *Working paper*.
- HEATH, D., R. JARROW, AND A. MORTON (1992): “Bond pricing and the term structure of interest rates: A new methodology for contingent claim valuation,” *Econometrica*, 60(1), 77–105.
- HINNERICH, M. (2008): “Inflation-indexed swaps and swaptions,” *Journal of Banking and Finance*, 32, 2293–2306.
- HUGHSTON, L. (1998): “Inflation Derivatives,” *Working Paper, Merrill Lynch*.
- ISDA (2009): “Derivative Usage Survey,” *International Swaps and Derivatives Association*.
- JACOBS, K., AND L. KAROUI (2009): “Conditional Volatility in Affine Term Structure Models: Evidence from Treasury and Swap Markets,” *Journal of Financial Economics*, 91, 288–318.
- JARROW, R., H. LI, AND F. ZHAO (2007): “Interest Rate Caps ‘Smile’ Too! But Can the LIBOR Market Models Capture It?,” *Journal of Finance*, 62 (1), 345–382.
- JARROW, R., AND Y. YILDIRIM (2003): “Pricing Treasury Inflation Protected Securities and Related Derivatives using an HJM Model,” *Journal of Financial and Quantitative Analysis*, 38 (2), 337–358.

- JOHANNES, M. (2004): “The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models,” *Journal of Finance*, pp. 227–260.
- JOHANNES, M., AND N. POLSON (2003): “MCMC Methods for Financial Econometrics,” *Forthcoming, Handbook of Financial Econometrics*.
- KIM, D. H. (2007): “Challenges in Macro-Finance Modelling,” *BIS Working Papers*, 240.
- KLUGE, W. (2005): “Time-inhomogeneous Lévy processes in interest rate and credit risk models,” Ph.D. thesis, Albert-Ludwigs-Universität Freiburg im Breisgau.
- KRISHNAMURTHY, A. (2010): “How debt markets have malfunctioned in the crisis,” *Journal of Economic Perspectives*, 24, 3–28.
- LIEPPOLD, M., AND L. WU (2002): “Asset Pricing under the Quadratic class,” *Journal of Financial and Quantitative Analysis*, 37, 271–295.
- LITTERMAN, R., AND J. SCHEINKMAN (1991): “Common Factors Affecting Bond Returns,” *The Journal of Fixed Income*, June, 54–61.
- LUND, J. (1997): “Econometric Analysis of continuous-time arbitrage free models of the term structure of interest rates,” *Working paper, Aarhus School of Business*.
- MADAN, D., P. CARR, AND E. CHANG (1998): “The Variance Gamma Process and Option Pricing,” *European Finance Review*, 2, 79–105.
- MADAN, D., AND E. SENETA (1990): “The Variance Gamma (VG) model for share market returns,” *Journal of Business*, 63, 511–524.
- MERCURIO, F. (2005): “Pricing Inflation-indexed derivatives,” *Quantitative Finance*, 5 (3), 289–302.
- MERCURIO, F., AND N. MORENI (2006): “Inflation with a smile,” *Risk*, March 19(3), 70–75.
- (2009): “Inflation modelling with SABR dynamics,” *Risk*, June, 106–111.
- METROPOLIS, N., M. ROSENBLUTH, A. ROSENBLUTH, A. TELLER, AND T. E. (1953): “Equations of State Calculations by fast Computing Machines,” *Journal of Chemical Physics*, 21, 1087–1091.

- MILTERSEN, K., K. SANDMANN, AND D. SONDERMANN (1997): “Closed form solutions for term structure derivatives with log-normal interest rates,” *The Journal of Finance*, 52, 409–430.
- MONROE, I. (1978): “Processes that can be embedded in Brownian motion,” *Annals of Probability*, 6, 42–56.
- NEWKEY, W. K., AND K. D. WEST (1987): “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55(3), 703–708.
- PROTTER, P. (2005): *Stochastic Integration and Differential Equations*, Stochastic Modelling and Applied Probability. Springer Verlag, second edn.
- RAIBLE, S. (2000): “Lévy Processes in Finance: Theory, Numerics, and Empirical Facts,” Ph.D. thesis, Albert-Ludwigs-Universität Freiburg im Breisgau.
- ROBERT, C., AND G. CASELLA (2004): *Monte Carlo Statistical Methods*, Springer Texts in Statistics. Springer Verlag, second edn.
- TRISTANI, O., AND P. HÖRDAHL (2007): “Inflation risk premia in the term structure of interest rates,” *ECB working paper No. 734*.
- (2010): “Inflation risk premia in the US and Euro area,” *ECB working paper No. 1260*.
- TROLLE, A. B., AND E. SCHWARTZ (2010): “An Empirical Analysis of the Swaption Cube,” *Working paper*.
- TROLLE, A. B., AND E. S. SCHWARTZ (2009): “A general stochastic volatility model for the pricing of interest rate derivatives,” *Review of Financial Studies*, 22, 2007–2057.
- VASICEK, O. (1977): “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics*, 5, 177–188.