INSTITUT NATIONAL DE LA STATISTIQUE ET DES ETUDES ECONOMIQUES
Série des Documents de Travail du CREST
(Centre de Recherche en Economie et Statistique)

n° 2010-49

An Analysis of the Ultra Long-Term Yields

S. DUBECQ[1]
C. GOURIÉROUX[2]

The second author gratefully acknowledges financial support of NSERC, Canada, and of the chair AXA/Risk Foundation : “Large Risks in Insurance”. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Banque de France.

1. Banque de France and CREST. simon.dubecq@banque-france.fr

2. CREST and University of Toronto, Canada. christian.gourieroux@ensae.fr

An Analysis of the Ultra Long-Term Yields

The discounting of very long-term cash-flows is crucial for the valuation of long-term investment projects. In this paper, we analyze the market prices of US government bonds with very long-term time-to-maturity, and emphasize some statistical specificities of very long-term zero-coupon rates, that standard Gaussian affine term structure models do not account for. In addition, we describe and estimate three Gaussian Nelson-Siegel affine term structure models, and highlight the model characteristics, which are necessary to match the dynamics of very long-term interest rates.

Keywords : Long Term Interest Rates, Affine Term Structure Model, Nelson-Siegel, STRIPS market.

1 Introduction

The growing exposure of financial institutions (pension funds and insurance companies) to longevity risk, or the assessment of the future costs of climate change calls for a better understanding of the discounting of very long-term cash-flows. In this paper, we analyze the so-called "ultra (or super)" long-term interest rates whose time-to-maturity exceeds the common long-term (usually 10 year) bonds. The most relevant ultra long-term securities which exist in fixed income markets are zero-coupon bonds whose time-to-maturity ranges from 30 years to 50 years[3]. We compare two sources of ultra long-term yields: US Treasury STRIPS bonds, and the zero-coupon yield curves [released by Gurkaynak, Sack, Wright (2007)], which are extracted from US Treasury coupon bonds. In this paper, we focus exclusively on 30 year zero-coupon bonds for the following reasons. First, the 50 year coupon bonds (and their associated STRIPS bonds) are few in the market (they do not exist in the US for instance) and raise higher liquidity concerns than 30 year zero-coupon bonds. Second, default risk starts being a significant concern in this maturity region, even for securities usually considered as risk-free by rating agencies[4], and *a fortiori* more for 50 year bonds than for 30 year bonds. In this respect, we decide to focus the analysis on 30 year zero-coupon bonds as a compromise between the very high time-to-maturity that we want to study and the correction for (perceived) default whose impact grows with the bond's maturity.

The term structure literature already provides theoretical insights on the pricing of very long-term zero-coupon bonds. Let us denote $B(t,h)$ the price of a zero-coupon bond at date t which pays one money unit at date $t+h$, and $r(t,h)$ the corresponding (geometric) yield $r(t,h) = -\frac{1}{h}\log B(t,h)$. Under

[3]We do not consider coupon bonds with infinite maturity, called "perpetuities", whose prices rely heavily on the shortest-term coupons.

[4]Rating agencies distinguish long term and short term rating. For Moody's for instance, "long term ratings are opinions on the relative credit risk of financial obligations with an original maturity of one year or more" [See Moody's (2010)]. Note also that the best ratings (AAA or Aaa) signal minimal credit risk, not zero credit risk. For instance, since the 16th century, the Spanish government has defaulted on its debt every 50 year on average. However, the exact long-term definition of rating agencies is quite imprecise and seems, according to professionals, to depend on the class of risk of the securities considered. It is also noteworthy to mention that regulators only consider short-term ratings in the computation of regulatory reserves' level.

no-arbitrage assumptions, the yield of an hypothetic zero coupon bond with infinite time-to-maturity $r(t,\infty) = \lim_{h\to\infty} r(t,h)$, if this limit exists, cannot decrease over time [see Dybvig, Ingersoll, Ross (1996) and Hubalek, Klein, Teichman (2002)]. Thus we get :

$$r(t_1,\infty) \leq r(t_2,\infty),\ \forall t_1 \leq t_2.$$

Moreover, most of the standard arbitrage-free term structure models imply the "infinite" zero-coupon rate to be deterministic and constant in time [see for instance Vasicek (1977), Cox, Ingersoll, Ross (1985) in continuous time]. In practice, we do not know how large is the distance between the longest time-to-maturity zero-coupon yield existing in the market (10 year, 30 year, 50 year ...) with the "infinite" time-to-maturity rate. Nonetheless we can expect from this constant "infinite" rate that ultra long-term zero-coupon rates would be less volatile than shorter maturity zero-coupon rates and very close in value to long-term zero-coupon yields. We would also expect the spread between ultra long-term and long-term zero-coupon yields to be mainly driven by the latter one.

The function linking at each date the spot zero-coupon rates to their time-to-maturity ($h \to r(t,h), \forall t$) is called the term structure of zero-coupon interest rates, or the zero-coupon yield curve. The standard statistical analysis on the 1Y-10Y segment reveals that at each date, the term structure of zero-coupon rates can be summarized by three components: the general level of zero-coupon yields, the slope of the zero-coupon yield curve (broadly speaking the difference between long-term zero-coupon yield and short-term zero-coupon yield) and the yield curve's curvature (term structure of zero-coupon rates exhibits usually a hump in the 1Y-10Y segment, whose size and location vary at each date t) [see Litterman, Scheinkman (1991)]. The term structure literature also highlights a second hump at the yield curve's long end (for zero-coupon yields with more than 20 year time-to-maturity): according to Sack (2000), zero-coupon yields peaks usually at maturities between 17Y and 25Y and then decline.
Historically in the US, we have observed high (in the early 1980's) and low (between 2000 and 2010) zero-coupon yield curve's level, steep upward (1992), flat (2005-2006) or negative (1981, 2007) yield curve's slope, and small and big hump in the term structure of zero-coupon interest rates. These different term structure patterns seem connected to the economic environment: the level of zero-coupon yields displays a positive correlation with the

US inflation while the zero-coupon yield curve's slope seems related to the economic activity [see Diebold, Rudebusch, Aruoba (2006)]. In particular, inverted yield curves are considered as good predictors of future recession in the US [see Estrella, Hardouvelis (1991)]. On the other hand, the economic interpretation of the zero-coupon yield curve's curvature remains less clear. The unconditional distribution of yields is such that zero-coupon yield curve exhibits on average a positive slope (the yields $r(t,h)$ tend to increase with the maturity h) and its volatility decreases with time-to-maturity (long-term yields are less volatile than short-term ones). Moreover, zero-coupon yields tend to display a high correlation with the zero-coupon yields with close time-to-maturity. In addition, long-term yields tend to be more persistent than short-term yields [see Piazzesi (2010), Dai, Singleton (2003) and Monfort, Pegoraro (2007) for a review of zero-coupon yield curve's stylized facts). Finally, the zero-coupon yield curve's slope predicts future yields movements with the wrong sign regarding the expectation theory: a positive slope predicts future drop in zero-coupon yields. The discrepancy between empirical result and expectation theory increases with the yield's maturity [see Campbell, Schiller (1991)].

The aim of our paper is to analyze the specificities of the ultra long-term zero-coupon rates, compared to the other segments of the zero-coupon yield curve. To do this, we emphasize some stylized facts regarding ultra long-term zero-coupon yields and estimate standard affine no-arbitrage term structure models on a sample of zero-coupon yields with 28 year time-to-maturity. Our goal is to stress the challenges that ultra long maturities raise for the usual term structure models. Nelson, Siegel (1987) formula is one of these term structure models commonly used by practitioners to fit the yield curve[5]. However, it has been shown [see e.g. Bjork, Christensen (1999), Filipovic (1999)] that the Nelson-Siegel term structure model is not arbitrage-free. The class of Nelson-Siegel model has been recently slightly modified to become consistent with no-arbitrage assumption [see Christensen, Diebold, Rudebusch (2010)]. However this extension has the drawback to feature divergent zero-coupon rates at the limit ($r(t,\infty) \to -\infty$) [see Dubecq, Gourieroux (2010)]. In this paper, we propose and estimate three different Affine Nelson-Siegel (ANS henceforth) term structure models, which are con-

[5]The Nelson-Siegel formula is originally a time-invariant formula, designed to fit the yield curve in a parsimonious way at a given date t.

sistent with no-arbitrage assumptions, and discuss their respective ability to model the dynamics of ultra long term rates. Finally, we present and compare the estimation results obtained for both datasets.

The rest of the paper is organized as follows: Section 2 presents the data and the stylized facts regarding ultra long-term zero-coupon yields and the difference between 28 year zero-coupon yields extracted from STRIPS bonds and the ones extracted from coupon bond prices. Section 3 describes the term structure models estimated on our yield sample: the original Nelson-Siegel model, and the extended Nelson-Siegel models. Section 4 recapitulates the estimation results, for both types of ultra long yields. Finally, Section 5 concludes. The proofs and description of the estimation techniques are gathered in appendices.

2 Data and Stylized Facts

2.1 Data

We consider two samples of 4 zero-coupon yields of US Treasury bonds with different maturities: 1 year (1Y), 5 year (5Y), 10 year (10Y) and 28 year (28Y). The samples differ in the computation of the 28Y yield, which is obtained *via* two different datasets. The first dataset is obtained from Datastream quotes of Principal Only (PO thereafter) STRIPS bonds. The second one is obtained from the work of Gurkaynak, Sack, Wright (2007) [GSW henceforth] available on the website of the Board of Governors of the Federal Reserve System. Both datasets differ regarding their origins and their computation methods.

PO STRIPS dataset

STRIPS bonds (STRIPS stands for Separate Trading of Registered Interest and Principal Securities) are zero-coupon bonds which pay to the STRIPS bond's holder a single payment (equal to the face value of the underlying bond) at the maturity of the STRIPS bond. STRIPS bonds are not issued directly by the US Treasury, but any holder of a sufficient amount of a same coupon bond can ask for the stripping of his/her bonds to the New-York Federal Reserve [see Jordan, Jorgensen, Kuipers (2000) and Sack (2000) for

details on the functioning of the STRIPS program][6]. Each separated (or stripped) coupon bond's cash-flows (interest and principal payment) can then be sold individually as new zero-coupon bonds. For instance, a newly issued 30Y coupon bond with semi-annual coupon payments generates 61 STRIPS securities, 60 Coupon STRIPS and 1 Principal Only (PO) STRIPS for the final payment of the bond's face value.

PO STRIPS and Coupon STRIPS with same payment date and same face value should have the same price. The literature, however, reports significant differences in the PO and Coupon STRIPS prices and explains this discrepancy by the stronger constraints put on PO STRIPS. Indeed, Coupon STRIPS with same payment date, and which are stripped from different coupon bonds, can be aggregated to reconstitute the coupon payments of any coupon bond, while the reconstitution of the payment of the coupon bond's face value require specifically the PO STRIPS initially stripped from the security to reconstitute[7]. However this price difference is not observed for STRIPS whose maturities exceeds 20 years, as the ones we want to study [see Figure 3 in Sack (2000)]. Moreover, the outstanding value of PO STRIPS, with more than 25Y residual time-to-maturity, is necessarily higher than for Coupon STRIPS with same payment dates. Therefore, in this application, we consider only STRIPS associated with the *in fine* payment of the principal of Treasury bonds.

The US Treasury distinguishes 3 types of marketable fixed income securities that he issues: Treasury Bills with less than 1 year time-to-maturity, Treasury Notes, whose time-to-maturities range from 2 years to 10 years, and Treasury Bonds, whose maturities exceeds 10 years. Figure 1, from GSW (2007), graphically shows the structure of the US public debt from 1962 to 2005. In particular, we see from Figure 1 that the US Treasury has gradually increased the maximal time-to-maturity of its issued bonds: 10Y notes were firstly introduced in the 1960's and are regularly issued since the mid-1970's, 15Y bonds were introduced in the early 1970's, and replaced by 20Y bonds in the early 1980's. 30Y bonds took the place of 20Y bonds from 1985 to 2001. In 2001, the US Treasury decided to stop the issuance of 30Y bonds, arguing that the gradual disappearance (at that time) of public US debt,

[6]The STRIPS program started in January 1985. Until 1997, only bonds, whose initial time-to-maturities were more or equal to 10 years, were eligible to be stripped. The reconstitution of coupon bonds from STRIPS is possible since January 1987.

[7]This constraint insures a minimal degree of liquidity for all coupon T-bonds, or in other words, that investors cannot modify the relative supply of coupon T-bonds.

together with the lower demand for very long-term bonds, made the issuance of 30Y bonds useless. The issuance of 30Y bonds restarted in 2006 for the converse reasons: the US government budget's deficit had become large and the demand of investors for very long-term bonds had renewed[8].
Table 1 present the repartition, in January 1999, of the stock of notes and bonds whose maturities exceeds 2Y. At that time, 30Y bonds represented 31% of the total of notes and bonds issued. Statistics from the *Monthly Statement of the Public Debt* reported in Table 2 show that Treasury bonds with more than 10 year maturities represented between 15% and 21% of all the marketable securities issued by the US Treasury (including Treasury bills). Treasury notes had a share ranging from 55% to 61% over the 1990-1999 period. In addition, the average time-to-maturity of the US public debt was 5.7 years in 1999 [see OECD (2009)], which is more than twice the average time-to-maturity of the US public debt in the 1970's (which reached a minimum of 30 months in 1974, see Sill (1994)).
We also report in Table 1 that the size of STRIPS bonds' market was significant in 1999, especially for coupon bonds whose time-to-maturity exceeds 10Y: about 35% of these bonds were held in stripped form, compared to about 10% for the 10Y coupon bonds. Besides, STRIPS and coupon bonds markets are highly integrated. Sack (2000) reports significant flows of stripping and reconstitution activity between the markets and finds very few stripping/reconstitution arbitrage opportunities. Moreover, Sack (2000) refers to anecdotal reports suggesting that the liquidity in the STRIPS market was analogous to the one of off-the-run bonds, except for the less than 10Y maturities.
From January 1988 to February 1998, we extract each Friday the (annual) arithmetic yield of the longest time-to-maturity STRIPS bonds quoted on the market. In our sample period, the US Treasury bonds which had the longest time-to-maturity were 30Y bonds (see Figure 1). From these zero-coupon yields, we compute weekly observations of the constant time-to-maturity 28Y[9] yields as a weighted average of the yields of STRIPS bonds with the

[8]The frequency of issuance differs with the bond time-to-maturity: Treasury notes, whose maturities range from 2Y to 7Y, are issued monthly, while 10Y notes are issued 8 times a year and 30Y bonds issuance is quarterly.

[9]The choice of 28Y time-to-maturity zero-coupon yields (rather than longer-term yields) excludes *de facto* "on-the-run" and "first off the run" issues that often trade at a premium, due to their higher liquidity, compared to other securities. This approach is also followed by GSW (2007).

closest time-to-maturity. For instance, we compute the 28Y yield at January 8th, 1988 as a weighted average of the (observed) yields on STRIPS bonds with maturities 11/15/2015 (27 years 11 months residual maturity) and 02/15/2016 (29 years 1 month residual maturity). We did not compute the constant time-to-maturity 30Y yield since we never observe STRIPS bonds with more than 30 years residual maturity. Finally, we obtained 528 weekly observations from January 8, 1988 to February 13, 1998[10].

GSW dataset

GSW provide an extensive dataset of zero-coupon yields on US Treasury bonds from 1961 to 2010. The data frequency is business-daily. The zero-coupon yields are computed in order to fit the observed coupon bonds' prices whose maturities range from 1Y to 30Y (the 21Y to 30Y series starts in November 1985).

GSW procedure relies upon the yield formula from Nelson-Siegel [NS] (1987) for zero-coupon bonds from 1961 to 1980, and Svensson [S] (1994) for zero-coupon bonds after 1980. These formulas are respectively:

$$r^{NS}(t,h) = X_{1t} + X_{2t}\frac{1-\exp\{-hX_{4t}\}}{hX_{4t}} + X_{3t}\left(\frac{1-\exp\{-hX_{4t}\}}{hX_{4t}} - \exp\{-hX_{4t}\}\right) \tag{1}$$

and

$$\begin{aligned} r^{S}(t,h) &= X_{1t} + X_{2t}\frac{1-\exp\{-hX_{4t}\}}{hX_{4t}} + X_{3t}\left(\frac{1-\exp\{-hX_{4t}\}}{hX_{4t}} - \exp\{-hX_{4t}\}\right) \\ &+ X_{5t}\left(\frac{1-\exp\{-hX_{6t}\}}{hX_{6t}} - \exp\{-hX_{6t}\}\right) \end{aligned} \tag{2}$$

$\forall h \in \mathbb{R}$, where $r^{NS}(t,h)$ (resp. $r^{S}(t,h)$) are the zero-coupon rates with time-to-maturity h at each date t according to the Nelson-Siegel (resp. Svensson) formula, and the $(X_{it})_{i=1...6}$ are time dependent parameters (see Appendix 2 for a detailed description of GSW procedure).

These functions are widely used by practitioners to smooth the term structure of (zero-coupon) yields at a given date [see for instance the BIS (2005) report

[10] Our sample stops in 1998 while the US Treasury continued to issue 30Y bonds until 2001. Unfortunately, Datastream does not provide quotes for PO STRIPS with more than 28Y time-to-maturity after February 13, 1998.

on the popularity of Nelson-Siegel/Svensson model among central banks]. In their application, GSW estimate at each quoting date (note that the NS model is originally presented in a cross-sectional framework, without time index) the $[X_{1t}, X_{2t}, X_{3t}, X_{4t}, X_{5t}, X_{6t}]$, which best fit the observed prices of off-the-run, non-callable US government bonds. GSW show that this model provide an excellent fit of the observed coupon bond prices. Nevertheless, by imposing a specific structure on the zero-coupon yield curve, NS formula may induce errors in the measurement of the underlying zero-coupon yields, in particular for long-term zero-coupon yields [see Appendix 2]. GSW recognize indeed that "the largest fitting errors tend to be seen at the longest maturities" [see Gurkaynak, Sack, Wright (2007) p. 2298].

Therefore, the remaining zero-coupon yields (1Y, 5Y, 10Y) in both samples come from the Gurkaynak, Sack, Wright dataset. We keep in the STRIPS sample the GSW zero-coupon yields with less than 10Y maturities because of the low liquidity of STRIPS bonds with these maturities [see Sack (2000) and above] and because the measurement errors problem discussed above is less a concern for the 1Y-10Y segment [see Appendix 2].

2.2 On the specificities of ultra long-term yields

In this section, we provide a preliminary analysis of the statistical properties of zero-coupon yields which emphasizes the specificities of the 28Y zero-coupon yield compared to the other zero-coupon yield curve's segments. Since the 1Y, 5Y and 10Y zero-coupon yields are the same in both sample, we present two sets of results only for the 28Y zero-coupon yields. Figure 2 plots the evolution of the zero-coupon yields since the beginning of our sample, i.e.January 1988. For the sake of completeness, we also plot the GSW zero-coupon yields until October 2010 without the STRIPS 28Y zero-coupon yields (whose sample ends in February 1998). However, our summary statistics will be based on the sample period common to both datasets, i.e. from January 1988 to February 1998. We observe a slight decreasing trend common to all the zero-coupon yields, with cyclical variations whose magnitude is higher for short-term zero-coupon yields. Figure 2 also shows that this difference in the magnitude of the yields' cyclical variation is not a feature of the whole sample: several times, for instance before June 1990 and between 1994 and 2000, the long-term zero-coupon yields seem as volatile as the short-term ones. In spite of the decreasing trend in zero-coupon yields, we plot in Figure

3 the autocorrelation functions of the zero-coupon yields (up to 120 lags, i.e. about 2 years lags), which exhibits a high persistence for all yields due to the trend effect. Moreover, Figure 3 reveals lower autocorrelation at higher lags (lags are in weeks) for each zero-coupon yields, in particular for the 10Y and 28Y zero-coupon yields. This feature reverses after 70 lags (about 1 year and 4 months) where the autocorrelation function of 10Y and 28Y starts increasing and peaks at 110 lags, while the autocorrelation of shorter-term yields continue to decrease. This reversal in the autocorrelation function emphasizes a common cycle, with about 2 years and 1 month periodicity, which affects mainly the long-term and ultra long-term zero-coupon yields.
Table 3 recapitulates the first four moments of the unconditional distribution of zero-coupon yields, with the 5% and 95% quantiles (Table 4 presents the correlation matrix of the zero-coupon yields), while we plot in Figure 4 the empirical (unconditional) distribution of the 1Y, 5Y, 10Y and 28Y zero-coupon yields for both datasets. Table 3 and Figure 4 show that the unconditional mean of yields, including the 28Y, increases with the time-to-maturity. In addition, the zero-coupon yield's volatility is a decreasing function of the time-to-maturity (in particular the 28Y yield volatility is about 75% of the 10Y yield volatility). These first results are in line with the stylized facts mentioned in the introduction. The specificity of the 28Y zero-coupon yields arises at higher order moments: contrary to shorter-term zero-coupon yields, 28Y are negatively skewed, and while the yield kurtosis tends to decrease on the 1Y-10Y segment, the 28Y zero-coupon yield kurtosis is higher than the 5Y one.
Statistics on the spreads 5Y-1Y, 10Y-5Y and 28Y-10Y (this last one for both datasets) are given in Table 5. The 28Y-10Y spread is on average smaller than the 10Y-5Y one, but significant (30 basis points on average) and about as volatile as the 10Y-5Y spread. Moreover, the 28Y-10Y spread is less rightly-skewed than the other spreads and exhibits more extreme negative values compared to the 10Y-5Y and the 5Y-1Y spreads. The correlation matrix of the spreads, reported in Table 6, highlights the relative disconnection of the 28Y-10Y spread with the other (10Y-5Y, 5Y-1Y) spreads: the correlation of the 28Y-10Y with the 10Y-5Y and 5Y-1Y is about 0.6 while the 10Y-5Y and 5Y-1Y exhibit a 0.9 correlation coefficient. We report in Figure 5 the evolution of the 5Y-1Y, 10Y-5Y and 28Y-10Y spreads over our samples for both datasets. Figure 5 highlights the specificity of the 28Y-10Y spread compared to the 10Y-5Y and 5Y-1Y spreads. While the 10Y-5Y and 5Y-1Y exhibits common cyclical variations (more pronounced for the 5Y-1Y

spread) and a slight decreasing trend, the 28Y-10Y spread moves sometimes in opposite directions (in particular before October 1991 and between July 1994 and December 1996) and displays an upward trend. In addition, we show in Figure 6 the cross-correlogram of the 28Y-10Y spread versus the 10Y or 28Y yields for both samples. Interestingly, the correlation between the 28Y-10Y spread and the 28Y yield is -0.57 for both datasets which is another indicator of the non-constancy of the ultra long-term yields.
Finally, let us focus on the pattern of the zero-coupon yield curve and emphasize the regularities in the ordering of the 1Y, 5Y, 10Y and 28Y zero-coupon yields. The number of possible orderings is $4! = 24$, but only 8 different types of yield curve shapes are observed more than 1% of the time in our sample. The ordering with their underlying frequency are presented in Table 8 for both datasets. About 70% of the observed slope of yield curve's long end is positive (1Y yield $<$ 5Y yield $<$ 10Y yield $<$ 28Y yield). The second most frequent observed pattern displays a hump at the 10Y time-to-maturity (1Y yield $<$ 5Y yield $<$ 28Y yield $<$ 10Y yield), while we observe rare events of negative slope (28Y yield $<$ 10Y yield $<$ 5Y yield $<$ 1Y yield) and very rare episodes of double hump (28 Y yield $<$ 5Y yield $<$ 10Y yield $<$ 1Y yield)).
To sum up, the broad picture drawn by these preliminary results seems to reject the assumption of constant ultra long-term rates. Besides, the specific analysis of the 28Y-10Y yields spread emphasizes the relative singularity of the 28Y yield *vis à vis* the shorter maturity segments of the yield curve which justifies a specific modeling of specific ultra long-term zero-coupon yields.

2.3 Comparison of GSW and STRIPS 28Y yields

At first sight, GSW and STRIPS 28Y yields series appear very close (see Figure 2), with similar autocorrelation functions (see Figure 3). Again we note the effect of the decreasing trend effect in the autocorrelation functions. As it can be shown from Tables 3, 4, 5, and Figure 4, both datasets induce the same conclusions regarding the historical moments of the yields and the spread between ultra long-term yields and long-term yields.
Nevertheless, a closer look at the STRIPS and GSW series show that the discrepancy between the datasets can be quite significant, ranging from -40 basis points to +50 basis points (see Figure 7). The estimated autocorrelation function in Figure 8 of the difference STRIPS - GSW 28Y zero-coupon yields emphasizes the persistence of this difference, at a lower degree of persistence than yields *per se* (see Figure 3).

Several reasons can explain the discrepancy between the yields extracted from the STRIPS bond prices and the ones from the GSW dataset. First, differences in the liquidity of coupon bond and STRIPS bond markets could provide a valid explanation for the differences in STRIPS and GSW yields. However, it has been shown in Section 2.1 that the liquidity of STRIPS bonds with more than 10Y maturities was actually fairly high and that STRIPS and coupon bonds markets were highly integrated.
Second, the exclusion of 20Y coupon bonds from the GSW dataset after 1996 (because of their relative cheapness to other coupon bonds) could explain partly the STRIPS-GSW discrepancy.
Third, zero-coupon yields extracted from the GSW procedure are smoothed yields (via Nelson-Siegel/Svensson formula, see above) and not raw yields as the one extracted from STRIPS bonds prices. GSW (2007) provides graphical evidences that the pricing errors, in absolute values, of coupon bond prices were very small with their procedure, whatever the coupon bond maturity. In particular, they show that, at each observation date, the average of the pricing errors in absolute value for coupon bonds whose maturity ranges from 1Y to 30Y is very small (less than 0.2 basis points for all the 1Y-5Y, 5Y-10Y, 10Y-20Y and 20Y-30Y yield curve segments after 1988). Nevertheless, small pricing errors on coupon bonds do not prevent for large errors on the related zero-coupon bonds, especially for the longest maturities (see Appendix 2). In particular, the decreasing zero-coupon yields at the yield curve's long end highlighted in Sack (2000) contradicts the assumption of the NS formula, according to which zero-coupon yields asymptote to a constant level. Moreover, Nelson-Siegel (and Svensson) formula are not consistent with the no-arbitrage assumption [see Filipovic (1999) and Section 3 below]. Therefore, the zero-coupon yield curves obtained by GSW (2007) are not arbitrage-free neither, while PO STRIPS bond yields are directly quoted in (relatively) active markets [see Section 2.1]. The model based extraction contains thus measurement errors, and these errors can be very different for short and long maturities. Finally, part of the STRIPS-GSW difference could be related to differences in the taxation of these two securities [see Sack(2000)].
These reasons lead us to focus mainly our subsequent analysis on the STRIPS sample[11] which includes 1Y, 5Y, 10Y zero-coupon yields from GSW dataset and 28Y zero-coupon yield from quotes of PO STRIPS bonds.

[11] We will still report the estimation results for the GSW 28Y yields as a robustness test.

3 The Term Structure Models

Let us introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration (I_t) which defines the sequence of information sets, generated by a (multivariate) random variable X_t, with dimension N, called factor or state vector. $\mathbb{P}$ is the so-called historical probability measure. We also consider the risk-neutral distribution $\mathbb{Q}$, equivalent to the $\mathbb{P}$ distribution, and such that :

$$B(t,h) = E_t^{\mathbb{Q}}[B(t,1)B(t+1,h-1)]. \tag{3}$$

Standard asset pricing theory proves the existence of a risk-neutral distribution under the Absence of Arbitrage Opportunity (AAO) [see Hansen, Richard (1987), Harrison, Kreps (1979)].
In this application, we only consider term structure models which belongs to the widely used affine class, in which interest rates are affine functions of the factors [see Appendix 1 for a presentation of Affine Term Structure Models, ATSM henceforth]:

$$r(t,h) = \alpha(h)X_t + \beta(h), \tag{4}$$

where $\alpha(h)$ is 1xN, X_t is Nx1, $\beta(h)$ is a scalar. Thus, in a N-dimensional ATSM, the term structure is a combination of the $N+1$ baseline term structures corresponding to the components of $\alpha(h)$ with time dependent (stochastic) coefficients and to $\beta(h)$, which accounts for time independent risk premium due to the stochastic fluctuations of factors X_t.

3.1 The Filipovic-Nelson-Siegel Term Structure Model

We first consider as a benchmark the basic Nelson-Siegel (NS henceforth) formula. In their 1987 paper, Nelson and Siegel propose a parsimonious term structure function [see equation (1)]. The NS model is originally presented in a cross-sectional framework, without time index, but is generally estimated daily (or monthly) on observed yield curves as if its parameters $X_{1t}, X_{2t}, X_{3t}, X_{4t}$ were actually time-dependent factors. With this practice, the NS model becomes a complicated 4-factor model, which is not affine in X_{4t}. It has been shown in [Filipovic (1999), theorem 4.1] that a necessary condition for this dynamic NS model to be arbitrage free is that X_{4t} is a time-independent parameter. Then, for constant $X_{4t} = \mu$, the NS model becomes a Filipovic-Nelson-Siegel (FNS thereafter) affine term structure model

with factors X_{1t}, X_{2t}, X_{3t} and coefficients

$$\alpha^{FNS}(h) = \left(\begin{array}{ccc} 1 & \frac{1-\exp[-\mu h]}{\mu h} & \frac{1-\exp[-\mu h]}{\mu h} - \exp[-\mu h] \end{array} \right), \beta^{FNS}(h) = 0.$$

The shape of the baseline term structures $\alpha^{FNS}(h)$ leads to interpret the factors X_{1t}, X_{2t}, X_{3t} as level, slope and curvature factors -respectively- in reference to the principal components analysis of the yield curve [see Litterman, Scheinkman (1991), Jones (1991)]. The first baseline term structure ($\alpha_1^{FNS}(h)$) is flat and moves by parallel shift. The second baseline term structure is a monotonic function of the yield-to-maturity, such as the factor's value X_{2t} mainly affects the short maturities (small h). Therefore the second baseline directly impact the slope of the term structure. Finally, the third baseline term structure is able to generate "humped" yield curve [see Figure 10]. The FNS term structure model is thus able to generate zero-coupon yield curves with high or low level of zero-coupon yields, upward, flat or downward slope and one hump[12]. However, even if the size of this hump can vary over time, its location is time-independent due to the constancy of parameter μ, in contradiction with observed zero-coupon yield curve's dynamics [see the third part of the introduction above].

Nevertheless, the FNS model (and FS model) is incomplete because of the absence of time independent risk-adjustment term $\beta^{FNS}(h) = \beta^{FS}(h) = 0$ and is still not compatible with no-arbitrage [see Bjork, Christensen(1999) and Filipovic (1999)].

3.2 Affine Nelson-Siegel Term Structure Model

In this section, we present different Gaussian Affine Term Structure Models (GATSM henceforth) [see Appendix 1 for details on GATSM], which are arbitrage-free and share some baseline term structures with the basic Filipovic-Nelson-Siegel model. As the FNS model, we consider essentially 3-factor models (i.e. GATSM with 4 baseline term structures).

[12]The Filipovic-Svensson (FS) extension of the FNS formula yield the following factor loadings :
$\alpha^{FS}(h) = \left(\begin{array}{cccc} 1 & \frac{1-\exp[-\mu_1 h]}{\mu_1 h} & \frac{1-\exp[-\mu_1 h]}{\mu_1 h} - \exp[-\mu_1 h] & \frac{1-\exp[-\mu_2 h]}{\mu_2 h} - \exp[-\mu_2 h] \end{array} \right)$, taking $X_{4t} = \mu_1$ and $X_{6t} = \mu_2$ in equation (2). This extension allows for a second hump in the original Nelson-Siegel formula thanks to the fourth component of $\alpha^{FS}(h)$.

More precisely, we assume that the short-term rate is an affine function of the 3-dimensional factor :

$$r(t,1) = \alpha(1)X_t + \beta(1), \tag{5}$$

and we assume that the short-term rate is a 3-variate Gaussian autoregressive process of order 1 under the risk-neutral distribution $\mathbb{Q}$:

$$X_{t+1} = \Phi X_t + \Sigma\epsilon_t, \text{ with } \epsilon_t \sim \mathcal{N}(0, Id). \tag{6}$$

The underlying factors are defined up to a one-to-one linear transformation. Therefore, without loss of generality, we can assume that factors have mean zero, and that the autoregressive matrix is written under its Jordan representation, that is, with its eigenvalues on the main diagonal, 0 or 1 on the diagonal just above the main diagonal, and zero elements anywhere else.
In the next subsections, we present 3 variants of this triangular GATSM with different constraints on the eigenvalues of the matrix Φ under the risk-neutral distribution.

3.2.1 Affine NS Model of type 1 (ANS1)

We first assume that all the eigenvalues of the (3x3) matrix Φ are distinct with modulo less than 1. In this case :

$$\Phi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix} \text{ with } \begin{array}{rcl} |\lambda_1| & < & 1 \\ |\lambda_2| & < & 1 \\ |\lambda_3| & < & 1 \end{array} \tag{7}$$

We derive from Proposition A2 in Appendix 1 the yield formula in this specific GATSM :

$$\begin{aligned} \alpha^{ANS1}(h) &= \begin{pmatrix} \alpha_1^{ANS1}(h) & \alpha_2^{ANS1}(h) & \alpha_3^{ANS1}(h) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\alpha_1^{ANS1}(1)(1-\lambda_1^h)}{h(1-\lambda_1)} & \frac{\alpha_2^{ANS1}(1)(1-\lambda_2^h)}{h(1-\lambda_2)} & \frac{\alpha_3^{ANS1}(1)(1-\lambda_3^h)}{h(1-\lambda_3)} \end{pmatrix} \\ \beta^{ANS1}(h) &= \beta^{ANS1}(1) \\ &- \frac{1}{2h}\sum_{k=1}^{h-1} \alpha^{ANS1}(1)(I-\Phi)^{-1}(I-\Phi^k)\Sigma\Sigma'(I-\Phi^k)'(I-\Phi)'^{-1}\alpha'^{ANS1}(1) \end{aligned} \tag{8}$$

We plot in the upper panel of Figure 11 the baseline term structure components $\alpha^{ANS1}(h)$ as a function of the time-to-maturity h. We observe that

these baseline term structures are monotonous functions of the yields maturity, which mainly affect the short-end of the yield curve, as the slope baseline term structure in the FNS model.
All the factor loadings $\alpha^{ANS1}(h)$ tend to 0 when the yield time-to-maturity increases. Therefore, the ANS1 model will induce more fluctuations in the short-term yields than in the long-term yields. Moreover, we see from Figure 11 that the higher the eigenvalue λ_k, the slower is the decrease in the associated factor loading $\alpha_k^{ANS1}(h)$. Therefore, a higher eigenvalue leads to baseline term structure which affects more uniformly the yield curve, and thus influences more the yield curve's long end than the other baseline term structures do. Finally, this specification combines to the time dependent components $\alpha^{ANS1}(h)X_t$ one time independent baseline $\beta^{ANS1}(h)$, which is absent from the previous FNS model.

3.2.2 Affine NS Model of type 2 (ANS2)

In this specification, we assume that the matrix Φ has at most 2 distinct eigenvalues, with modulus less than 1 and cannot be diagonalized. Formally :

$$\Phi = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 1 \\ 0 & 0 & \lambda_2 \end{pmatrix} \text{ with } \begin{array}{rcl} |\lambda_1| & < & 1 \\ |\lambda_2| & < & 1 \end{array} \tag{9}$$

The power of Φ is easily derived

$$\Phi^h = \begin{pmatrix} \lambda_1^h & 0 & 0 \\ 0 & \lambda_2^h & h\lambda_2^{h-1} \\ 0 & 0 & \lambda_2^h \end{pmatrix}$$

The underlying zero-coupon yields are given by :

$$\begin{aligned} \alpha^{ANS2}(h) &= \begin{pmatrix} \alpha_1^{ANS2}(h) & \alpha_2^{ANS2}(h) & \alpha_3^{ANS2}(h) \end{pmatrix} \\ &= \begin{pmatrix} \frac{\alpha_1^{ANS2}(1)(1-\lambda_1^h)}{h(1-\lambda_1)} \\ \frac{\alpha_2^{ANS2}(1)(1-\lambda_2^h)}{h(1-\lambda_2)} \\ \frac{1-\lambda_2^h}{h(1-\lambda_2)}\left(\frac{\alpha_2^{ANS2}(1)}{(1-\lambda_2)} + \alpha_3^{ANS2}(1)\right) - \frac{\alpha_2^{ANS2}(1)\lambda_2^h}{\lambda_2(1-\lambda_2)} \end{pmatrix}' \end{aligned} \tag{10}$$

$$\begin{aligned} \beta^{ANS2}(h) &= \beta^{ANS1}(1) \\ &- \frac{1}{2h}\sum_{k=1}^{h-1} \alpha^{ANS2}(1)(I-\Phi)^{-1}(I-\Phi^k)\Sigma\Sigma'(I-\Phi^k)'(I-\Phi)'^{-1}\alpha'^{ANS2}(1) \end{aligned}$$

The middle panel of Figure 11 displays the factor loadings $\alpha^{ANS2}(h)$ for the ANS2 model. We observe that the two first factor loadings are continuously decreasing as the three ones in the ANS1 model, and as the slope factor loading in the FNS model. However, this specification implies a "'humped"' baseline term structure due to a difference between the multiplicity order of eigenvalue λ_2 and the dimension of its associated eigenspace[13], which is similar to the third baseline in the FNS model.

3.2.3 Affine NS model of type 3 (ANS3)

Finally, we constrain the matrix Φ to have only one distinct eigenvalue with modulus less than 1, and a dimension of the eigenspace equal to 1.. In this case,

$$\Phi = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \text{ with } |\lambda| < 1 \text{ and } \Phi^h = \begin{pmatrix} \lambda^h & h\lambda^{(h-1)} & h(h-1)\lambda^{(h-1)}/2 \\ 0 & \lambda^h & h\lambda^{(h-1)} \\ 0 & 0 & \lambda^h \end{pmatrix}. \tag{11}$$

which induces the following baseline term structure coefficients :

$$\alpha^{ANS3}(h) = \begin{pmatrix} \frac{\alpha_1^{ANS3}(1)(1-\lambda^h)}{h(1-\lambda)} \\ \frac{1-\lambda^h}{h(1-\lambda)}\left(\frac{\alpha_1^{ANS3}(1)}{(1-\lambda)} + \alpha_2^{ANS3}(1)\right) - \frac{\alpha_1^{ANS3}(1)\lambda^h}{\lambda(1-\lambda)} \\ \frac{1-\lambda^h h}{h(1-\lambda)}\left(\frac{\alpha_1^{ANS3}(1)}{(1-\lambda)^2} + \frac{\alpha_2^{ANS3}(1)}{(1-\lambda)} + \alpha_3^{ANS3}(1)\right) \\ -\frac{\lambda^h}{\lambda(1-\lambda)}\left(\frac{\alpha_1^{ANS3}(1)(3\lambda-1)}{2\lambda(1-\lambda)} + \alpha_2^{ANS3}(1)\right) - \frac{\alpha_1^{ANS3}(1)h\lambda^h}{2\lambda^2(1-\lambda)} \end{pmatrix}' \tag{12}$$

$$\begin{aligned} \beta(h)^{ANS3} &= \beta^{ANS3}(1) \\ &- \frac{1}{2h}\sum_{k=1}^{h-1} \alpha^{ANS3}(1)(I-\Phi)^{-1}(I-\Phi^k)\Sigma\Sigma'(I-\Phi^k)'(I-\Phi)'^{-1}\alpha'^{ANS3}(1) \end{aligned}$$

[13]This difference generates "resonance" among the factors: their cross correlation function reaches a maximum at intermediate lags [see Gourieroux, Jasiak (2001)].

We highlight in the lower panel of Figure 11 the shape of the baseline term structures $\alpha^{ANS3}(h)$ as a function of the yield maturity h. The ANS3 specification introduces a second "humped" dynamic baseline in the yield curve modeling. This second hump does not exist in the previous ANS1, ANS2 and FNS models[14] and makes the ANS3 model able to generate yield curves with 2 humps, whose sizes are time dependent. As the other Affine Nelson-Siegel models, the ANS3 model introduce a time independent baseline $\beta^{ANS3}(h)$ to correct for the risk of factors fluctuations.

Therefore, all three Affine NS models presented in this section rely on three dynamic baseline term structures and one time independent baseline. For some models (ANS2 and ANS3), these dynamic baselines feature humps, when the multiplicity order of some eigenvalues is not equal to the dimension of the associated eigenspace.
In all ANS models, these dynamic baseline term structures become null at high yields time-to-maturities, which implies that, at the limit, the very long-term zero-coupon yields are exclusively defined by the time independent baseline $\beta(h)$.

Proposition 1

In ANS1, ANS2 and ANS3, where each eigenvalue has a modulus less than one, the limit of zero-coupon yields $\lim_{h\to\infty} r(t,h) = r(t,\infty)$ exists and is equal to :

$$r(t,\infty)^j = \beta^j(1) - -\frac{1}{2}\alpha^j(1)(I-\Phi)^{-1}\Sigma\Sigma'(I-\Phi)'^{-1}\alpha'(1)^j \qquad (13)$$

where j can be ANS1, ANS2, ANS3.

3.2.4 A Limiting Case

It turns out, as emphasized in Christensen, Diebold, Rudebusch (2010), that a modified ANS2 model with $\lambda_1 = 1$ is able to reproduce the (first) FNS dynamic baseline term structure, which makes the whole zero-coupon yield curve move by parallel shifts. However, this modified ANS2 model has the

[14]The FS extension also introduces a second hump in a 4-factor model.

drawback of infinite long-term rate due to the divergence of the time independent risk-adjustment term $\lim_{h\to\infty} \beta(h) = -\infty$ [see Dubecq, Gourieroux (2010)].

Proposition 2

In the modified ANS2 model with $\lambda_1 = 1$:

i) the baseline dynamic term structure associated with the unitary eigenvalue is flat and moves by parallel shifts.

ii) The limit of the zero-coupon yields $\lim_{h\to\infty} r(t,h)$ does not exist.

4 Estimation Results

In this section, we present the results of the estimation of the term structure models described in Section 3.
In the FNS model, we have to estimate 1 time independent parameter μ and 3x528[15] time dependent factors. We will apply a nonlinear least squares approach (see Appendix 2). Since, for given μ, the rates $r(t,h)$ are affine functions of the factors X_t, we can concentrate the least-squares criterion with respect to μ. More precisely, the factors can be estimated *via* OLS conditionally on μ. Then, we look for the value of the parameter μ, which minimizes globally the sum of squared errors on the yields, the factors X_t being estimated via OLS conditionally on the μ chosen.
The other ANS models are estimated in the same way. We look for the value of the time independent parameters (λ[16], Σ and $\beta(1)$)[17], which minimize the sum of squared errors, the time dependent factors X_t being estimated *via* OLS conditionally on the value of the time independent parameters.

The time-independent parameters, estimated on the STRIPS dataset, are presented in Table 9 for each model[18], while the underlying baseline term

[15]There are 528 weekly observations in each sample, see Section 2.1

[16]More precisely, the ANS1 (resp. ANS2) models involves λ_1, λ_2, and λ_3 (resp. λ_1 and λ_2).

[17]For identification purpose, $\alpha(1)$ is set to $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ for all ANS models.

[18]To save space, the standard deviations of parameters estimates are not reported. There

structures $\alpha(h)$ and $\beta(h)$ are displayed in Figures 12 and 13. We report the same results for the GSW dataset in Table 10, and Figures 14 and 15. The estimated shape of the time dependent baseline term structures $\alpha(h)$ confirms the conclusions of Section 3. We observe that the first baseline term structure decrease very slowly, in all ANS models, due to the very high first eigenvalue. All eigenvalues are close to 1, but this can be due to the rather high frequency of weekly observations. For the sake of comparison with other term structure models estimated at lower frequency, Table 9 reports the monthly, quarterly and yearly equivalent of the estimated λ's, i.e. λ^4, λ^{13} and λ^{52}, respectively. After this adjustment for the observation frequency, we observe that in all ANS models, the largest eigenvalue is indeed close to 1, but not the other ones in models ANS1 and ANS2. The slow decay in the first baseline term structure makes all the estimated ANS models able to generate 10Y and 30Y rates, which are sufficiently varying. For ANS3 model, this slow decay exists also in the second and third baseline term structures. This allows humps at rather long time-to-maturity. This feature is not shared by the other ANS1 and ANS2 models, due to the very rapid decay of their second and third baseline term structures. For instance, the third baseline term structure in ANS2 is humped at about 8 months maturity. Thus, the estimated models seem to have very different features. At this stage, it is important to note that the three models are nonnested by definition of their Jordan representation.
The eigenvalues are not the only determinants of the behavior of the long-term rates, which also depend on the time independent baseline term structure $\beta(h)$. This function is determined by all the time independent parameters (λ, Σ and $\beta(1)$) [see the rates formulas in Section 3.2]. For instance, in the ANS3 model, the $\beta(h)$ baseline is almost flat around $\beta(1)$, due to the small risk-neutral volatility Σ of the factors.

We also report the Root Mean Squared Errors (RMSE) as well as summary statistics on the errors on the zero-coupon rates in Tables 11 and 12 (13 and 14 for the GSW sample). We plot the time evolution of these errors, and their empirical distributions, in Figures 16 and 17 (18 and 19 for the GSW results). These Tables and Figures emphasize the differentiated performance of the models, regarding the maturity of the rates. The ANS3 (and the non coherent FNS) model fits well better the 10Y and 28Y rates than the ANS1 and

are available upon request to the authors.

ANS2 models do, while these latter ones performed better on the short-term rates. Since the first eigenvalue in model ANS2 is very close to 1, this means in particular that the model introduced in Christensen, Diebold, Rudebush (2010) does not seem relevant for ultra long maturities. In fact,a triangular autoregressive form with twice a 1 on the diagonal above the main diagonal seem necessary to capture the double hump in the long and ultra-long maturities,but a simpler Jordan form is useful to analyze short term rates. This means that an autoregressive matrix of at least dimension 4 would be necessary, i.e. at least 4-risk-neutral factors [19].

Finally, let us focus on the estimated dynamics of the factors (X_t). We report in Figures 20 and 21 (22 and 23 for the GSW dataset) their time evolution and their empirical distributions, respectively. As mentioned in Appendix 1, for term structure models, the physical and risk-neutral factors distributions are weakly linked. This is a consequence of the finite lifetime of zero-coupon bonds and of its regular replacement by a new issued one. For the sake of simplicity, we have specified the factors as autoregressive Gaussian processes of order 1 under the risk-neutral measure, to get simple analytical formulas for the rates. The equivalence between physical and risk-neutral probability measures implies that the support of the physical and risk-neutral factor distributions are the same. However, arbitrage pricing theory does not require the factors in the estimated models to be also autoregressive Gaussian of order 1 under the physical distribution, even with different mean, volatility and autoregressive parameters [see Le, Singleton, Dai (2010)]. In our estimation procedure, we do not impose any restriction on the physical dynamics of the factors, since the cross-sectional estimation method involves the risk-neutral parameters only.
Figure 21 highlights the difference between the physical and risk-neutral factor distributions: their historical distributions are clearly not Gaussian. They are all skewed, with several modes, which reveals underlying switching regimes in the physical world, whereas they were assumed absent in the risk-neutral one. Moreover, we observe that these historical factor distributions are consistent with the zero-coupon yields distributions (reported in Figure 4), which are also skewed with several modes. This feature comes from the affine relationship between rates and factors.

[19] But such a 4-factor model cannot be applied in our framework with 4 rates only for identification and robustness reasons.

Finally, we estimate a VAR(1) model on the estimated factors $(\hat{X}_t)$ from the ANS1, ANS2 and ANS3 models, in order to compare the $\mathbb{P}$ conditional historical factor distribution to the $\mathbb{Q}$ risk-neutral one[20]. We report the vector of eigenvalues ($\Lambda_{\mathbb{P}}$, say) of the autoregressive matrix ($\Phi_{\mathbb{P}}$, say), together with their monthly ($\Lambda_{\mathbb{P}^4}$), quarterly ($\Lambda_{\mathbb{P}^{13}}$) and yearly ($\Lambda_{\mathbb{P}^{52}}$) equivalents, in Table 15 (resp. 16) for the STRIPS (resp. GSW) dataset. All the models present very similar results, with the two highest eigenvalues being very close to 1, which highlights the difference with the conditional risk-neutral factor distributions in the ANS1, ANS2 and ANS3 models. This shows that the three risk-neutral factors are cointegrated in the physical world, with only one cointegration direction. In particular, the number of unit roots for these factors in the physical and risk-neutral world are not necessarily the same. Consistently with arbitrage pricing theory, both physical and risk-neutral factors distributions have the same infinite support in these models. As a consequence, the Gaussian class of Affine Term Structure Models (to which the ANS1, ANS2, ANS3 models belong) has the drawback to potentially generate negative zero-coupon rates. This problem has limited implications in terms of fitting, when the observed rates are relatively far from 0, as in our sample[21]. However, this feature implies that the Gaussian Affine Term Structure Models could forecast unrealistic negative interest rates, in particular at long horizons, when the forecasts are less influenced by the current term structure. This problem is of particular concern for ultra long-term rates, since the analysis of the evolution of ultra long-term rates (and of their associated risk premia) requires realistic forecasts, at a very long horizon, of the future short term rates.

5 Concluding Remarks

This paper is concerned with the analysis of ultra long-term rates, whose time-to-maturity exceeds the usual long-term (10Y) time-to-maturity. Relying on two different samples of US ultra long-term zero-coupon yields, one obtained from STRIPS bonds, the other one extracted from coupon bonds, we document several specificities of ultra long-term rates.

[20]Formally, $\hat{X}_t = \mu_{\mathbb{P}} + \Phi_{\mathbb{P}}\hat{X}_{t-1} + \epsilon_{\mathbb{P},t}$, where $\hat{X}_t$ are the estimated factors and $\epsilon_{\mathbb{P},t}$ are i.i.d. Gaussian under the physical measure.

[21]In this regard, the problem may be more critical for the estimation of GATSM on current term structure.

First, on our 1988-1998 sample, ultra long-term rates are higher on average and less volatile than short-term ones, consistently with the usual stylized facts on the 1Y-10Y term structure's segment. However, they are also left-skewed and less platykurtic than the short-term rates (they exhibit relatively more, usually negative, "extreme" values). Second, the 28Y-10Y segment of the term structure moves somehow independently with respect to the other yield curve spreads.

These stylized facts call for a specific modeling of the dynamics of the ultra long-term rates. This is usually not done in standard Gaussian affine term structure models, which imply very strong connection between the ultra long-term zero-coupon yields (28Y time-to-maturity, say) and the long term ones (10Y time-to-maturity, say). In this paper, we estimate 3-factor Gaussian Affine Nelson-Siegel Term Structure Models, which differ by the assumptions on the Jordan representation of the risk-neutral autoregressive matrix of the factor process, to show the importance of these assumptions when fitting a term structure model. The estimation results emphasize the need of several "humped" baseline term structures (with one of the hump located at the long end of the term structure) in order to fit the dynamics of the long-term zero-coupon yields and likely the need for at least a fourth risk-neutral factor.

The assessment of the benefits of economic projects with very long-term maturity (such as tunnels, nuclear plants, or transport infrastructures), and thus the decision to implement or not these projects, relies crucially on the discounting of the very long-term future cash flows. Therefore, a dynamic market for ultra long-term bonds issued by relatively creditworthy institutions, such as public governments, is a public good and should be promoted. The scarcity of the existing market for ultra long-term bonds with more than 30Y time-to-maturity raises the question of how this ultra long-term public fixed income markets could be developed and organized, in parallel with the currently growing private market for ultra long-term products (such as longevity bonds, or Insurance Linked Securities, ILS). The securitization of bonds issued specifically for very long-term infrastructure project, potentially supported by the local government could be an avenue to explore further, but raises questions which lie beyond the scope of this paper.

R E F E R E N C E S

Bank of International Settlements (2005) : "Zero-coupon Yield Curves: Technical Documentation", BIS Working Paper, 25.

Bjork, T., and P., Christensen (1999) : "Interest Rate Dynamics and Consistent Forward Rate Curves", Mathematical Finance, 9, 323-348.

Campbell, J., Y. and R., J., Schiller (1991) : "Yields Spreads and Interest Rate Movements: A Bird's Eye View", The Review of Economic Studies, 58, 3, 495-514.

Christensen, J., Diebold, f., and G., Rudebusch (2010) : "The Affine Arbitrage Free Class of Nelson-Siegel Term Structure Model", Journal of Econometrics, forthcoming.

Cox, J., Ingersoll, J., and S. Ross (1985) : "A Theory of the Term Structure of Interest Rates", Econometrica, 53, 385-408.

Dai, Q. and K., Singleton (2003) : "Term Structure Modeling in Theory and Reality", Review of Financial Studies, 16, 631-678.

Darolles, S., Gourieroux, C., and J., Jasiak (2006) : "Structural Laplace Transform and Compound Autoregressive Models", Journal of Time Series Analysis, 27, 477-503.

Diebold, F., X., Rudebusch, G., D. and Aruoba, B. (2006) : "The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach", Journal of Econometrics, 131, 309-338.

Dubecq S., and C., Gourieroux (2010) : "A Term Structure Model with Level Factor Cannot Be Realistic and Arbitrage Free", CREST-DP.

Duffie, D., Filipovic, D., and W., Schachermayer (2003) : "Affine Processes and Applications in Finance", Annals of Applied Probability, 13, 984-1053.

Duffie, D., R. Kan (1996) : "A Yield-factor Model of Interest Rates", Mathematical Finance, 6, 379-406.

Dybvig, P., Ingersoll, J., and S., Ross (1996) : "Long Forward and Zero Coupon Rates Can Never Fall", Journal of Business, 69, 1-25.

El Karoui, N., Frachot, A., and H., Geman (1998) : "On the Behaviour of Long Zero Coupon Rates in a No-Arbitrage Framework", Review of Derivatives Research, 1, 351-389.

Estrella, A. and G. A. Hardouvelis (1991) : "The Term Structure as a Predictor of Real Economic Activity", The Journal of Finance, 46, 555-576.

Fama, E. F., and R., R., Bliss (1987) : "The Information in Long-Maturity Forward Rates", American Economic Review, 77, 680-692.

Fisher, M. , Nychka, D. and D. Zervos (1995) : "Fitting the Term Structure of Interest Rates with Smoothing Splines", Finance and Economics Discussion Series, 1995-1.

Filipovic, D. (1999) : "A Note on the Nelson-Siegel Family" Mathematical Finance, 9, 349-35.

Gourieroux, C. (2009) : "CaR and Affine Processes", in Econometric Forecasting and High Frequency Data Analysis, IMS Lecture Notes Series.

Gourieroux, C. and J. Jasiak (2001) : "Financial Econometrics", Princeton Series in Finance.

Gourieroux, C. and J. Jasiak (2006) : "Autoregressive Gamma processes", Journal of Forecasting, 25, 129-152.

Gourieroux, C., Monfort, A., and V., Polimenis (2006) : "Affine Models for Credit Risk Analysis", Journal of Financial Econometrics, 4, 494-530..

Gourieroux, C. and R. Sufana (2006) : "Classification of Affine Term Structure Models", Journal of Financial Econometrics, 4, 31-52.

Gourieroux, C. and R. Sufana (2007) : "Wishart Quadratic Term Structure", CREST-DP.

Gurkaynak, R. S., Sack, B. and J. H. Wright (2007) : "The U.S. Treasury Yield Curve: 1961 to the Present", Journal of Monetary Economics, 54, 8,

2291-2304.

Hansen, L., and S. Richard (1987) : "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing", Econometrica, 55, 587-614.

Harrison, M., and D. Kreps (1979) : "Martingales and Arbitrage in Multiperiod Securities Markets", Journal of Economic Theory, 20, 381-408.

Hubalek, F., Klein, I., and J., Teichmann (2002) : "A General Proof of the Dybvig-Ingersoll-Ross Theorem : Long Forward Rates Can Never Fall", Mathematical Finance, 12, 447-451.

Ingersoll,J., J. Skelton and R. L. Weil (1978) : "Duration Forty Years Later", Journal of Financial Quantitative Analysis, 13, 627-650.

Jones, F. (1991) : "Yield Curve Strategies", Journal of Fixed Income, 43-51.

Jordan, B., D., Jorgensen, R., D., and D., R., Kuipers (2000) : "The Relative Pricing of US Treasury STRIPS: Empirical Evidence", Journal of Financial Economics, 56, 89-123.

Langetieg, T. (1980) : "A Multivariate Model of the Term Structure", Journal of Finance, 35, 71-97.

Leh, A., Singleton, K., and Q. Dai (2010) :"Discrete-Time Dynamic Term Structure Models with Generalized Market Prices of Risk", Review of Financial Studies, 23, 2184-2227.

Litterman, R., and J., Scheinkman (1991) : "Common Factors Affecting Bond Returns", Journal of Fixed Income, 1, 54-61.

Macaulay, F. R. (1938) : "Some Theoretical Problems Suggested by the Movements of Interest Rates, Bonds Yields, and Stock Prices in the United States Since 1856", New York, Columbia University Press.

McCulloch, J., H. (1971) : "Measuring the Term Structure of Interest Rates", Journal of Business, 44, 19-31.

Moody's (2010) : "Rating Symbols and Definitions", Moody's Investors Service, October 2010.

Monfort, A. and F., Pegoraro (2007) : "Switching VARMA Term Structure Models", Journal of Financial Econometrics, 5, 103-151.

Nelson, C., and A., Siegel (1987) : "Parsimonious Modeling of Yield Curves", Journal of Business, 60, 473-489.

OECD (2009) : "Central Government Debt", Statistical Yearbook 1999-2008.

Piazzesi, M. (2010) : "Affine Term Structure Models", Handbook of Financial Econometrics, Chapter 12.

Rogers, L. (1977) : "The Potential Approach to the Term Structure of Interest Rates and Foreign Exchange Rates", Mathematical Finance, 7, 157-176.

Sack, B. (2000) : "Using Treasury STRIPS to Measure the Yield Curve", Federal Reserve Working Paper, 200042.

Sill, K. (1994) : "Managing the Public Debt", Federal Reserve Bank of Philadelphia Business Review Articles, July/August 1994.

Svensson, L. (1995) : "Estimating Forward Interest Rates with the Extended Nelson and Siegel Model", Quarterly Review, Sveriges Riskbank, 3, 13-26.

Vasicek, O., (1977) : "An Equilibrium Characterization of the Term Structure", Journal of Financial Economics, 5, 177-188.

Vasicek, O., and F., Fong (1982) : "Term Structure Modeling Using Exponential Splines", The Journal of Finance, 37, 339-348.

Tables

Original Maturity	Total Issued	Stripped	Not Stripped	Percent Stripped
2Y Note	260,770	722	260.048	0.3
3Y Note	38,721	0	38,721	0.0
5Y Note	154,516	342	154,174	0.2
10Y Note	475,526	57,353	418,173	12.1
20Y Bond	23,706	8,987	14,719	37.9
30Y Bond	440,448	153,861	286,587	34.9
Total	1,393,687	221,265	1,172,422	15.9

Table 1 [from Sack (2000)]. Total amount of coupon notes and coupon bonds issued by the US Treasury in January 31, 1999. Numbers are in millions of US dollars, except in the last column in which the percentage of coupon notes and bonds held in stripped form is reported.

Original Maturity	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000
Less than 1Y (Bills)	0.24	0.24	0.24	0.24	0.23	0.23	0.22	0.20	0.21	0.23	0.22
Between 2Y and 10Y (Notes)	0.58	0.58	0.58	0.59	0.60	0.61	0.61	0.61	0.59	0.56	0.55
More than 10Y (Bonds)	0.17	0.17	0.17	0.16	0.16	0.15	0.16	0.17	0.18	0.20	0.21

Table 2. Repartition of all the marketable securities issued by the US Treasury (except Inflation Indexed bonds) between less than 1Y time-to-maturity bills, notes, whose time-to-maturities range between 2Y and 10Y, and more than 10Y time-to-maturity bonds. Numbers are percentages. Source: US Monthly Statement of Public Debt.

Maturity	1Y	5Y	10Y	28Y GSW	28Y STRIPS
Mean	0.0598	0.069	0.074	0.0768	0.0766
Median	0.0570	0.0670	0.0740	0.0780	0.0780
Standard Deviation	0.0162	0.0118	0.0104	0.0077	0.0080
Skewness	0.1523	0.1556	0.0576	−0.2162	−0.1570
Kurtosis	2.2122	1.9250	1.8397	2.1455	2.1270
$\text{Quantile}_{5\%}$	0.0344	0.0516	0.0581	0.0646	0.0635
$\text{Quantile}_{95\%}$	0.0868	0.0885	0.0906	0.0883	0.0891
$\frac{\text{Quantile}_{95\%}-\text{Quantile}_{5\%}}{\text{Median}}$	0.9183	0.5516	0.4404	0.3044	0.3278

Table 3. Summary statistics on the distribution of the zero-coupon yields for both datasets (528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset.

Maturity	1Y	5Y	10Y	28Y GSW	28Y STRIPS
Mean	0.0598	0.069	0.074	0.0768	0.0766
Median	0.0570	0.0670	0.0740	0.0780	0.0780
Standard Deviation	0.0162	0.0118	0.0104	0.0077	0.0080
Skewness	0.1523	0.1556	0.0576	−0.2162	−0.1570
Kurtosis	2.2122	1.9250	1.8397	2.1455	2.1270
$\text{Quantile}_{5\%}$	0.0344	0.0516	0.0581	0.0646	0.0635
$\text{Quantile}_{95\%}$	0.0868	0.0885	0.0906	0.0883	0.0891
$\frac{\text{Quantile}_{95\%}-\text{Quantile}_{5\%}}{\text{Median}}$	0.9183	0.5516	0.4404	0.3044	0.3278

Table 4. Summary statistics on the distribution of the zero-coupon yields for both datasets (528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset.

Maturity	1Y	5Y	10Y	28Y GSW	28Y STRIPS
1Y	1	0.9111	0.7553	0.5812	0.5789
5Y	0.9111	1	0.9513	0.8470	0.8451
10Y	0.7553	0.9513	1	0.9528	0.9625
28Y GSW	0.5812	0.8470	0.9528	1	0.9886
28Y STRIPS	0.5789	0.8451	0.9625	0.9886	1

Table 5. Correlation matrix of the zero-coupon yields for both datasets (528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset.

Spread	5Y - 1Y	10Y - 5Y	28Y(GSW) - 10Y	28Y(STRIPS) - 10Y
Mean	0.0091	0.0051	0.0028	0.0026
Median	0.0069	0.0042	0.0028	0.0028
Standard Deviation	5.291×10^{-5}	1.398×10^{-5}	1.484×10^{-5}	1.193×10^{-5}
Skewness	0.4526	0.6445	0.2461	0.3604
Kurtosis	2.1821	2.8898	2.4834	2.4052
$\text{Quantile}_{5\%}$	−0.0012	−0.0000	−0.0033	−0.0025
$\text{Quantile}_{95\%}$	0.0228	0.0119	0.0099	0.0091
$\frac{\text{Quantile}_{95\%}-\text{Quantile}_{5\%}}{\text{Median}}$	3.4580	2.8343	4.7022	4.1091

Table 6. Summary statistics on the distribution of the spreads 5Y - 1Y, 10Y - 5Y, 28Y - 10Y for both datasets (528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset.

Spread	5Y - 1Y	10Y - 5Y	28Y(GSW) - 10Y	28Y(STRIPS) - 10Y
5Y - 1Y	1	0.8949	0.5284	0.6019
10Y - 5Y	0.8949	1	0.5547	0.6921
28Y(GSW) - 10Y	0.5284	0.5547	1	0.9492
28Y(STRIPS) - 10Y	0.6019	0.6921	0.9492	1

Table 7. Correlation matrix of the spreads 5Y - 1Y, 10Y - 5Y, 28Y - 10Y for both datasets (528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset.

Ordering	Frequency GSW dataset	Frequency STRIPS dataset
$1Y < 5Y < 10Y < 28Y$	0.70	0.70
$1Y < 5Y < 28Y < 10Y$	0.18	0.18
$28Y < 10Y < 5Y < 1Y$	0.04	0.04
$28Y < 5Y < 10Y < 1Y$	0.02	0.02
$1Y < 28Y < 5Y < 10Y$	0.02	0.01
$28Y < 5Y < 1Y < 10Y$	0.01	0.01
$5Y < 1Y < 10Y < 28Y$	0.01	0.01
$1Y < 28Y < 10Y < 5Y$	0.01	0.01

Table 8. Observed orderings of 1Y, 5Y, 10Y 28Y zero-coupon yields for both datasets, ranked by frequency(528 weekly observations between January 8, 1988 and February 13, 1998). "GSW" and "STRIPS" precise the corresponding dataset. For the sake of parsimony, we only report the orderings with more than 1% frequency.

FNS	ANS1							
$\exp\{-\mu\}$	Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.9963	0.9989	0.0406	0	0	−0.1028	0.9956	0.9858	0.9444
	0.9902	−0.5259	0.0135	0		0.9614	0.8798	0.5992
	0.9055	−0.0531	0.0071	0.0288		0.6723	0.2751	0.0057

ANS2							
Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.999	0.0378	0	0	−1.1688	0.996	0.9871	0.9493
0.95	−0.5160	0.0079	0		0.8145	0.5133	0.0694
0.95	−0.0366	0.0127	0.0307		0.8145	0.5133	0.0694

ANS3							
Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.9988	$0.3778.10^{-9}$	0	0	0.1916	0.9952	0.9845	0.9395
0.9988	0	$0.0795.10^{-9}$	0		0.9952	0.9845	0.9395
0.9988	0	0	$0.0101.10^{-9}$		0.9952	0.9845	0.9395

Table 9. Parameter estimates for the FNS, ANS1, ANS2, ANS3 models. The column Λ reports the eigenvalues of the matrix Φ in the ANS1, ANS2, ANS3 models. STRIPS dataset (528 weekly observations).

FNS	ANS1							
$\exp\{-\mu\}$	Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.9965	0.9989	0.0406	0	0	0.0510	0.9956	0.9858	0.9444
	0.9902	−0.5264	0.0103	0		0.9614	0.8998	0.5992
	0.9004	−0.0520	−0.0018	0.0223		0.6573	0.2557	0.0043

ANS2							
Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.9992	0.0391	0	0	−0.6519	0.9968	0.9896	0.9592
0.9501	−0.4788	0.0736	0		0.8148	0.5140	0.0698
0.9501	−0.0373	−0.0020	0.0339		0.8148	0.5140	0.0698

ANS3							
Λ	Σ			$\beta(1)$	Λ^4	Λ^{13}	Λ^{52}
0.9988	$0.3906.10^{-9}$	0	0	0.2084	0.9952	0.9845	0.9395
0.9988	0	$0.7360.10^{-9}$	0		0.9952	0.9845	0.9395
0.9988	0	0	$0.0124.10^{-9}$		0.9952	0.9845	0.9395

Table 10. Parameter estimates for the FNS, ANS1, ANS2, ANS3 models. The column Λ reports the eigenvalues of the matrix Φ in the ANS1, ANS2, ANS3 models. GSW dataset (528 weekly observations).

Model	1Y	5Y	10Y	28Y
FNS	1.7075	5.6475	5.1953	1.2553
ANS1	0.0564	2.4007	6.1097	5.2329
ANS2	0.0001	2.1540	6.4063	5.8767
ANS3	3.1226	8.0710	5.8662	0.9688

Table 11. RMSE on the 1Y, 5Y, 10Y, 28Y zero-coupon yields for all models in basis points. STRIPS dataset.

	Model	1Y	5Y	10Y	28Y(STRIPS)
Mean	FNS	$0.2983.10^{-4}$	$-0.9867.10^{-4}$	$0.9077.10^{-4}$	$-0.2193.10^{-4}$
	ANS1	$0.0079.10^{-7}$	$-0.3368.10^{-7}$	$0.8571.10^{-7}$	$-0.7341.10^{-7}$
	ANS2	$-0.0007.10^{-7}$	$0.2067.10^{-7}$	$-0.6241.10^{-7}$	$0.5892.10^{-7}$
	ANS3	$-0.0883.10^{-8}$	$0.2282.10^{-8}$	$-0.1659.10^{-8}$	$0.0274.10^{-8}$
Median	FNS	$0.1334.10^{-4}$	$-0.4412.10^{-4}$	$0.4058.10^{-4}$	$-0.0981.10^{-4}$
	ANS1	$0.0027.10^{-4}$	$-0.1149.10^{-4}$	$0.2924.10^{-4}$	$-0.2504.10^{-4}$
	ANS2	$0.0001.10^{-4}$	$-0.2218.10^{-4}$	$0.6694.10^{-4}$	$-0.6322.10^{-4}$
	ANS3	$0.2614.10^{-4}$	$-0.6756.10^{-4}$	$0.4910.10^{-4}$	$-0.0811.10^{-4}$
Standard Deviation	FNS	$0.1683.10^{-3}$	$0.5566.10^{-3}$	$0.5120.10^{-3}$	$0.1237.10^{-3}$
	ANS1	$0.0056.10^{-3}$	$0.2403.10^{-3}$	$0.6116.10^{-3}$	$0.5238.10^{-3}$
	ANS2	$0.0001.10^{-3}$	$0.2265.10^{-3}$	$0.6834.10^{-3}$	$0.6454.10^{-3}$
	ANS3	$0.3126.10^{-3}$	$0.8079.10^{-3}$	$0.5872.10^{-3}$	$0.0970.10^{-3}$
Skewness	FNS	0.1323	−0.1323	0.1323	−0.1323
	ANS1	0.1966	−0.1966	0.1966	−0.1966
	ANS2	0.1317	−0.1285	0.1285	−0.1285
	ANS3	−0.7539	0.7539	−0.7539	0.7539
Kurtosis	FNS	3.2396	3.2396	3.2396	3.2396
	ANS1	3.2713	3.2713	3.2713	3.2713
	ANS2	3.2034	3.2046	3.2046	3.2046
	ANS3	3.5753	3.5753	3.5753	3.5753
$\text{Quantile}_{5\%}$	FNS	−0.0002	−0.0011	−0.0007	−0.0002
	ANS1	−0.0000	−0.0004	−0.0010	−0.0009
	ANS2	−0.0001	−0.0004	−0.0011	−0.0012
	ANS3	−0.0006	−0.0011	−0.0011	−0.0001
$\text{Quantile}_{95\%}$	FNS	0.0003	0.0008	0.0010	0.0002
	ANS1	0.0000	0.0004	0.0011	0.0009
	ANS2	0.0001	0.0004	0.0012	0.0011
	ANS3	0.0004	0.0016	0.0008	0.0002
$\frac{\text{Quantile}_{95\%}-\text{Quantile}_{5\%}}{\text{Median}}$	FNS	42.2464	−42.2464	42.2464	−42.2464
	ANS1	71.3563	−71.3553	71.3553	−71.3553
	ANS2	46.0845	−35.0713	35.0713	−35.0713
	ANS3	39.1715	−39.1715	39.1715	−39.1715

Table 12. Summary statistics on the distribution of the errors on the zero-coupon yields, for the FNS, ANS1, ANS2, and ANS3 models. STRIPS dataset.

Model	1Y	5Y	10Y	28Y
FNS	2.0176	6.4888	5.8164	1.3453
ANS1	0.0764	3.2630	8.3069	7.1161
ANS2	0.0001	2.9604	8.4507	7.0855
ANS3	3.2872	8.4547	6.1086	0.9906

Table 13. RMSE on the 1Y, 5Y, 10Y, 28Y zero-coupon yields for all models in basis points. GSW dataset.

	Model	1Y	5Y	10Y	28Y(STRIPS)
Mean	FNS	$0.1121.10^{-4}$	$-0.3606.10^{-4}$	$0.3232.10^{-4}$	$-0.0748.10^{-4}$
	ANS1	$-0.0058.10^{-8}$	$0.2595.10^{-8}$	$-0.6605.10^{-8}$	$0.5661.10^{-8}$
	ANS2	$-0.0202.10^{-9}$	$-0.0661.10^{-9}$	$0.1825.10^{-9}$	$-0.1618.10^{-9}$
	ANS3	$-0.0864.10^{-7}$	$0.2222.10^{-7}$	$-0.1606.10^{-7}$	$0.0260.10^{-7}$
Median	FNS	$-0.0576.10^{-4}$	$0.1854.10^{-4}$	$-0.1662.10^{-4}$	$0.0384.10^{-4}$
	ANS1	$-0.0032.10^{-4}$	$0.1371.10^{-4}$	$-0.3489.10^{-4}$	$0.2989.10^{-4}$
	ANS2	$-0.0001.10^{-4}$	$0.0356.10^{-4}$	$-0.1017.10^{-4}$	$0.0853.10^{-4}$
	ANS3	$0.3409.10^{-4}$	$-0.8767.10^{-4}$	$0.6334.10^{-4}$	$-0.1027.10^{-4}$
Standard Deviation	FNS	$0.2016.10^{-3}$	$0.6485.10^{-3}$	$0.5813.10^{-3}$	$0.1344.10^{-3}$
	ANS1	$0.0076.10^{-3}$	$0.3266.10^{-3}$	$0.8315.10^{-3}$	$0.7123.10^{-3}$
	ANS2	$0.0001.10^{-3}$	$0.2963.10^{-3}$	$0.8459.10^{-3}$	$0.7092.10^{-3}$
	ANS3	$0.3290.10^{-3}$	$0.8463.10^{-3}$	$0.6114.10^{-3}$	$0.0992.10^{-3}$
Skewness	FNS	0.1860	−0.1860	0.1860	−0.1860
	ANS1	0.4820	−0.4820	0.4820	−0.4820
	ANS2	0.5301	−0.5294	0.5294	−0.5294
	ANS3	−0.7500	0.7500	−0.7500	0.7500
Kurtosis	FNS	3.4066	3.4066	3.4066	3.4066
	ANS1	3.2013	3.2013	3.2013	3.2013
	ANS2	3.6410	3.6370	3.6370	3.6370
	ANS3	3.5630	3.5630	3.5630	3.5630
$\text{Quantile}_{5\%}$	FNS	−0.0003	−0.0013	−0.0009	−0.0003
	ANS1	−0.0001	−0.0006	−0.0013	−0.0013
	ANS2	−0.0001	−0.0005	−0.0013	−0.0012
	ANS3	−0.0006	−0.0012	−0.0012	−0.0001
$\text{Quantile}_{95\%}$	FNS	0.0004	0.0010	0.0011	0.0002
	ANS1	0.0001	0.0005	0.0016	0.0011
	ANS2	0.0001	0.0004	0.0015	0.0011
	ANS3	0.0004	0.0016	0.0008	0.0002
$\frac{\text{Quantile}_{95\%}-\text{Quantile}_{5\%}}{\text{Median}}$	FNS	−121.8842	121.8842	−121.8842	121.8842
	ANS1	−80.9247	80.9247	−80.9246	80.9246
	ANS2	−174.3155	269.0035	−269.0036	269.0038
	ANS3	31.5799	−31.5799	31.5799	−31.5799

Table 14. Summary statistics on the distribution of the errors on the zero-coupon yields, for the FNS, ANS1, ANS2, and ANS3 models. GSW dataset.

$ANS1$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9969	0.9876	0.9602	0.8501
0.9928	0.9714	0.9101	0.6859
0.9486	0.8097	0.5036	0.0643

$ANS2$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9969	0.9876	0.9602	0.8501
0.9928	0.9714	0.9101	0.6859
0.9486	0.8097	0.5036	0.0643

$ANS3$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9969	0.9876	0.9602	0.8501
0.9928	0.9714	0.9101	0.6859
0.9486	0.8097	0.5036	0.0643

Table 15. Set of Eigenvalues $\Lambda_{\mathbb{P}}$ of the autoregressive matrix of the factor process (X_t), estimated as a VAR(1) process, under the physical $\mathbb{P}$ factor distribution. STRIPS dataset.

$ANS1$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9960	0.9840	0.9489	0.8106
0.9934	0.9739	0.9176	0.7090
0.9294	0.7462	0.3861	0.0222

$ANS2$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9960	0.9840	0.9489	0.8106
0.9934	0.9739	0.9176	0.7090
0.9294	0.7462	0.3861	0.0222

$ANS3$			
$\Lambda_{\mathbb{P}}$	$\Lambda_{\mathbb{P}}^{4}$	$\Lambda_{\mathbb{P}}^{13}$	$\Lambda_{\mathbb{P}}^{52}$
0.9960	0.9840	0.9489	0.8106
0.9934	0.9739	0.9176	0.7090
0.9294	0.7462	0.3861	0.0222

Table 16. Set of Eigenvalues $\Lambda_{\mathbb{P}}$ of the autoregressive matrix of the factor process (X_t), estimated as a VAR(1) process, under the physical $\mathbb{P}$ factor distribution. GSW dataset.

Figures

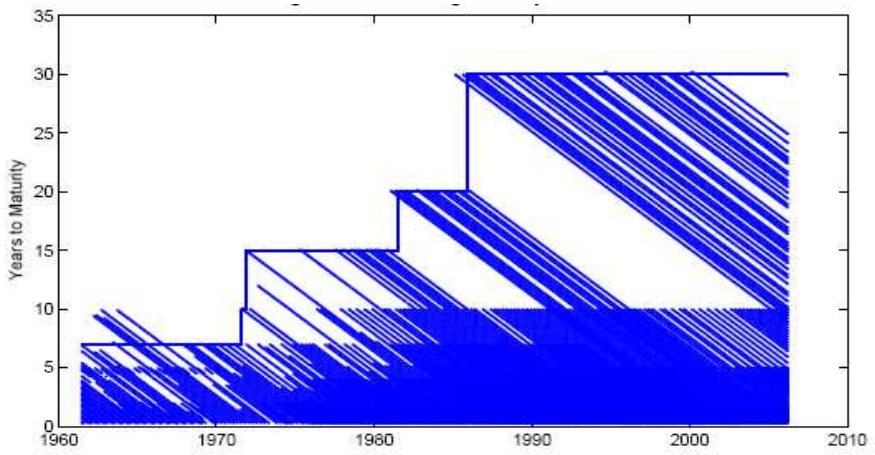

Figure 1 [from Gurkaynak, Sack, Wright (2007)]. Outstanding of US Treasury securities between 1962 and 2005. Each line starts at the issuance of a specific bond (or note) and decreases with the residual time-to-maturity at each date.

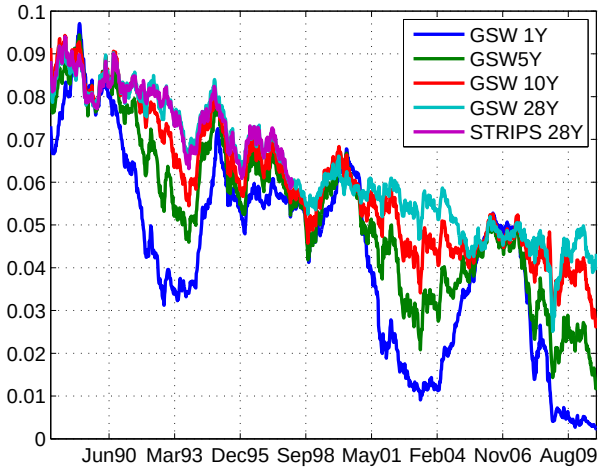

Figure 2. US Treasury zero-coupons yield curve from both datasets between January 1988 and October 2010. Except between mid-1989 and June 1990, the lowest zero-coupon yield is the 1Y yield, and the second lowest one is the 5Y zero-coupon yield. "STRIPS" and "GSW" precise the corresponding dataset for the 28Y zero-coupon yield. STRIPS 28Y zero-coupon yields are not available after February 1998.

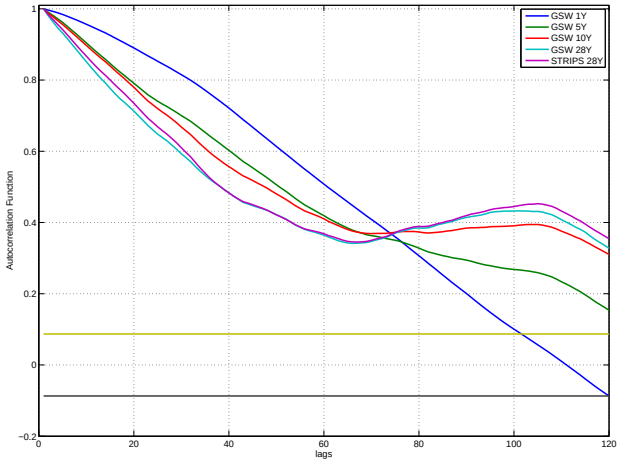

Figure 3. Autocorrelation function of the 1Y, 5Y, 10Y, 28Y yields for both datasets. "GSW" and "STRIPS" precise the dataset. Lags are in weeks. From lag 1 to lags 70, the lowest autocorrelation function is the 28Y zero-coupon yields' one.

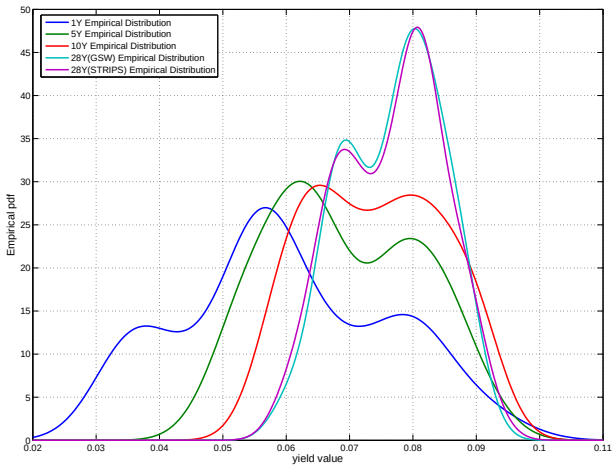

Figure 4. Empirical distribution of the 1Y, 5Y, 10Y and 28Y zero-coupon yields for both "GSW" and "STRIPS" datasets (528 weekly observations between January 8, 1988 and February 13, 1998). The empirical distribution's mean of the zero-coupon yields moves to the right when the time-to-maturity increases.

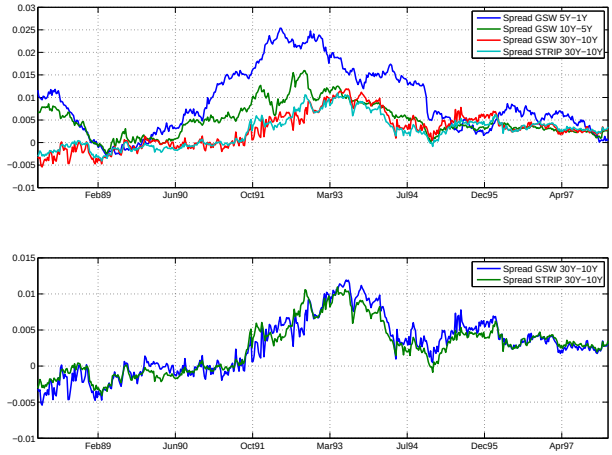

Figure 5. 5Y-1Y, 10Y-5Y, 28Y-10Y zero-coupon yields spreads for both datasets over the whole sample. "GSW" and "STRIPS" precise the dataset. The lowest spreads are generally the GSW and STRIPS 28Y-10Y spreads, the highest spread being the 5Y-1Y spread. The bottom panel focuses specifically on the 28Y-10Y spread for both datasets.

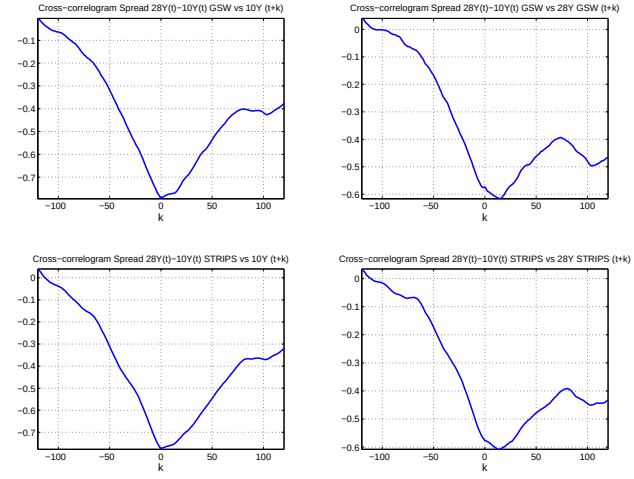

Figure 6. Crosscorrelation function of the 28Y-10Y spread with lead/lag values of the 10Y (left panels) and 28Y (right panels) zero-coupon yields for GSW (upper panels) and STRIPS (bottom panels) samples. "GSW" and "STRIPS" precise the dataset. Lags ("k") are in weeks.

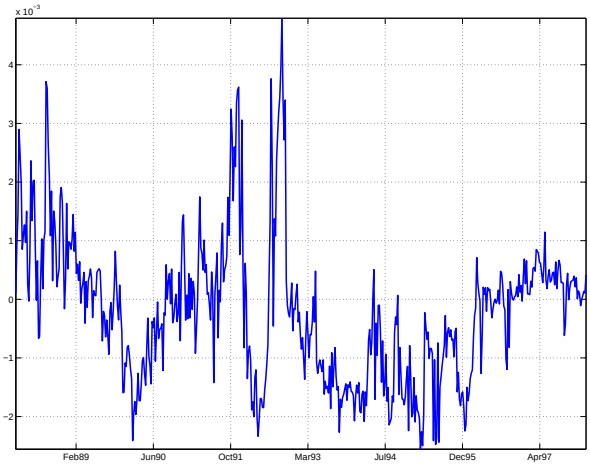

Figure 7. Difference STRIPS-GSW 28Y zero-coupon yields between January 1988 and February 1998.

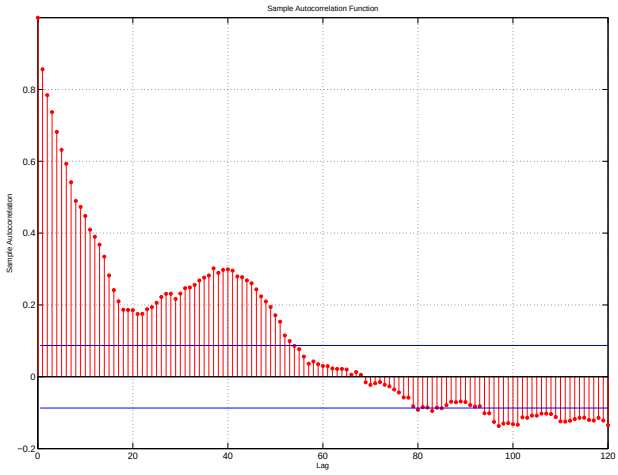

Figure 8. Autocorrelation function of the difference STRIPS-GSW for 28Y zero-coupon yields. Lags are in weeks.

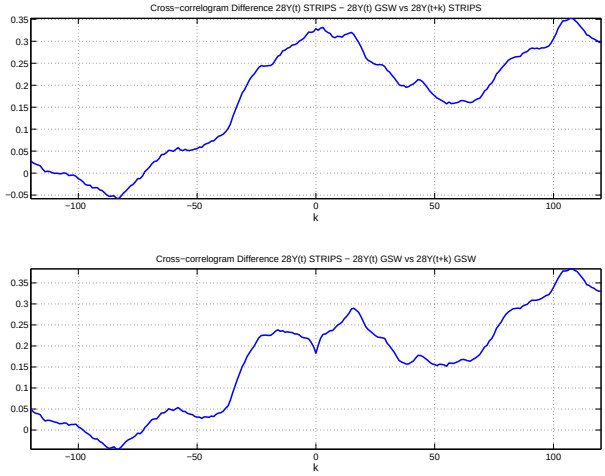

Figure 9. Crosscorrelation function of the difference STRIPS-GSW for 28Y zero-coupon yields with lead/lag values of the 28Y STRIPS (upper panel) and the 28Y GSW (bottom panel) zero-coupon yields. Lags are in weeks.

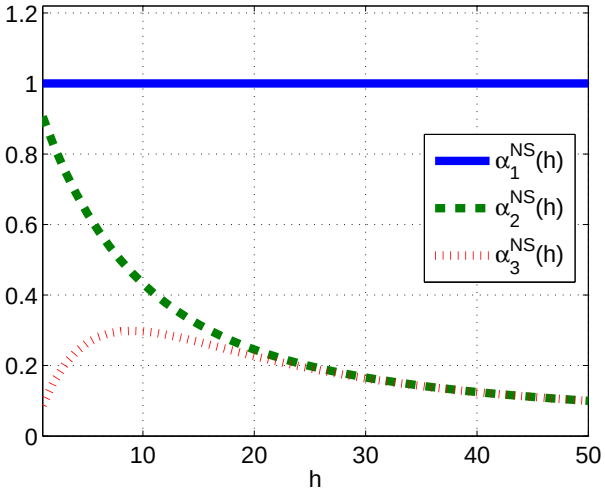

Figure 10. Baseline term structures for Nelson-Siegel (NS) model. "h" stands for the zero-coupon yields time-to-maturity.

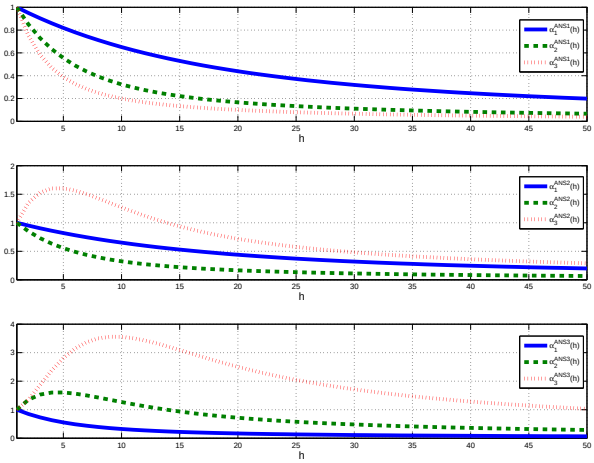

Figure 11. Factor loadings of the Arbitrage Free Nelson-Siegel models presented in Section 3.2: ANS1 (upper panel), ANS2 (middle panel) and ANS3 (lower panel). The underlying parameter values are the following: $\lambda_1 = 0.9$, $\lambda_2 = 0.7$ **and** $\lambda_3 = 0.5$ **for ANS1,** $\lambda_1 = 0.9$, $\lambda_2 = 0.7$ **for ANS2 and** $\lambda = 0.7$ **for ANS3. For all models,** $\alpha(1) = [1\ 1\ 1]$**.**

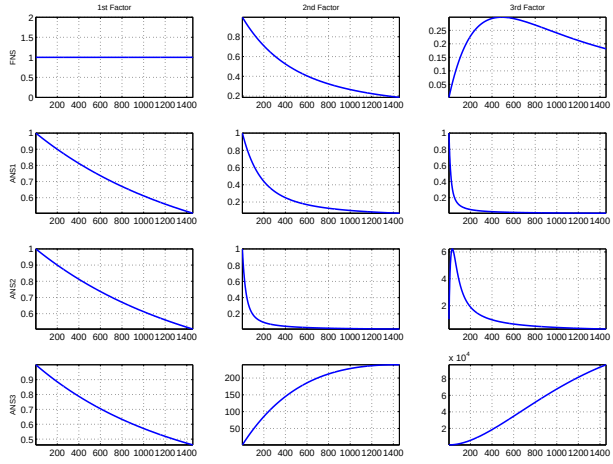

Figure 12. Factor loadings $\alpha(h)$ **for all models as a function of the time-to-maturity of zero-coupon yield, estimated on the STRIPS dataset. The x-axis stands for the time-to-maturity** h**.**

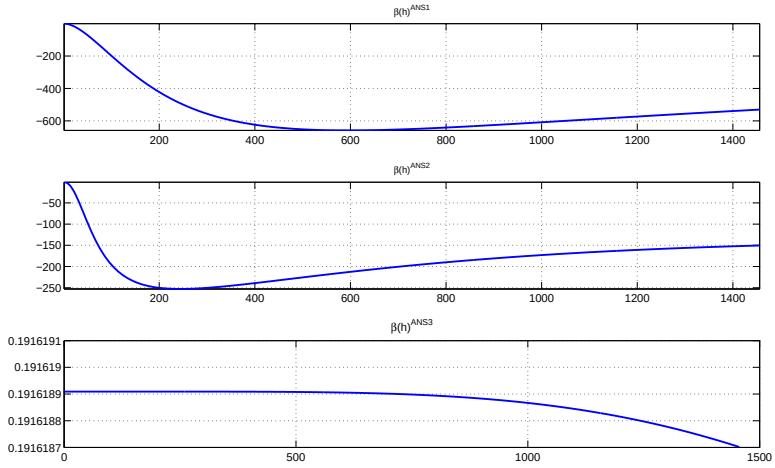

Figure 13. Time independent baseline term structure $\beta(h)$ for the three ANS models ($\beta(h) = 0\ \forall h$) in the FNS model, estimated on the STRIPS dataset. The x-axis stands for the time-to-maturity h.

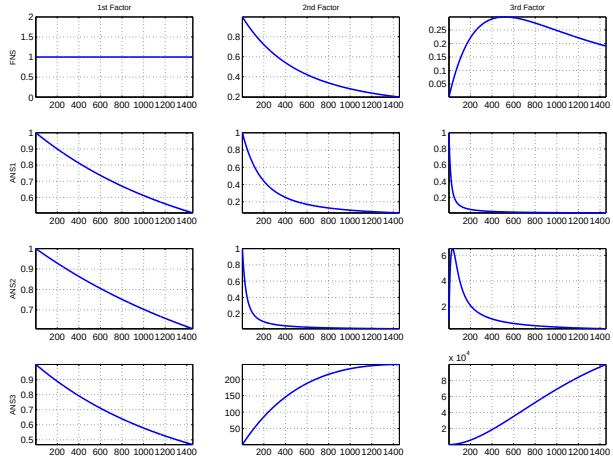

Figure 14. Factor loadings $\alpha(h)$ for all models as a function of the time-to-maturity of zero-coupon yield, estimated on the GSW dataset. The x-axis stands for the time-to-maturity h.

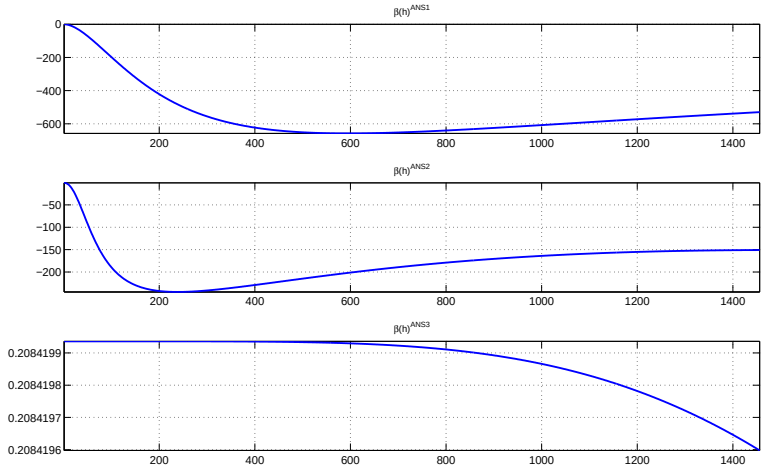

Figure 15. Time independent baseline term structure $\beta(h)$ for the three ANS models ($\beta(h) = 0\ \forall h$) in the FNS model, estimated on the GSW dataset. The x-axis stands for the time-to-maturity h.

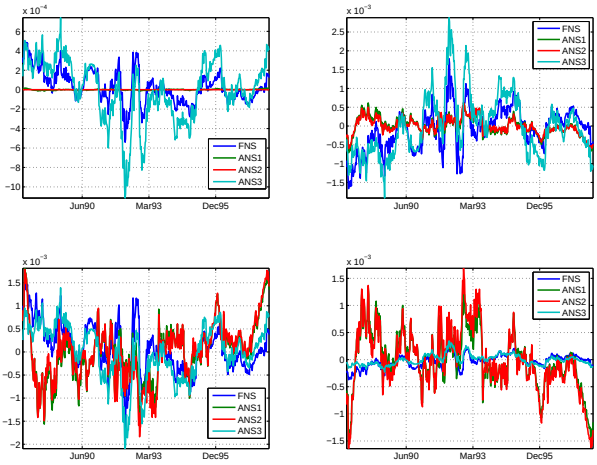

Figure 16. Pricing errors on the 1Y (upper left panel), 5Y (upper right panel), 10Y (lower left panel), 28Y (lower right panel) zero-coupon yields, for each term structure model. The lowest errors on the short term yields are obtained via the ANS1 and ANS2 models, and conversely for the long-term rates. STRIPS dataset.

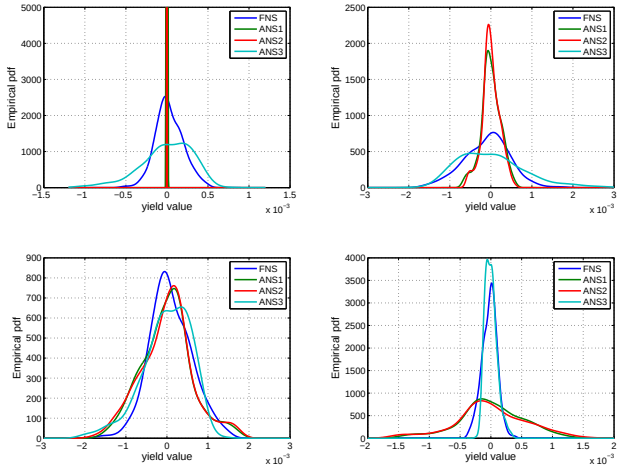

Figure 17. Empirical (unconditional) distribution of the errors on the 1Y (upper left panel), 5Y (upper right panel), 10Y (lower left panel), 28Y (lower right panel) zero-coupon yields, for each term structure model. STRIPS dataset.

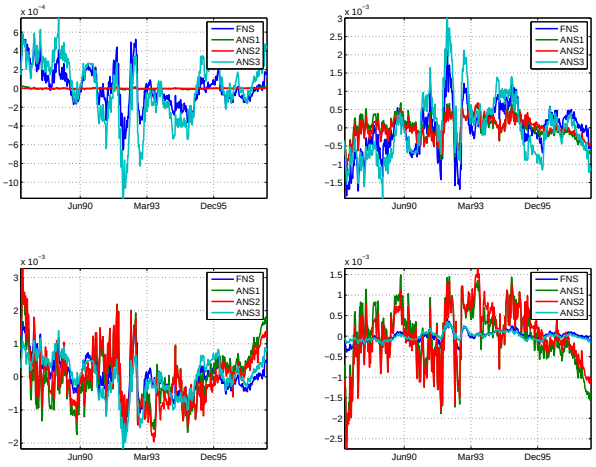

Figure 18. Pricing errors on the 1Y (upper left panel), 5Y (upper right panel), 10Y (lower left panel), 28Y (lower right panel) zero-coupon yields, for each term structure model. The lowest errors on the short term yields are obtained via the ANS1 and ANS2 models, and conversely for the long-term rates. GSW dataset.

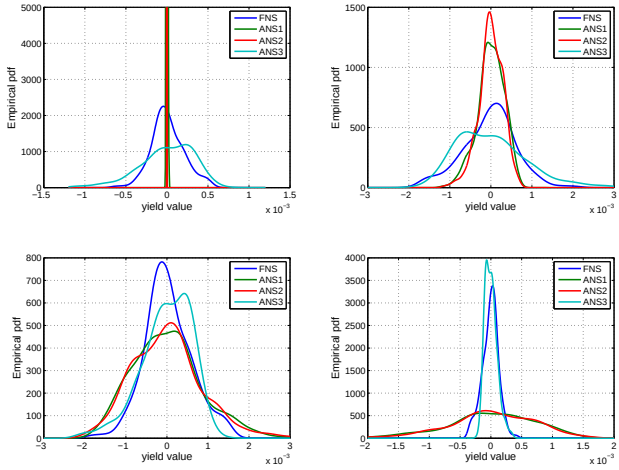

Figure 19. Empirical (unconditional) distribution of the errors on the 1Y (upper left panel), 5Y (upper right panel), 10Y (lower left panel), 28Y (lower right panel) zero-coupon yields, for each term structure model. GSW dataset.

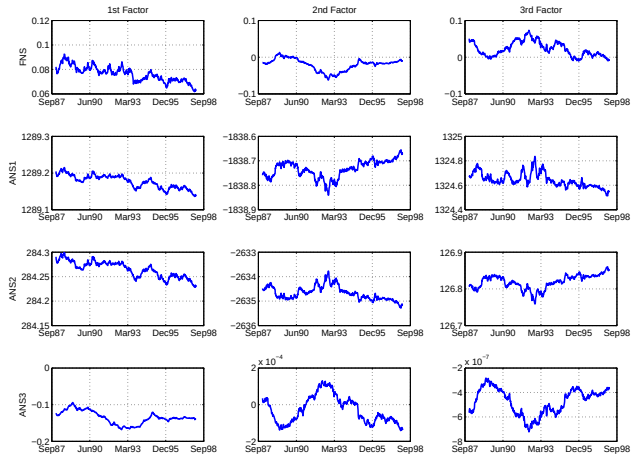

Figure 20. Time evolution of the factor X_t estimates on the STRIPS sample for all models.

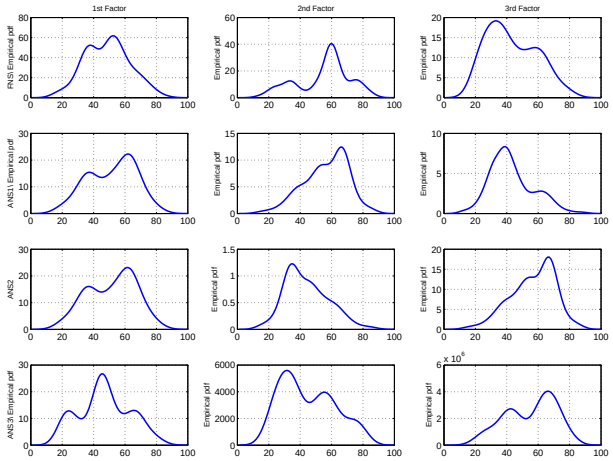

Figure 21. Empirical (unconditional) distribution of the three factors (X_t), for each term structure model. STRIPS dataset.

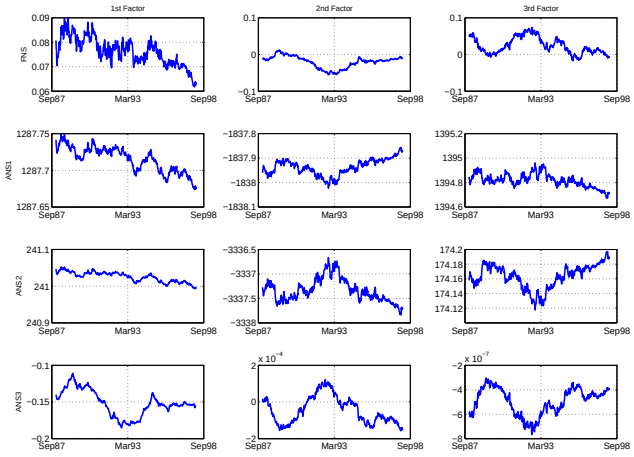

Figure 22. Time evolution of the factor X_t estimates on the GSW sample for all models.

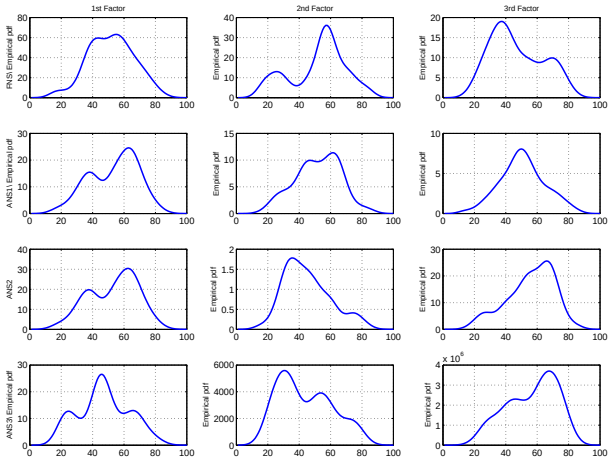

Figure 23. Empirical (unconditional) distribution of the three factors (X_t), for each term structure model. GSW dataset.

Appendix 1
Affine Term Structure Models

i) Definition

We introduce a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with a filtration (I_t) which defines the sequence of information sets, generated by (X_t) as in the beginning of Section 3. $\mathbb{P}$ is the historical probability measure, and $\mathbb{Q}$ the risk-neutral one.
Affine term structure models have become increasingly popular in the term structure literature since the seminal work of Duffie, Kan (1996). This type of (reduced-form) model constitutes a class of the arbitrage-free factor models which make the yields affine functions of the factors [see Gourieroux, Monfort, Polimenis (2006) for a complete description of affine term structure models in discrete time]. Pricing formula in discrete-time ATSM are usually obtained in the following way.

Definition A1

i) The sequence of information sets is generated by the factors (X_t), $t \in \mathbb{N}$

ii) The short term rate $r(t,1)$ is an affine function of the latent factors (X_t) $\forall t$:

$$r(t,1) = \alpha(1)X_t + \beta(1); \tag{14}$$

iii) The risk-neutral factor dynamics is such that the conditional Laplace transform is exponential affine under $\mathbb{Q}$:

$$E^{\mathbb{Q}}[\exp\{-uX_{t+1}\}\,|\underline{X}_t] = \exp\{-a(u)'X_t - b(u)\}\,\text{say},$$

where $\underline{X}_t = (X_t, X_{t-1}, ..., X_0)$.
In other words, factors are CAR processes under $\mathbb{Q}$ [see Darolles, Gourieroux, Jasiak (2006)][22].

Proposition A1

For the Affine Term Structure Models (ATSM), zero-coupon yields are affine functions of the latent factors

$$r(t,h) = \alpha(h)X_t + \beta(h). \tag{15}$$

Moreover, the factor loadings $\alpha(h)$ and $\beta(h)$ are such that :

$$h\alpha(h) = a\left((h-1)\alpha(h-1)\right) + \alpha(1) \tag{16}$$

$$h\beta(h) = b\left((h-1)\alpha(h-1)\right) + \beta(1) + (h-1)\beta(h-1), \tag{17}$$

[22]For sake of simplicity, we consider that the factors X_t are Markov under $\mathbb{Q}$: $\mathbb{Q}(X_{t+1}|\underline{X}_t) = \mathbb{Q}(X_{t+1}|X_t)$.

Proof

We have :

$$\begin{aligned} B(t,h) &= E_t^{\mathbb{Q}}[\exp\{-r(t,1)\}\, B(t+1,h-1)], \\ \exp\{-h\alpha(h)X_t - h\beta(h)\} &= E_t^{\mathbb{Q}}[\exp\{-\alpha(1)X_t - \beta(1) \\ &- (h-1)\alpha(h-1)X_{t+1} - (h-1)\beta(h-1)\}], \end{aligned}$$

$$\begin{aligned} h\alpha(h)X_t + h\beta(h) &= (a\,((h-1)\alpha(h-1)) + \alpha(1))\, X_t \\ &+ b\,((h-1)\alpha(h-1)) + \beta(1) + (h-1)\beta(h-1), \end{aligned}$$

by identifying the terms, we get the recursive equations of Proposition A1. QED.

Conversely, it is also known that, under weak regularity conditions, when all the rates $r(t,h)$, including the short-term rate $r(t,1)$, are affine functions of the factors $\forall h$, the factor dynamics is necessarily CAR under $\mathbb{Q}$[on this topic, see Duffie, Filipovic, Schachermayer (2003), and Gourieroux, Sufana (2006)].

Equations (16), (17) and (18) emphasize the constraints imposed on the term structure coefficients $\alpha(h)$, $\beta(h)$ by an Affine Term Structure Model. It is known that historical ($\mathbb{P}$) and risk-neutral ($\mathbb{Q}$) distributions are weekly linked [see Harrison, Kreps (1979) and Rogers (1997)]. For instance, some factors can feature stationarity under $\mathbb{P}$, nonstationarity under $\mathbb{Q}$, or conversely. Nevertheless, in spite of the weak restrictions on the risk-neutral distribution of the factors, yield curve formula cannot be chosen freely due to the recursive equations, that is, the coefficients $\alpha(h)$ and $\beta(h)$ have to be compatible with a CAR risk-neutral dynamic of the factors.
Researchers on ATSM in discrete time have investigated several $\mathbb{Q}$-dynamics of the factors (Z_t), such as autoregressive Gaussian process [see Langetieg (1980)], autoregressive Gamma process[23] [see Gourieroux, Jasiak (2006)], and Wishart process [see Gourieroux, Sufana (2007)].
In the next subsection, we describe the particular case of Gaussian Affine Term Structure Models (GATSM).

ii) Gaussian ATSM

Definition A2

A Gaussian Affine Term Structure Model is an ATSM with Gaussian autoregressive factors under the risk-neutral distribution $\mathbb{Q}$:

$$X_t = \mu + \Phi X_{t-1} + \Sigma\epsilon_t, \text{ where } \epsilon_t \sim \mathcal{N}(0, I), \tag{18}$$

under $\mathbb{Q}$. Their corresponding conditional Laplace transform under the risk-neutral distribution is such that :

[23]Term structure models with autoregressive Gamma factors are discrete equivalent of the Cox, Ingersoll, Ross (1985) term structure model [see Gourieroux, Jasiak (2006)].

$$\log E_{t-1}^{\mathbb{Q}}\left(\exp\left\{-uX_t\right\}\right) = -u\Phi X_{t-1} - u\mu + \frac{1}{2}u\Sigma\Sigma' u', \tag{19}$$

which is exponential affine.

Proposition A2

In a Gaussian ATSM, the rates are given by

$$\begin{aligned} r(t,h) &= \beta(1) + \frac{1}{h}\alpha(1)\left(\sum_{k=0}^{h-2}\Phi^k\right)\mu + \frac{1}{h}\alpha(1)\left(\sum_{k=0}^{h-1}\Phi^k\right)X_t \\ &- \frac{1}{2h}\sum_{j=1}^{h-1}\alpha(1)\left(\sum_{k=0}^{h-1-j}\Phi^k\right)\Sigma\Sigma'\left(\sum_{k=0}^{h-1-j}\Phi^k\right)'\alpha'(1). \end{aligned} \tag{20}$$

In particular, the rates are Gaussian under the risk-neutral distribution.

Proof

As shown in Proposition A1, the term structure coefficients are defined by the conditional Laplace transform of the factors (given by (19) in the Gaussian autoregressive case). Therefore :

$$\begin{aligned} h\alpha(h) &= (h-1)\alpha(h-1)\Phi + \alpha(1) \qquad (21) \\ h\beta(h) &= (h-1)\alpha(h-1)\mu + \frac{1}{2}(h-1)^2\alpha(h-1)\Sigma\Sigma'\alpha'(h-1) + \beta(1) + (h-1)\beta(h-1) \end{aligned}$$

from which is derived the closed form formula in Proposition A2. QED

Corollary A2

If all the eigenvalues of Φ have modulus less than 1, then $\left(\sum_{k=0}^{h}\Phi^k\right)$ converges and $(I-\Phi)$ is invertible. In that case, the zero-coupon yields are given by :

$$\begin{aligned} r(t,h) &= \frac{1}{h}\alpha(1)(Id-\Phi)^{-1}(Id-\Phi^h)X_t \\ &+ \beta(1) + \frac{1}{h}\alpha(1)\sum_{k=1}^{h-1}(I-\Phi)^{-1}(I-\Phi^k)\mu \qquad (22) \\ &- \frac{1}{2h}\sum_{k=1}^{h-1}\alpha(1)(I-\Phi)^{-1}(I-\Phi^k)\Sigma\Sigma'(I-\Phi^k)'(I-\Phi)'^{-1}\alpha'(1). \end{aligned}$$

Appendix 2
Zero Coupon Term Structure Estimation Techniques

i) Market incompleteness and lack of identification

Let us denote $P(t, C^j)$ the price of a h_j time-to-maturity coupon bond, denoted j, where C^j represents the sequence of future cash-flows $C^j_{t+k_j}$ paid at dates $t + k_j$ $(1 \leq k_j \leq h_j)$ to the holder of the j coupon bond. Under no-arbitrage, the price of any coupon bond $j = 1, ..., J$, actively traded on the market, can be written as :

$$P(t, C^j) = \sum_{j=1}^{J} B(t, k_j) C^j_{t+k_j}, \tag{23}$$

where $B(t, k_j)$ represents an admissible price at date t for a zero-coupon bond maturing at $t + k_j$, with $(1 \leq k_j \leq h_j)$.
Let us now assume that the coupons are paid monthly and we want to derive an admissible yield curve at monthly maturities. In general, the bonds, which are actively traded, are coupon bonds, even if a limited number of zero-coupon bonds can be. Moreover, there is a limited number of bonds actively traded daily, with varying design, and this number J_t, which depends on date t is much smaller than the number H of zero-coupon bond maturities, we are interested in (equal to 360 if we want to analyze up to 30Y). Thus, we cannot identify the underlying zero-coupon yield curve in a unique way from the observed coupon prices. In other words, the bond market is incomplete.
This problem of identification can be reduced in two ways :
First by increasing the number of bonds actively traded. This is the purpose of the increased issuing of 30Y T-bonds, or the reason for creating the STRIPS market. Second, by introducing some assumed structure on the zero-coupon yields. Then a yield curve will be obtained by mixing appropriately the market data and the assumptions of the model.

ii) Model based approaches

One can usually circumvent the identification problem by approximating zero-coupon bond prices, either by a parametric form [see Nelson, Siegel (1987)], or by a spline representation [see for instance McCulloch (1971)].

Cross-sectional Calibration

Let us first consider a calibration approach performed daily, say. The parametric specification of the zero-coupon prices can be written as :

$$B(t, h) = g(X_t, h) + \epsilon_{t,h}, \text{ say,}$$

where $B(t, h)$ is the price of a zero-coupon bond with time-to-maturity h, $g(X_t, h)$ is the selected specification parametrized by time dependent parameters X_t, and $\epsilon_{t,h}$ are the

errors terms.

For instance, a spline approach approximates the zero-coupon bond prices by combining polynomial functions. Formally, the term structure of zero-coupon bond prices is separated into L different segments, such as :

$$g(X_t, h) = \sum g_l(X_t)\mathbb{I}(h \geq k_l),$$

where $(k_l)_{l=1...L}$ are the nodes of the spline, assumed fixed a priori, and the $g_l(X_t)$ are polynomial functions satisfying regularity conditions at the nodes. In the literature, it has been suggested to use splines of order 0 (in this case, the zero-coupon bond prices are approximated by a step function), or order 1 (which leads to piecewise linear functions for zero-coupon bond prices).
For instance, the unsmoothed Fama-Bliss (1987) approach relies on splines of order 0, which tend to generate very erratic term structures. On the other hand, McCulloch (1971) proposes to use either quadratic or cubic splines. Recent splines approach [see for instance Fisher, Nychka, Zervos (1995)] slightly modify McCulloh's method by imposing an additional parameter, which penalizes excess curvature in the fitted zero-coupon yield curve. These approach are called smoothed cubic spline methods. It has also been proposed [see Vasiceck, Fong (1982)] to approximate the zero-coupon yields rather than the zero-coupon bond prices. This approach is equivalent to approximate the zero-coupon bond prices by exponential splines.

Alternatively, one can use a Nelson-Siegel approximation, such as :

$$g(X_t, h) = \exp\{-hf(X_t, h)\},$$

where $f(X_t, h)$ is the approximated zero-coupon rate with time-to-maturity k. In the Nelson-Siegel case [see equation (1) in Section 2.1] :

$$f(X_t, h) = X_{1t} + X_{2t}\frac{1 - \exp\{-hX_{4t}\}}{hX_{4t}} + X_{3t}\left(\frac{1 - \exp\{-hX_{4t}\}}{hX_{4t}} - \exp\{-hX_{4t}\}\right).$$

Once a daily parametric specification is selected, the price of an observed coupon bond can be written as :

$$P(t, C^j) = \sum_{k_j}\left(g(X_t, k_j)C^j_{t+k_j} + \epsilon_{t,k_j}C^j_{t+k_j}\right).$$

Thus, the error on the price of a coupon bond :

$$\eta_{t,j} = \sum_{k_j} C^j_{t+k_j}\epsilon_{t,k_j},$$

is a combination of the errors on the zero-coupon bonds, weighted by the coupons. Any least squares calibration method, will have to account for these coupon effects: it has to be weighted in an appropriate way, and the estimated parameters will be defined as :

$$\hat{X}_t = \arg\min_{X_t} \sum_j (w_{t,j}\eta_{t,j})^2.$$

Usually, the weights are chosen such as the corresponding errors on zero-coupon yields have a same variance. By considering zero-coupon, we account for the coupon effects, and by considering the yield instead of the prices, we get an adjustment for time-to-maturity. For computational reason, GSW(2007) have proposed to approximate the approach above, by applying on coupon bond prices a weighted nonlinear least squares method with weights the inverse of the durations of the observed bonds.
The concept of duration of a bond has been introduced by MacCaulay (1938) and defined as :

$$D(t,j) = \frac{\sum_{j=1}^{J} k_j B(t,k_j) C^j_{t+k_j}}{P(t,C^j)},$$

where $D(t,j)$ is the duration at date t of the bond j, and C^j represents the future cash-flow sequence $C^j_{t+k_j}$ paid at dates $t+k_j$ $(1 \leq k_j \leq h_j)$ to the holder of the j bond [as in i) above].
The duration of a bond increases with its time-to-maturity and its coupon rate; thus the GSW (2007) procedure underweights the pricing errors of long-term bonds paying large coupons.

Global calibration

It is also possible to partially relate the models of the different days. This is usually done by considering a model including both time varying parameters X_t, and time independent parameters θ, say. The parametric specification becomes:

$$B(t,h) = g(X_t, \theta, h) + \epsilon_{t,h}.$$

Then, the calibration is done jointly on all observations with respect to both types of parameters. The least square criterion involves a double sum on both time and maturities :

$$\left(\hat{\theta}, \hat{X}_t\right) = \arg\min_{\theta, X_t} \sum_t \sum_j (w_{t,j}\eta_{t,j})^2.$$