Can Spanned Term Structure Factors Drive Stochastic Volatility?[†]

Jens H. E. Christensen

Jose A. Lopez

Glenn D. Rudebusch

Federal Reserve Bank of San Francisco

101 Market Street, Mailstop 1130

San Francisco, CA 94105

Abstract

The ability of the usual factors from empirical arbitrage-free representations of the term structure—that is, spanned factors—to account for interest rate volatility dynamics has been much debated. We estimate new arbitrage-free Nelson-Siegel (AFNS) term structure specifications that allow for stochastic volatility to be linked to one or more of the spanned AFNS yield curve factors. Our results with three separate daily data sets—U.S. Treasury yields, U.K. gilt yields, and U.S. dollar swap and LIBOR rates—suggest that much observed stochastic volatility cannot be associated with spanned term structure factors in terms of time-series correlations at high frequency. However, some of the AFNS models with stochastic volatility do provide a close fit to our measure of realized yield volatility in addition to providing a good fit to the yield term structure.

1   Introduction

Understanding and predicting the variability of interest rates plays a crucial role in derivatives pricing and portfolio risk management, so creating good empirical models of interest rate stochastic volatility has been a key research priority. Unfortunately, while the canonical affine arbitrage-free term structure models have been widely applied to price bonds, the ability of these popular models to capture the changing volatility of interest rates has been seriously questioned. Indeed, using U.S. swap rate data, Collin-Dufresne et al. (2009) find that a standard three-factor affine model, in which one of the factors drives volatility, produces very poor volatility estimates. They suggest that an unspanned volatility factor (i.e., a factor that does not influence the model's conditional mean dynamics under the pricing measure) has to be added to the affine term structure model in order to capture the stochastic volatility observed in U.S. dollar swap rates. However, Jacobs and Karoui (2009) argue the conclusions of Collin-Dufresne et al. (2009) depend in large part on the particular sample period analyzed, and they strongly recommend further research to examine whether (spanned) factors in an affine arbitrage-free model can capture the conditional volatility in U.S. Treasury and swaps data. In this paper, we conduct such research by examining to what extent spanned factors can generate stochastic interest rate volatility.

In previous work, analysis of multiple factor sources for spanned volatility have been hampered by problems in estimating the parameters of multifactor affine models.[1] To avoid these difficulties, we incorporate spanned stochastic volatility into the class of affine, arbitrage-free Nelson-Siegel (AFNS) term structure models developed by Christensen, Diebold, and Rudebusch (CDR, 2007). These models are characterized by imposing the level, slope and curvature factors used in the original Nelson-Siegel yield curve model and observed in principal components analysis of interest rates. This class of models captures both the cross section of yields and their time-series dynamics quite well and can be readily estimated. We introduce five new specifications of AFNS models that incorporate stochastic volatility. The first two specifications allow for one factor—either the level or curvature factor—to generate stochastic volatility, and these are denoted $AFNS_1$-L and $AFNS_1$-C, respectively.[2,3] The third and fourth specifications allow for two factors to generate stochastic volatility. These are denoted as $AFNS_2$-L,C when the level and curvature factors generate stochastic volatility, and $AFNS_2$-

---

[1]The latent nature of the factors and the over-parameterization of the models make estimation quite difficult. See Kim and Orphanides (2005) and Duffee (2008).

[2]As explained in Section 3, it is not possible to specify a stochastic volatility model based on the slope factor within the AFNS framework.

[3]Our nomenclature draws on Dai and Singleton (2000). Our $AFNS_n$ models are members of their $A_n(3)$ class of models, which have three state variables and n square-root processes.

S,C when the slope and curvature factors generate stochastic volatility.[4] Finally, the fifth specification, denoted $AFNS_3$, allows all three factors to generate stochastic volatility. A key advantage of our approach to modeling stochastic volatility is that the factors remain well-defined as level, slope, and curvature and do not change for any admissible parameter set. This structure makes the results comparable across model classes and allows us to detail which factors are able to generate stochastic yield volatility most similar to that observed in the data. This feature distinguishes our approach from the existing literature on affine models where the optimal parameters for any unconstrained affine model only implicitly reveal which factor(s) generate(s) stochastic volatility.[5]

In the existing literature, a few papers have tried to incorporate stochastic volatility into the dynamic Nelson-Siegel models introduced by Diebold and Li (2006). Hautsch and Ou (2009) incorporate stochastic volatility into the dynamic Nelson-Siegel model by including three additional state variables that are the drivers of the stochastic volatility in the level, slope, and curvature, respectively. They find that the stochastic volatility factors of the slope and curvature factor, in particular, contain important information that help forecast excess holding period returns on Treasury bonds. Koopman et al. (2008) incorporate one common factor driving volatility in the fitted errors across all maturities in their sample. They also try to allow for a common volatility factor that directly causes stochastic volatility in the three state variables, but they find limited gains from that specification. Unlike the approach we detail in this paper, neither of these papers address the problem of eliminating the existence of arbitrage opportunities inherent in the standard dynamic Nelson-Siegel model.

In the empirical part of the paper, we examine the performance of these new model classes on three datasets. We first examine daily U.S. Treasury yields from the Gürkaynak et al. (2007) database over the period from January 2, 1985 to March 1, 2010 for eight maturities. Focusing on the most parsimonious specification in which the three factors are independent, we find that the introduction of stochastic volatility does not weaken the models' in-sample fit of the term structure relative to the $AFNS_0$ model with constant volatility. With respect to the models' fitted stochastic volatility measured in terms of standard deviations, we find that different magnitudes of variation and correlations with our measure of realized standard deviations based on the daily data are induced. The correlation between the fitted and realized bond yield standard deviations is quite low and often negative over the full sample. However,

---

[4]The third possible specification in which the level and slope factors generate stochastic volatility is not compatible with the AFNS framework as detailed in Section 3.

[5]We only use bond yields in the model estimation and leave the issue of "unspanned stochastic volatility", a condition where bond prices are unaffected by changes in interest rate volatility, as per Collin-Dufresne and Goldstein (2002) and Collin-Dufresne et al. (2009), for future research.

Jacobs and Karoui (2009) find that, for U.S. Treasury yields, affine term structure models are much better able to generate stochastic volatility measures that correspond to the observed data in the period prior to 1992 for as of yet unclear reasons. Our correlations confirm this result as they increased markedly to roughly 30% for several models in the pre-1992 part of our sample. If, instead of focusing on time-series correlations at high frequency, we use the root-mean-squared errors (RMSE) between the fitted and realized standard deviations as model validation, the results are more favorable to the affine models and the spanned factors. In particular the $AFNS_3$ model, which exhibits the most variation in the fitted volatility measure, performs well based on this measure with RMSEs below 15 basis points at all maturities in addition to providing a good fit to the cross section of yields.

In the second empirical exercise, we examine daily U.K. gilt yields downloaded from the website of the Bank of England covering the same period and the same eight maturities as the U.S. Treasury data. In general, the results accord with the findings from the U.S. data. First, the introduction of stochastic volatility has little effect on the in-sample yield fit as compared to the Gaussian $AFNS_0$ model. Second, the time-series correlations between the fitted and realized measures of yield volatility are weak and sample dependent with the 1992-2002 period exhibiting the largest, positive correlations, but still not exceeding 42%. Third, and more importantly, the $AFNS_3$ model provides the closest fit to the measure of realized yield volatility with RMSEs of around 19 basis points for the shortest maturities down to 11 basis points for the longest maturities.

For robustness, our third and last empirical exercise looks at the daily U.S. dollar swap and LIBOR data examined by Collin-Dufresne et al. (2009). As with the U.S. Treasury and U.K. gilt data, we do not find important differences between the in-sample performance of the stochastic volatility models and the $AFNS_0$ model. However, with respect to the models' fitted stochastic volatility series, we find again that the $AFNS_3$ model induces a reasonable degree of variation and provides a close fit to the realized volatility measure with RMSEs ranging from 16 basis points for the six-month LIBOR down to 10 basis points for the ten-year swap rate. This advantage does not translate as readily into superior performance with respect to correlation with the realized standard deviations based on daily data as other models generate higher correlations than the $AFNS_3$ model. The correlations between the fitted and realized standard deviations for the full sample are high for the $AFNS_1$-L and $AFNS_2$-L,C models, with values above 60% for the six-month LIBOR rate and roughly 20% for the ten-year swap rate. The $AFNS_3$ model generates quite low and even negative correlations over the full period, but the correlations increase to nearly 50% for the short-term rates in the pre-1992 period

highlighted by Jacobs and Karoui (2009). Yet, the other two models based on the stochastic level factor perform quite well with high correlations at the longer maturities. These models also have relatively high correlations in the post-1991 sample, whereas the AFNS$_3$ model has negative correlations for seven of the eight maturities considered.

To summarize, our results suggest that, while incapable of matching the high-frequency time variation of realized yield volatilities, three-factor affine models can be relied upon to provide a close fit to both the cross section of yields and their realized volatility for U.S. and U.K. government bond yields as well as U.S. dollar swap and LIBOR rates. In this sense, spanned yield curve factors can be said to be able to capture a large part of the realized yield volatility. However, the important question as to whether the non-explained part of yield volatility can be profitably exploited and hence require the introduction of unspanned stochastic volatility as advocated by Collin-Dufresne et al. (2009), is beyond the scope of this paper and we leave it for future research.

The rest of the paper is structured as follows. Section 2 presents a short summary of the AFNS model of the term structure. Section 3 presents our five classes of modified AFNS models with volatility dynamics. Section 4 presents empirical results for the daily U.S. Treasury yields data, Section 5 reports results for the daily U.K. gilt yields, and Section 6 presents the results for the weekly U.S. swaps and LIBOR data. Section 7 concludes. An appendix contains additional technical details.

2   The AFNS Model with Constant Volatility

In this section, we briefly review the AFNS model with constant volatility (that is, the AFNS$_0$ specification).[6] We start from a standard continuous-time affine arbitrage-free structure (Duffie and Kan, 1996) that underlies all the models in this paper. To represent an affine diffusion process, define a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), Q)$, where the filtration $(\mathcal{F}_t) = \{\mathcal{F}_t : t \geq 0\}$ satisfies the usual conditions; see Williams (1997). The state variable X_t is assumed to be a Markov process defined on a set $M \subset \mathbf{R}^n$ that solves the following stochastic differential equation (SDE):

$$dX_t = K^Q(t)[\theta^Q(t) - X_t]dt + \Sigma(t)D(X_t, t)dW_t^Q, \tag{1}$$

---

[6]This model has been shown to exhibit both good in-sample fit and out-of-sample forecast accuracy for various yield curves. The empirical analysis conducted in Christensen et al. (2007) is based on unsmoothed Fama-Bliss data for nominal Treasury yields. Christensen et al. (2010) examine yields for nominal and real Treasuries as per Gürkaynak et al. (2007, 2010). Christensen et al. (2009) examine short-term LIBOR and highly-rated financial firms' corporate bond rates, while Christensen and Lopez (2008) examines corporate bond rates from a broad set of industrial sectors and credit ratings.

where W^Q is a standard Brownian motion in $\mathbf{R}^n$, the information of which is contained in the filtration $(\mathcal{F}_t)$.[7] The drift terms $\theta^Q : [0, T] \to \mathbf{R}^n$ and $K^Q : [0, T] \to \mathbf{R}^{n \times n}$ are bounded, continuous functions.[8] Similarly, the volatility matrix $\Sigma : [0, T] \to \mathbf{R}^{n \times n}$ is assumed to be a bounded, continuous function, while $D : M \times [0, T] \to \mathbf{R}^{n \times n}$ is assumed to have the following diagonal structure:

$$
\begin{pmatrix}
\sqrt{\gamma^1(t) + \delta^1(t)X_t} & \cdots & 0 \\
\vdots & \ddots & \vdots \\
0 & \cdots & \sqrt{\gamma^n(t) + \delta^n(t)X_t}
\end{pmatrix},
$$

where

$$
\gamma(t) = \begin{pmatrix} \gamma^1(t) \\ \vdots \\ \gamma^n(t) \end{pmatrix}, \quad
\delta(t) = \begin{pmatrix} \delta_1^1(t) & \cdots & \delta_n^1(t) \\ \vdots & \ddots & \vdots \\ \delta_1^n(t) & \cdots & \delta_n^n(t) \end{pmatrix},
$$

$\gamma : [0, T] \to \mathbf{R}^n$ and $\delta : [0, T] \to \mathbf{R}^{n \times n}$ are bounded, continuous functions, and $\delta^i(t)$ denotes the ith row of the $\delta(t)$-matrix. Finally, the instantaneous risk-free rate is assumed to be an affine function of the state variables

$$
r_t = \rho_0(t) + \rho_1(t)' X_t,
$$

where $\rho_0 : [0, T] \to \mathbf{R}$ and $\rho_1 : [0, T] \to \mathbf{R}^n$ are bounded, continuous functions.

Duffie and Kan (1996) prove that zero-coupon bond prices in this framework are exponential-affine functions of the state variables

$$
P(t, T) = E_t^Q \big[\exp \big(- \int_t^T r_u du \big) \big] = \exp \big(B(t, T)' X_t + A(t, T) \big),
$$

where $B(t, T)$ and $A(t, T)$ are the solutions to the following system of ordinary differential equations (ODEs)

$$
\frac{dB(t, T)}{dt} = \rho_1 + (K^Q)' B(t, T) - \frac{1}{2} \sum_{j=1}^n (\Sigma' B(t, T) B(t, T)' \Sigma)_{j,j} (\delta^j)', \quad B(T, T) = 0, \quad (2)
$$

$$
\frac{dA(t, T)}{dt} = \rho_0 - B(t, T)' K^Q \theta^Q - \frac{1}{2} \sum_{j=1}^n (\Sigma' B(t, T) B(t, T)' \Sigma)_{j,j} \gamma^j, \quad A(T, T) = 0, \quad (3)
$$

---

[7]Note that the affine property applies to bond prices; therefore, affine models only impose structure on the factor dynamics under the pricing measure.

[8]Stationarity of the state variables is ensured if all the eigenvalues of $K^Q(t)$ are positive. If the eigenvalues are complex, the real component should be positive; see Ahn et al. (2002). However, stationarity is not a necessary requirement for the process to be well defined.

and the possible time-dependence of the parameters is suppressed in the notation. These pricing functions imply that the zero-coupon yields are given by affine functions of X_t

$$y(t,T) = -\frac{1}{T-t}\log P(t,T) = -\frac{B(t,T)'}{T-t}X_t - \frac{A(t,T)}{T-t}.$$

In the AFNS model with constant volatility, the instantaneous risk-free rate is defined by

$$r_t = X_t^1 + X_t^2.$$

In addition, the three state variables in the model $X_t = (X_t^1, X_t^2, X_t^3)$ are described by the following system of SDEs under the risk-neutral Q-measure:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix}\left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}\right]dt + \Sigma\begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}, \quad \lambda > 0.$$

In matrix notation, this system is denoted as

$$dX_t = K^Q(\theta^Q - X_t)dt + \Sigma^Q dW_t^Q.$$

CDR (2007) show that this specification implies that zero-coupon bond yields are given by

$$y(t,T) = X_t^1 + \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)}\right)X_t^2 + \left(\frac{1 - e^{-\lambda(T-t)}}{\lambda(T-t)} - e^{-\lambda(T-t)}\right)X_t^3 - \frac{A(t,T)}{T-t}.$$

Importantly, the factor loadings in this yield function match the level, slope, and curvature loadings introduced in Nelson and Siegel (1987) with a final yield-adjustment term, which represents convexity effects due to Jensen's inequality.

The model is completed with a risk premium specification that connects the factor dynamics to the dynamics under the real-world (or historical) P-measure. It is important to note that there are no restrictions on the dynamic drift components under the empirical P-measure beyond the requirement of constant volatility. To facilitate empirical implementation, we use the extended affine risk premium developed by Cheridito et al. (2007). In the Gaussian framework, this specification implies that the risk premiums Γ_t depend on the state variables; that is,

$$\Gamma_t = \gamma^0 + \gamma^1 X_t,$$

where $\gamma^0 \in \mathbf{R}^3$ and $\gamma^1 \in \mathbf{R}^{3\times3}$ contain unrestricted parameters. The relationship between

real-world yield curve dynamics under the P-measure and risk-neutral dynamics under the Q-measure is given by

$$dW_t^Q = dW_t^P + \Gamma_t dt.$$

Thus, the P-dynamics of the state variables are

$$dX_t = K^P(\theta^P - X_t)dt + \Sigma dW_t^P,$$

where both K^P and θ^P are allowed to vary freely relative to their counterparts under the Q-measure. Following CDR, we identify this class of models by fixing the θ^Q means under the Q-measure at zero without loss of generality. Furthermore, CDR show that Σ cannot be more than a triangular matrix for the model to be identified. Thus, the maximally flexible specification of the original AFNS model has Q-dynamics given by

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix},$$

while its P-dynamics are given by

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

The main limitation of the AFNS class of models above is the constant volatility matrix Σ. The purpose of this paper is to modify the AFNS model in a straightforward fashion in order to incorporate stochastic volatility. The key assumption to preserving the desirable Nelson-Siegel factor loading structure in the zero-coupon bond yield function is to maintain the K^Q mean-reversion matrix under the Q-measure. Furthermore, all model classes will be characterized by an instantaneous risk-free rate defined as the sum of the first two factors

$$r_t = X_t^1 + X_t^2.$$

3 Five AFNS Specifications with Stochastic Volatility

In this section, we present five AFNS specifications with stochastic volatility that vary depending on whether they contain one, two, or three stochastic volatility factors and on the

identity of those factors. For each model class, we derive the maximally flexible specification that can be obtained using the extended affine risk premium specification introduced in Cheridito et al. (2007).

3.1   AFNS Models with One Stochastic Volatility Factor

There are two AFNS stochastic volatility specifications that allow just one factor to exhibit stochastic volatility. The first, denoted as the AFNS_1-L model, allows only the level factor to exhibit stochastic volatility. The state variables in this specification follow this system of stochastic differential equations under the risk-neutral Q-measure:

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_t^1} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix},
$$

where the X_t^1 level factor is a square-root process with stochastic volatility that affects the instantaneous volatility of the two other factors through the β_{21} and β_{31} volatility sensitivity parameters.[9]

For the factor loadings in the zero-coupon bond prices, $B^1(t,T)$ is the solution to

$$
\begin{aligned}
\frac{dB^1(t,T)}{dt} &= 1 + \varepsilon B^1(t,T) - \frac{1}{2}\sigma_{11}^2 B^1(t,T)^2 - \frac{1}{2}\sigma_{21}^2 B^2(t,T)^2 - \frac{1}{2}\sigma_{31}^2 B^3(t,T)^2 \\
&\quad - \sigma_{21}\sigma_{11}B^1(t,T)B^2(t,T) - \sigma_{31}\sigma_{11}B^1(t,T)B^3(t,T) - \sigma_{21}\sigma_{31}B^2(t,T)B^3(t,T) \\
&\quad - \frac{1}{2}\beta_{21}\left[\sigma_{22}^2 B^2(t,T)^2 + \sigma_{32}^2 B^3(t,T)^2 + 2\sigma_{22}\sigma_{32}B^2(t,T)B^3(t,T)\right] - \frac{1}{2}\beta_{31}\sigma_{33}^2 B^3(t,T)^2,
\end{aligned}
$$

while $B^2(t,T)$ and $B^3(t,T)$ are given by

$$
B^2(t,T) = -\left(\frac{1-e^{-\lambda(T-t)}}{\lambda}\right),
$$
$$
B^3(t,T) = (T-t)e^{-\lambda(T-t)} - \left(\frac{1-e^{-\lambda(T-t)}}{\lambda}\right).
$$

The last two factor loadings match exactly the factor loadings of the slope and curvature

---

[9]Note that we cannot set κ_{11}^Q to zero as that would eliminate the drift of X_t^1 and cause this process to remain at zero once it hits zero, which it will P-a.s. Instead, we fix this parameter at a small, but positive, $\varepsilon = 10^{-6}$, to get close to the unit-root property imposed in the AFNS_0 model.

factors in the Nelson-Siegel zero-coupon yield function, while the ODE for $B^1(t,T)$ contains quadratic elements related to the stochastic volatility of X_t^1. The $A(t,T)$-function in the yield-adjustment term in this class of models must solve the following ODE:

$$\frac{dA(t,T)}{dt} = -B(t,T)'K^Q\theta^Q - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2 - \frac{1}{2}(\sigma_{32}^2 + \sigma_{33}^2)B^3(t,T)^2 - \sigma_{22}\sigma_{32}B^2(t,T)B^3(t,T).$$

To estimate this model, we specify the dynamics under the real-world P-measure as the measure change $dW^Q = dW_t^P + \Gamma_t dt$. Note that we are limited to the essentially affine risk premium structure introduced by Duffee (2002) for this particular model class.[10] Given this limitation, the maximally flexible affine P-dynamics are, in general, given by

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix}\left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}\right] dt$$
$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix}\begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_t^1} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

For the first factor with stochastic volatility, there is a restriction on the mean parameter θ_1^P that we implement as[11]

$$\theta_1^P = \frac{\varepsilon \cdot \theta_1^Q}{\kappa_{11}^P}.$$

Furthermore, for this process to be well-defined under both probability measures, we require that

$$\kappa_{11}^P\theta_1^P > 0 \quad \text{and} \quad \varepsilon \cdot \theta_1^Q > 0.$$

These two inequalities are satisfied provided $\kappa_{11}^P > 0$ and $\theta_1^Q > 0$. These restrictions ensure that the X_t^1-process will move into positive territory whenever it hits the lower zero-boundary. Finally, we identify this class of models by fixing $\theta_2^Q = \theta_3^Q = 0$, eliminating the Q-means of the unconstrained processes as in CDR (2007). These restrictions allow the corresponding means under the P-measure to be determined in the estimation. There are 19 parameters

---

[10] We cannot use the extended affine risk premium specification for this particular specification because of the restriction imposed on κ_{11}^Q to obtain a level factor structure as similar as possible to the one in the Nelson-Siegel model. If we were to do so, the Feller condition for X_t^1 could not reasonably be expected to be satisfied under the Q-measure as X_t^1 approaches a unit-root process. Please see the technical appendix for further details on this point.

[11] A similar approach is used in the other model classes with stochastic volatility generated by the level factor.

in the maximally flexible specification of this class of models. In contrast, if we assume the factors are independent for the sake of parsimony, the number of parameters is reduced to 12.

The natural next AFNS one-factor stochastic volatility specification would allow the slope factor to exhibit stochastic volatility. However, examination of the matrix

$$K^Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix},$$

shows that X_t^2 cannot be a square-root process with X_t^3 as an unconstrained process, if the important off-diagonal element κ_{23}^Q is to remain equal to $-\lambda$, which generates the unique factor loading of the curvature factor in the AFNS model. Thus, there is no admissible AFNS$_1$-S model. Instead, we turn to the AFNS$_1$-C model by allowing the curvature factor to be a stochastic volatility factor. This approach preserves the properties of the level and slope factors, allows the curvature factor to continue serving as the stochastic mean of the slope factor under the pricing measure, and designates the curvature factor to be the source of stochastic volatility in the model.

For the AFNS$_1$-C model, we assume that the state variables X_t are described under the risk-neutral Q-measure as:

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$

$$+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1+\beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.$$

The curvature factor here is a square-root process that induces stochastic volatility in the other two factors through the β_{13} and β_{23} volatility sensitivity parameters.

In this model class, the first two factor loadings are identical to those in the $A_0(3)$ model, while $B^3(t,T)$ is the solution to:

$$\begin{aligned} \frac{dB^3(t,T)}{dt} &= -\lambda B^2(t,T) + \lambda B^3(t,T) - \frac{1}{2}\sigma_{13}^2 B^1(t,T)^2 - \frac{1}{2}\sigma_{23}^2 B^2(t,T)^2 - \frac{1}{2}\sigma_{33}^2 B^3(t,T)^2 \\ &\quad - \sigma_{13}\sigma_{23}B^1(t,T)B^2(t,T) - \sigma_{13}\sigma_{33}B^1(t,T)B^3(t,T) - \sigma_{23}\sigma_{33}B^2(t,T)B^2(t,T) \\ &\quad - \frac{1}{2}\beta_{13}\sigma_{11}^2 B^1(t,T)^2 - \frac{1}{2}\beta_{23}\left[\sigma_{12}^2 B^1(t,T)^2 + \sigma_{22}^2 B^2(t,T)^2 + 2\sigma_{12}\sigma_{22}B^1(t,T)B^2(t,T)\right]. \end{aligned}$$

10

The $A(t,T)$-function in the yield-adjustment term in this class of models solves the ODE:

$$\frac{dA(t,T)}{dt} = -B(t,T)'K^Q\theta^Q - \frac{1}{2}(\sigma_{11}^2 + \sigma_{12}^2)B^1(t,T)^2 - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2 - \sigma_{12}\sigma_{22}B^1(t,T)B^2(t,T).$$

We estimate this model using the extended affine risk premium specification such that the measure change is $dW^Q = dW_t^P + \Gamma_t dt$. The maximally flexible affine P-dynamics are, in general, given by

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ 0 & 0 & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$
$$+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1 + \beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

To keep the model arbitrage-free, X_t^3 cannot be allowed to hit the zero-boundary. This outcome is prevented by requiring that the parameters for the X_t^3-process satisfy the Feller condition under both probability measures; i.e.,

$$\kappa_{33}^P\theta_3^P > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.$$

Finally, we identify this class of models by fixing $\theta_1^Q = \theta_2^Q = 0$, which allows the means under the P-measure of the unconstrained factors to vary freely and be determined in the estimation. In total, there are 20 free parameters in the maximally flexible specification of this model class and 13 for the independent factor specification.

3.2 AFNS Models with Two Stochastic Volatility Factors

Our second class of stochastic volatility models allows for two stochastic volatility factors. Although there are three potential specifications, the specification with just the level and slope factors exhibiting stochastic volatility is not admissible because it does not permit the important off-diagonal element κ_{23}^Q to equal $-\lambda$, which is the unique characteristic of the curvature factor in the original AFNS model. Instead, stochastic volatility is associated with either level and curvature or slope and curvature. The first of these specifications, denoted

11

AFNS$_2$-L,C, has factor dynamics under the risk-neutral Q-measure given by[12]

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{21} X_t^1 + \beta_{23} X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

The X_t^1 and X_t^3 factors exhibit stochastic volatility and induce volatility in the X_t^2 factor via the β_{21} and β_{23} volatility sensitivity parameters.

The factor loadings in the zero-coupon bond price function are the unique solutions to the following set of ODEs:

$$
\begin{aligned}
\frac{dB^1(t,T)}{dt} &= 1 + \varepsilon B^1(t,T) - \frac{1}{2}\sigma_{11}^2 B^1(t,T)^2 - \frac{1}{2}\sigma_{21}^2 B^2(t,T)^2 \\
&\quad - \sigma_{11}\sigma_{21}B^1(t,T)B^2(t,T) - \frac{1}{2}\beta_{21}\sigma_{22}^2 B^2(t,T)^2, \\
\frac{dB^2(t,T)}{dt} &= 1 + \lambda B^2(t,T), \\
\frac{dB^3(t,T)}{dt} &= -\lambda B^2(t,T) + \lambda B^3(t,T) - \frac{1}{2}\sigma_{33}^2 B^3(t,T)^2 - \frac{1}{2}\sigma_{23}^2 B^2(t,T)^2 \\
&\quad - \sigma_{23}\sigma_{33}B^2(t,T)B^3(t,T) - \frac{1}{2}\beta_{23}\sigma_{22}^2 B^2(t,T)^2,
\end{aligned}
$$

where we note that the solution to $B^2(t,T)$ is simply

$$
B^2(t,T) = -\frac{1 - e^{-\lambda(T-t)}}{\lambda}.
$$

Hence, X_t^2 preserves its role as a slope factor. The $A(t,T)$-function is the solution to:

$$
\frac{dA(t,T)}{dt} = -B(t,T)' K^Q \theta^Q - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2.
$$

Using the extended affine risk premium structure, the maximally flexible affine P-dynamics

<hr />

[12]Note that, as before, we fix $\varepsilon = 10^{-6}$ to approximate the unit-root property imposed in the standard AFNS$_0$ model.

are given by

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & 0 & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{21} X_t^1 + \beta_{23} X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

For the level factor, the condition $\varepsilon \cdot \theta_1^Q = \kappa_{11}^P \theta_1^P$ must be satisfied. Furthermore, to keep this model class arbitrage-free, X_t^3 cannot hit the zero-boundary. This outcome is prevented by requiring that the parameters for the X_t^3 process satisfy the Feller condition under both probability measures; i.e.,[13]

$$
\kappa_{31}^P \theta_1^P + \kappa_{33}^P \theta_3^P > \frac{1}{2} \sigma_{33}^2 \quad \text{and} \quad \lambda \theta_3^Q > \frac{1}{2} \sigma_{33}^2.
$$

Finally, to have a well-defined X_t^3 process, the effect of the level factor on the drift of the curvature factor must be positive, which we impose with the $\kappa_{31}^P \leq 0$ constraint. This condition implies that the two square-root processes cannot be negatively correlated. To identify this model class, we fix the θ_2^Q mean at zero. There are 18 parameters in the maximally flexible specification of this class of models and 13 in the independent factors specification.

The second AFNS specification with two volatility factors allows the slope and curvature factors to be square-root processes while the level factor remains unconstrained. The factor dynamics of the AFNS$_2$-S,C model under the Q-measure are

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1 + \beta_{12} X_t^2 + \beta_{13} X_t^3} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

Note that the X_t^2 and X_t^3 square-root processes are positively correlated through the off-diagonal element $\kappa_{23}^Q = -\lambda < 0$. Beyond generating their own stochastic volatility, these two

---

[13] For X_t^1, we just need to ensure that the process does not turn negative, which is assured provided that $\varepsilon \cdot \theta_1^Q > 0$ and $\kappa_{11}^P \theta_1^P > 0$.

factors induce instantaneous volatility for X_t^1 via the β_{12} and β_{13} volatility sensitivities.

For the first factor loading in the zero-coupon bond price function, this structure implies that

$$B^1(t,T) = -(T-t),$$

which preserves the role of the level factor. The next two factor loadings are the unique solutions to:

$$
\begin{aligned}
\frac{dB^2(t,T)}{dt} &= 1 + \lambda B^2(t,T) - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2 - \frac{1}{2}\sigma_{12}^2 B^1(t,T)^2 \\
&\quad - \sigma_{12}\sigma_{22}B^1(t,T)B^2(t,T) - \frac{1}{2}\beta_{12}\sigma_{11}^2 B^1(t,T)^2, \\
\frac{dB^3(t,T)}{dt} &= -\lambda B^2(t,T) + \lambda B^3(t,T) - \frac{1}{2}\sigma_{33}^2 B^3(t,T)^2 - \frac{1}{2}\sigma_{13}^2 B^1(t,T)^2 \\
&\quad - \sigma_{13}\sigma_{33}B^1(t,T)B^3(t,T) - \frac{1}{2}\beta_{13}\sigma_{11}^2 B^1(t,T)^2.
\end{aligned}
$$

The $A(t,T)$-function in the yield-adjustment term is the solution to

$$
\frac{dA(t,T)}{dt} = -B(t,T)'K^Q\theta^Q - \frac{1}{2}\sigma_{11}^2 B^1(t,T)^2.
$$

Using the extended affine risk premium specification, the maximally flexible affine P-dynamics can be written as

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} =
\begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ 0 & \kappa_{22}^P & \kappa_{23}^P \\ 0 & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix}
\left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix}
\begin{pmatrix} \sqrt{1+\beta_{12}X_t^2+\beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix}
\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

To keep this class of models arbitrage-free, the X_t^2 and X_t^3 factors must avoid hitting the zero-boundary. This outcome is ensured by imposing the Feller condition on their parameters as follows:

$$
\kappa_{22}^P\theta_2^P + \kappa_{23}^P\theta_3^P > \frac{1}{2}\sigma_{22}^2; \quad \lambda\theta_2^Q - \lambda\theta_3^Q > \frac{1}{2}\sigma_{22}^2; \quad \kappa_{33}^P\theta_3^P + \kappa_{32}^P\theta_2^P > \frac{1}{2}\sigma_{33}^2; \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.
$$

Furthermore, for X_t^2 and X_t^3 to be well-defined, the sign of the effect they have on each other must be positive, which we impose using the constraints $\kappa_{23}^P \leq 0$ and $\kappa_{32}^P \leq 0$. This implies

that the two square-root processes cannot be negatively correlated. Finally, we identify this class of models by fixing $\theta_1^Q = 0$, which allows θ^P to vary freely. In total, there are 20 free parameters in the maximally flexible specification and 13 for the independent factors specification.

3.3   AFNS Models with Three Stochastic Volatility Factors

In the fifth and last AFNS$_3$ specification, all three factors exhibit stochastic volatility. The dynamics of X_t are described under the Q-measure as[14]

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \varepsilon & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

In this model class, the factor loadings in the zero-coupon bond price function are given by the unique solution to

$$
\begin{aligned}
\frac{dB^1(t,T)}{dt} &= 1 + \varepsilon B^1(t,T) - \frac{1}{2}\sigma_{11}^2 B^1(t,T)^2, \\
\frac{dB^2(t,T)}{dt} &= 1 + \lambda B^2(t,T) - \frac{1}{2}\sigma_{22}^2 B^2(t,T)^2, \\
\frac{dB^3(t,T)}{dt} &= -\lambda B^2(t,T) + \lambda B^3(t,T) - \frac{1}{2}\sigma_{33}^2 B^3(t,T)^2,
\end{aligned}
$$

while the $A(t,T)$-function in the yield-adjustment term is given by the solution to:

$$
\frac{dA(t,T)}{dt} = -B(t,T)'K^Q\theta^Q.
$$

---

[14]Note that, we again fix $\varepsilon = 10^{-6}$ to approximate the unit-root property imposed in the AFNS$_0$ model.

Applying the extended affine risk premium specification, the maximally flexible affine P-dynamics are given by

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

For X_t^1, the constraint $\varepsilon \cdot \theta_1^Q = \kappa_{11}^P \theta_1^P$ must be satisfied. The limited risk premium specification due to the near unit-root property of X_t^1 also implies that X_t^2 and X_t^3 cannot impact the drift of X_t^1 once κ_{12}^Q and κ_{13}^Q have been fixed at zero. We need these restrictions in order to match the Nelson-Siegel factor loading structure as closely as possible.

To keep this model class arbitrage-free, X_t^2 and X_t^3 must not hit the zero-boundary. We ensure this by imposing the Feller condition on their parameters under both probability measures, i.e.,[15]

$$
\kappa_{21}^P \theta_1^P + \kappa_{22}^P \theta_2^P + \kappa_{23}^P \theta_3^P > \frac{1}{2}\sigma_{22}^2; \quad \lambda\theta_2^Q - \lambda\theta_3^Q > \frac{1}{2}\sigma_{22}^2; \quad \kappa_{31}^P \theta_1^P + \kappa_{32}^P \theta_2^P + \kappa_{33}^P \theta_3^P > \frac{1}{2}\sigma_{33}^2; \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.
$$

Furthermore, to have well-defined processes for X_t^2 and X_t^3, the sign of the effect that the factors have on each of these two factors must be positive, which we impose with the restrictions $\kappa_{21}^P \leq 0$, $\kappa_{23}^P \leq 0$, $\kappa_{31}^P \leq 0$, and $\kappa_{32}^P \leq 0$. Note that these restrictions imply that the three square-root processes cannot be negatively correlated. In total, there are 16 parameters in the maximally flexible specification of this class of models and 10 in the independent factors specification.[16]

---

[15] For X_t^1, we just need to ensure that the process does not become negative, which is assured if $\varepsilon \cdot \theta_1^Q > 0$ and $\kappa_{11}^P \theta_1^P > 0$.

[16] It turns out that θ_3^Q is difficult to estimate for our data sets. It is consistently estimated at the boundary of the Feller condition for X_t^2 under the Q-measure, which must be satisfied in order to use the extended affine risk premium structure. Our solution is to fix θ_3^Q so that the Feller condition is satisfied by $\varepsilon = 10^{-6}$. We impose restrictions such that

$$
\theta_3^Q = \frac{\lambda\theta_2^Q - \frac{1}{2}\sigma_{22}^2}{\lambda} - \varepsilon.
$$

We caution that this is a property specific to our data sets and is not necessarily of general validity.

3.4 Estimation Methodology

The stochastic volatility models described above are estimated using the Kalman filter algorithm. In term structure models, zero-coupon yields are affine functions of the state variables, such that

$$y_t(\tau) = -\frac{1}{\tau}B(\tau)'X_t - \frac{1}{\tau}A(\tau) + \varepsilon_t(\tau),$$

where $\varepsilon_t(\tau)$ is i.i.d. Gaussian white noise measurement errors. The conditional mean for multi-dimensional affine diffusion processes is given by

$$E^P[X_T|X_t] = (I - \exp(-K^P(T-t)))\theta^P + \exp(-K^P(T-t))X_t, \qquad (4)$$

where $\exp(-K^P(T-t))$ is a matrix exponential. In general, the conditional covariance matrix for affine diffusion processes is given by

$$V^P[X_T|X_t] = \int_t^T \exp(-K^P(T-s))\Sigma D(E^P[X_s|X_t])D(E^P[X_s|X_t])'\Sigma' \exp(-(K^P)'(T-s))ds. \quad (5)$$

Stationarity of the system under the P-measure is ensured if the real components of all the eigenvalues of K^P are positive, and this condition is imposed in all estimations. For this reason, we can start the Kalman filter at the unconditional mean and covariance matrix[17]

$$\widehat{X}_0 = \theta^P \quad \text{and} \quad \widehat{\Sigma}_0 = \int_0^\infty e^{-K^P s}\Sigma D(\theta^P)D(\theta^P)'\Sigma' e^{-(K^P)'s}ds.$$

However, the introduction of stochastic volatility implies that the factors are no longer simply Gaussian. We chose to approximate the true probability distribution of the state variables using the first and second moments and use the Kalman filter algorithm *as if* the state variables were Gaussian.[18] Under these assumptions, the Kalman filter only provides quasi-maximum likelihood estimation. The discretized state equation is given by

$$X_t = (I - \exp(-K^P\Delta t))\theta^P + \exp(-K^P\Delta t)X_{t-1} + \eta_t, \quad \eta_t \sim N(0, V_{t-1}),$$

where Δt is the time between observations and V_{t-1} is the conditional covariance matrix given in Equation (5). Furthermore, the discretization can cause the square-root processes to

---

[17]In the estimation, we calculate the conditional and unconditional covariance matrices using the analytical solutions provided in Fisher and Gilles (1996).

[18]A few notable examples of papers that follow this approach include Duffee (1999), Driessen (2005), and Feldhütter and Lando (2008). Jacobs and Karoui (2009) show that use of the extended Kalman filter, which allows all yields to be measured with error, does not change their qualitative results. In contrast, Collin-Dufresne et al. (2009) use Bayesian estimation methods for their stochastic volatility models.

become negative despite the fact that the parameter sets are forced to satisfy Feller conditions and other non-negativity restrictions. Whenever this happens, we follow the literature and simply truncate those processes at zero; see Duffee (1999) for example.

In the Kalman filter estimations, the error structure is given by

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} V_{t-1} & 0 \\ 0 & H \end{pmatrix} \right],$$

where H is assumed to be a diagonal matrix of the measurement error standard deviations, $\sigma_\varepsilon(\tau_i)$, that are specific to each yield maturity in the data set. The linear least-squares optimality of the Kalman filter requires that the white noise transition and measurement errors be orthogonal to the initial state; i.e., $E[f_0 \eta_t'] = 0$ and $E[f_0 \varepsilon_t'] = 0$. Finally, the standard deviations of the estimated parameters are calculated as

$$\Sigma(\widehat{\psi}) = \frac{1}{T} \left[\frac{1}{T} \sum_{t=1}^{T} \frac{\partial \log l_t(\widehat{\psi})}{\partial \psi} \frac{\partial \log l_t(\widehat{\psi})'}{\partial \psi} \right]^{-1},$$

where $\widehat{\psi}$ denotes the optimal parameter set.

4   Empirical Results with Daily U.S. Treasury Yields

We first estimate our AFNS models with stochastic volatility using U.S. Treasury zero-coupon bond yields from the Gürkaynak et al. (2007) database.

4.1   Data Description

The specific U.S. Treasury bond yields we use are zero-coupon yields constructed by the method described in Gürkaynak at al. (2007)[19] and briefly detailed here. For each business day a zero-coupon yield curve of the Svensson (1994)-type

$$y(\tau) = \beta_0 + \frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} \beta_1 + \left[\frac{1 - e^{-\lambda_1 \tau}}{\lambda_1 \tau} - e^{-\lambda_1 \tau} \right] \beta_2 + \left[\frac{1 - e^{-\lambda_2 \tau}}{\lambda_2 \tau} - e^{-\lambda_2 \tau} \right] \beta_3$$

is fitted to price a large pool of underlying off-the-run U.S. Treasury bonds. Thus, for each business day we have the fitted values of the four factors $(\beta_0(t), \beta_1(t), \beta_2(t), \beta_3(t))$ and the two parameters $(\lambda_1(t), \lambda_2(t))$. From this data set zero-coupon yields for any relevant maturity

---

[19]The Board of Governors in Washington DC frequently updates the factors and parameters of this method, see the related website http://www.federalreserve.gov/pubs/feds/2006/index.html
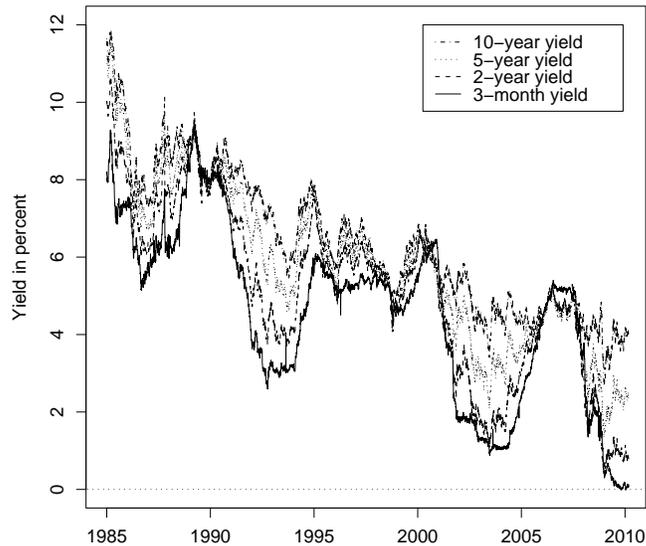
Figure 1: **Time Series of U.S. Treasury Bond Yields.**
Illustration of the daily U.S. Treasury zero-coupon bond yields covering the period from January 2, 1985 to March 1, 2010. The yields shown have maturities in three months, two years, five years and ten years, respectively.

can be calculated. As demonstrated by Gürkaynak et al. (2007), this model fits the underlying pool of bonds extremely well. By implication, the zero-coupon yields derived from this approach constitute a very good approximation to the true underlying Treasury zero-coupon yield curve. From this data set we construct zero-coupon bond yields with the following maturities: 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, and 10-year. We use daily data and limit our sample to the period from January 2, 1985 to March 1, 2010. The summary statistics are provided in Table 1, while Figure 1 illustrates the constructed time series of the three-month, two-year, five-year, and ten-year U.S. Treasury zero-coupon yields.

Researchers have typically found that three factors are sufficient to model the time-variation in the cross section of U.S. Treasury bond yields (e.g., Litterman and Scheinkman, 1991). Indeed, for our daily U.S. Treasury bond yield data, 99.96% of the total variation is accounted for by three factors. Table 2 reports the eigenvectors that correspond to the first three principal components of our data. The first principal component accounts for 95.2% of the variation in the Treasury bond yields, and its loading across maturities is uniformly negative. Thus, like a level factor, a shock to this component changes all yields in the same direction irrespective of maturity. The second principal component accounts for 4.5% of the

Maturity in months	No. obs.	Mean in %	Std. dev. in %	Skewness	Kurtosis
3	6,269	4.59	2.22	-0.22	2.33
6	6,269	4.69	2.26	-0.23	2.30
12	6,269	4.87	2.28	-0.21	2.30
24	6,269	5.16	2.24	-0.13	2.35
36	6,269	5.40	2.17	-0.02	2.41
60	6,269	5.77	2.05	0.20	2.46
84	6,269	6.06	1.94	0.35	2.49
120	6,269	6.38	1.83	0.48	2.55

Table 1: **Summary Statistics for the U.S. Treasury Bond Yields.**
Summary statistics for the sample of daily U.S. Treasury zero-coupon bond yields covering the period from January 2, 1985 to March 1, 2010.

Maturity in months	Loading on		
	First P.C.	Second P.C.	Third P.C.
3	-0.36	-0.45	0.52
6	-0.37	-0.39	0.20
12	-0.38	-0.27	-0.20
24	-0.38	-0.05	-0.47
36	-0.37	0.10	-0.42
60	-0.34	0.31	-0.12
84	-0.32	0.43	0.17
120	-0.29	0.53	0.45
% explained	95.24	4.52	0.18

Table 2: **Eigenvectors of the First Three Principal Components in U.S. Treasury Bond Yields.**
The loadings of yields of various maturities on the first three principal components are shown. The final row shows the proportion of all bond yield variability accounted for by each principal component. The data consist of daily U.S. Treasury zero-coupon bond yields from January 2, 1985 to March 1, 2010.

variation in these data and has sizable negative loadings for the shorter maturities and sizable positive loadings for the long maturities. Thus, like a slope factor, a shock to this component steepens or flattens the yield curve. Finally, the third component, which accounts for only 0.2% of the variation, has a U-shaped factor loading as a function of maturity, which is naturally interpreted as a curvature factor. This motivates our use of the Nelson-Siegel model with its level, slope, and curvature factor for modeling this sample of U.S. Treasury yields.

4.2 Conditional mean results

We first examine the in-sample estimation results for the five model specifications introduced in Section 3 in addition to the $AFNS_0$ model. We only present results for the diagonal, independent-factors specification for each AFNS model class. For example, the $AFNS_1$-L
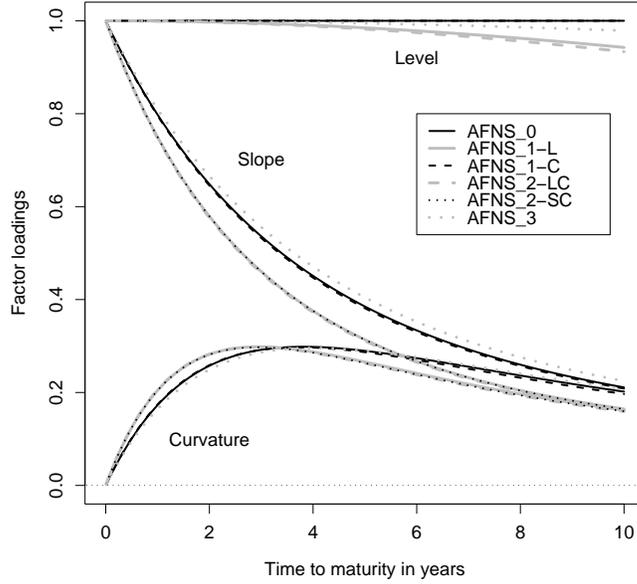
Figure 2: **Factor Loadings in the AFNS$_i$ Models.**
The factor loadings on the three state variables in the zero-coupon bond yield function in the AFNS$_i$ models are shown. The parameters for each model are taken from Tables 3 and 4.

model has P-dynamics given by

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ 0 & \kappa_{22}^P & 0 \\ 0 & 0 & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$

$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{21} X_t^1} & 0 \\ 0 & 0 & \sqrt{1 + \beta_{31} X_t^1} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

We use an independent-factors specification because the AFNS models deliver essentially identical decompositions of the data into level, slope, and curvature factors independent of the specification of the P-dynamics. Since it is this factor decomposition that determines the shape and form of the model-implied stochastic volatility, at least at the short one-month horizon we focus on in this paper, this restriction comes at a minimal loss of generality. Furthermore, it makes the results more readily comparable across model classes.[20]

---

[20]Results summarizing the estimation of the maximally flexible specifications of the models are available upon request.

Figure 2 illustrates the factor loadings in the zero-coupon bond yield function in all six AFNS$_i$ models. As mentioned in the technical Section 3, the inclusion of stochastic volatility into the AFNS model prevents us from obtaining the exact Nelson-Siegel factor loadings unlike what is the case for the AFNS$_0$ model class. Importantly, though, the NS factor loading structure is approximately preserved in all five new model classes, as desired by construction, independent of the differences in the models' ability to generate stochastic volatility.

Tables 3 and 4 present our parameter estimates of the six models. The parameter estimates exhibit similarities across the model specifications, especially for the K^P matrix. The estimated K^P parameter for the level factor indicates the most persistence, while the curvature factor is the least persistent, in all specifications. As for both the mean parameters in θ^P and the σ volatility parameters, we see some notable differences across the various models depending on whether the factor in question is generating stochastic volatility or not. In general, in any of the AFNS$_i$ models with stochastic volatility, if a factor is not generating stochastic volatility, its associated estimated σ value is close to the corresponding estimate in the AFNS$_0$ model. For the θ^P parameters, the variation in the estimated values is tied to differences in the scale at which each factor operates. Since the factors are latent, this level varies and depends on which factors generate stochastic volatility and therefore have to be bound away from the zero-boundary. Finally, the β volatility sensitivity parameters suggest that the level factor plays a role in generating stochastic volatility for both the slope and the curvature factor, whereas there is little evidence that slope and curvature play a role for the volatility of the level factor or for the volatility of each other in this sample of U.S. Treasury yields.

If we turn to a performance comparison of the various AFNS$_i$ specifications, we can start by comparing the obtained maximum log likelihood values reported in Table 4. Even though all AFNS$_i$ models are non-nested and therefore not directly comparable, the relatively large differences in likelihood values still suggest that the AFNS$_1$-L model provides the overall best fit to the cross-sectional and time-series variation of the data relatively closely followed by the AFNS$_1$-L,C model. On the other hand, the AFNS$_3$ model obtains a markedly lower maximum likelihood value than any of the other models. This model is restricted by the fact that all three factors have to remain non-negative, and one or more of these restrictions are binding periodically, not least during the last 18 months of the sample with the low interest rate environment in the wake of the financial crisis of 2008 and 2009. Duffee (2002), in his analysis of general affine $A_i(3)$ term structure models, also find that the $A_1(3)$ model class performs the best, and the $A_3(3)$ model class the poorest. Note, though, that he only uses

Parameters	AFNS models with independent factors					
	AFNS_0	$\text{AFNS}_1\text{-L}$	$\text{AFNS}_1\text{-C}$	$\text{AFNS}_2\text{-LC}$	$\text{AFNS}_2\text{-SC}$	AFNS_3
κ_{11}^P	0.0269	0.0503	0.0149	0.0600	0.0097	0.0496
	(0.0436)	(0.0456)	(0.0499)	(0.0433)	(0.0431)	(0.00991)
κ_{22}^P	0.0799	0.1830	0.1006	0.1577	0.1349	0.3771
	(0.0941)	(0.152)	(0.0935)	(0.141)	(0.0974)	(0.0556)
κ_{33}^P	0.7552	1.0662	0.8649	0.9036	1.3099	1.2717
	(0.199)	(0.250)	(0.102)	(0.104)	(0.121)	(0.135)
θ_1^P	0.0895	0.0618	0.0746	0.0565	-0.0067	0.0213
	(0.0217)	–	(0.0526)	–	(0.0649)	–
θ_2^P	-0.0410	-0.0199	-0.0341	-0.0179	0.0533	0.0278
	(0.0288)	(0.0200)	(0.0202)	(0.0199)	(0.0151)	(0.00380)
θ_3^P	-0.0158	-0.0028	0.0709	0.0824	0.0680	0.0410
	(0.00746)	(0.00622)	(0.00682)	(0.00707)	(0.00493)	(0.00378)
σ_{11}	0.0057	0.0608	0.0054	0.0657	0.0053	0.0362
	(0.00004)	(0.00012)	(0.00006)	(0.00023)	(0.00015)	(0.00038)
σ_{22}	0.0092	0.0111	0.0086	0.0107	0.0351	0.0359
	(0.00008)	(0.00019)	(0.00014)	(0.00028)	(0.00028)	(0.00028)
σ_{33}	0.0294	0.0306	0.0961	0.0914	0.1084	0.1239
	(0.00015)	(0.00050)	(0.00067)	(0.00066)	(0.00066)	(0.00137)
β_{11}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{12}	–	–	–	–	0.0000	–
	–	–	–	–	(1.30)	–
β_{13}	–	–	0.0000	–	0.0000	–
	–	–	(0.451)	–	(0.555)	–
β_{21}	–	6.3275	–	3.5858	–	–
	–	(0.729)	–	(1.09)	–	–
β_{22}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{23}	–	–	0.0000	–	–	–
	–	–	(0.568)	–	–	–
β_{31}	–	0.9532	–	0.0000	–	–
	–	(0.542)	–	(0.619)	–	–
β_{32}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{33}	–	–	–	–	–	–
	–	–	–	–	–	–

Table 3: **Parameter Estimates of the P-Dynamics for AFNS_i Models with the Independent-Factors Specification for U.S. Treasury Data.**
The table contains the estimated K^P matrix, θ^P vector, Σ matrix, and β volatility sensitivity parameters for the independent-factors specification of the P-dynamics in the AFNS_i models. Estimated standard deviations for the parameter estimates are given in parentheses. The estimations are based on daily observations from January 2, 1985 to March 1, 2010.

essentially affine risk premium specifications, which are less general than the extended affine risk premium specifications applied in this paper, in particular for $A_3(3)$ models.

Another way to assess the performance of the different AFNS specifications of stochastic

Parameters	AFNS models with independent factors					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
θ_1^Q	–	3,105	–	3,390	–	1,060
	–	(0.718)	–	(11.8)	–	(2.88)
θ_2^Q	–	–	–	–	0.08	0.0493
	–	–	–	–	–	(0.00022)
θ_3^Q	–	–	0.08	0.08	0.0790	0.0478
	–	–	–	–	(0.00017)	–
λ	0.4697	0.6067	0.4757	0.6127	0.6063	0.4381
	(0.00121)	(0.00104)	(0.00138)	(0.00119)	(0.00139)	(0.00080)
Max log L	305,776.3	316,191.6	300,973.5	313,826.0	299,689.1	280,342.7

Table 4: **Parameter Estimates of the Q-Dynamics for $AFNS_i$ Models with the Independent-Factors Specification for U.S. Treasury Data.**
The table contains the estimated θ^Q vector and λ parameters for the independent-factors specification of the P-dynamics in the $AFNS_i$ models. Estimated standard deviations for the parameter estimates are given in parentheses. The estimations are based on daily observations from January 2, 1985 to March 1, 2010. The maximum log-likelihood values are reported, although the models are non-nested.

Maturity in months	RMSE for AFNS models with independent factors					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	21.42	19.89	20.69	18.87	8.86	10.78
6	9.53	8.54	9.11	7.90	0.56	2.38
12	0.10	0.00	0.20	0.56	3.86	6.66
24	2.41	1.76	2.62	1.58	1.33	6.79
36	0.00	0.00	1.93	0.66	2.56	5.58
60	2.82	1.58	2.99	1.54	1.96	6.72
84	1.83	0.50	2.09	0.87	7.87	10.15
120	9.79	4.92	9.86	5.16	22.78	15.46

Table 5: **RMSE of Fitted Yields for the $AFNS_i$ Models for U.S. Treasury Data.**
The table presents the root-mean-squared errors for the fitted yields across the 8 maturities under the independent-factors specification of the AFNS model with different stochastic volatility specifications. The sample covers the period from January 2, 1985 to March 1, 2010. All numbers are expressed in basis points.

volatility is to examine the cross-sectional fit of the yield curve, as shown in Table 5 using root-mean-squared-error for the models' fitted errors. Relative to the $AFNS_0$ specification, the introduction of stochastic volatility reduces the RMSE of the fitted yields for the short-term three- and six-month maturities. However, for the remaining maturities, the stochastic volatility specifications do not necessarily insure a reduction in the RMSE measure with one exception, the $AFNS_1$-L model does deliver a uniform improvement in model fit over the $AFNS_0$ model. A more detailed comparison of the five new $AFNS_i$ models shows that three of the models ($AFNS_1$-L, $AFNS_1$-C, and $AFNS_2$-L,C) fit the short-term yields relatively

poorly, while they deliver a very good fit for the remaining maturities. On the other hand, the AFNS$_2$-S,C model has the opposite ranking with a very good fit for the first seven maturities, but a poorer fit of the ten-year yield. Finally, the AFNS$_3$ model falls in between with a decent fit for all eight maturities. Thus, based on the evidence so far, there is no basis for either disqualifying or preferring any particular of the five new AFNS$_i$ model classes with stochastic volatility over the others or the original AFNS$_0$ model. Of course, from a mechanical point of view, the AFNS$_3$ specification has the ability to induce the greatest degree of in-sample stochastic volatility of all the specifications and thus should be best suited ex ante to closely match the observed data characteristics in terms of yield volatility in addition to providing a good in-sample fit to the cross section of yields. However, before analyzing whether that is the case, we will discuss some identification issues that appear in the estimation of three of the AFNS$_i$ model classes with stochastic volatility.

4.3 Identification Issues Related to θ^Q Parameters

The fact that the θ^Q parameters are not statistically identifiable and fixed at zero in the Gaussian AFNS$_0$ model is in itself a warning that the θ^Q parameters in the AFNS models with stochastic volatility should be treated with caution. Against this background, it is not all that surprising that θ_3^Q turn out to be hard to identify in the AFNS$_1$-C and AFNS$_2$-L,C models and a similar problem pertains to the value of θ_2^Q in the AFNS$_2$-S,C model. However, as we will show in the following, the specific value of these θ^Q parameters significantly affect the size of the generated stochastic yield volatility. As a result, there is an important identification problem to address. Here, we exemplify this problem and how we deal with it for the AFNS$_1$-C model. In the empirical analysis, we apply a similar approach to the other two model classes mentioned above.

Figure 3 illustrates the fitted one-month standard deviation of the two-year U.S. Treasury yield in the independent-factors specification of the AFNS$_1$-C model when θ_3^Q is left unconstrained (estimated value of 0.545) and when we fix it at two much lower values, 0.08 and 0.06, respectively. We note that, for the unconstrained model, the generated yield volatility is almost a flat line for this yield maturity, even though this is the part of the covered maturity range where the curvature factor has its peak effect (see Figure 2) and, by implication, this particular yield and its conditional volatility should have close to the maximum possible sensitivity to variation in the curvature factor. On the other hand, when we restrict θ_3^Q at the low values, the variation in the fitted yield volatility is much larger. Also, with θ_3^Q restricted in this way, the estimated values and standard deviations of some of the key parameters, in
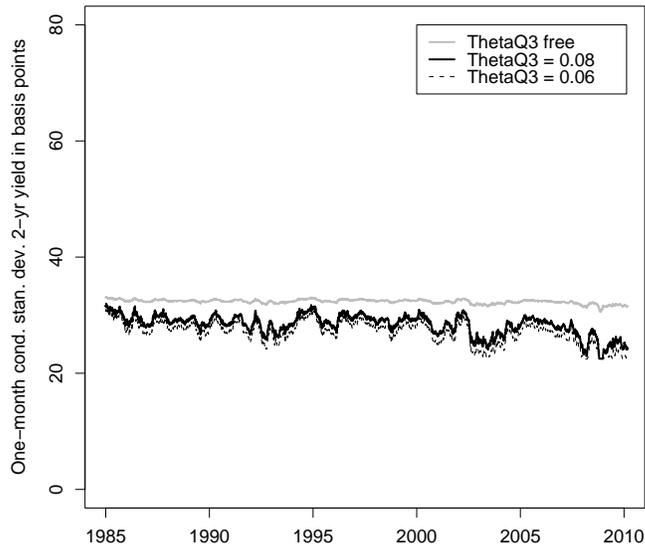
Figure 3: **Fitted One-Month Conditional Standard Deviations of the Two-Year U.S. Treasury Yield from AFNS$_1$-C Models.**

particular the K^P parameters, look better identified and closer to those obtained in the other AFNS$_i$ models as can be seen in Table 6, which reports the estimated parameters for the three AFNS$_1$-C models analyzed here.

Thus, based on the above evidence, this appears to be a useful restriction that allows the AFNS$_1$-C model class to generate more meaningful levels of yield volatility in addition to being better identified. To analyze whether there are any drawbacks to this kind of restriction, we first study the impact on the estimated factors. Figure 4 shows the estimated level and slope factors from the three models. The minimum correlation between the level factors is 99.5%. For the slope factors the minimum pairwise correlation is 99.9%. Thus, the decomposition into level and slope is completely unaffected by restrictions on θ_3^Q. Figure 5 shows the estimated paths for the curvature factor that generates the stochastic volatility in this model class. We notice the difference in the estimated value of this factor which changes almost one-for-one with the size of θ_3^Q. However, importantly, the time variation is almost identical. For the model with θ_3^Q fixed at 0.08, its correlation for the affected curvature factor is 99.2% with the unrestricted model, while it is 99.5% with the model with θ_3^Q fixed at 0.06. Of course, the correlation between the unrestricted model and that with θ_3^Q fixed at 0.06 is lower, but still high, at 97.8%. In summary, the models deliver qualitatively identical decompositions into

26

Parameters	Independent-factors AFNS$_1$-C models		
	θ_3^Q free	$\theta_3^Q = 0.08$	$\theta_3^Q = 0.06$
κ_{11}^P	0.0364	0.0149	0.0128
	(0.0543)	(0.0499)	(0.0568)
κ_{22}^P	0.0001	0.1006	0.1284
	(0.0873)	(0.0935)	(0.0945)
κ_{33}^P	0.0437	0.8649	0.7531
	(0.135)	(0.102)	(0.119)
θ_1^P	0.0874	0.0746	0.0624
	(0.0837)	(0.0526)	(0.0931)
θ_2^P	1.695	-0.0341	-0.0284
	(0.0321)	(0.0202)	(0.0160)
θ_3^P	0.6237	0.0709	0.0543
	(0.00881)	(0.00682)	(0.00749)
σ_{11}	0.0053	0.0054	0.0054
	(0.00009)	(0.00006)	(0.00006)
σ_{22}	0.0086	0.0086	0.0085
	(0.00020)	(0.00014)	(0.00012)
σ_{33}	0.0408	0.0961	0.1006
	(0.00110)	(0.00067)	(0.00096)
β_{13}	0.3002	0.0000058	0.0000082
	(0.208)	(0.451)	(0.655)
β_{23}	0.1956	0.0000096	0.0000242
	(0.282)	(0.568)	(0.658)
θ_3^Q	0.5453	0.08	0.06
	(0.00357)	—	—
λ	0.4687	0.4757	0.4694
	(0.00120)	(0.00138)	(0.00154)
Max log L	305,679.5	300,973.5	296,667,0

Table 6: **Parameter Estimates of AFNS$_1$-C Models for U.S. Treasury Data.**
The table contains the estimated dynamic parameters for the independent-factors specification of the
P-dynamics in AFNS$_1$-C models with varying restrictions on θ_3^Q. Estimated standard deviations for
the parameter estimates are given in parentheses. The estimations are based on daily observations
from January 2, 1985 to March 1, 2010.

level, slope, and curvature independent of restrictions imposed on θ_3^Q or lack thereof.

Second, we analyze the yield fit across the three AFNS$_1$-C specifications. Table 7 reports
the mean and RMSEs of the fitted errors for the three specifications. We note that the fit is
identical, despite the large difference in likelihood values, but consistent with the very high
correlation between the three estimated factors. Based on this we conclude that the dramatic
loss in likelihood value is not matched by a corresponding decline in model fit in any of
the restricted models. Also, this suggests that the significant differences in the maximum
likelihood values across the various AFNS$_i$ models we observed in the previous section should
be interpreted with caution as they are not necessarily matched by a corresponding decline
in model performance.
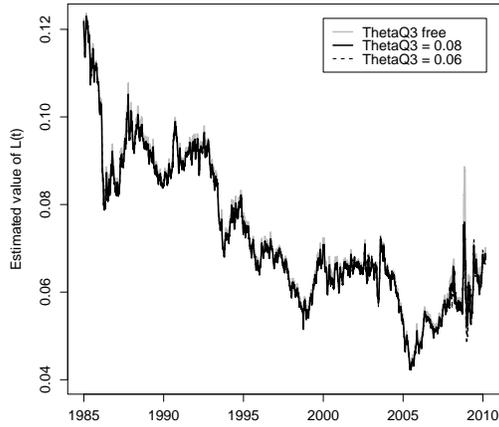
(a) Estimated level factors.
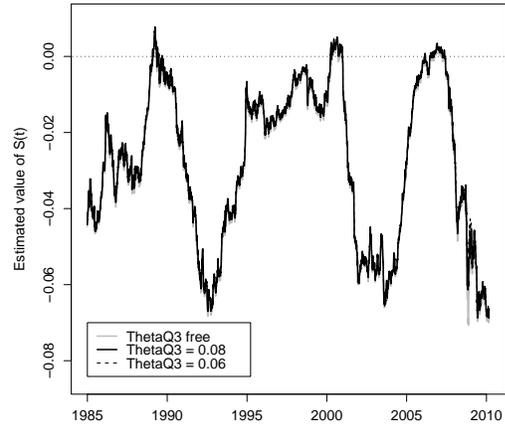
(b) Estimated slope factors.

Figure 4: **Estimated Level and Slope Factors from AFNS$_1$-C Models for U.S. Treasury Data.**
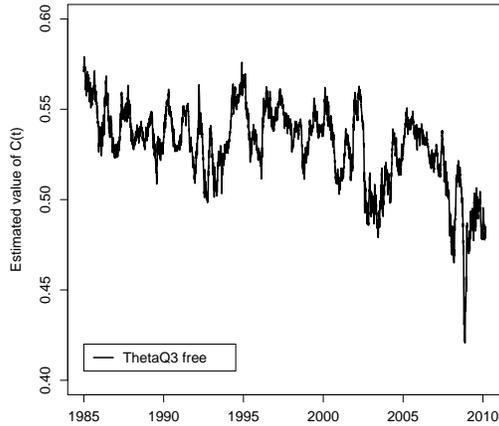
(a) θ_3^Q unrestricted.
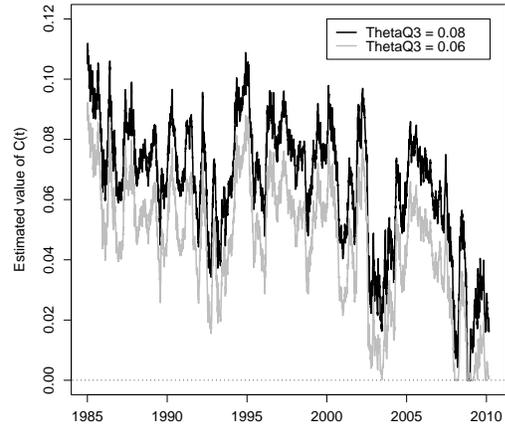
(b) θ_3^Q fixed.

Figure 5: **Estimated Curvature Factors from AFNS$_1$-C Models for U.S. Treasury Data.**

Ex ante it would be natural to expect that these different specifications of the same model class should generate about the same stochastic yield volatility. The high correlation between the respective factor paths in each specification supports this view and also explains the very similar in-sample fit reported in Table 7. But how does this square with the difference in the estimated conditional yield volatility in Figure 3? To understand the difference, note that the range of variation for each factor is the same across the three specifications and about 0.09 for

Maturity in months	AFNS$_1$-C models with independent factors					
	θ_3^Q free		$\theta_3^Q = 0.08$		$\theta_3^Q = 0.06$	
	Mean	RMSE	Mean	RMSE	Mean	RMSE
3	-6.10	21.43	-5.73	20.69	-5.37	20.68
6	-3.40	9.53	-3.19	9.11	-2.99	9.17
12	-0.01	0.10	-0.02	0.20	-0.02	0.24
24	1.43	2.41	1.25	2.62	1.12	3.36
36	0.00	0.01	-0.20	1.93	-0.30	3.21
60	-1.92	2.82	-2.00	2.99	-2.00	3.28
84	0.11	1.83	0.14	2.09	0.17	2.51
120	7.17	9.83	7.27	9.86	7.25	10.09

Table 7: **Mean and RMSE of Fitted Yields for AFNS$_1$-C Models for U.S. Treasury Data.**
The table presents the mean and root-mean-squared errors for the fitted yields across the 8 maturities for the independent-factors specification of the AFNS$_1$-C model with varying restrictions on θ_3^Q. The sample covers the period from January 2, 1985 to March 1, 2010. All numbers are expressed in basis points.

the curvature factors. This similarity is driven by the fact that these factors affect yields in the same way due to the imposed Nelson-Siegel factor loading structure. By implication, they will exhibit approximately the same range of variation in order to deliver approximately the same fit to the cross section of yields as also documented in Table 7. However, importantly, the absolute level of this factor is very different across the three specifications, which turns out to have dramatic consequences for the size of the generated stochastic volatility. For the AFNS$_1$-C model with θ_3^Q fixed at 0.08, the curvature factor varies in the range $(0, 0.11)$. Given the estimated value of σ_{33}, this translates into the following variation of the stochastic volatility generated by the curvature factor

$$\sigma_{33}\sqrt{X_t^3} \in (0, 0.032).$$

On the other hand, for the unrestricted AFNS$_1$-C model, the curvature factor varies in the range $(0.42, 0.58)$. This translates into the following variation in the stochastic volatility generated by X_t^3 in that model[21]

$$\sigma_{33}\sqrt{X_t^3} \in (0.027, 0.031).$$

Thus, in the AFNS$_1$-C model with θ_3^Q fixed at 0.08, the conditional volatility at the largest values of X_t^3 is multiples of the conditional volatility at the lowest values of X_t^3. In the unrestricted AFNS$_1$-C model, on the other hand, there is much less variation in the generated

---

[21]This is based on an estimate of σ_{33} of 0.041.

stochastic volatility as we move from the highest to the lowest values of X_t^3. Equally important, the range of generated stochastic volatility in the restricted specification spans that of the unrestricted specification. These observations combined leads us to focus on a value for θ_3^Q in the neighborhood of 0.08. At this value, the curvature factor only hits the lower zero-boundary for a very brief period around the very peak of the financial crisis in 2008.[22] Thus, we are close to maximizing the range of generated volatility. If we raise θ_3^Q above 0.08, we will start to approach the unrestricted case that delivered a narrow range of generated volatility, and if we go below 0.08, we will reduce the top of the range of generated volatility, while being restricted by zero at the bottom of the range.

The only drawback of this type of restriction appears to be a slight downward bias due to the lower estimated values of the curvature factor which is only partially offset by a higher estimated value of σ_{33} as is evident from Figure 3. However, as we will see later in the analysis of the U.K. gilt yields, this is not always the case.

We conclude that there is only a very limited downside to imposing restrictions on the θ^Q parameters in the three model classes discussed here. Thus, these are, in our view, very innocent restrictions on parameters that are not all that well identified to begin with. Furthermore, as the results later will show, even this extra 'helping' hand of restricting the θ^Q parameters to a useful range does not allow any of these three model classes (AFNS$_1$-C, AFNS$_2$-L,C, and AFNS$_2$-S,C) to generate stochastic yield volatility that is more consistent with our measure of realized volatility than the competing AFNS$_i$ model classes. Thus, the restrictions discussed here do not affect the conclusions we draw later on, rather they underscore that these three model classes might suffer both from estimation problems and relatively poorer fit to the aspects of the yield data that we focus on in this paper.

4.4   Stochastic Volatility Results for the AFNS Model Specifications

Collin-Dufresne et al. (2009) demonstrate that there is a tension in affine models between fitting the cross section of yields and capturing their stochastic volatility. They further argue that to allow only one factor to generate stochastic volatility in a three-factor affine model is too restrictive to fit both aspects of the data. By allowing for more factors to generate stochastic volatility in our AFNS specification, we hope to mitigate this tension. As indicated in the discussion of the results above, the AFNS specifications with stochastic volatility do not differ markedly in terms of fitting the observed U.S. Treasury yield curve. However, their fitted volatility measures, which we define here as the fitted standard deviation, of

---

[22]This happens for a number of days in the period from October 20 to December 16, 2008, but not outside this very short time window.
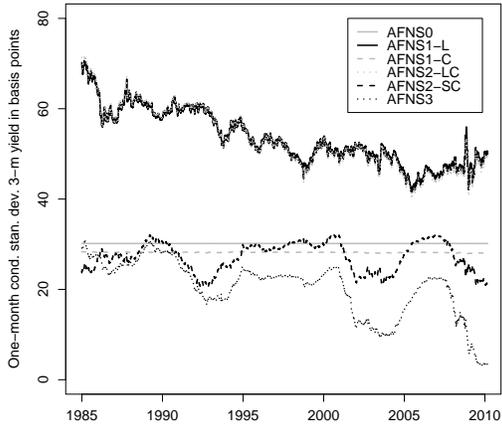
(a) Three-month Treasury yield
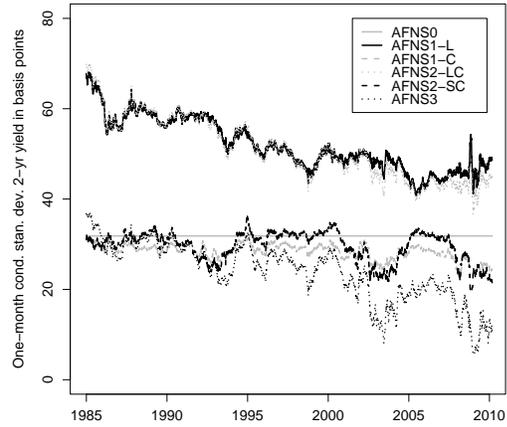
(b) Two-year Treasury yield
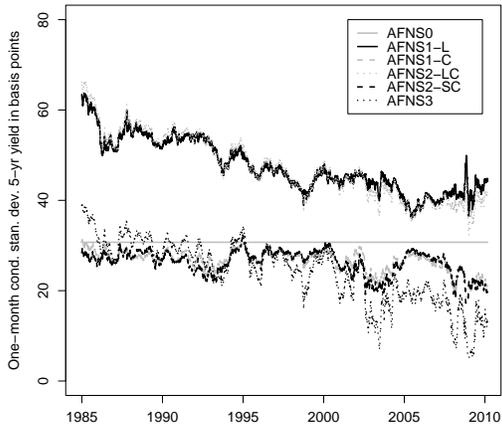
(c) Five-year Treasury yield
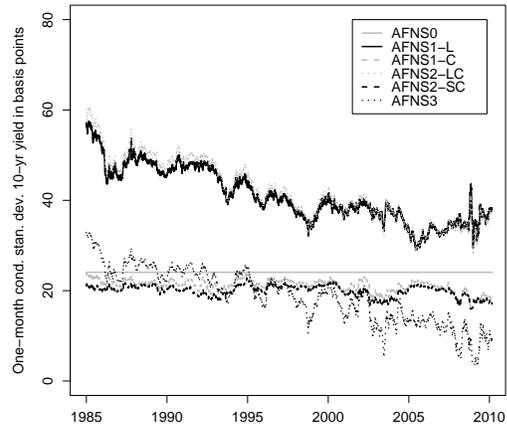
(d) Ten-year Treasury yield

Figure 6: **Fitted One-Month Conditional Standard Deviations of U.S. Treasury Yields from the AFNS$_i$ Models.**

these specifications do differ greatly from each other and from measures of the data's realized volatility.[23]

Figure 6 shows the fitted one-month conditional yield volatility for four different maturities based on the six AFNS$_i$ models. We note the Gaussian AFNS$_0$ model with its flat fitted

---

[23]The fitted one-month conditional volatility measures are given by the square root of

$$V_t^P[y_T(\tau)] = \frac{1}{\tau^2} B(\tau)' V_t^P[X_T] B(\tau),$$

where $V_t^P[X_T]$ is the conditional covariance matrix of the state variables, $T - t = \frac{1}{12}$, and τ is the yield maturity in years.

Correlation	Three-month U.S. Treasury yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.440	1.000	-0.221	0.578
$AFNS_1$-C			1	0.457	0.340	0.677
$AFNS_2$-LC				1	-0.212	0.588
$AFNS_2$-SC					1	0.636
$AFNS_3$						1

Correlation	Two-year U.S. Treasury yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.434	0.987	0.149	0.767
$AFNS_1$-C			1	0.569	0.833	0.857
$AFNS_2$-LC				1	0.291	0.856
$AFNS_2$-SC					1	0.722
$AFNS_3$						1

Correlation	Five-year U.S. Treasury yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.431	0.989	0.239	0.810
$AFNS_1$-C			1	0.558	0.906	0.859
$AFNS_2$-LC				1	0.372	0.886
$AFNS_2$-SC					1	0.725
$AFNS_3$						1

Correlation	Ten-year U.S. Treasury yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.433	0.997	0.250	0.904
$AFNS_1$-C			1	0.492	0.912	0.761
$AFNS_2$-LC				1	0.310	0.931
$AFNS_2$-SC					1	0.598
$AFNS_3$						1

Table 8: **Pairwise Correlations of the One-Month Conditional Standard Deviation of Four U.S. Treasury Yields Across the $AFNS_i$ Models.**
The table contains the pairwise correlations between the one-month conditional standard deviations of the three-month, the two-year, the five-year, and the ten-year U.S. Treasury yields estimated by the $AFNS_i$ models. The estimations are based on daily data from January 2, 1985 to March 1, 2010.

volatility lines. All the models where the level factor is allowed to generate stochastic volatility exhibit a downward trending pattern in the fitted volatility at all maturities until 2005 at which point a slow upward trend starts. However, in the $AFNS_3$ model, the recent upward trend is more than offset by the extremely low interest rates we have seen since the credit crisis of 2008 and 2009. This depresses the slope and curvature in all the $AFNS_i$ models with consequences for the fitted conditional yield volatility in the $AFNS_1$-C, $AFNS_2$-S,C, and $AFNS_3$ models.

Table 8 reports the pairwise correlations of the fitted conditional yield volatility series

for four maturities across all six AFNS$_i$ models. There is a large dispersion in the correlations across models with some natural clustering. For example, the fitted yield volatility of the AFNS$_1$-L and AFNS$_2$-L,C models are highly correlated at all maturities. Also, the AFNS$_1$-C and AFNS$_2$-S,C models tend to produce highly correlated fitted volatilities with the exception of the very short maturities where the difference in the role of the slope factor is most pronounced. Overall, though, the AFNS$_3$ model appears as a reasonable compromise candidate which has a high, positive correlation with the fitted yield volatility from the other four AFNS$_i$ models with stochastic volatility.

To evaluate the in-sample fit of these one-month-ahead conditional standard deviations, we compare them to a standard measure of realized volatility based on the same daily data used in the estimations. We generate the realized standard deviation of the daily changes in the interest rates for the 31-day period ahead on a rolling basis. The realized variance measure is used by Andersen and Benzoni (2010), Collin-Dufresne et al. (2009), as well as Jacobs and Karoui (2009) in their assessments of stochastic volatility models. This measure is fully nonparametric and has been shown to converge to the underlying realization of the conditional variance as the sampling frequency increases; see Andersen et al. (2003) for details. The square root of this measure retains these properties. For each observation date t we determine the number of trading days N during the subsequent 31-day time window (where N is most often 21 or 22).[24] We then generate the realized standard deviation as

$$RV_{t,\tau}^{STD} = \sqrt{\sum_{n=1}^{N} \Delta y_{t+n/N}^2(\tau)},$$

where $\Delta y_{t+n/N}(\tau)$ is the change in yield $y_t(\tau)$ from trading day $(n-1)$ to trading day n.[25]

Figure 7 plots the realized 31-day ahead volatility series over the full sample period for four maturities: 3 months, 2 years, 5 years, and 10 years. In each chart we include the corresponding fitted yield volatility from four AFNS models: AFNS$_0$, AFNS$_1$-L, AFNS$_2$-S,C, and AFNS$_3$. The figure highlights three empirical features of the realized volatilities. First, the realized volatility series become less volatile as the maturity increases. Table 9 shows that the standard deviation of the realized standard deviation for the changes in the three-month yield is almost 1.5 times greater than that of the ten-year yield. The degree of variation declines sharply with the standard deviation of this volatility measure for the one-

---

[24] As a consequence, the realized volatility measure can be calculated for the period from January 2, 1985 to January 29, 2010.

[25] Note that other measures of realized volatility have been used in the literature, such as the realized mean absolute deviation measure as well as fitted GARCH estimates. Collin-Dufresne et al. (2009) also use option-implied volatility as a measure of realized volatility.
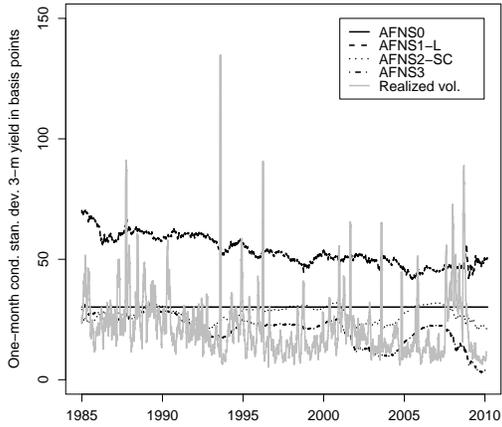
(a) Three-month Treasury yield
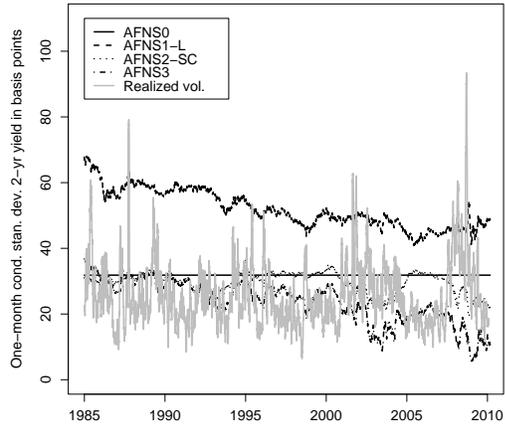
(b) Two-year Treasury yield
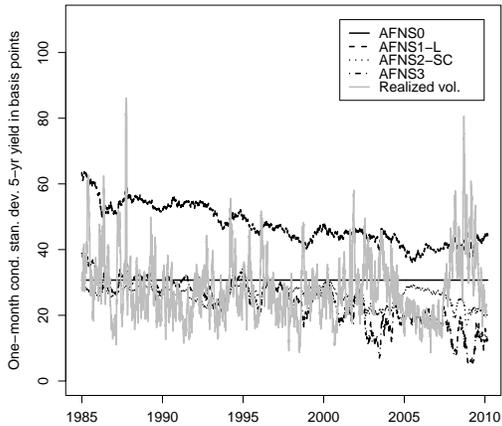
(c) Five-year Treasury yield
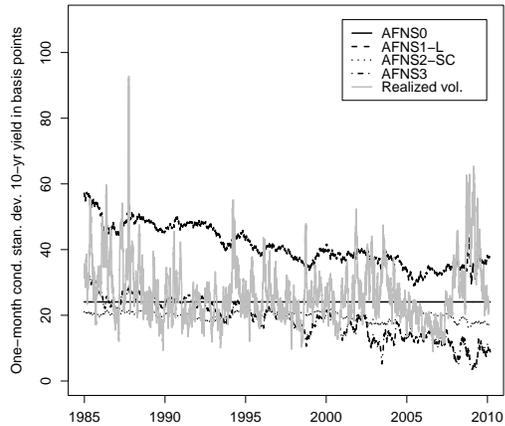
(d) Ten-year Treasury yield

Figure 7: **Fitted Standard Deviations from Four AFNS$_i$ Models for the U.S. Treasury Data.**

year series falling to just 1.06 times greater than that of the ten-year series.[26] This pattern of declining variation as maturity increases suggests that the standard deviations generated by all the model specifications should exhibit better fit as maturity increases, which is, in general, the pattern observed in Table 10, which contains the summary statistics of the fitted errors between the model-implied one-month conditional standard deviations and the 31-day-ahead realized volatility for all eight maturities in the U.S. Treasury data.

---

[26]Please note that this pattern is similar to the one presented by Jacobs and Karoui (2009) for monthly Treasury yields, although their measures decline at a slower rate as maturity increases. The differences may be due to the longer sample period from 1970 to 2003 that they use.

Maturity in months	Mean in bps	Std. dev. in bps	Std. dev. ratio
3	22.56	14.44	1.49
6	19.20	11.57	1.19
12	21.63	10.26	1.06
24	26.05	10.21	1.05
36	27.88	10.25	1.06
60	28.45	10.09	1.04
84	28.00	9.93	1.02
120	27.29	9.71	1.00

Table 9: **Summary Statistics for the 31-Day Realized Standard Deviation Series based on the Daily U.S. Treasury Data.**
The summary statistics are for the 31-day rolling realized standard deviations based on the daily U.S. Treasury data from January 2, 1985 to March 1, 2010. The standard deviation ratio is calculated as the standard deviation in question divided by the standard deviation for the ten-year maturity.

| Maturity in months | RMSE for fitted standard deviations | | | | | | | | | | | |
| | $AFNS_0$ | | $AFNS_1$-L | | $AFNS_1$-C | | $AFNS_2$-LC | | $AFNS_2$-SC | | $AFNS_3$ | |
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
3	7.61	16.32	30.94	34.07	5.63	15.49	30.73	33.93	4.72	15.47	-2.04	14.71
6	10.66	15.73	33.54	35.53	8.55	14.39	33.37	35.43	8.02	14.41	1.47	12.10
12	8.64	13.41	30.62	32.58	6.07	12.00	30.28	32.41	6.50	12.79	0.08	11.73
24	5.78	11.74	25.89	28.26	2.33	10.91	25.19	27.93	3.59	12.00	-2.10	12.84
36	4.52	11.20	23.01	25.64	0.56	10.97	22.32	25.43	1.48	11.70	-2.86	13.40
60	2.27	10.34	19.09	22.08	-1.79	11.03	18.98	22.42	-2.19	11.40	-4.18	13.59
84	-0.18	9.94	16.46	19.73	-3.83	11.23	16.93	20.52	-4.91	11.73	-5.94	13.71
120	-3.22	10.23	13.81	17.35	-6.13	11.78	14.74	18.41	-7.31	12.44	-8.46	14.27

Table 10: **Mean and RMSE for the Fitted Errors of the One-Month Conditional Standard Deviations from the AFNS Models for the U.S. Treasury Data.**
The table presents the mean and RMSE values for the fitted error of the monthly model-based fitted standard deviations relative to the 31-day realized standard deviations based on the daily U.S. Treasury data over the period from January 2, 1985 to March 1, 2010. All numbers are measured in basis points.

Second, note that the $AFNS_0$, $AFNS_1$-C, and $AFNS_2$-S,C models produce consistently low RMSE values between fitted and realized standard deviations for all maturities. However, as shown in Table 11, the degree of variation exhibited by these fitted standard deviations is quite low relative to the $AFNS_3$ specification. As our objective is to best capture the stochastic volatility of these interest rate series, the $AFNS_3$ specification stands out as a model that delivers a reasonable fit to both the cross section of yields as well as to the cross section of realized yield volatilities.

Third, aside from measures of fit, the correlations between the fitted and realized standard deviations are important measures of how well the various specifications are able to capture the stochastic volatility observed in the data. Table 12 presents the correlations across the model specifications and the maturities we examine over the full sample period. The correlations are

Maturity	Ratios of variation for the fitted AFNS standard deviations					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0.00	0.98	0.01	1.06	0.51	1.00
6	0.00	1.01	0.05	1.11	0.50	1.00
12	0.00	1.00	0.14	1.13	0.49	1.00
24	0.00	0.89	0.27	1.05	0.50	1.00
36	0.00	0.83	0.33	1.00	0.48	1.00
60	0.00	0.84	0.34	1.00	0.39	1.00
84	0.00	0.91	0.30	1.06	0.30	1.00
120	0.00	0.99	0.23	1.12	0.21	1.00

Table 11: **Ratios of Variation between AFNS$_i$ Fitted Standard Deviations.**
The table presents the ratios of variation between AFNS fitted standard deviations, which are calculated as the standard deviation of a model's fitted standard deviations for a given maturity divided by the standard deviation of fitted standard deviations from the AFNS$_3$ model.

Maturity in months	Correlations between fitted and realized standard deviation series					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0	0.241	0.039	0.238	0.014	0.194
6	0	0.242	-0.007	0.235	-0.011	0.184
12	0	0.142	-0.079	0.125	-0.124	0.024
24	0	0.082	-0.170	0.039	-0.244	-0.107
36	0	0.085	-0.214	0.027	-0.284	-0.135
60	0	0.111	-0.238	0.055	-0.296	-0.122
84	0	0.137	-0.227	0.095	-0.274	-0.080
120	0	0.176	-0.195	0.148	-0.229	-0.006

Table 12: **Correlations Between Fitted and Realized Standard Deviation Series for the Full Sample of U.S. Treasury Data.**
The table presents the correlations between the 31-day fitted and realized standard deviations for the U.S. Treasury yield data over the full sample period from January 2, 1985 to March 1, 2010.

relatively low, reaching a maximum of just 0.242, and often being negative with the lowest value being -0.296. While these low values suggest that the model specifications are not capable of capturing the stochastic volatility in the data very well, the subsample correlation results reported by Jacobs and Karoui (2009) suggest that sample periods play a key, but as of yet not well understood, role in this analysis. For their monthly and weekly U.S. Treasury yields, they found that term structure models do not generate stochastic volatility measures that fit the data well for the post-1991 period.[27] For this reason we split our sample into three periods. The first period covers the seven-year period from January 2, 1985 to December 31, 1991. The second period covers the eleven years from January 2, 1992 to December 31, 2002. Finally, the third period covers the seven years from January 2, 2003 to January 29, 2010, which is the last day for which we can calculate the 31-day ahead realized volatility measure.

---

[27]Please note that our correlation values are not directly comparable to the correlations reported by Jacobs and Karoui (2009) as they smooth their logged realized variance series using an ARMA(1,1) filter.

Correlations between fitted and realized standard deviation series						
Maturity	January 2, 1985 to December 31, 1991 sample					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0	0.177	0.247	0.175	0.070	0.199
6	0	0.228	0.242	0.231	0.013	0.227
12	0	0.227	0.205	0.236	0.137	0.329
24	0	0.248	0.264	0.272	0.257	0.361
36	0	0.265	0.306	0.296	0.273	0.356
60	0	0.262	0.337	0.288	0.282	0.349
84	0	0.251	0.335	0.268	0.289	0.333
120	0	0.245	0.313	0.252	0.301	0.203
Maturity	January 2, 1992 to December 31, 2002 sample					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0	0.074	-0.056	0.071	-0.062	-0.023
6	0	0.130	-0.130	0.122	-0.166	-0.085
12	0	0.084	-0.138	0.059	-0.305	-0.273
24	0	0.081	-0.140	0.029	-0.326	-0.248
36	0	0.097	-0.158	0.035	-0.305	-0.200
60	0	0.069	-0.183	0.015	-0.276	-0.175
84	0	0.003	-0.167	-0.030	-0.228	-0.167
120	0	-0.075	-0.107	-0.089	-0.146	-0.152
Maturity	January 2, 2003 to January 29, 2010 sample					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0	0.120	-0.207	0.105	0.039	0.108
6	0	0.035	-0.242	-0.002	0.085	0.117
12	0	0.145	-0.296	0.068	-0.111	-0.057
24	0	0.315	-0.442	0.084	-0.384	-0.327
36	0	0.380	-0.546	0.033	-0.512	-0.461
60	0	0.421	-0.654	0.085	-0.625	-0.573
84	0	0.427	-0.697	0.187	-0.667	-0.605
120	0	0.417	-0.721	0.266	-0.686	-0.595

Table 13: **Correlations Between Fitted and Realized Standard Deviation Series for Three Subsample Periods in the U.S. Treasury Data.**
The table presents the correlations between the 31-day fitted and realized standard deviations for the U.S. Treasury yield data over three sample periods. The top panel is based on the period from January 2, 1985 to December 31, 1991 (1747 daily observations). The middle panel is based on the period from January 2, 1992 to December 31, 2002 (2731 daily observations). The bottom panel is based on the period January 2, 2003 to January 29, 2010 (1771 daily observations).

Our empirical results for the three subperiods suggest a result close to those of Jacobs and Karoui (2009). The top panel of Table 13 shows these correlations for the seven years from January 2, 1985 to December 31, 1991. Clearly, these correlations are all positive and much higher, reaching a maximum of 0.361. The $AFNS_1$-C and $AFNS_3$ specifications generate the highest correlations with the realized volatility series, but as mentioned above, the greater degree of variation in the fitted standard deviations generated by the $AFNS_3$ specification makes this model preferable over the $AFNS_1$-C model even for this period. However, the

Maturity	Loading on		
in months	First P.C.	Second P.C.	Third P.C.
3	-0.41	-0.65	0.46
6	-0.38	-0.38	-0.14
12	-0.36	-0.11	-0.43
24	-0.36	0.13	-0.41
36	-0.36	0.23	-0.26
60	-0.34	0.32	0.07
84	-0.32	0.35	0.30
120	-0.28	0.36	0.50
% explained	73.58	18.94	5.44

Table 14: **Eigenvectors of the First Three Principal Components of the 31-Day Realized Standard Deviation Series in U.S. Treasury Data.**
The loadings of yields of various maturities on the first three principal components of the realized standard deviation series are shown. The final row shows the proportion of all realized volatility variability accounted for by each principal component. The underlying data consist of daily U.S. Treasury zero-coupon bond yields from January 2, 1985 to March 1, 2010.

Correlation	$AFNS_0$			$AFNS_3$		
	L_t	S_t	C_t	L_t	S_t	C_t
P.C. 1	-0.212	0.220	0.146	-0.169	0.195	0.048
P.C. 2	0.037	0.349	0.243	0.090	0.331	0.162
P.C. 3	0.151	-0.086	-0.018	0.111	-0.074	0.100

Table 15: **Correlations Between Principal Components of the Realized Volatility Series and the Estimated Factors in the $AFNS_0$ and $AFNS_3$ Models for U.S. Treasury Data.**
The table presents the pairwise correlations between the first three principal components of the eight 31-day realized yield standard deviation series based on daily U.S. Treasury yields and the three estimated factors in the $AFNS_0$ and $AFNS_3$ models, respectively.

middle and bottom panels of Table 13 presents the low and mainly negative correlations for the subsequent two subsample periods. Note, though, that the $AFNS_1$-L model stands out for the most recent seven-year period with correlations above 0.4 for the five- to ten-year maturity range. Also, these values are reasonable given the low overall correlations reported in Table 12 for the full sample.

In summary, AFNS models of the term structure can be expanded to incorporate stochastic volatility, and the empirical results suggest that the $AFNS_3$ specification generates fitted volatility measures that exhibit a high degree of variation and simultaneously provide a close fit to the realized volatility measures in this sample of U.S. Treasury data. However, at a daily frequency, the correlation between any of the $AFNS_i$ model-implied yield standard deviations and the realized yield volatility is rather low and frequently negative. A potential explanation for this result is suggested by an examination of the realized volatility series. Table 14 reports

the loadings across yield maturities of the first three principal components in the eight series of realized yield standard deviations. The analysis shows that three factors are needed to encompass 98% of the variation in the data. In Table 15, we correlate these three principal components with the estimated factors from the most diverse AFNS_i models, namely the AFNS_0 and AFNS_3 models. There is weak positive correlation between the AFNS slope factors and the second principal component, but beyond that the principal components of the yield volatility series are close to being uncorrelated with the spanned yield curve factors. Hence, the relatively diffuse nature of the volatility dynamics observed in the U.S. Treasury data may require one or more additional factors to model it adequately at the high, daily frequency.

5   Empirical Results with Daily U.K. Gilt Yields

In this section, we estimate our AFNS models with stochastic volatility using U.K. gilt zero-coupon bond yields downloaded from the website of the Bank of England.[28]

5.1   Data Description

The specific U.K. gilt yields we use are zero-coupon yields constructed from the cubic-spline method described in Anderson and Sleath (1999, 2001).[29] The underlying data contain prices on U.K. gilt securities[30] with a significant amount outstanding and at least three months to maturity.[31] Furthermore, only after March 1997, general collateral repo rates with maturities up to six months are included in the data. As a consequence, the short end of the yield curve is sparsely populated prior to this date. We use gilt zero-coupon yields with the following maturities: 3-month, 6-month, 1-year, 2-year, 3-year, 5-year, 7-year, and 10-year. We use daily data and limit our sample to the period from January 2, 1985 to March 1, 2010 to have the sample of U.K. yields match that of the U.S. Treasury yields analyzed in the previous section. The summary statistics are provided in Table 16, while Figure 8 illustrates the time series of the three-month, two-year, five-year, and ten-year gilt zero-coupon bond yields.

As we saw in the previous section, three factors are sufficient to model the time variation

---

[28]The data is publicly available at the website of the Bank of England at the following link:
http://www.bankofengland.co.uk/statistics/yieldcurve/index.htm

[29]This method is an adaptation to U.K. yields of the method originally presented in Waggoner (1997) for U.S. Treasury yields.

[30]These are securities issued and guaranteed by the U.K. government with fixed coupon rates paid semi-annually and no embedded options.

[31]Note that there is no distinction between recently issued on-the-run securities and more seasoned off-the-run securities, a distinction frequently made in analysis of U.S. Treasury yields.
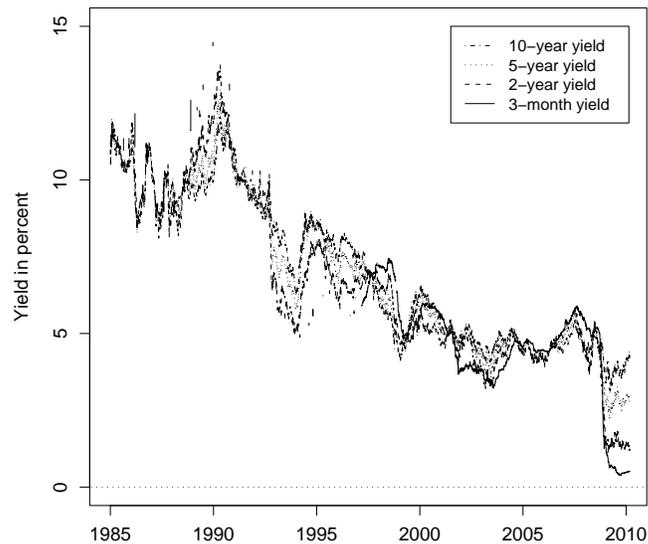
Figure 8: **U.K. Gilt Yields.**
Illustration of the U.K. zero-coupon gilt yields. The sample covers daily data for the period from January 2, 1985 to March 1, 2010. The yields shown have maturities in three months, two years, five years, and ten years.

Maturity in months	No. obs.	Mean in %	Std. dev. in %	Skewness	Kurtosis
3	3,377	4.77	1.89	0.15	6.14
6	5,534	6.32	3.00	0.59	3.04
12	6,306	6.71	3.02	0.41	2.65
24	6,323	6.76	2.78	0.33	2.35
36	6,323	6.83	2.65	0.31	2.13
60	6,323	6.92	2.53	0.31	1.88
84	6,323	6.97	2.48	0.31	1.74
120	6,323	6.98	2.42	0.29	1.60

Table 16: **Summary Statistics for U.K. Gilt Yields.**
Summary statistics of the zero-coupon U.K. gilt yields. The sample covers daily data for the period from January 2, 1985 to March 1, 2010.

in the cross section of U.S. Treasury bond yields. Here, we make a similar observation for the sample of U.K. gilt yields. Indeed, for the most recent ten years of our sample of U.K. gilt yields where all eight maturities are fully represented, 99.95% of the total variation is accounted for by three factors. Table 17 reports the eigenvectors that correspond to the first three principal components for this subsample of our data. The first principal component accounts for 96.4% of the variation in the gilt yields, and its loading across maturities is

Maturity	Loading on		
in months	First P.C.	Second P.C.	Third P.C.
3	-0.46	-0.49	0.34
6	-0.47	-0.35	0.10
12	-0.45	-0.08	-0.23
24	-0.39	0.23	-0.42
36	-0.33	0.35	-0.31
60	-0.25	0.41	0.05
84	-0.19	0.40	0.36
120	-0.12	0.36	0.65
% explained	96.40	3.22	0.33

Table 17: **Principal Component Analysis of the U.K. Gilt Yields.**
The principal component analysis of the U.K. gilts yields with maturities from three months to ten years. The sample covers daily data for the period from January 4, 2000 to March 1, 2010.

uniformly negative. Thus, like a level factor, a shock to this component changes all yields in the same direction irrespective of maturity. The second principal component accounts for 3.2% of the variation in these data and has sizable negative loadings for the shorter maturities and sizable positive loadings for the long maturities. Thus, like a slope factor, a shock to this component steepens or flattens the yield curve. Finally, the third component, which accounts for only 0.3% of the variation, has a U-shaped factor loading as a function of maturity, which is naturally interpreted as a curvature factor. These results motivate our use of the Nelson-Siegel model with its level, slope, and curvature factor for modeling this sample of U.K. gilt yields.

5.2   Conditional mean results

We first examine the in-sample estimation results for the five model specifications introduced in Section 3 in addition to the $AFNS_0$ model. Similar to the analysis of the U.S. yields, we only present results for the diagonal, independent-factors specification for each AFNS model class. We use the independent-factors specification because the AFNS models deliver essentially identical decompositions of the data into level, slope, and curvature factors independent of the specification of the P-dynamics. Thus, this restriction comes at a minimal loss of generality.[32]

Tables 18 and 19 present our parameter estimates of the six models. Similar to what we observed in the U.S. Treasury yield data, the level factor is the most persistent, and the curvature factor the least persistent factor in all six AFNS model classes. Also, in both the

---

[32]Results summarizing the estimation of the maximally flexible specifications of the models are available upon request.

Parameters	AFNS models with independent factors					
	AFNS_0	$\text{AFNS}_1\text{-L}$	$\text{AFNS}_1\text{-C}$	$\text{AFNS}_2\text{-LC}$	$\text{AFNS}_2\text{-SC}$	AFNS_3
κ_{11}^P	0.0666	0.0031	0.0667	0.0033	0.0706	$5.7 \cdot 10^{-6}$
	(0.0831)	(0.0215)	(0.0827)	(0.0204)	(0.0810)	(0.00090)
κ_{22}^P	0.2179	0.1842	0.2095	0.1765	0.1978	0.1196
	(0.139)	(0.130)	(0.136)	(0.121)	(0.105)	(0.0596)
κ_{33}^P	1.4522	1.3884	1.6409	1.6941	1.6113	1.6503
	(0.392)	(0.270)	(0.344)	(0.257)	(0.340)	(0.326)
θ_1^P	0.0828	0.0273	0.0828	0.0297	0.0029	0.00003
	(0.0197)	–	(0.0195)	–	(0.0185)	–
θ_2^P	-0.0156	-0.0226	-0.0159	-0.0229	0.0613	0.0369
	(0.0190)	(0.0198)	(0.0196)	(0.0199)	(0.0173)	(0.0171)
θ_3^P	-0.0111	-0.0092	0.0680	0.0680	0.0656	0.0322
	(0.00531)	(0.00550)	(0.00509)	(0.00470)	(0.00535)	(0.00512)
σ_{11}	0.0098	0.0388	0.0097	0.0381	0.0076	0.0447
	(0.00005)	(0.00018)	(0.00017)	(0.00019)	(0.00022)	(0.00025)
σ_{22}	0.0159	0.0013	0.0157	0.0011	0.0581	0.0472
	(0.00009)	(0.00001)	(0.00015)	(0.00001)	(0.00028)	(0.00037)
σ_{33}	0.0338	0.0206	0.1327	0.1376	0.1321	0.1751
	(0.00027)	(0.00024)	(0.00109)	(0.00116)	(0.00226)	(0.00158)
β_{11}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{12}	–	–	–	–	7.8959	–
	–	–	–	–	(0.962)	–
β_{13}	–	–	0.0007	–	0.4398	–
	–	–	(0.497)	–	(0.767)	–
β_{21}	–	2,041	–	2,536	–	–
	–	(0.499)	–	(0.238)	–	–
β_{22}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{23}	–	–	0.0009	–	–	–
	–	–	(0.226)	–	–	–
β_{31}	–	25.00	–	0.0118	–	–
	–	(0.611)	–	(0.371)	–	–
β_{32}	–	–	–	–	–	–
	–	–	–	–	–	–
β_{33}	–	–	–	–	–	–
	–	–	–	–	–	–

Table 18: **Parameter Estimates of the P-dynamics for AFNS_i Models with the Independent-Factors Specification for U.K. Gilt Data.**
The table contains the estimated K^P matrix, θ^P vector, Σ matrix, and β volatility sensitivity parameters for the independent-factors specification of the P-dynamics in the AFNS_i models. Estimated standard deviations for the parameter estimates are given in parentheses. The estimations are based on daily observations from January 2, 1985 to March 1, 2010.

U.S. and the U.K. data, the level factor is close to being a unit-root process. The slope factor has approximately the same rate of mean-reversion in both currency areas, about 0.15, while the curvature factor is slightly more rapidly mean-reverting in the U.K. data than in the U.S.

Parameters	AFNS models with independent factors					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
θ_1^Q	−	85.81	−	98.94	−	0.0002
	−	(0.344)	−	(0.285)	−	(0.0371)
θ_2^Q	−	−	−	−	0.08	0.0459
	−	−	−	−	−	(0.00003)
θ_3^Q	−	−	0.08	0.08	0.0775	0.0444
	−	−	−	−	(0.00161)	−
λ	0.6736	0.6554	0.6650	0.6495	0.6644	0.7333
	(0.00182)	(0.00174)	(0.00175)	(0.00167)	(0.00177)	(0.00096)
max. log L	283,576.1	284,344.5	283,346.8	284,041.3	283,246.9	269,593.2

Table 19: **Parameter Estimates of the Q-dynamics for $AFNS_i$ Models with the Independent-Factors Specification for U.K. Gilt Data.**
The table contains the estimated θ^Q vector and λ parameters for the independent-factors specification of the P-dynamics in the $AFNS_i$ models. Estimated standard deviations for the parameter estimates are given in parentheses. The estimations are based on daily observations from January 2, 1985 to March 1, 2010. The maximum log-likelihood values are reported, although the models are non-nested.

data. For the mean parameters in θ^P, we get close to the same estimated values when we compare each $AFNS_i$ model across the two data sets. However, for the volatility parameters, we start to see some differences between the U.S. and the U.K. yield curves. If we compare the results for the $AFNS_0$ models, which provide a proxy for the unconditional volatility of each factor, the volatility of the U.K. level factor is about twice that of the U.S. level factor, and the volatility of the U.K. slope factor is about 50% larger than the corresponding estimate for the U.S. slope factor. These differences could be a reflection of the fact that there has been several monetary policy regime shifts in the U.K. during this 25-year period, in particular the U.K. departure from the EMS in 1992 and the independence of the Bank of England in May 1997 come to mind, while events of that nature are absent in the U.S. Treasury data. These events are likely to have caused uncertainty to be elevated for extended periods of the sample. Finally, as for the β volatility sensitivity parameters, there is a role for the level factor in the stochastic volatility of the slope and curvature factor and, possibly, a role for the slope factor in the stochastic volatility of the level factor. However, as in the U.S. Treasury data, there does not appear to be any role for the curvature factor in the stochastic volatility of either the level or the slope factor.

Despite the fact that the $AFNS_i$ models with stochastic volatility are non-nested, we can still use the obtained maximum log likelihood values as a crude measure of model performance as noted earlier. The ranking of the six $AFNS_i$ models for the U.K. gilt data is identical to the ranking in the U.S. Treasury data with the $AFNS_1$-L and $AFNS_2$-L,C models delivering

Maturity in months	RMSE for AFNS models with independent factors					
	AFNS_0	AFNS_1-L	AFNS_1-C	AFNS_2-LC	AFNS_2-SC	AFNS_3
3	23.95	24.72	23.88	24.51	23.76	4.52
6	14.88	15.03	14.84	14.98	14.80	9.46
12	0.42	0.21	0.51	0.31	0.63	10.04
24	2.77	2.81	2.76	2.80	2.76	4.44
36	0.00	0.00	0.00	0.00	0.00	1.22
60	2.62	2.62	2.62	2.61	2.61	3.87
84	0.87	0.88	0.93	0.95	1.04	5.43
120	10.15	9.46	10.15	9.50	10.13	12.02

Table 20: **RMSE of the Fitted Errors for U.K. Gilt Yields in the AFNS_i Models.** The table presents the root-mean-squared errors for the fitted U.K. gilt yields across the 8 maturities under the independent-factors specification of the AFNS model with different stochastic volatility specifications. The sample covers the period from January 2, 1985 to March 1, 2010. All numbers are expressed in basis points.

the highest maxima, and the AFNS_3 model obtaining the lowest maximum likelihood value. Note, though, that the differences in likelihood value across models are smaller in the U.K. data.

Another way to assess the performance of the different AFNS specifications of stochastic volatility is to examine the cross-sectional fit of the yield curve, as shown in Table 20 using root-mean-squared-error for the models' fitted errors. Unlike the U.S. Treasury data, we do not see a clear improvement in model fit from the introduction of stochastic volatility relative to the AFNS_0 specification, rather the contrary, five of the AFNS_i models deliver essentially identical fit to the cross section of yields. It is only the AFNS_3 model that produces a different, more even distribution of the fitted errors across maturities. The relatively poor fit of the three- and six-month yields in most of the AFNS_i models could be a consequence of the fact that these two maturities are only observed periodically prior to March 1997 and largely reflect rates on general collateral repo contracts rather than yields on gilt securities.

Overall, though, all AFNS_i models deliver a satisfactory fit to the data and no single model stands out based on either the RMSEs of the fitted errors or the obtained likelihood values. Thus, like with the U.S. Treasury data, we will be using the fit of the model-implied yield volatility to the realized yield volatility as a way of model validation. However, before turning to that task, we will briefly repeat our discussion of the identification issues pertaining to certain θ^Q parameters as they appear in the U.K. data.
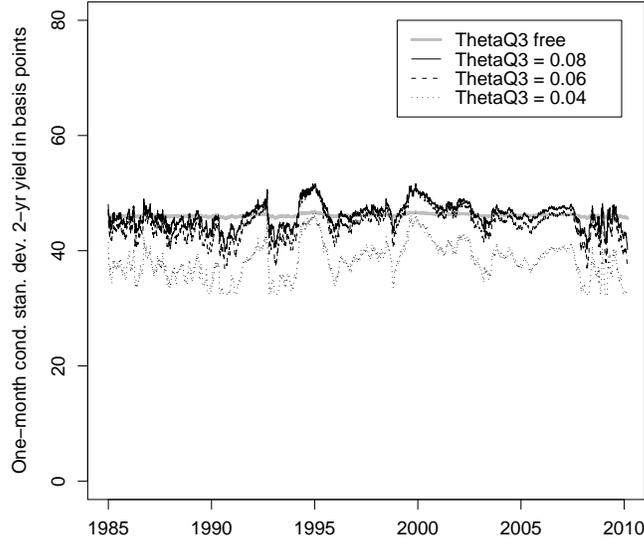
Figure 9: **Fitted One-Month Conditional Standard Deviations of the Two-Year U.K. Gilt Yield From AFNS$_1$-C Models.**

5.3 Identification Problems Related to θ^Q Parameters

Given that some of the θ^Q parameters were hard to identify in the U.S. Treasury data, it is not all that surprising that we encounter similar problems in the U.K. gilt data, i.e., θ_3^Q remains hard to identify in the AFNS$_1$-C and AFNS$_2$-L,C models and a similar problem pertains to the value of θ_2^Q in the AFNS$_2$-S,C model. Also, similar to what we observed in the U.S. Treasury data, the specific value of these θ^Q parameters significantly affect the size of the generated stochastic yield volatility. Figure 9 illustrates this for the AFNS$_1$-C model. Here, we demonstrate how restrictions on θ_3^Q affect the results obtained in the AFNS$_1$-C model. Similar results are obtained, but not reported, for the other two model classes mentioned above.

In Figure 10, we compare the estimated level and slope factors from four specifications of the AFNS$_1$-C model one of which is the unrestricted model, while the three other specifications have θ_3^Q fixed at low values. Similar to the U.S. Treasury data, we find that the decomposition into level and slope is entirely unaffected by restrictions on θ_3^Q. The minimum correlation between the estimated level factors is 0.9998, and the minimum correlation between the estimated slope factors is also 0.9998.

Figure 11 compares the estimated curvature factors with the result of the unrestricted
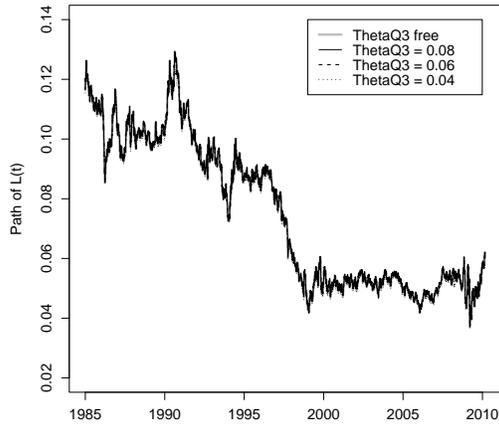
(a) Estimated level factors.
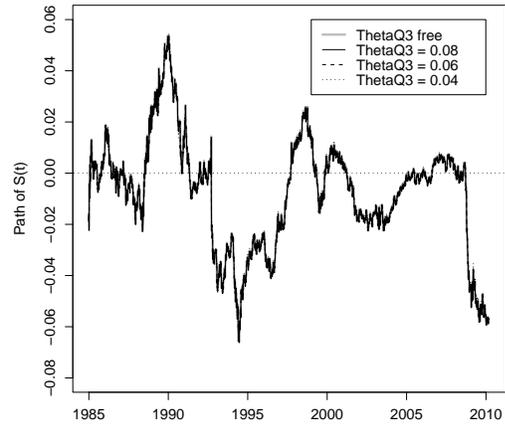
(b) Estimated slope factors.

Figure 10: **Estimated Level and Slope Factors in AFNS$_1$-C Models for U.K. Gilt Data.**
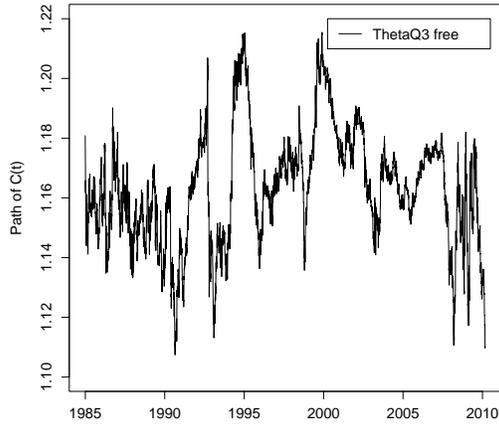
(a) θ_3^Q unrestricted.
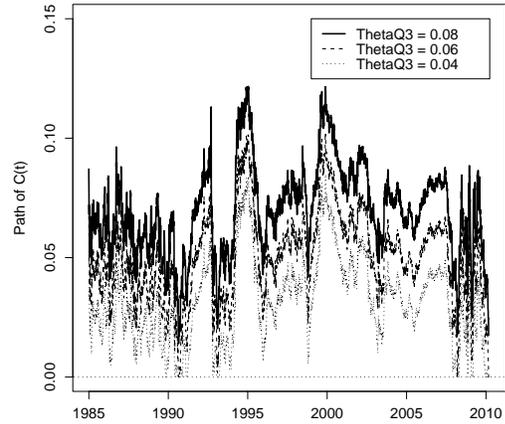
(b) θ_3^Q fixed.

Figure 11: **Estimated Curvature Factors in AFNS$_1$-C Models for U.K. Gilt Data.**

model shown in Figure 11(a), while Figure 11(b) illustrates the corresponding results from the three restricted specifications. We note that the curvature factor in the unrestricted model has values close to the estimated value of θ_3^Q (1.17), and in the restricted models they are also close to the corresponding restricted values of θ_3^Q. Thus, the choice of θ_3^Q determines the level at which the curvature factor operates, which in turn affects the size of the generated yield volatility through the mechanics explained earlier in the discussion of the U.S. Treasury data. However, the time series correlation remains very high with the smallest correlation

46

Maturity in months	AFNS$_1$-C models with independent factors							
	θ_3^Q free		$\theta_3^Q = 0.08$		$\theta_3^Q = 0.06$		$\theta_3^Q = 0.04$	
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
3	8.58	23.95	8.47	23.88	8.48	23.64	8.95	22.03
6	5.64	14.88	5.60	14.84	5.59	14.77	6.14	15.00
12	-0.06	0.42	-0.07	0.51	-0.10	0.77	-0.11	2.62
24	-0.30	2.77	-0.30	2.76	-0.32	2.75	-0.55	2.66
36	0.00	0.00	0.00	0.00	0.00	0.05	0.00	1.23
60	-0.45	2.62	-0.44	2.62	-0.41	2.61	-0.11	2.42
84	-0.01	0.87	0.00	0.93	0.01	1.09	0.07	1.51
120	0.96	10.15	0.89	10.15	0.69	10.17	-0.32	9.97

Table 21: **RMSE of the Fitted Errors for U.K. Gilt Yields in AFNS$_1$-C Models.** The table presents the root-mean-squared errors for the fitted yields across the 8 maturities under the independent-factors specification of the AFNS$_1$-C model with varying restrictions on θ_3^Q. The sample covers the period from January 2, 1985 to March 1, 2010. All numbers are expressed in basis points.

being 0.9905.

In terms of model fit, Table 21 reports the mean and RMSE of the fitted errors in the four specifications discussed here. We note that the fitted error statistics are indistinguishable for the unrestricted model and the two models with θ_3^Q fixed at 0.08 and 0.06, respectively. However, for the model with θ_3^Q fixed at 0.04, we start to see some deviations related to the fact that the curvature is bound by the lower zero-boundary on several occasions due to the low value of θ_3^Q as can be seen in Figure 11(b). Furthermore, as observed in Figure 9, a very low value for θ_3^Q like 0.04 induces a downward bias in the fitted yield volatility that is not apparent for values of θ_3^Q around 0.08. As in the U.S. data, this makes us fix θ_3^Q at 0.08 in the empirical analysis.

Finally, when θ_3^Q is unrestricted and, as a consequence, the general level of X_t^3 is not that well identified, the estimation has problems identifying σ_{11}, σ_{22}, θ_3^P, and θ_3^Q. Table 22, which reports the estimated dynamic model parameters for the four specifications discussed here, provides evidence of this. On the other hand, when we restrict θ_3^Q at the low values, the σ and θ^P parameters are better identified. However, there are still uncertainty about the values of the β parameters, even though in most estimations they are insignificant based on likelihood ratio tests.[33] Thus, overall, it looks like the entire set of parameters is better identified when we restrict θ_3^Q to a reasonable, but low value, at little costs in terms of model fit.

In summary, we fix the relevant θ^Q parameters at low values in an attempt to maximize the size of the generated yield volatility in the AFNS$_1$-C, AFNS$_2$-L,C, and AFNS$_2$-S,C models with essentially no effect on their model fit. Still, as we will see in the next section and as we

---

[33]Note, though, the Kalman filter is only a QML estimator in the AFNS$_i$ models with stochastic volatility. As a consequence, the asymptotics of the LR tests are not known.

Parameters	AFNS$_1$-C models with independent factors			
	θ_3^Q free	$\theta_3^Q = 0.08$	$\theta_3^Q = 0.06$	$\theta_3^Q = 0.04$
κ_{11}^P	0.0696	0.0667	0.0685	0.0746
	(0.0834)	(0.0827)	(0.0804)	(0.0709)
κ_{22}^P	0.2239	0.2095	0.1945	0.1453
	(0.140)	(0.136)	(0.131)	(0.122)
κ_{33}^P	1.4758	1.6409	1.979	1.407
	(0.403)	(0.344)	(0.219)	(0.249)
θ_1^P	0.0826	0.0828	0.0805	0.0675
	(0.0190)	(0.0195)	(0.0187)	(0.0198)
θ_2^P	-0.0154	-0.0159	-0.0163	-0.0212
	(0.0187)	(0.0196)	(0.0203)	(0.0253)
θ_3^P	1.1623	0.0680	0.0498	0.0429
	(0.645)	(0.00509)	(0.00425)	(0.00670)
σ_{11}	0.0093	0.0097	0.0094	0.0081
	(0.00312)	(0.00017)	(0.00013)	(0.00009)
σ_{22}	0.0094	0.0157	0.0151	0.0135
	(0.00527)	(0.00015)	(0.00012)	(0.00014)
σ_{33}	0.0314	0.1327	0.1534	0.1424
	(0.00872)	(0.00109)	(0.00111)	(0.00132)
β_{13}	0.0902	0.0007418	0.1861	3.333
	(0.613)	(0.497)	(0.532)	(0.785)
β_{23}	1.6122	0.000926	0.00000305	0.0000313
	(2.085)	(0.226)	(0.261)	(0.471)
θ_3^Q	1.1736	0.08	0.06	0.04
	(0.645)	−	−	−
λ	0.6731	0.6650	0.6689	0.6514
	(0.00188)	(0.00175)	(0.00133)	(0.000772)
Max log L	283,569.3	283,346.8	283,009.2	278,844.5

Table 22: **Parameter Estimates of AFNS$_1$-C Models for the U.K. Gilt Data.**
The table contains the estimated dynamic parameters for the independent-factors specification of the P-dynamics in AFNS$_1$-C models with varying restrictions on θ_3^Q. Estimated standard deviations for the parameter estimates are given in parentheses. The estimations are based on daily observations from January 2, 1985 to March 1, 2010. The maximum log likelihood values are reported in the last row.

saw for the U.S. Treasury data, this 'helping hand' does not allow any these three models to outperform the other AFNS$_i$ models in terms of fitting yield volatility. Thus, none of these restrictions affect the conclusions we draw later on.

5.4 Stochastic volatility results for the AFNS model specifications

Collin-Dufresne et al. (2009) demonstrate that there is a tension in affine models between fitting the cross section of yields and capturing their stochastic volatility. In this section, we analyze how severe that tension is in our sample of U.K. gilt yields.

As indicated in the discussion above of the in-sample results, the AFNS model specifica-
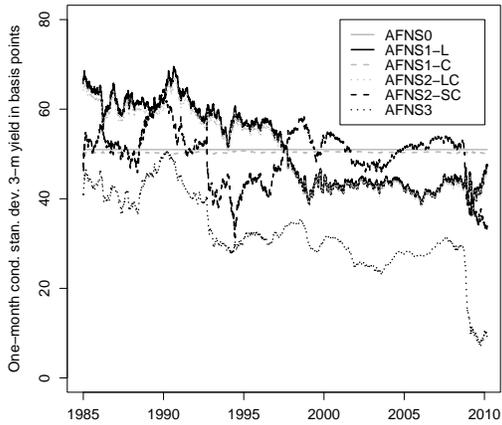
(a) Three-month gilt yield
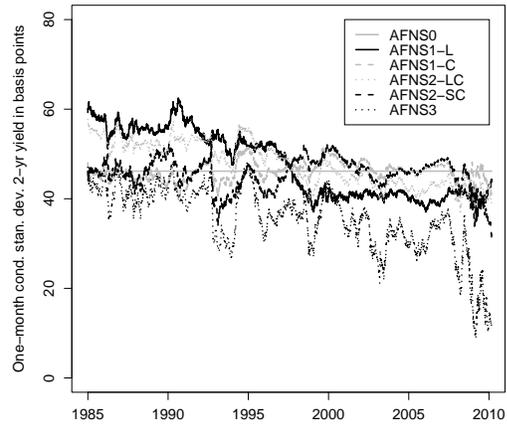
(b) Two-year gilt yield
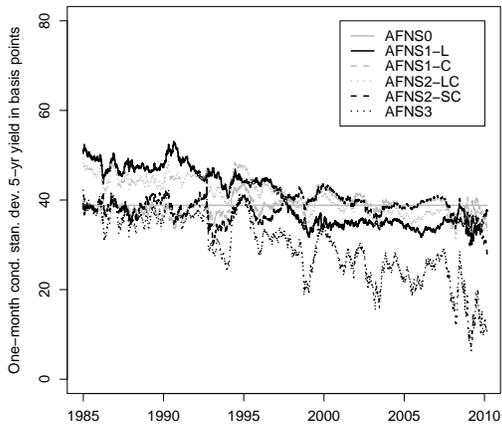
(c) Five-year gilt yield
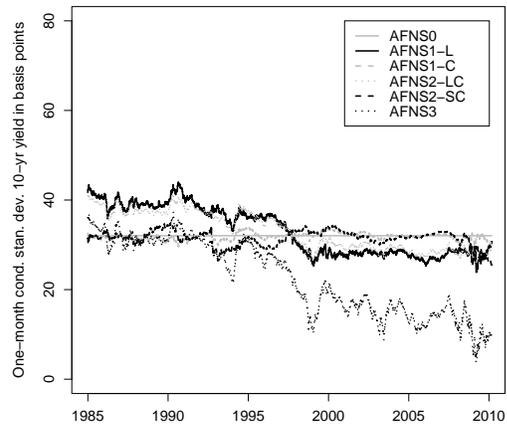
(d) Ten-year gilt yield

Figure 12: **Fitted One-Month Conditional Standard Deviations of Bond Yields from the AFNS$_i$ Models for U.K. Gilt Data.**

tions with stochastic volatility do not differ markedly in terms of fitting the observed U.K. gilt yield curve. However, as in the U.S. Treasury data, their fitted volatility measures do differ greatly from each other.[34] Figure 12 shows this for four of the eight maturities in our sample.

---

[34]The fitted one-month conditional volatility measures are given by the square root of

$$V_t^P[y_T(\tau)] = \frac{1}{\tau^2} B(\tau)' V_t^P[X_T] B(\tau),$$

where $V_t^P[X_T]$ is the conditional covariance matrix of the state variables, $T - t = \frac{1}{12}$, and τ is the yield maturity in years.

Correlation	Three-month U.K. gilt yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	-0.326	1.000	0.111	0.776
$AFNS_1$-C			1	-0.311	-0.054	-0.181
$AFNS_2$-LC				1	0.112	0.778
$AFNS_2$-SC					1	0.688
$AFNS_3$						1

Correlation	Two-year U.K. gilt yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	-0.331	0.917	-0.078	0.629
$AFNS_1$-C			1	0.071	0.539	0.382
$AFNS_2$-LC				1	0.152	0.831
$AFNS_2$-SC					1	0.634
$AFNS_3$						1

Correlation	Five-year U.K. gilt yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	-0.331	0.894	-0.150	0.802
$AFNS_1$-C			1	0.124	0.712	0.223
$AFNS_2$-LC				1	0.185	0.954
$AFNS_2$-SC					1	0.405
$AFNS_3$						1

Correlation	Ten-year U.K. gilt yield					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	-0.329	0.983	-0.033	0.962
$AFNS_1$-C			1	-0.154	0.408	-0.093
$AFNS_2$-LC				1	0.049	0.990
$AFNS_2$-SC					1	0.142
$AFNS_3$						1

Table 23: **Pairwise Correlations of the One-Month Conditional Standard Deviation of Four U.K. Gilt Yields in the $AFNS_i$ Models.**
The table contains the pairwise correlations between the one-month conditional standard deviations of the three-month, two-year, five-year, and ten-year U.K. gilt yields estimated by the $AFNS_i$ models. The estimation is based on daily observations from January 2, 1985 to March 1, 2010.

First, the $AFNS_0$ model produces a flat line close to the average across the six $AFNS_i$ specifications. Second, the three models where the level factor is allowed to generate stochastic volatility exhibits a declining trend in the fitted yield volatility at all maturities throughout the entire sample. Third, the $AFNS_3$ model stands out in that it produces fitted yield volatilities that are clearly lower than those in any of the other models.

In Table 23, we calculate the correlations between the fitted yield volatilities from the six $AFNS_i$ models for the same four maturities depicted in Figure 12. Again, we observe some natural clustering. The fitted measures from the $AFNS_1$-L and $AFNS_2$-L,C models

are very highly correlated for all maturities. In turn, the AFNS_3 model is highly correlated with both of those models. At the other extreme, the AFNS_1-C model exhibits rather low, and frequently negative, correlations in its generated yield volatility relative to those of the other models. Finally, the fitted volatility of the AFNS_2-S,C model is hardly correlated with those from the AFNS_1-L and AFNS_2-L,C models. However, it does exhibit a high degree of correlation with the AFNS_3 model for the short- and medium-term yield maturities, where the slope and curvature factors have their maximum effect on the yield curve. Thus, as in the U.S. Treasury data, the AFNS_3 model emerges as a strong representative specification whose fitted yield volatility is highly correlated with those of the other AFNS_i models in exactly the maturity ranges where each of these other AFNS_i models can be expected to produce the closest fit to the actual yield volatility.[35]

To evaluate the in-sample fit of these monthly standard deviations, we compare them to a standard measure of realized volatility based on the high-frequency daily data. This measure is fully nonparametric and has been shown to converge to the underlying realization of the conditional variance as the sampling frequency increases. The square root of this measure retains these properties. For a given month t with N trading days (where N is most often 21 or 22), we generate the realized standard deviation as

$$RV_{t,\tau}^{STD} = \sqrt{\sum_{n=1}^{N} \Delta y_{t+n/N}^2(\tau)},$$

where $\Delta y_{t+n/N}(\tau)$ is the change in yield $y_t(\tau)$ from trading day $(n-1)$ to trading day n. Note that, due to the limited availability of the three- and six-month yields in the data, it is not possible to reliably calculate the realized volatility measure for these two maturities prior to April 1997.

Figure 13 plots the realized yield volatility measure for four maturities and compares it to the fitted yield volatility from the four most diverse AFNS_i models. We note that, on average, the realized volatility is below the fitted volatility from the AFNS_i models.

Table 24 reports the summary statistics for the realized yield standard deviations based on the U.K. gilt data. If we compare it to the statistics for the U.S. Treasury data in Table 9, it follows that the mean realized volatility is about the same in the two datasets for the one- to ten-year maturities where the sample periods are overlapping. However, the variability in the realized volatility is larger in the U.K. data as measured by the standard deviation of the

---

[35]For example, the AFNS_2-S,C model can be expected to fit the volatility of short- and medium-term yields closely while having little to bear on the volatility of long-term yields due to the decay in the factor loading of the slope and the curvature factor in the AFNS_i models.
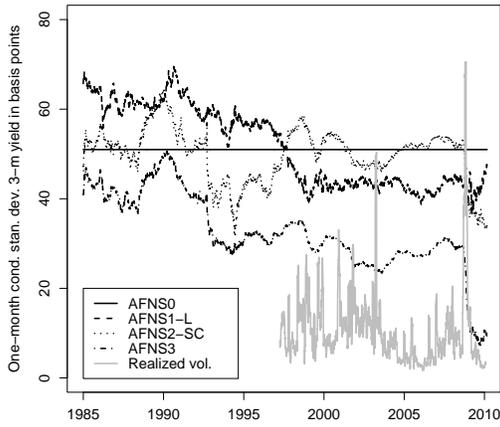
(a) Three-month gilt yield
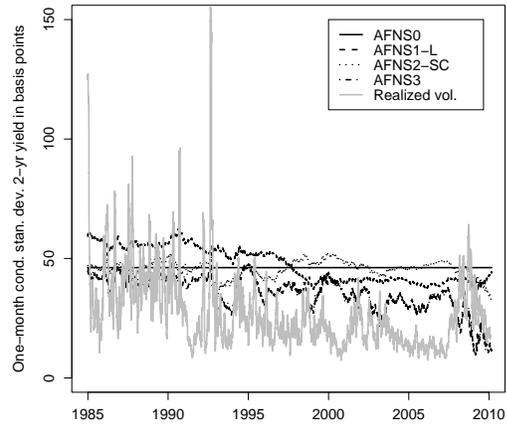
(b) Two-year gilt yield
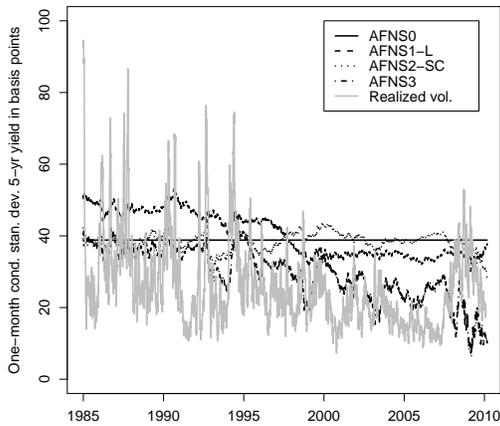
(c) Five-year gilt yield
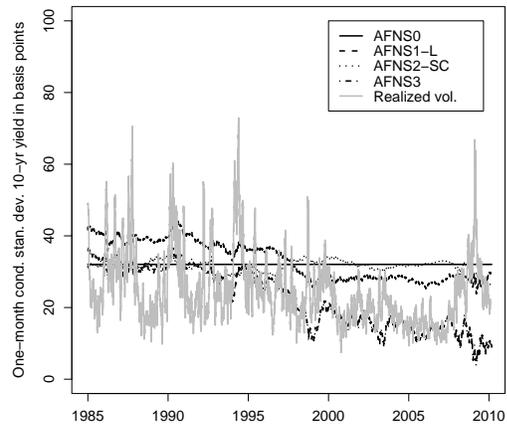
(d) Ten-year gilt yield

Figure 13: **Comparison of the Fitted versus Realized One-Month Conditional Standard Deviations for U.K. Gilt Yields.**

realized yield volatilities.

Table 25 reports the mean and RMSE of the erors of the fitted yield volatilities from the six AFNS$_i$ models relative to the measure of realized yield volatility. The AFNS$_3$ model is the only model that is consistently close to fitting the realized volatility measure. In fact, it produces the lowest RMSEs amongst the six AFNS$_i$ models for all eight maturities in the data ranging from 20 basis points for the variable short-term maturities down to just 11 basis points for the longest, less variable maturities.

In terms of the high-frequency time-series correlations between the realized and fitted

Maturity in months	Mean in bps	Std. dev. in bps	Std. dev. ratio
3	10.89	8.73	0.87
6	13.03	8.50	0.84
12	25.75	19.43	1.92
24	27.45	16.35	1.62
36	27.32	14.58	1.44
60	26.20	12.23	1.21
84	25.55	10.94	1.08
120	24.98	10.10	1.00

Table 24: **Summary Statistics for the 31-Day Realized Standard Deviation Series based on Daily U.K. Gilt Data.**
The summary statistics are for the 31-day rolling realized standard deviations based on the daily U.K. gilt yield data over the period from January 2, 1985 through March 1, 2010. The standard deviation ratio is calculated as the standard deviation in question divided by the standard deviation for the ten-year maturity.

Maturity in months	RMSE for fitted standard deviations											
	$AFNS_0$		$AFNS_1$-L		$AFNS_1$-C		$AFNS_2$-LC		$AFNS_2$-SC		$AFNS_3$	
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
3	40.09	41.03	32.53	33.72	39.50	40.46	32.03	33.23	39.17	40.44	16.21	19.21
6	36.41	37.39	29.35	30.54	36.01	37.01	29.36	30.57	35.66	36.92	14.53	17.62
12	22.16	29.47	23.26	28.70	21.87	29.41	22.93	28.64	21.39	28.79	8.56	19.66
24	18.72	24.85	20.20	24.93	18.70	25.18	20.34	25.34	18.17	24.70	8.03	18.08
36	16.55	22.05	18.25	22.51	16.62	22.46	18.53	23.01	16.07	22.05	6.82	16.24
60	12.62	17.57	14.40	18.26	12.64	17.82	14.61	18.50	12.14	17.55	3.39	12.74
84	9.54	14.52	11.13	15.10	9.47	14.61	11.23	15.11	9.04	14.49	0.37	11.23
120	7.06	12.32	8.17	12.52	6.89	12.30	8.17	12.43	6.51	12.35	-2.36	11.01

Table 25: **RMSE for the 31-Day Fitted Conditional Standard Deviations.**
The table presents the RMSE values for the monthly model-based fitted standard deviations relative to the 31-day realized standard deviations based on the daily U.K. gilt data over the period from January 2, 1985 to March 1, 2010. Note that the three- and six-month maturities are missing prior to April 1997.

yield volatility measures, Table 26 reports those for all six $AFNS_i$ models for the full sample. Three models stand out: the $AFNS_1$-L, $AFNS_2$-L,C, and $AFNS_3$ models exhibit consistent high, positive correlations in the range from 23% to 51% for the one- to ten-year maturity range.

In Table 27, we study the high frequency time-series correlations between the realized and fitted yield volatility series for three subperiods: 1985-1991, 1992-2002, and 2003-2010. From the table it follows that the high correlations are primarily observed during the period from the beginning of 1992 to the end of 2002, while the early 1985-1991 period is characterized by positive, but low correlations not exceeding 0.23. For the most recent seven years, most correlations have been negative in all the models with the exception of the $AFNS_1$-L model.

Maturity	Correlations between fitted and realized standard deviation series					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
3	0	0.077	-0.033	0.075	0.009	0.077
6	0	0.163	-0.148	0.129	0.034	0.079
12	0	0.509	-0.189	0.498	0.147	0.415
24	0	0.448	-0.192	0.391	-0.008	0.243
36	0	0.430	-0.162	0.358	-0.067	0.228
60	0	0.409	-0.106	0.376	-0.085	0.289
84	0	0.387	-0.078	0.380	-0.111	0.313
120	0	0.378	-0.081	0.378	-0.165	0.332

Table 26: **Correlations between Fitted and Realized Standard Deviation Series for the Full Sample of U.K. Gilt Yields.**
The table presents the correlations between the 31-day fitted and realized standard deviations for the U.K. gilt yield dataset over the full sample period from January 2, 1985 to March 1, 2010.

This model exhibits systematically positive correlations throughout the entire 25-year sample. However, from Table 25 it follows that the fitted yield volatilities from this model are not particularly close to the realized yield volatilities.

We can now summarize our results for the U.K. gilt data. First, based on correlations, which is a measure widely used in the literature on spanned and unspanned stochastic volatility, three-factor affine models, in general, have difficulties generating the right time variation because the three spanned yield factors do not respond to the short-term, or high frequency, variation in the realized yield volatility measures. This conclusion is independent of the number of factors that are allowed to generate stochastic volatility, and whichever combination of the spanned factors is allowed to be the source of the stochastic volatility. However, it is not obvious that all realized volatility should be priced and therefore reflected in the spanned yield curve factors. If a brief one-day spike in volatility caused, for example, by the historical 20-minute 700 point drop in the Dow Jones Industrial index on May 6, 2010 is not expected to repeat itself going forward, should it really be reflected in the fitted volatility measure even though such shocks surely causes spikes in the realized volatility? If, instead, we rely on the statistics for the fitted errors between the fitted and realized measures of yield volatility for the purpose of model validation, the conclusion is much more favorable towards the spanned factors. Also, given that measures of fitted errors are widely accepted in terms of judging whether a term structure model is able to fit the cross section of yields, it is not obvious that we should not also use this kind of model evaluation for judging a model's ability to fit the term structure of yield volatilities. Along that dimension, we find that the $AFNS_3$ model, which allows all three spanned factors to generate stochastic volatility, delivers a reasonable fit to both the cross section of yields as well as the term structure of yield volatilities for both

Correlations between fitted and realized standard deviation series						
Maturity	January 2, 1985 to December 31, 1991					
in months	AFNS$_0$	AFNS$_1$-L	AFNS$_1$-C	AFNS$_2$-LC	AFNS$_2$-SC	AFNS$_3$
3	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
6	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
12	0	0.031	0.046	0.051	0.045	0.110
24	0	0.032	0.004	0.037	0.069	0.146
36	0	0.057	0.021	0.071	0.061	0.168
60	0	0.118	0.036	0.139	0.063	0.203
84	0	0.169	0.040	0.196	0.075	0.227
120	0	0.213	0.002	0.225	0.090	0.230
Maturity	January 2, 1992 to December 31, 2002					
in months	AFNS$_0$	AFNS$_1$-L	AFNS$_1$-C	AFNS$_2$-LC	AFNS$_2$-SC	AFNS$_3$
3	0	-0.255	-0.048	-0.258	-0.066	-0.191
6	0	0.019	-0.206	-0.011	0.129	0.093
12	0	0.277	0.037	0.296	-0.029	0.279
24	0	0.288	0.057	0.319	-0.036	0.204
36	0	0.322	0.088	0.365	-0.034	0.246
60	0	0.340	0.143	0.420	-0.039	0.355
84	0	0.335	0.154	0.412	-0.095	0.377
120	0	0.368	0.110	0.411	-0.223	0.389
Maturity	January 2, 2003 to January 29, 2010					
in months	AFNS$_0$	AFNS$_1$-L	AFNS$_1$-C	AFNS$_2$-LC	AFNS$_2$-SC	AFNS$_3$
3	0	0.231	-0.240	0.219	-0.058	0.031
6	0	0.256	-0.298	0.193	-0.039	0.012
12	0	0.280	-0.418	-0.052	-0.218	-0.231
24	0	0.279	-0.538	-0.419	-0.450	-0.485
36	0	0.277	-0.545	-0.466	-0.507	-0.522
60	0	0.236	-0.554	-0.476	-0.577	-0.571
84	0	0.159	-0.547	-0.450	-0.634	-0.613
120	0	0.086	-0.522	-0.340	-0.654	-0.580

Table 27: **Correlations between Fitted and Realized Standard Deviation Series for Three Subsample Periods in the U.K. Gilt Data.**
The table presents the correlations between the 31-day fitted and realized standard deviations for the U.K. gilt yield dataset over three sample periods. The top panel is based on the period from January 2, 1985 to December 31, 1991 (1770 daily observations). The middle panel is based on the period from January 2, 1992 to December 31, 2002 (2744 daily observations). The bottom panel is based on the period from January 2, 2003 to January 29, 2010 (1791 daily observations).

U.S. Treasury and U.K. gilt yields during the 25-year sample period analyzed here.

Finally, if we want to refine any of these models, we need to incorporate some of the structure in the realized yield volatilities. Table 28 reports the results of a principal component analysis based on the realized volatility for the six maturities in the U.K. gilt data that are represented throughout the sample period. The analysis reveals that the first three prinicpal components explain 99.68% of the variation. However, as shown in Table 29, the three principal components are not particularly highly correlated with the estimated AFNS factors,

Maturity	Loading on		
in months	First P.C.	Second P.C.	Third P.C.
12	-0.55	-0.57	0.57
24	-0.49	-0.18	-0.40
36	-0.44	0.06	-0.49
60	-0.35	0.34	-0.17
84	-0.29	0.47	0.19
120	-0.23	0.55	0.46
% explained	86.97	11.46	1.25

Table 28: **Eigenvectors of the First Three Principal Components of the 31-Day Realized Standard Deviation Series in the U.K. Gilt Data.**
The loadings of yields of various maturities on the first three principal components of the realized standard deviation series are shown. The final row shows the proportion of all realized volatility variability accounted for by each principal component. The underlying data consist of daily U.K. gilt zero-coupon bond yields from January 2, 1985 to March 1, 2010.

Correlation	$AFNS_0$			$AFNS_3$		
	L_t	S_t	C_t	L_t	S_t	C_t
P.C. 1	0.475	0.068	-0.155	0.473	0.082	-0.174
P.C. 2	-0.171	0.235	-0.017	-0.172	0.218	0.059
P.C. 3	0.243	-0.130	0.039	0.249	-0.129	-0.020

Table 29: **Correlations Between Principal Components of the Realized Volatility Series and the Estimated Factors in the $AFNS_0$ and $AFNS_3$ Models in the U.K. Gilt Data.**
The table presents the pairwise correlations between the first three principal components of the six 31-day realized yield standard deviation series based on daily U.K. gilt yields for which the full sample period is available and the three estimated factors in the $AFNS_0$ and $AFNS_3$ models, respectively.

only the fairly high positive correlation between the first principal component and the AFNS level factor appear worth mentioning, but it is still below 50%. Thus, additional factors are required to match the realized yield volatility series more closely at high frequency.

6 Empirical Results with Daily U.S. Dollar Swap and LIBOR Rates

Recent research by Collin-Dufresne et al. (CGJ, 2009) as well as Jacobs and Karoui (2009) have examined the stochastic volatility present in U.S. dollar swap and LIBOR rates. These studies use weekly data encompassing the period from January 1988 through December 2005. In this section, we estimate our proposed AFNS models with stochastic volatility on this dataset, but at daily frequency.[36] The data set we examine consists of zero-coupon yields

---

[36] We thank Chris Jones for sharing these data with us.
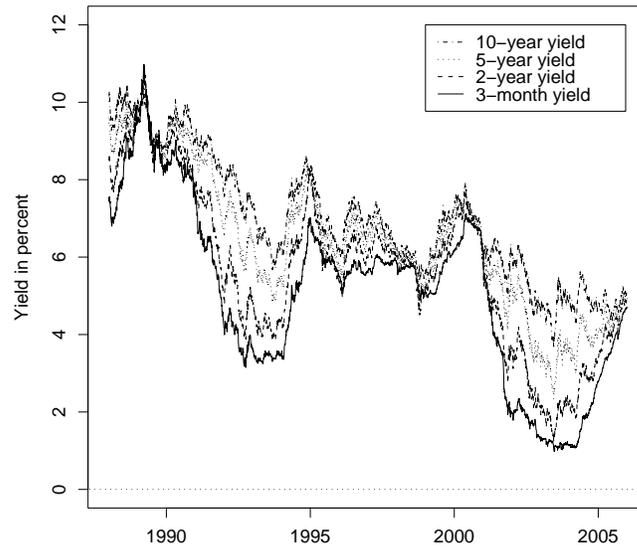
Figure 14: **U.S. Dollar Swap and LIBOR Rates.**
Illustration of the U.S. swap and LIBOR rates. The sample covers daily data for the period from January 4, 1988 to December 29, 2005. The yields shown have maturities in six months, two years, five years, and ten years.

generated from daily swap and LIBOR rates from January 4, 1988 through December 29, 2005. For each observation date, the yield curves are constructed by bootstrapping all the available swap and LIBOR rates. The eight maturities in the data set are 6 months and 1-, 2-, 3-, 4-, 5-, 7- and 10-years. These maturities were chosen because actual yield quotes were observed for each day in the sample, which should ensure that the bootstrapped yields are particularly accurate. For further details on this data set, please see the description in CGJ (2009).

As we saw in the previous sections, three factors are sufficient to model the time variation in the cross section of U.S. and U.K. government bond yields. Here, we make a similar observation for the sample of U.S. swap and LIBOR rates. For this sample, 99.98% of the total variation is accounted for by three factors. Table 31 reports the eigenvectors that correspond to the first three principal components for this data. The first principal component accounts for 96.5% of the variation in the swap and LIBOR rates, and its loading across maturities is uniformly negative. Thus, like a level factor, a shock to this component changes all yields in the same direction irrespective of maturity. The second principal component accounts for 3.3% of the variation of these data and has sizable negative loadings for the shorter maturities

Maturity in months	No. obs.	Mean in %	Std. dev. in %	Skewness	Kurtosis
6	4,456	5.10	2.27	0.01	2.40
12	4,456	5.34	2.23	0.02	2.43
24	4,456	5.71	2.09	0.03	2.41
36	4,456	6.00	1.98	0.05	2.38
48	4,456	6.23	1.89	0.09	2.33
60	4,456	6.42	1.83	0.12	2.28
84	4,456	6.70	1.72	0.20	2.20
120	4,456	6.98	1.62	0.29	2.16

Table 30: **Summary Statistics for U.S. Dollar Swap and LIBOR Rates.**
Summary statistics of the U.S. swap and LIBOR rates. The sample covers daily data for the period from January 4, 1988 to December 29, 2005.

Maturity in months	Loading on		
	First P.C.	Second P.C.	Third P.C.
6	-0.40	-0.51	0.60
12	-0.40	-0.41	-0.03
24	-0.38	-0.16	-0.41
36	-0.36	0.03	-0.38
48	-0.35	0.17	-0.25
60	-0.33	0.28	-0.10
84	-0.31	0.41	0.17
120	-0.28	0.52	0.47
% explained	96.54	3.28	0.16

Table 31: **Principal Component Analysis of U.S. Dollar Swap and LIBOR Rates.**
The principal component analysis of the U.S. swap and LIBOR rates with maturities from six months to ten years. The sample covers daily data for the period from January 4, 1988 to December 29, 2005.

and sizable positive loadings for the long maturities. Thus, like a slope factor, a shock to this component steepens or flattens the yield curve. Finally, the third component, which accounts for only 0.2% of the variation, has a U-shaped factor loading as a function of maturity, which is naturally interpreted as a curvature factor. Again, these results motivate our use of the Nelson-Siegel model with its level, slope, and curvature factor for modeling this sample of U.S. dollar swap and LIBOR rates.

6.1   Conditional mean results

Table 32 presents the estimated parameters of our AFNS stochastic volatility models for the U.S. dollar swap market. As expected, these estimates have important similarities with the estimated parameters for the U.S. Treasury data presented earlier. In particular, the persistence of each factor is close to the same across the two data samples. Also, the estimated mean vector θ^P in each AFNS$_i$ specification is hardly distinguishable from the corresponding

Parameters	AFNS models with independent factors					
	AFNS_0	AFNS_1-L	AFNS_1-C	AFNS_2-LC	AFNS_2-SC	AFNS_3
κ_{11}^P	0.0335	0.0490	0.0370	0.0473	0.0357	0.0610
	(0.0737)	(0.0477)	(0.0769)	(0.0526)	(0.0759)	(0.0338)
κ_{22}^P	0.2275	0.2923	0.2424	0.2851	0.4274	0.3218
	(0.130)	(0.102)	(0.147)	(0.178)	(0.147)	(0.0978)
κ_{33}^P	1.5156	2.5499	1.5874	2.5508	1.5494	1.8953
	(0.247)	(0.110)	(0.356)	(0.475)	(0.354)	(0.387)
θ_1^P	0.0825	0.0520	0.0809	0.0536	0.0005	0.0234
	(0.0215)	−	(0.0217)	−	(0.0231)	−
θ_2^P	-0.0287	-0.0157	-0.0278	-0.0156	0.0493	0.0357
	(0.0165)	(0.0169)	(0.0156)	(0.0180)	(0.0120)	(0.00983)
θ_3^P	-0.0073	-0.0000	0.0742	0.0805	0.0681	0.0569
	(0.00399)	(0.00313)	(0.00400)	(0.00307)	(0.00494)	(0.00529)
σ_{11}	0.0059	0.0583	0.0060	0.0584	0.0059	0.0392
	(0.00006)	(0.00013)	(0.00020)	(0.00012)	(0.00023)	(0.00046)
σ_{22}	0.0135	0.0036	0.0092	0.0016	0.0688	0.0544
	(0.00012)	(0.00003)	(0.00039)	(0.00001)	(0.00062)	(0.00051)
σ_{33}	0.0264	0.0038	0.0979	0.1097	0.1004	0.1464
	(0.00021)	(0.00004)	(0.00084)	(0.00137)	(0.00256)	(0.00133)
β_{11}	−	−	−	−	−	−
	−	−	−	−	−	−
β_{12}	−	−	−	−	0.0000	−
	−	−	−	−	(0.717)	−
β_{13}	−	−	0.0000	−	0.0000	−
	−	−	(0.878)	−	(1.03)	−
β_{21}	−	303.3	−	1,494	−	−
	−	(0.113)	−	(0.838)	−	−
β_{22}	−	−	−	−	−	−
	−	−	−	−	−	−
β_{23}	−	−	15.66	0.0050	−	−
	−	−	(2.55)	(0.773)	−	−
β_{31}	−	1,039	−	−	−	−
	−	(0.0921)	−	−	−	−
β_{32}	−	−	−	−	−	−
	−	−	−	−	−	−
β_{33}	−	−	−	−	−	−
	−	−	−	−	−	−
θ_1^Q	−	2,547	−	2,537	−	1,425
	−	(0.0981)	−	(0.658)	−	(43.4)
θ_2^Q	−	−	−	−	0.08	0.0509
	−	−	−	−	−	(0.00041)
θ_3^Q	−	−	0.08	0.08	0.0737	0.0462
	−	−	−	−	(0.00271)	−
λ	0.3890	0.4925	0.3791	0.4808	0.3760	0.3126
	(0.00131)	(0.00141)	(0.00131)	(0.00140)	(0.00127)	(0.00066)
Max log L	232,737.6	236,659.5	232,592.9	236,509.8	232,165.1	218,807.3

Table 32: **Parameter Estimates for AFNS$_i$ Models with the Independent-Factors Specification for U.S. Dollar Swap Data.**
The table contains the estimated K^P matrix, θ^P vector, Σ matrix, θ^Q vector, as well as the estimated λ parameters for the independent-factors specification of the P-dynamics in the AFNS$_i$ models, all estimated for the U.S. dollar swaps and LIBOR rates. Estimated standard deviations of the parameter estimates are given in parentheses. The maximum log-likelihood values are reported, although the models are non-nested.

Maturity	RMSE for fitted standard deviations					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
6	28.26	27.40	28.29	27.49	28.20	16.02
12	9.99	9.27	10.04	9.34	10.05	0.67
24	0.33	0.08	0.38	0.13	0.46	5.02
36	0.74	0.63	0.74	0.63	0.73	2.72
48	0.00	0.00	0.00	0.00	0.00	0.31
60	0.60	0.42	0.60	0.42	0.60	2.06
84	1.00	0.51	0.98	0.50	0.99	5.04
120	5.21	4.04	5.19	4.00	5.16	10.02

Table 33: **RMSE of the Fitted Yields from AFNS$_i$ Models for U.S. Dollar Swap Data.**
The table presents the RMSE values for the daily fitted yields from the AFNS models with stochastic volatility estimated on the U.S. dollar swap dataset.

estimates for U.S. Treasury yields. As for the σ volatility parameters, they are harder to compare across data sets, even for identical AFNS$_i$ specifications, as their estimated value is impacted by the size of the β volatility sensitivity parameters, which do vary quite a bit across the two samples although they still indicate that there is no role for the slope or curvature in the volatility of the level factor, while the level factor does have a role in the volatility of the other two factors. Finally, the λ parameters have lower estimated values in the U.S. dollar swap data relative to both the U.S. Treasury and U.K. gilt data. This implies a slower decay in the factor loading of the slope and curvature factor, which appears reasonable as the constellation of maturities in the U.S. swaps data is skewed slightly towards longer maturities relative to the samples of U.S. Treasury and U.K. gilt yields analyzed previously.

With respect to the in-sample fit of the yield curve presented in Table 33, the RMSE values of the fitted yields across the AFNS$_i$ models are not that different from each other. This result suggests that, as in the U.S. Treasury data, the introduction of stochastic volatility factors does not affect the overall performance of AFNS models. All six AFNS$_i$ models fit the two- to ten-year range very well, while they all underperform for the six-month LIBOR. Note also that the fit of the AFNS$_3$ specification preferred in the analysis of the U.S. and U.K. government bond yields is on par with that of the other models. Hence, the performance of the models is better judged with respect to their fitted volatility measures.

6.2   Conditional variance results

If we examine the models' performance with respect to the data's realized volatility, we again find that the AFNS$_3$ specification generates reasonable fitted standard deviations, but other models also perform reasonably.
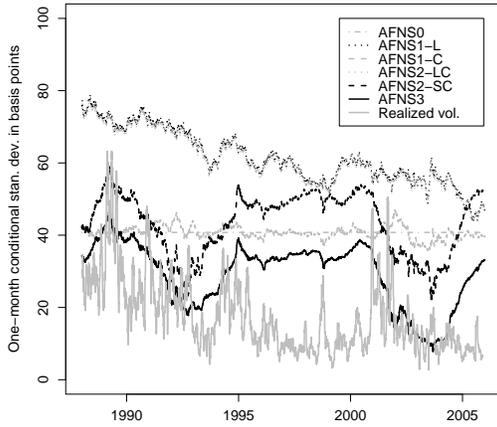
(a) Six-month LIBOR rate
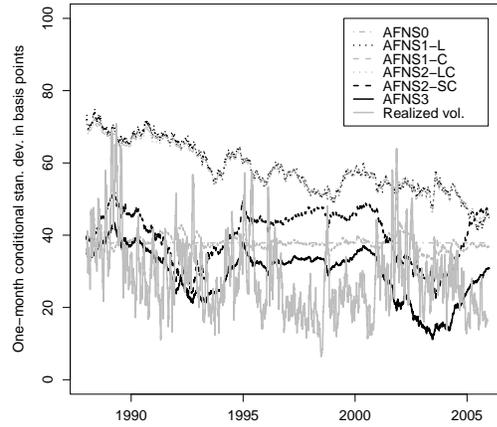
(b) Two-year swap rate
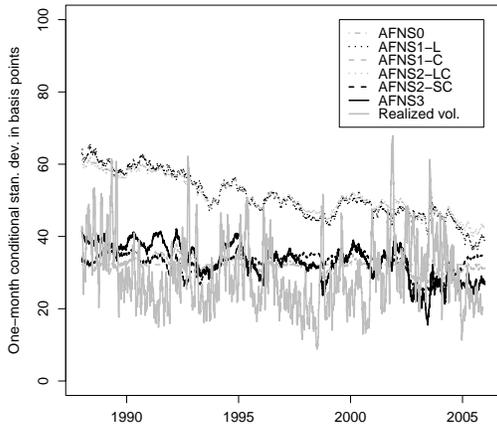
(c) Five-year swap rate
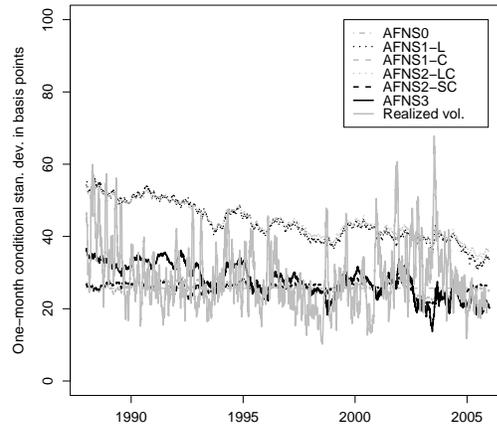
(d) Ten-year swap rate

Figure 15: **Fitted Standard Deviations from the AFNS$_i$ Models for U.S. Dollar Swap and LIBOR Data.**

Figure 15 presents the fitted one-month conditional standard deviations of the six-month LIBOR rate and the two-, five- and ten-year swap rates calculated based on the full sample estimation for each of the six AFNS$_i$ models.[37] We note that the AFNS$_3$ model produces the lowest fitted yield volatility, while the AFNS$_1$-L and AFNS$_2$-L,C models systematically

[37]The figure shows the square root of

$$V_t^P[y_T(\tau)] = \frac{1}{\tau^2} B(\tau)' V_t^P[X_T] B(\tau),$$

where $V_t^P[X_T]$ is the conditional covariance matrix of the state variables, $T - t = \frac{1}{12}$, and $\tau = 0.5, 2, 5,$ and 10 years, respectively.

Correlation	Six-month LIBOR					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.396	1.000	0.000	0.389
$AFNS_1$-C			1	0.398	0.015	0.195
$AFNS_2$-LC				1	0.000	0.390
$AFNS_2$-SC					1	0.909
$AFNS_3$						1

Correlation	Two-year swap rate					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.396	0.999	0.016	0.493
$AFNS_1$-C			1	0.428	0.057	0.301
$AFNS_2$-LC				1	0.032	0.508
$AFNS_2$-SC					1	0.867
$AFNS_3$						1

Correlation	Five-year swap rate					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.398	0.989	0.203	0.787
$AFNS_1$-C			1	0.521	0.529	0.804
$AFNS_2$-LC				1	0.314	0.866
$AFNS_2$-SC					1	0.633
$AFNS_3$						1

Correlation	Ten-year swap rate					
	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
$AFNS_0$	1	0	0	0	0	0
$AFNS_1$-L		1	0.397	0.996	0.292	0.880
$AFNS_1$-C			1	0.467	0.743	0.764
$AFNS_2$-LC				1	0.365	0.914
$AFNS_2$-SC					1	0.614
$AFNS_3$						1

Table 34: **Pairwise Correlations of the One-Month Conditional Standard Deviation of One U.S. LIBOR and three U.S. Swap Rates in the $AFNS_i$ Models.**
The table contains the pairwise correlations between the one-month conditional standard deviations of the six-month LIBOR and two-year, five-year, and ten-year U.S. swap rates as estimated by the $AFNS_i$ models. The estimation is based on daily observations from January 4, 1988 to December 29, 2005.

produce higher fitted yield volatilities than any of the other models.

Table 34 reports the pairwise correlations of the fitted yield volatilities across the six models for those same four yield maturities. Again, the $AFNS_1$-L and $AFNS_2$-L,C are essentially perfectly correlated. Also, the $AFNS_2$-S,C and $AFNS_3$ models exhibit a higher degree of covariation between their fitted yield volatilities than in the previous data sets for U.S. and U.K. government bond yields. Finally, the $AFNS_1$-C model is characterized by fitted yield volatilities that have a relatively low correlation with the fitted measures from the other

Maturity	Ratios of variation for the fitted AFNS standard deviations					
	AFNS$_0$	AFNS$_1$-L	AFNS$_1$-C	AFNS$_2$-LC	AFNS$_2$-SC	AFNS$_3$
6	0.00	0.87	0.20	0.85	0.96	1.00
12	0.00	0.91	0.21	0.88	0.96	1.00
24	0.00	1.02	0.24	0.95	0.97	1.00
36	0.00	1.25	0.32	1.07	0.93	1.00
48	0.00	1.34	0.37	1.10	0.80	1.00
60	0.00	1.30	0.37	1.06	0.57	1.00
84	0.00	1.25	0.34	1.04	0.44	1.00
120	0.00	1.22	0.27	1.07	0.32	1.00

Table 35: **Ratios of Variation between AFNS$_i$ Fitted Standard Deviations for U.S. Dollar Swap and LIBOR Rates.**
The table presents the ratios of variation between AFNS$_i$ fitted standard deviations, which are calculated as the standard deviation of a model's fitted yield volatility for a given maturity divided by the standard deviation of the corresponding fitted yield volatility from the AFNS$_3$ model.

Maturity in months	Mean in bps	Std. dev. in bps	Std. dev. ratio
6	17.73	10.38	1.16
12	23.95	10.40	1.16
24	28.81	10.44	1.16
36	29.70	9.92	1.11
48	29.74	9.65	1.08
60	29.60	9.48	1.06
84	29.25	9.20	1.03
120	29.04	8.96	1.00

Table 36: **Summary Statistics for the 31-Day Realized Standard Deviation Series based on Daily U.S. Dollar Swap and LIBOR Data.**
The summary statistics are for the 31-day rolling realized standard deviations based on daily U.S. dollar swap and LIBOR rates over the period from January 6, 1988 to December 31, 2002. The standard deviation ratio is calculated as the standard deviation in question divided by the standard deviation for the ten-year maturity.

models.

In terms of variation in the generated yield volatility, Table 35 reveals that several of the AFNS$_i$ models produce fitted yield volatilities with at least as much variation as that of the AFNS$_3$ model. Equally visible is the very low variation in the fitted volatility of the AFNS$_1$-C model.

Figure 15 also presents the 31-day-ahead realized standard deviations of the six-month LIBOR rate and the two-, five- and ten-year swap rates calculated from the sample of daily data. Table 36 contains the summary statistics of the realized volatility measures and shows that the unconditional volatility of the realized volatility of the swap and LIBOR rates has a hump-shaped pattern that peaks at the one-year maturity but declines steadily up to the

Maturity	Correlations between fitted and realized standard deviation series					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
6	0	0.633	0.229	0.633	0.065	0.298
12	0	0.546	0.218	0.546	0.023	0.229
24	0	0.408	0.131	0.406	-0.113	0.078
36	0	0.303	0.048	0.289	-0.201	0.019
48	0	0.230	-0.024	0.202	-0.250	-0.007
60	0	0.191	-0.075	0.154	-0.264	0.015
84	0	0.174	-0.118	0.140	-0.271	0.019
120	0	0.235	-0.093	0.212	-0.235	0.107

Table 37: **Correlations between Fitted and Realized Standard Deviation Series for the Full Sample of U.S. Dollar Swaps Data.**
The table presents the correlations between the monthly, fitted and realized standard deviations for the U.S. dollar swaps dataset over the full sample period from January 4, 1988 through December 29, 2005.

ten-year maturity. Furthermore, we note that the means of the realized volatility measure are slightly higher for the one- to ten-year maturity, while the mean realized volatility of the six-month LIBOR is about 1.5 basis points lower than the corresponding number for the six-month U.S. Treasury yield. Of course, this comparison does not correct for the differences in sample periods and should be interpreted with caution.

With respect to the correlations between the fitted and realized standard deviations, Table 37 shows that several models, including the $AFNS_3$ model, exhibit low and often negative values as before. However, the $AFNS_1$-L and the $AFNS_2$-L,C models generate very high correlations at the short maturities (on the order of 50%) and reasonably high correlations for longer maturities (on the order of 20%). This good performance suggests that the level factor plays an important role in the stochastic volatility exhibited by this dataset over the full sample. Table 38 shows the correlations for three subsample periods: 1988-1991, 1992-2002, and 2003-2005. As noted by Jacobs and Karoui (2009), most of the models' correlations deteriorate for the middle 1992-2002 sample period for as yet unclear reasons. In the early 1988-1991 period, we see correlations as high as 50% for $AFNS_1$-L, $AFNS_2$-L,C, $AFNS_3$ models, and the first two models preserve those high correlations in 2003-2005 period, whereas the $AFNS_3$ continues to exhibit negative correlations in that period.

If, instead of simple correlations, we focus on how close the fitted yield volatilities are to the series of realized yield volatility, the results look more favorable for the spanned factors. Table 39 reports the RMSEs of the fitted yield volatilities in the six $AFNS_i$ models. We note that the $AFNS_3$ model again stands out with the closest fit to the volatility of the short- and medium-term swap rates with RMSEe starting at 16 basis points for the six-month LIBOR

Correlations between fitted and realized standard deviation series						
Maturity	January 6, 1988 to December 31, 1991					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
6	0	0.024	-0.528	0.021	0.407	0.428
12	0	0.085	-0.551	0.077	0.497	0.523
24	0	0.182	-0.558	0.156	0.431	0.467
36	0	0.261	-0.547	0.192	0.355	0.354
48	0	0.329	-0.551	0.230	0.287	0.180
60	0	0.386	-0.553	0.278	0.117	-0.112
84	0	0.465	-0.532	0.388	-0.072	-0.122
120	0	0.536	-0.477	0.493	-0.291	0.006
Maturity	January 2, 1992 to December 31, 2002					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
6	0	0.338	0.256	0.339	-0.328	-0.238
12	0	0.284	0.290	0.286	-0.351	-0.291
24	0	0.248	0.253	0.253	-0.381	-0.316
36	0	0.213	0.200	0.215	-0.386	-0.239
48	0	0.165	0.147	0.157	-0.376	-0.137
60	0	0.118	0.101	0.103	-0.325	-0.014
84	0	0.056	0.047	0.038	-0.286	-0.014
120	0	0.051	0.047	0.038	-0.247	0.015
Maturity	January 3, 2003 to November 28, 2005					
in months	$AFNS_0$	$AFNS_1$-L	$AFNS_1$-C	$AFNS_2$-LC	$AFNS_2$-SC	$AFNS_3$
6	0	0.338	0.026	0.339	-0.350	-0.324
12	0	0.371	-0.115	0.370	-0.406	-0.415
24	0	0.608	-0.233	0.603	-0.636	-0.610
36	0	0.658	-0.298	0.624	-0.682	-0.540
48	0	0.674	-0.346	0.600	-0.696	-0.371
60	0	0.670	-0.385	0.570	-0.680	-0.099
84	0	0.635	-0.448	0.542	-0.674	-0.033
120	0	0.564	-0.492	0.501	-0.650	-0.009

Table 38: **Correlations between Fitted and Realized Standard Deviation Series for Three Subsample Periods in the U.S. Dollar Swap and LIBOR Data.**
The table presents the correlations between the monthly, fitted and realized standard deviations for the U.S. dollar swap and LIBOR dataset over three sample periods. The top panel is based on the period from January 4, 1988 to December 31, 1991 (995 daily observations). The middle panel is based on the period from January 2, 1992 to December 31, 2002 (2721 daily observations). The bottom panel is based on the period from January 2, 2003 to November 28, 2005 (719 daily observations).

down to 11 basis points for the three-year swap rate. For the longer-term swap rates it performance is on par with the second best model, the $AFNS_2$-S,C model, both producing RMSEs around 10 basis points for the five- to ten-year maturity range. On the other hand, the $AFNS_1$-L and $AFNS_2$-L,C models that had relatively high positive correlations with the realized volatility series produce disappointingly high RMSEs as they imply fitted yield volatility that is systematically too high as can be seen in Figure 15

In summary, AFNS models that incorporate stochastic volatility seem able to generate

Maturity in months	RMSE for fitted standard deviations											
	AFNS$_0$		AFNS$_1$-L		AFNS$_1$-C		AFNS$_2$-LC		AFNS$_2$-SC		AFNS$_3$	
	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE	Mean	RMSE
6	23.03	25.26	43.81	44.55	22.98	25.11	43.31	44.06	25.63	28.71	11.42	16.16
12	15.61	18.76	36.22	37.30	15.59	18.61	35.79	36.86	18.10	22.24	5.01	12.72
24	9.08	13.83	29.54	31.18	9.08	13.77	29.20	30.81	11.28	17.33	0.56	12.17
36	6.35	11.78	26.56	28.47	6.35	11.84	26.34	28.17	7.96	14.49	1.50	11.41
48	5.07	10.90	24.82	26.93	5.07	11.10	24.71	26.71	6.20	12.96	2.95	11.34
60	2.59	9.83	21.40	23.78	2.61	10.13	21.52	23.76	3.14	11.03	3.67	11.23
84	0.10	9.20	18.45	21.01	0.16	9.52	18.76	21.18	0.39	9.99	2.40	10.56
120	-3.36	9.57	14.79	17.50	-3.27	9.72	15.31	17.88	-3.23	9.96	-1.11	9.66

Table 39: **RMSE for the Monthly Fitted Conditional Standard Deviations.**
The table presents the RMSE values for the monthly model-based fitted standard deviations relative to the 31-day realized standard deviations based on the daily U.S. dollar swap and LIBOR rates over the period from January 4, 1988 to December 29, 2005.

fitted standard deviations that match some part of the realized standard deviations. If we use simple correlations as a performance measure, the AFNS$_1$-L model performs well with strictly positive correlations for all maturities for the full sample as well as for the three subsample periods. For the six-month LIBOR, which was the focus of attention in CGJ (2009), this model produces a full sample correlation as high as 63.3%. If, instead, performance is measured by the closeness of the fitted yield volatility to the corresponding realized measure, the AFNS$_3$ model performs well with low RMSEs at all maturities. Thus, it is feasible for spanned factors to match important aspects of the realized yield volatility in this sample of U.S. dollar swap and LIBOR rates, and the common element of the best fitting models is the presence of the level factor as a driver of the volatility dynamics. Further analysis is required to determine why this factor appears to play a greater role in the swaps data than in the Treasury data. However, a principal components analysis of the eight realized yield standard deviation series shows that the first three components account for 99% of the variation in the data, see Table 40, and the AFNS level factor is the spanned factor with the highest correlation with all three of those principal components as evidenced in Table 41.

7   Conclusion

In this paper, we extend the AFNS model introduced by CDR (2007) to incorporate stochastic volatility. We do so by proposing five new specifications whose sources of stochastic volatility are different permutations of the AFNS model's three spanned factors. Our empirical exercises show that the introduction of these volatility factors does not have a significant impact on the models' fitted yield values relative to the constant volatility AFNS$_0$ model. Furthermore,

Maturity	Loading on		
in months	First P.C.	Second P.C.	Third P.C.
6	-0.32	-0.61	0.48
12	-0.37	-0.45	-0.03
24	-0.40	-0.17	-0.39
36	-0.38	0.05	-0.38
48	-0.37	0.21	-0.24
60	-0.35	0.30	-0.05
84	-0.33	0.37	0.27
120	-0.31	0.34	0.58
% explained	83.39	13.02	2.51

Table 40: **Eigenvectors of the First Three Principal Components of the 31-Day Realized Standard Deviation Series for U.S. Dollar Swap and LIBOR Rates.**
The loadings of yields of various maturities on the first three principal components of the realized standard deviation series are shown. The final row shows the proportion of all realized volatility variability accounted for by each principal component. The underlying data consist of daily U.S. swap and LIBOR zero-coupon rates from January 4, 1988 to December 29, 2005.

Correlation	AFNS_0			AFNS_3		
	L_t	S_t	C_t	L_t	S_t	C_t
P.C. 1	-0.387	0.168	-0.046	-0.364	0.147	-0.080
P.C. 2	0.409	0.346	0.357	0.459	0.348	0.133
P.C. 3	0.323	0.031	-0.094	0.343	0.029	-0.125

Table 41: **Correlations Between Principal Components of the Realized Volatility Series and the Estimated Factors in the AFNS_0 and AFNS_3 Models for U.S. Dollar Swap and LIBOR Data.**
The table presents the pairwise correlations between the first three principal components of the eight 31-day realized yield standard deviation series based on daily U.S. dollar swap and LIBOR rates and the three estimated factors in the AFNS_0 and AFNS_3 models, respectively.

our results suggest that certain of these models, particularly the AFNS_3 model based on all three factors exhibiting stochastic volatility, are able to generate a reasonable amount of volatility dynamics in sample. In particular, for our daily U.S. Treasury and U.K. gilt yields datasets, the AFNS_3 model generates the most variation in its fitted standard deviations and provides the closest fit to our realized volatility measures in addition to exhibiting the best correlations for the pre-1992 period. For the daily U.S. dollar swap and LIBOR dataset, the AFNS_3 model also produces the overall closest fit to the realized yield volatility measures. However, in terms of correlations, the two models with stochastic level factor, the AFNS_1-L and the AFNS_2-L,C models, stand out with positive, and frequently high, correlations with the realized measures. In conclusion, we find evidence that the extended AFNS modeling framework captures an important fraction of the stochastic volatility observed in all three data sets in addition to preserving the good in-sample yield fit and ease of estimation that

is the advantage of the original Gaussian AFNS$_0$ model class. Still, at daily frequency, parts of the observed volatility in interest rates is only weakly associated with any of the spanned term structure factors.

Appendix

This appendix contains additional details of the five model specifications.

The AFNS$_1$ model with stochastic volatility through the level factor

In this model class, the Q-dynamics are assumed to be

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_t^1} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

This structure implies that γ and δ in the system of ODEs provided in Equations (2) and (3) are given by

$$
\gamma = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \quad \text{and} \quad \delta = \begin{pmatrix} 1 & 0 & 0 \\ \beta_{21} & 0 & 0 \\ \beta_{31} & 0 & 0 \end{pmatrix},
$$

and $B^1(t,T)$, $B^2(t,T)$, and $B^3(t,T)$ are the unique solutions to the following system of ODEs

$$
\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}
$$
$$
- \frac{1}{2} \sum_{j=1}^{3} \left[\begin{pmatrix} \sigma_{11} & \sigma_{21} & \sigma_{31} \\ 0 & \sigma_{22} & \sigma_{32} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} (B^1)^2 & B^1B^2 & B^1B^3 \\ B^1B^2 & (B^2)^2 & B^2B^3 \\ B^1B^3 & B^2B^3 & (B^3)^2 \end{pmatrix} \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \right]_{j,j} (\delta^j)'.
$$

To detail the essentially affine risk premium specification first introduced in Duffee (2002) for this class of models, start by defining the matrices $D(X_t)$ and $D^{-1}(X_t)$ as

$$
D(X_t) = \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1} & 0 \\ 0 & 0 & \sqrt{1+\beta_{31}X_t^1} \end{pmatrix} \quad \text{and} \quad D^{-1}(X_t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{1+\beta_{21}X_t^1}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{1+\beta_{31}X_t^1}} \end{pmatrix}.
$$

In the essentially affine risk premium specification for the $A_1(3)$ class of models considered here, Γ_t is given by

$$
\Gamma_t = D(X_t) \begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \end{pmatrix} + D^{-1}(X_t) \begin{pmatrix} 0 & 0 & 0 \\ \gamma_{21}^2 & \gamma_{22}^2 & \gamma_{23}^2 \\ \gamma_{31}^2 & \gamma_{32}^2 & \gamma_{33}^2 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}.
$$

By implication, there is a total of 9 free parameters in the essentially affine risk premium specification of Γ_t

in this case. From the structure of Γ_t, it follows that the product $D(X_t)\Gamma_t dt$ is given by

$$D(X_t)\Lambda_t dt = \begin{pmatrix} X_t^1 \gamma_1^1 \\ (1 + \beta_{21} X_t^1)\gamma_2^1 \\ (1 + \beta_{31} X_t^1)\gamma_3^1 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \gamma_{21}^2 X_t^1 + \gamma_{22}^2 X_t^2 + \gamma_{23}^2 X_t^3 \\ \gamma_{31}^2 X_t^1 + \gamma_{32}^2 X_t^2 + \gamma_{33}^2 X_t^3 \end{pmatrix} dt.$$

By deducting the term $\Sigma D(X_t)\Gamma_t dt$ from the SDE for the Q-dynamics and replacing dW_t^Q with dW_t^P, we obtain the SDE for the P-dynamics which, without loss of generality, can be written as

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$
$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & 0 \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{21} X_t^1} & 0 \\ 0 & 0 & \sqrt{1 + \beta_{31} X_t^1} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

From the structure of the product $D(X_t)\Lambda_t dt$ it is clear that all drift parameters for the last two factors, X_t^2 and X_t^3, are allowed to vary freely when we move from the Q-dynamics to the P-dynamics detailed above. However, for the first factor with stochastic volatility there is a restriction on the value of θ_1^P given by the equation

$$10^{-6} \cdot \theta_1^Q = \kappa_{11}^P \theta_1^P.$$

By implication, θ_1^P and θ_1^Q cannot both vary freely when we switch from the Q-measure to the P-measure under the essentially affine risk premium structure.

If we use the extended affine risk premium specification, the X_t^1-process has to satisfy the Feller condition under both probability measures, i.e.

$$\kappa_{11}^P \theta_1^P > \frac{1}{2}\sigma_{11}^2 \quad \text{and} \quad 10^{-6} \cdot \theta_1^Q > \frac{1}{2}\sigma_{11}^2.$$

Here, it is obvious that with our requirement of $\kappa_{11}^Q = 10^{-6}$ needed to obtain a level factor structure as similar as possible to the one in the Nelson-Siegel model, the Feller condition for X_t^1 cannot reasonably be expected to be satisfied under the Q-measure as X_t^1 is close to a unit-root process. As a consequence, we limit ourselves to the essentially affine risk premium specification for this model class. Thus, we have to maintain the above restriction on the value of θ_1^P.

The AFNS$_1$ model with stochastic volatility through the curvature factor

In this model class, the Q-dynamics are assumed to be

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$
$$+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1 + \beta_{13} X_t^3} & 0 & 0 \\ 0 & \sqrt{1 + \beta_{23} X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.$$

This structure implies that γ and δ in the system of ODEs provided in Equations (2) and (3) are given by

$$\gamma = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \delta = \begin{pmatrix} 0 & 0 & \beta_{13} \\ 0 & 0 & \beta_{23} \\ 0 & 0 & 1 \end{pmatrix},$$

and $B^1(t,T)$, $B^2(t,T)$, and $B^3(t,T)$ are the unique solutions to the following system of ODEs

$$\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}$$

$$- \frac{1}{2} \sum_{j=1}^{3} \left[\begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{12} & \sigma_{22} & 0 \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} \begin{pmatrix} (B^1)^2 & B^1 B^2 & B^1 B^3 \\ B^1 B^2 & (B^2)^2 & B^2 B^3 \\ B^1 B^3 & B^2 B^3 & (B^3)^2 \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \right]_{j,j} (\delta^j)'.$$

To detail the extended affine risk premium specification for this class of models, start out by defining the matrices $D(X_t)$ and $D^{-1}(X_t)$ by

$$D(X_t) = \begin{pmatrix} \sqrt{1+\beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \quad \text{and} \quad D^{-1}(X_t) = \begin{pmatrix} \frac{1}{\sqrt{1+\beta_{13}X_t^3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{1+\beta_{23}X_t^3}} & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

For this model class the extended affine specification of Γ_t is given by

$$\Gamma_t = D(X_t) \begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \end{pmatrix} + D^{-1}(X_t) \begin{pmatrix} \gamma_{11}^2 & \gamma_{12}^2 & \gamma_{13}^2 \\ \gamma_{21}^2 & \gamma_{22}^2 & \gamma_{23}^2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \gamma_3^3 \end{pmatrix}.$$

By implication, there is a total of 10 free parameters in this specification of Γ_t. Now, the product $D(X_t)\Gamma_t dt$ is given by

$$D(X_t)\Gamma_t dt = \begin{pmatrix} (1+\beta_{13}X_t^3)\gamma_1^1 \\ (1+\beta_{23}X_t^3)\gamma_2^1 \\ X_t^3 \gamma_3^1 \end{pmatrix} dt + \begin{pmatrix} \gamma_{11}^2 X_t^1 + \gamma_{12}^2 X_t^2 + \gamma_{13}^2 X_t^3 \\ \gamma_{21}^2 X_t^1 + \gamma_{22}^2 X_t^2 + \gamma_{23}^2 X_t^3 \\ 0 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 0 \\ \gamma_3^3 \end{pmatrix} dt.$$

By deducting $\Sigma D(X_t)\Gamma_t dt$ from the SDE for the Q-dynamics and replacing dW_t^Q with dW_t^P, we obtain the

71

SDE for the P-dynamics which, without loss of generality, can be written as

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ 0 & 0 & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$

$$
+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1+\beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

From the specification of Γ_t it is clear that if X_t^3 hits the zero boundary, the term $\frac{1}{\sqrt{X_t^3}}$ in the extended affine risk premium specification explodes. In order to keep the model arbitrage-free, this has to be prevented by requiring that the parameters for the X_t^3-process satisfy the Feller condition under both measures, i.e.

$$
\kappa_{33}^P \theta_3^P > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.
$$

The AFNS$_2$ model with stochastic volatility through the level and curvature factor

In this model class, the Q-dynamics are assumed to be

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$

$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

This structure implies that γ and δ in the system of ODEs provided in Equations (2) and (3) are given by

$$
\gamma = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \delta = \begin{pmatrix} 1 & 0 & 0 \\ \beta_{21} & 0 & \beta_{23} \\ 0 & 0 & 1 \end{pmatrix},
$$

and $B^1(t,T)$, $B^2(t,T)$, and $B^3(t,T)$ are the unique solutions to the following system of ODEs

$$
\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}
$$

$$
- \frac{1}{2} \sum_{j=1}^{3} \left[\begin{pmatrix} \sigma_{11} & \sigma_{21} & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & \sigma_{23} & \sigma_{33} \end{pmatrix} \begin{pmatrix} (B^1)^2 & B^1B^2 & B^1B^3 \\ B^1B^2 & (B^2)^2 & B^2B^3 \\ B^1B^3 & B^2B^3 & (B^3)^2 \end{pmatrix} \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix} \right]_{j,j} (\delta^j)'.
$$

To detail the extended affine risk premium specification for this class of models, start by defining the matrices $D(X_t)$ and $D^{-1}(X_t)$ by

$$D(X_t) = \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \quad \text{and} \quad D^{-1}(X_t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{1+\beta_{21}X_t^1+\beta_{23}X_t^3}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix}.$$

The reason why element $(1,1)$ in $D^{-1}(X_t)$ cannot equal $\frac{1}{\sqrt{X_t^1}}$ is that this term is only well defined if X_t^1 is positive $a.s.$ However, with the near unit-root property $\kappa_{11}^Q = 10^{-6}$ imposed, X_t^1 is likely to hit zero under the Q-measure. Hence, the above is the appropriate specification of $D^{-1}(X_t)$ in this model class.

The maximally flexible extended affine specification of Γ_t in this class of models is given by

$$\Gamma_t = D(X_t)\begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \end{pmatrix} + D^{-1}(X_t)\begin{pmatrix} 0 & 0 & 0 \\ \gamma_{21}^2 & \gamma_{22}^2 & \gamma_{23}^2 \\ \gamma_{31}^2 & 0 & 0 \end{pmatrix}\begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix}\begin{pmatrix} 0 \\ 0 \\ \gamma_3^3 \end{pmatrix}.$$

This implies that $D(X_t)\Gamma_t dt$ is given by

$$D(X_t)\Gamma_t dt = \begin{pmatrix} \gamma_1^1 X_t^1 \\ \gamma_2^1(1+\beta_{21}X_t^1+\beta_{23}X_t^3) \\ \gamma_3^1 X_t^3 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \gamma_{21}^2 X_t^1 + \gamma_{22}^2 X_t^2 + \gamma_{23}^2 X_t^3 \\ \gamma_{31}^2 X_t^1 \end{pmatrix} dt + \begin{pmatrix} 0 \\ 0 \\ \gamma_3^3 \end{pmatrix} dt.$$

By deducting $\Sigma D(X_t)\Gamma_t dt$ from the SDE for the Q-dynamics and replacing dW_t^Q with dW_t^P, we obtain the SDE for the P-dynamics which, without loss of generality, can be written as

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & 0 & \kappa_{33}^P \end{pmatrix}\left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix}\right] dt$$

$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ 0 & 0 & \sigma_{33} \end{pmatrix}\begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{1+\beta_{21}X_t^1+\beta_{23}X_t^3} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix}\begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

Note that X_t^1 can have a different rate of mean-reversion under the P-measure relative to that under the Q-measure, but it is not possible to change the constant term through the measure change. Thus, the following equation has to be satisfied

$$10^{-6} \cdot \theta_1^Q = \kappa_{11}^P \theta_1^P.$$

Furthermore, the limited risk premium specification due to the near unit-root property of X_t^1 also implies that X_t^3 cannot impact the drift of X_t^1 ($\kappa_{13}^P = 0$) once κ_{13}^Q has been fixed at 0, which we need to get as close as possible to the desired Nelson-Siegel factor loading structure.[38]

From the specification of Γ_t it is clear that, if X_t^3 hits the zero boundary, the term $\frac{1}{\sqrt{X_t^3}}$ in the extended affine risk premium specification will explode. In order to keep this class of models arbitrage-free, such infinite profit opportunities must be eliminated which is done by requiring that the parameters for the X_t^3-process

---

[38]Note that κ_{12}^P and κ_{32}^P must be zero under all circumstances as the unconstrained process X_t^2 cannot be allowed to impact the drift of any of the two square-root processes.

satisfy the Feller condition under both measures, i.e.[39]

$$\kappa_{31}^P \theta_1^P + \kappa_{33}^P \theta_3^P > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.$$

The AFNS$_2$ model with stochastic volatility through the slope and curvature factor

In this model class, the Q-dynamics are assumed to be

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1+\beta_{12}X_t^2 + \beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.
$$

This structure implies that γ and δ in the system of ODEs provided in Equations (2) and (3) are given by

$$
\gamma = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \delta = \begin{pmatrix} 0 & \beta_{12} & \beta_{13} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},
$$

and $B^1(t,T)$, $B^2(t,T)$, and $B^3(t,T)$ are the unique solutions to the following system of ODEs

$$
\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}
$$
$$
- \frac{1}{2}\sum_{j=1}^3 \left[\begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{12} & \sigma_{22} & 0 \\ \sigma_{13} & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} (B^1)^2 & B^1 B^2 & B^1 B^3 \\ B^1 B^2 & (B^2)^2 & B^2 B^3 \\ B^1 B^3 & B^2 B^3 & (B^3)^2 \end{pmatrix} \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \right]_{j,j} (\delta^j)'.
$$

To detail the extended affine risk premium specification for this class of models start by defining the matrices $D(X_t)$ and $D^{-1}(X_t)$ by

$$
D(X_t) = \begin{pmatrix} \sqrt{1+\beta_{12}X_t^2+\beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \quad \text{and} \quad D^{-1}(X_t) = \begin{pmatrix} \frac{1}{\sqrt{1+\beta_{12}X_t^2+\beta_{13}X_t^3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{X_t^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix}.
$$

The maximally flexible extended affine specification of Γ_t in this class of models is given by

$$
\Gamma_t = D(X_t) \begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \end{pmatrix} + D^{-1}(X_t) \begin{pmatrix} \gamma_{11}^2 & \gamma_{12}^2 & \gamma_{13}^2 \\ 0 & 0 & \gamma_{23}^2 \\ 0 & \gamma_{32}^2 & 0 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{X_t^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix} \begin{pmatrix} 0 \\ \gamma_2^3 \\ \gamma_3^3 \end{pmatrix}.
$$

---

[39]For X_t^1 we just need to ensure that the process does not turn negative. This is assured provided that $10^{-6} \cdot \theta_1^Q > 0$ and $\kappa_{11}^P \theta_1^P > 0$.

This implies that $D(X_t)\Gamma_t dt$ is given by

$$D(X_t)\Gamma_t dt = \begin{pmatrix} \gamma_1^1(1 + \beta_{12}X_t^2 + \beta_{13}X_t^3) \\ \gamma_2^1 X_t^2 \\ \gamma_3^1 X_t^3 \end{pmatrix} dt + \begin{pmatrix} \gamma_{11}^2 X_t^1 + \gamma_{12}^2 X_t^2 + \gamma_{13}^2 X_t^3 \\ \gamma_{23}^2 X_t^3 \\ \gamma_{32}^2 X_t^2 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \gamma_2^3 \\ \gamma_3^3 \end{pmatrix} dt.$$

By deducting $\Sigma D(X_t)\Gamma_t dt$ from the SDE for the Q-dynamics and replacing dW_t^Q with dW_t^P, we obtain the SDE for the P-dynamics which, without loss of generality, can be written as

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & \kappa_{12}^P & \kappa_{13}^P \\ 0 & \kappa_{22}^P & \kappa_{23}^P \\ 0 & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$

$$+ \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{1 + \beta_{12}X_t^2 + \beta_{13}X_t^3} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.$$

From the specification of Γ_t it is clear that if either X_t^2 or X_t^3 hits the zero boundary the corresponding terms $\frac{1}{\sqrt{X_t^2}}$ or $\frac{1}{\sqrt{X_t^3}}$ in the extended affine risk premium specification will explode. In order to keep this class of models arbitrage-free such infinite profit opportunities must be eliminated. This is done by requiring that the parameters for the X_t^2- and X_t^3-processes satisfy the Feller condition under both measures, i.e.

$$\kappa_{22}^P\theta_2^P + \kappa_{23}^P\theta_3^P > \frac{1}{2}\sigma_{22}^2 \quad \text{and} \quad \lambda\theta_2^Q - \lambda\theta_3^Q > \frac{1}{2}\sigma_{22}^2,$$

and

$$\kappa_{33}^P\theta_3^P + \kappa_{32}^P\theta_2^P > \frac{1}{2}\sigma_{33}^2 \quad \text{and} \quad \lambda\theta_3^Q > \frac{1}{2}\sigma_{33}^2.$$

Furthermore, to have well-defined processes for X_t^2 and X_t^3, the sign of the effect they have on each other must be positive. Thus, we need to impose the following non-positive boundaries

$$\kappa_{23}^P \leq 0 \quad \text{and} \quad \kappa_{32}^P \leq 0.$$

This implies that the two square-root processes cannot be negatively correlated.

The AFNS$_3$ model with three stochastic volatility factors

In this model class, the Q-dynamics are assumed to be

$$\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ \theta_2^Q \\ \theta_3^Q \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt$$

$$+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,Q} \\ dW_t^{2,Q} \\ dW_t^{3,Q} \end{pmatrix}.$$

This structure implies that γ and δ in the system of ODEs provided in Equations (2) and (3) are given by

$$\gamma = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \delta = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and $B^1(t,T)$, $B^2(t,T)$, and $B^3(t,T)$ are the unique solutions to the following system of ODEs

$$\begin{pmatrix} \frac{dB^1(t,T)}{dt} \\ \frac{dB^2(t,T)}{dt} \\ \frac{dB^3(t,T)}{dt} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 10^{-6} & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & -\lambda & \lambda \end{pmatrix} \begin{pmatrix} B^1(t,T) \\ B^2(t,T) \\ B^3(t,T) \end{pmatrix}$$
$$- \frac{1}{2} \sum_{j=1}^{3} \left[\begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} (B^1)^2 & B^1B^2 & B^1B^3 \\ B^1B^2 & (B^2)^2 & B^2B^3 \\ B^1B^3 & B^2B^3 & (B^3)^2 \end{pmatrix} \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \right]_{j,j} (\delta^j)'.$$

To detail the extended affine risk premium specification for this class of models, start by defining the matrices $D(X_t)$ and $D^{-1}(X_t)$ by

$$D(X_t) = \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \quad \text{and} \quad D^{-1}(X_t) = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{X_t^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix}.$$

Similar to the previous models with stochastic volatility via the level factor, element $(1,1)$ in $D^{-1}(X_t)$ cannot equal $\frac{1}{\sqrt{X_t^1}}$ since this term is only well defined if X_t^1 is positive $a.s.$ and with the near unit-root property imposed via $\kappa_{11}^Q = 10^{-6}$, X_t^1 is likely to hit zero under the Q-measure. Hence, the above is the appropriate specification of $D^{-1}(X_t)$ in this model class.

The maximally flexible extended affine specification of Γ_t in this class of models is given by

$$\Gamma_t = D(X_t) \begin{pmatrix} \gamma_1^1 \\ \gamma_2^1 \\ \gamma_3^1 \end{pmatrix} + D^{-1}(X_t) \begin{pmatrix} 0 & 0 & 0 \\ \gamma_{21}^2 & 0 & \gamma_{23}^2 \\ \gamma_{31}^2 & \gamma_{32}^2 & 0 \end{pmatrix} \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{X_t^2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{X_t^3}} \end{pmatrix} \begin{pmatrix} 0 \\ \gamma_2^3 \\ \gamma_3^3 \end{pmatrix}.$$

This implies that $D(X_t)\Gamma_t dt$ is given by

$$D(X_t)\Gamma_t dt = \begin{pmatrix} \gamma_1^1 X_t^1 \\ \gamma_2^1 X_t^2 \\ \gamma_3^1 X_t^3 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \gamma_{21}^2 X_t^1 + \gamma_{23}^2 X_t^3 \\ \gamma_{31}^2 X_t^1 + \gamma_{32}^2 X_t^2 \end{pmatrix} dt + \begin{pmatrix} 0 \\ \gamma_2^3 \\ \gamma_3^3 \end{pmatrix} dt.$$

By deducting $\Sigma D(X_t)\Gamma_t dt$ from the SDE for the Q-dynamics and replacing dW_t^Q with dW_t^P, we obtain the

SDE for the P-dynamics which, without loss of generality, can be written as

$$
\begin{pmatrix} dX_t^1 \\ dX_t^2 \\ dX_t^3 \end{pmatrix} = \begin{pmatrix} \kappa_{11}^P & 0 & 0 \\ \kappa_{21}^P & \kappa_{22}^P & \kappa_{23}^P \\ \kappa_{31}^P & \kappa_{32}^P & \kappa_{33}^P \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} X_t^1 \\ X_t^2 \\ X_t^3 \end{pmatrix} \right] dt
$$
$$
+ \begin{pmatrix} \sigma_{11} & 0 & 0 \\ 0 & \sigma_{22} & 0 \\ 0 & 0 & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_t^1} & 0 & 0 \\ 0 & \sqrt{X_t^2} & 0 \\ 0 & 0 & \sqrt{X_t^3} \end{pmatrix} \begin{pmatrix} dW_t^{1,P} \\ dW_t^{2,P} \\ dW_t^{3,P} \end{pmatrix}.
$$

Note that X_t^1 can have a different rate of mean-reversion under the P-measure relative to that under the Q-measure, but it is not possible to change the constant term through the measure change. Thus, the following equation has to be satisfied

$$
10^{-6} \cdot \theta_1^Q = \kappa_{11}^P \theta_1^P.
$$

The limited risk premium specification due to the near unit-root property of X_t^1 also implies that X_t^2 and X_t^3 cannot impact the drift of X_t^1 once κ_{12}^Q and κ_{13}^Q have been fixed at 0, which we need to get as close as possible to the desired Nelson-Siegel factor loading structure.

References

Ahn, Dong-Hyun, Robert F. Dittmar, and A. Ronald Gallant, 2002, "Quadratic Term Structure Models: Theory and Evidence," *Review of Financial Studies*, Vol. 15, No. 1, 243-288.

Andersen, T., Bollerslev, T., Diebold, F.X. and Labys, P., 2003, "Modeling and Forecasting Realized Volatility," *Econometrica*, 71, 579-626.

Andersen, Torben G. and Luca Benzoni, 2010, "Do Bonds Span Volatility Risk in the U.S. Treasury Market? A Specification Test for Affine Term Structure Models," *Journal of Finance*, Vol. 65, No. 2, 603-653.

Anderson, Nicola and John Sleath, 1999, "New estimates of the UK real and nominal yield curves," Bank of England Quarterly Bulletin: November 1999, 384-392.

Anderson, Nicola and John Sleath, 2001, "New estimates of the UK real and nominal yield curves," Bank of England Working Paper No. 126.

Cheridito, Patrick, Damir Filipović, and Robert L. Kimmel, 2007, "Market Price of Risk Specifications for Affine Models: Theory and Evidence," *Journal of Financial Economics*, Vol. 83, 123-170.

Christensen, Jens H. E., Francis X. Diebold, and Glenn D. Rudebusch, 2007, "The Affine Arbitrage-free Class of Nelson-Siegel Term Structure Models," Working Paper # 2007-20, Federal Reserve Bank of San Francisco.

Christensen, Jens H. E. and Jose A. Lopez, 2008, "Common Risk Factors in the US Treasury and Corporate Bond Markets: An Arbitrage-free Dynamic Nelson-Siegel Modeling Approach," unpublished working paper, Federal Reserve Bank of San Francisco.

Christensen, Jens H. E., Jose A. Lopez, and Glenn D. Rudebusch, 2009, "Do Central Bank Liquidity Facilities Affect Interbank Lending Rates?," Working Paper #2009-13, Federal Reserve Bank of San Francisco.

Christensen, Jens H. E., Jose A. Lopez, and Glenn D. Rudebusch, 2010, "Inflation Expectations and Risk Premiums in an Arbitrage-Free Model of Nominal and Real Bond Yields," forthcoming *Journal of Money, Credit and Banking*.

Collin-Dufresne, P. and Goldstein, R.S., 2002, "Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility," *Journal of Finance*, 57, 1685-1730.

Collin-Dufresne, P., Goldstein, R.S. and Jones, C.S., 2009, "Can Interest Rate Volatility Be Extracted from the Cross-Section of Bond Yields?" *Journal of Financial Economics*, 94, 47-66.

Dai, Qiang and Kenneth J. Singleton, 2000, "Specification Analysis of Affine Term Structure Models," *Journal of Finance*, Vol. 55, 1943-1978.

Diebold, Francis X. and Canlin Li, 2006, "Forecasting the Term Structure of Government Bond Yields," *Journal of Econometrics*, Vol. 130, 337-364.

Driessen, Joost, 2005, "Is Default Event Risk Priced in Corporate Bonds?," *Review of Financial Studies*, Vol. 18, No. 1, 165-195.

Duffee, Gregory R., 1999, "Estimating the price of default risk, " *Review of Financial Studies*, Vol. 12, 197-226.

Duffee, Gregory R., 2002, "Term Premia and Interest Rate Forecasts in Affine Models," *Journal of Finance*, Vol. 57, 405-443.

Duffee, Gregory R., 2008, "Forecasting with the term structure: the role of no-arbitrage," unpublished working paper, Haas School of Business.

Duffie, Darrell and Rui Kan, 1996, "A Yield-factor Model of Interest Rates," *Mathematical Finance*, Vol. 6, 379-406.

Fama, Eugene F. and Robert R. Bliss, 1987, "The Information in Long-Maturity Forward Rates," *American Economic Review*, Vol. 77, 680-692.

Feldhütter, Peter and David Lando, 2008, "Decomposing Swap Spreads," *Journal of Financial Economics*, Vol. 88, 375-405.

Fisher, Mark and Christian Gilles, 1996, "Term Premia in Exponential-Affine Models of the Term Structure," unpublished manuscript, Board of Governors of the Federal Reserve System.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2007, "The U.S. Treasury Yield Curve: 1961 to the Present," *Journal of Monetary Economics*, Vol. 54, 2291-2304.

Gürkaynak, Refet S., Brian Sack, and Jonathan H. Wright, 2010. "The TIPS Yield Curve and Inflation Compensation." *American Economic Journal: Macroeconomics*, Vol. 2, No. 1, 7092.

Harvey, Andrew C., 1989, "Forecasting, structural time series models and the Kalman filter," Cambridge University Press.

Hautsch, Nikolaus and Yangguoyi Ou, 2009, "Analyzing Interest Rate Risk: Stochastic Volatility in the Term Structure of Government Bond Yields, " Working Paper # 2009-03, Center for Financial Studies, Goethe - Universität Frankfurt.

Jacobs, Kris and Lofti Karoui, 2009. "Conditional Volatility in Affine Term Structure Models: Evidence from Treasury and Swap Markets," *Journal of Financial Economics*, Vol. 91, 288-318.

Kim, Don H. and Athanasios Orphanides, 2005, "Term Structure Estimation with Survey Data on Interest Rate Forecasts," Finance and Economics Discussion Series, No. 48, Board of Governors of the Federal Reserve System.

Koopman, Siem J., Max I. P. Mallee, and Michel van der Wel, 2008, "Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters," forthcoming *Journal of Business and Economic Statistics.*

Litterman, R., and J. A. Scheinkman, 1991, "Common Factors Affecting Bond Returns," *Journal of Fixed Income*, Vol. 1, 62-74.

Nelson, Charles R. and Andrew F. Siegel, 1987, "Parsimonious Modeling of Yield Curves," *Journal of Business*, Vol. 60, 473-489.

Waggoner, Daniel F., 1997, "Spline methods for extracting interest rate curves from coupon bond prices," Working Paper No. 97-10, Federal Reserve Bank of Atlanta.

Williams, David, 1997, *Probability with Martingales*, Cambridge University Press.