

A Modified Arbitrage-Free Nelson Siegel Model: An Alternative Affine Term Structure Model of Interest Rates

Masamitsu Ohnishi

Graduate School of Economics & Center for the Study of Finance and Insurance
Osaka University

Dara Sim

Graduate School of Economics
Osaka University

1 Introduction

Term structure modeling has been a topic of research for many decades. This is because it plays a key role in measuring and managing interest rate risk, as well as in the valuation of fixed-income securities and their derivatives. It is of interest to many economic participants including corporations, insurance companies, pension funds, hedge funds, investment banks, commercial banks, and other financial institutions dealing with fixed-income assets and liabilities, swaps and other exotic interest rate products, credit derivatives, mortgage portfolios with prepayment options, and so on.

There are at least two approaches to modeling dynamic term structure of interest rates—the forward rate approach and the spot rate approach. One of the most popular dynamic term structure models (DTSMs) following the forward rate approach is the Heath-Jarrow-Morton model (also known as the HJM model) by Heath et al. (1992). On the other hand, the DTSMs following the spot rate approach find their history back to the work of Vasicek (1977). Other well-known one-factor affine DTSMs include the CIR model of Cox, Ingersoll, and Ross (1985), and the Hull and White (1990)'s model, an extension of the Vasicek model. Duffie and Kan (1996) generalize these literatures introducing a general form of continuous-time affine arbitrage-free DTSMs, while Dai and Singleton (2000) characterize classes of admissible affine DTSMs. Affine DTSMs are popular among practitioners and academics because of their theoretical tractability. Other classes of DTSMs are also introduced. The quadratic DTSMs, for instance, are introduced in order to ensure that the short rate is always positive.¹

On the hand, the well-known Nelson-Siegel model (later, NS) and its extensions are in fact non-dynamic term structure models. This class of models has been used by central banks for day-to-day yield curve fitting, and by fixed-income portfolio managers to measure interest rate risk as a consequence of non-parallel shifts in the yield curve. Attention has been shifted back to the NS model after Diebold and Li (2006) extended it to incorporate dynamic structure. The Dynamic Nelson-Siegel (later, DNS) model is empirically tractable and performs empirically well. Furthermore, since it inherits the properties of the NS, its factors can be interpreted as level, slope, and curvature of the yield curve. However, theoretically it does not satisfy the arbitrage-free condition. Christensen et al. (2011) solve this theoretical flaw by deriving a new class of models called the Arbitrage-Free Nelson-Siegel (AFNS) model. Since it is in fact belongs to the $A_0(3)$ class of affine DTSMs of Dai and Singleton (2000), it is theoretically rigorous, and since it is of the NS-type, it possesses the same properties as the DNS model.

This paper derives a modified version of the AFNS model in which the level factor in the AFNS model is replaced by a CIR process. There are three motivations for doing so. First of all, for modeling purpose, we sometimes require modeling the long-term yield as a positive process, and it is natural for doing so. Such a motivation appears when we attempt to model

¹See for example Lin and Wu (2010) for a survey on term structure models.

negative correlation between default intensity and interest rate factors in the reduced-form credit risk modeling framework, using AFNS as a term structure of interest rates model. Second, we happen to realize the fact that although changing the dynamic of the first factor would change its loading in the yield function, such loading can be approximated by one for short time-to-maturity, and is a slightly decreasing function of time-to-maturity, and so the resulted model can in fact be considered an approximation to AFNS. Finally, whereas AFNS is a subclass of the $\mathbb{A}_0(3)$ class of affine DTSMs of Dai and Singleton (2000), the modified model (MAFNS) is in fact a subclass of the $\mathbb{A}_1(3)$ class of affine DTSMs. This has important implication because in the same manner that AFNS has the edge, in terms of empirical performance and tractability, over any other models from the same $\mathbb{A}_0(3)$ class thanks to its NS-specified structures, it could happen that MAFNS has the edge over any other models that belong to the $\mathbb{A}_1(3)$ class in that regard.

Empirically, we explore the performance and properties of our model using two sets of interest rates data. One is the US treasury zero-coupon yield data based on Svensson's method, and the other one is the Japan Government Bond (JGB) zero-coupon yield data based on Steeley's method. We compare our proposed model, in terms of in-sample fitting, and out-of-sample forecasting performance, with the independent-factor DNS, and AFNS models. We also examine whether changing the first factor would change the economic interpretation of the latent state variable factors. For the US treasury data, we find that our model outperforms AFNS in terms of in-sample fit, but that is offset by poorer forecasting performance relative to AFNS. However, our model still maintains some good properties, which includes the fact that one of its factors can be interpreted as long-term yields, and one as slope, and that the long-term yield factor has the most persistent dynamic. For JGB data, we do not find significant difference between our MAFNS and AFNS, which means that both models have similar properties and performances. Finally, we give an example of using our model in joint credit risk pricing model within the reduced-form framework.

This paper contributes to the literature by proposing a DTSM that for modeling purpose, can be used as an alternative to the existing AFNS model. With our model as a DTSM of interest rates, we are able to model a joint credit risk pricing model that captures some important stylized facts and fulfills a theoretical requirement. In addition, to our knowledge, this is the first study to use the AFNS model on Japan data, and also the first paper to propose using an NS-type model in credit risk pricing.

The organization of the paper is as following. Section 2 provides an overview of the NS-type models. In particular, the NS, DNS, and AFNS, as well as other related models are briefly reviewed. Section 3 presents our modified AFNS (later, MAFNS) model. Its empirical performance and properties in comparison to the DNS and AFNS models, as well as economic interpretation of its factors, are examined in Section 4. Section 5 illustrates with example how our proposed model can be used in reduced-form credit risk pricing model to capture some important stylized facts. Section 6 concludes the paper. Proof of proposition 1 and other details are illustrated in the appendix.

2 Overview of the Nelson-Siegel type models

The Nelson-Siegel model has been extensively used by central banks and policy makers to estimate zero-coupon yield curve. The fixed-income portfolio manager uses these models to calculate the interest rate risk in order to immunize their portfolios. According to BIS (2005), the countries whose central bank uses NS-type models for zero-coupon yield curve fitting, include Belgium, Finland, France, Germany, Italy, Norway, Spain and Switzerland. Their popularities can probably be attributed to at least 3 important characteristics. First, they are easy to estimate.

Second, the factors in these models have intuitive economic interpretation, namely level, slope, and curvature. Finally, the most important thing is that these models exhibit good empirical performance.

In this section, the NS-type models are briefly reviewed. In the next section, we present a modified version of the AFNS model which is the model proposed in this paper.

The NS model

In the NS model of Nelson and Siegel (1987), the zero-coupon yield with time-to-maturity $\tau \geq 0$ is given by the functional form,

$$y(\tau) = \alpha_1 + \frac{1 - e^{-\lambda\tau}}{\lambda\tau} \alpha_2 + \left[\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right] \alpha_3,$$

which is correspondent to the forward rate of the form

$$f(\tau) = \alpha_1 + e^{-\lambda\tau} \alpha_2 + \lambda\tau e^{-\lambda\tau} \alpha_3.$$

The instantaneous short rate is then $r = f(0) = \alpha_1 + \alpha_2$, where α_1 is interpreted as the level parameter, which represents the long-term yield. This can be easily seen by letting the time-to-maturity τ go to infinity at which the value of the long-term yield is equal to α_1 . On the other hand, α_2 can be interpreted as negative slope of the zero-coupon yield curve, since it is the difference between the short rate and the long term rate ($\alpha_2 = r - \alpha_1$). Positive (resp. negative) slope ($\alpha_2 < 0$ resp. > 0), (i.e., the long term interest rate is higher (resp. lower) than the short term rate) corresponds to the normal (resp. inverted) yield curve which indicates investors' expectation of future economic growth (resp. downturn). Meanwhile, when the slope is zero (the long term rate is equal to the short term rate), the yield curve can be flat or humped, in either case there is high uncertainty about future economic condition. Also, α_3 is interpreted as the curvature of the term structure over the medium term. The term structure is concave when $\alpha_3 > 0$, and convex when $\alpha_3 < 0$. Finally, λ is the shape parameter which determines the speed of decay of the second and third loadings. It also determines at which maturity the factor loading on the curvature parameter reaches its maximum value. The higher the value of λ , the faster the speed of convergence towards zero of the second and third loadings, and thus leaves only the first parameter α_1 to describe the long term rate.

The DNS model

The NS model is static because level, slope, and curvature parameters are not time-varying. Diebold and Li (2006) extend this model by introducing time-varying factors instead of the parameters $\alpha_1, \alpha_2, \alpha_3$. The zero-coupon yield at time t with time-to-maturity τ is then of the functional form:

$$y_t(\tau) = L_t + \frac{1 - e^{-\lambda\tau}}{\lambda\tau} S_t + \left[\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right] C_t$$

where L_t, S_t, C_t can be interpreted as the level, the slope, and the curvature factors and can be assumed to have a Vector Auto-Regressive (VAR) dynamic. λ is a constant parameter. Its value is estimated to range between 0.5 and 1 when the time-to-maturity is measured in year for US data.

The AFNS model

Both NS and DNS show great empirical performance; nevertheless, as shown by Björk and Christensen (1999), there is no affine arbitrage-free model consistent with these models. In other word, they do not satisfy the theoretical arbitrage-free condition.²

Christensen et al. (2011) extend the result of Diebold and Li (2006) by deriving an arbitrage-free NS model (AFNS), where the yield equation takes the form of the DNS model with an additional term they refer to as an arbitrage-free adjustment term. In AFNS, the zero-coupon yield curve equation now becomes

$$y_t(\tau) = X_{1,t} + \frac{1 - e^{-\lambda\tau}}{\lambda\tau} X_{2,t} + \left[\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right] X_{3,t} - \frac{A(\tau)}{\tau},$$

where $-A(\tau)/\tau$ is the deterministic arbitrage-free adjustment term. Note that AFNS differs from DNS only by this term.

They show empirically that this model outperforms the DNS model for out-of-sample forecast. In addition, though the family of AFNS models is just a subfamily of the $\mathbb{A}_0(3)$ models as classified by Dai and Singleton (2000), by imposing parameter restrictions to obtain the structure of the Nelson-Siegel, it is not only empirically tractable, but also provides better out-of-sample performance.

Other NS-type models

The Svensson (1995) yield curve model is an extension of the NS model, and thus can be considered as a family of NS-type. Whereas the original NS model has only one curvature, the Svensson has two curvatures and so it fits the long-term yield better than the NS model. Christensen et al. (2009), in their attempt to derive another arbitrage-free version of Svensson (1995) models, derive a generalized AFNS (AFGNS) model in which two additional factors, one for slope and the other one for curvature, are included. They show empirically that their model provides better in sample fit than any other corresponding dynamic models that do not rule out the arbitrage opportunity.

3 A modified version of the AFNS model

3.1 The modified AFNS model (MAFNS)

In this section, we present a modified model of the AFNS model in which its Gaussian level factor is replaced by a CIR process. Following Christensen et al. (2011), we start from the standard arbitrage-free affine term structure of Duffie and Kan (1996). To represent an affine diffusion process, define a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{Q})$, where the filtration $(\mathcal{F}_t)_{t \geq 0}$ satisfies the usual condition. The short rate is assumed to be an affine function as follow:

$$r_t = \rho_0(t) + \boldsymbol{\rho}_1(t)' \mathbf{X}_t, \quad (3.1.1)$$

where $\rho_0 : [0, T] \rightarrow \mathbb{R}$, and $\boldsymbol{\rho}_1 : [0, T] \rightarrow \mathbb{R}^n$ are bounded continuous functions, and $T \geq 0$ is a finite time horizon. The state variable \mathbf{X} is an n -dimensional Markov process defined on a subset $M \subset \mathbb{R}^n$, satisfying the following Stochastic Differential Equations (SDEs):

$$d\mathbf{X}_t = \mathbf{K}^{\mathbb{Q}}(t)(\boldsymbol{\theta}^{\mathbb{Q}}(t) - \mathbf{X}_t)dt + \boldsymbol{\Sigma}(t)\sqrt{\mathbf{S}(t, \mathbf{X}_t)}d\mathbf{W}_t^{\mathbb{Q}}, \quad (3.1.2)$$

²However, Coroneo, Nyholm, and Vidova-Koleva (2008) show with 95% confident interval that the NS model is compatible with the arbitrage-freeness.

where $\mathbf{K}^{\mathbb{Q}} : [0, T] \rightarrow \mathbb{R}^{n \times n}$, $\boldsymbol{\theta}^{\mathbb{Q}} : [0, T] \rightarrow \mathbb{R}^n$, $\boldsymbol{\Sigma} : [0, T] \rightarrow \mathbb{R}^{n \times n}$ are bounded continuous functions, and $\mathbf{S} : [0, T] \times M \rightarrow \mathbb{R}^{n \times n}$ is diagonal with the i th element as an affine function of \mathbf{X}_t of the form $\delta_0^i(t) + \boldsymbol{\delta}_1^i(t)' \mathbf{X}_t$, $\delta_0^i : [0, T] \rightarrow \mathbb{R}$, $\boldsymbol{\delta}_1^i : [0, T] \rightarrow \mathbb{R}^n$ are bounded, continuous function. $\mathbf{W}_t^{\mathbb{Q}}$ is an n -dimensional standard Brownian Motion in \mathbb{R}^n . From now on, the square root of any diagonal matrices will denote diagonal matrices with square root on every diagonal elements.

To obtain the Nelson-Siegel structure in the sense that the factor loadings are identical to the NS model, Christensen et al. (2011) gives specification of a 3-factor model as follow:

$$r_t = X_{1,t} + X_{2,t}, \quad (3.1.3)$$

$$d \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{Q}} \\ \theta_2^{\mathbb{Q}} \\ \theta_3^{\mathbb{Q}} \end{pmatrix} - \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{pmatrix} \right] dt + \Sigma d \begin{pmatrix} W_{1,t}^{\mathbb{Q}} \\ W_{2,t}^{\mathbb{Q}} \\ W_{3,t}^{\mathbb{Q}} \end{pmatrix}. \quad (3.1.4)$$

This model is just a class of 3-factor Gaussian model. In this paper, we modify this model by replacing the first factor X_1 with a CIR process. In Proposition 1 below, we derive the zero-coupon yield curves of our model.

Proposition 1. (MAFNS)

Assume the short rate is given as in (3.1.3), and $\mathbf{X}_t = (X_{1,t}, X_{2,t}, X_{3,t})$ solve the following SDEs:

$$d \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{pmatrix} = \begin{pmatrix} \kappa_1^{\mathbb{Q}} & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \left[\begin{pmatrix} \theta_1^{\mathbb{Q}} \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} X_{1,t} \\ X_{2,t} \\ X_{3,t} \end{pmatrix} \right] dt + \begin{pmatrix} \sigma_{11} & 0 & 0 \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} \begin{pmatrix} \sqrt{X_{1,t}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} d \begin{pmatrix} W_{1,t}^{\mathbb{Q}} \\ W_{2,t}^{\mathbb{Q}} \\ W_{3,t}^{\mathbb{Q}} \end{pmatrix}, \quad (3.1.5)$$

where $\kappa_1^{\mathbb{Q}} > 0, \theta_1^{\mathbb{Q}} > 0, \lambda > 0$. Then the zero-coupon bond prices are given by

$$P(t, T) = \mathbb{E}^{\mathbb{Q}}[e^{-\int_t^T r_u du} | \mathcal{F}_t] = \exp\{A(T-t) + B_1(T-t)X_{1,t} + B_2(T-t)X_{2,t} + B_3(T-t)X_{3,t}\}, \quad (3.1.6)$$

where

$$B_1(\tau) = -\frac{2(1 - e^{-\eta\tau})}{\eta + \kappa_1^{\mathbb{Q}} + (\eta - \kappa_1^{\mathbb{Q}})e^{-\eta\tau}}, \quad (3.1.7)$$

$$B_2(\tau) = -\frac{1 - e^{-\lambda\tau}}{\lambda}, \quad (3.1.8)$$

$$B_3(\tau) = -\frac{1 - e^{-\lambda\tau}}{\lambda} + \tau e^{-\lambda\tau}, \quad (3.1.9)$$

$$A(\tau) = (\mathbf{K}^{\mathbb{Q}} \boldsymbol{\theta}^{\mathbb{Q}})' \int_0^\tau \mathbf{B}(t) dt + \frac{1}{2} \sum_{j=2}^3 \int_0^\tau (\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{jj} dt, \quad (3.1.10)$$

and the zero-coupon yield curves are

$$y(t, T) = \frac{2(1 - e^{-\eta(T-t)})}{(T-t)(\eta + \kappa_1^{\mathbb{Q}} + (\eta - \kappa_1^{\mathbb{Q}})e^{-\eta(T-t)})} X_{1,t} + \frac{1 - e^{-\lambda(T-t)}}{(T-t)\lambda} X_{2,t} + \left[\frac{1 - e^{-\lambda(T-t)}}{\lambda} - e^{-\lambda(T-t)} \right] X_{3,t} - \frac{A(T-t)}{T-t}, \quad (3.1.11)$$

with $\eta = \sqrt{(\kappa_1^{\mathbb{Q}})^2 + 2\sigma_{11}^2}$.

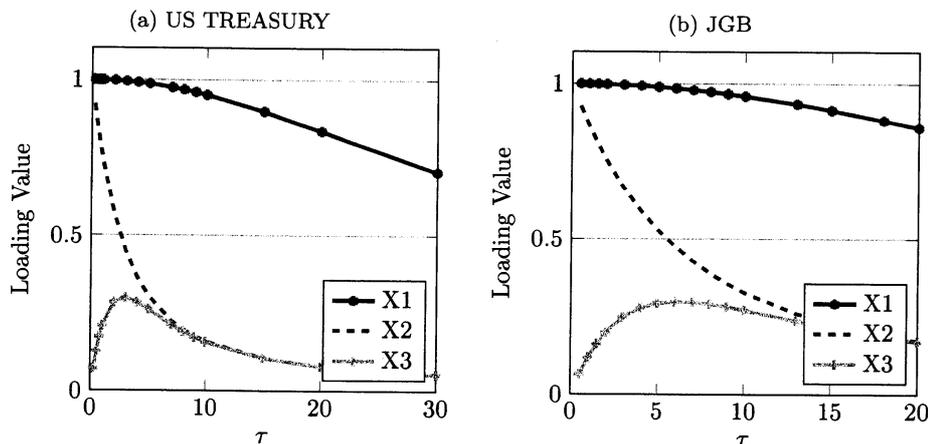
Proof. see Appendix A.1.

As we can see, the factor loadings on the 2nd and 3rd factors are identical to those of the AFNS model, while the loading on the 1st factor is no longer 1. It is equal to

$$\frac{2(1 - e^{-\eta(T-t)})}{(T-t)(\eta + \kappa_1^{\mathbb{Q}} + (\eta - \kappa_1^{\mathbb{Q}})e^{-\eta(T-t)})}. \quad (3.1.12)$$

However, when the time-to-maturity $T - t$ or η is small, (3.1.12) can be approximated by 1. Figure 1 plots the factor loadings of MAFNS for the value of parameter estimates of US treasury and JGB zero-coupon yield data. As we can see, for both sets of data, the 1st factor loading decays slightly from 1. Therefore, MAFNS may no longer be an NS-type model, rather it is an approximation of the NS-type model. Two things can be said about this. First, since it is an approximation, it may still preserve the property that the 1st factor is the long-term yield, the 2nd and 3rd are respectively the slope and the curvature of yield curve. Second, since (3.1.12) is more flexible than 1, it may help improve fitting performance.

Figure 1: Factor Loadings of MAFNS



Also note that there is one factor that enters the volatility structure of (3.1.5). This means that MAFNS is in fact a subclass of the $\mathbb{A}_1(3)$ class of affine DTSMs. In the same manner that AFNS is the NS-specified $\mathbb{A}_0(3)$ class of affine DTSMs, MAFNS can also be considered as the NS-specified $\mathbb{A}_1(3)$ class of affine DTSMs. In Appendix A.2, we show using an invariant transformation that imposing some parameter restrictions on the canonical representation of $\mathbb{A}_1(3)$ will give us the MAFNS model.

3.2 The change of measure

To estimate the model, we also need to know the distribution of the state variables under the real world probability measure \mathbb{P} . The link between the \mathbb{P} - and the \mathbb{Q} - measures is given by the measure change

$$d\mathbf{W}_t^{\mathbb{Q}} = d\mathbf{W}_t^{\mathbb{P}} + \Lambda_t dt,$$

where Λ_t represents the risk premium. Following Duffee (2002), we consider the essentially risk premium specification of the form

$$\Lambda_t = \sqrt{\mathbf{S}_t} \hat{\Gamma}_1 + \sqrt{\mathbf{S}_t^-} \hat{\Gamma}_2 \mathbf{X}_t,$$

where $\hat{\Gamma}_1$ is a 3-dim vector, $\hat{\Gamma}_2$ is a 3×3 matrix, and

$$\mathbf{S}_t^- = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus, we have

$$\begin{aligned} \Sigma \sqrt{\mathbf{S}_t} \Lambda_t &= \begin{pmatrix} \sigma_{11} X_{1,t} & 0 & 0 \\ \sigma_{21} X_{1,t} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} X_{1,t} & \sigma_{32} & \sigma_{33} \end{pmatrix} \hat{\Gamma}_1 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \sigma_{22} & \sigma_{23} \\ 0 & \sigma_{32} & \sigma_{33} \end{pmatrix} \hat{\Gamma}_2 \mathbf{X}_t \\ &= \Gamma_0 + \Gamma_1 \mathbf{X}_t, \end{aligned}$$

where Γ_0 is a 3-dim vector, and Γ_1 is a 3×3 matrix.

Under this specification, the dynamic under \mathbb{P} of the state variables \mathbf{X}_t is still affine of the form

$$d\mathbf{X}_t = \mathbf{K}^{\mathbb{P}} (\boldsymbol{\theta}^{\mathbb{P}} - \mathbf{X}_t) dt + \Sigma \sqrt{\mathbf{S}_t} d\mathbf{W}_t^{\mathbb{P}}, \quad (3.2.1)$$

with $\mathbf{K}^{\mathbb{P}} = \mathbf{K}^{\mathbb{Q}} - \Gamma_1$, and $\mathbf{K}^{\mathbb{P}} \boldsymbol{\theta}^{\mathbb{P}} = \mathbf{K}^{\mathbb{Q}} \boldsymbol{\theta}^{\mathbb{Q}} + \Gamma_0$.

3.3 The independent-factor model

In this study, we focus on the independent-factor models because they are easy to estimate and are proven to have outperformed correlated-factor models. In analogy with the independent-factor DNS and AFNS models, we specify the MAFNS model so that the factors are independent under \mathbb{P} -measure, that is, $\mathbf{K}^{\mathbb{P}}$, and Σ are specified as a diagonal matrix as follow:

$$\mathbf{K}^{\mathbb{P}} = \begin{pmatrix} \kappa_1^{\mathbb{P}} & 0 & 0 \\ 0 & \kappa_2^{\mathbb{P}} & 0 \\ 0 & 0 & \kappa_3^{\mathbb{P}} \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix},$$

where we assume $\mathbf{K}^{\mathbb{P}} > \mathbf{0}$, so that the process is stationary.

Under this specification, the deterministic term is given by

$$\begin{aligned} A(\tau) &= \sigma_2^2 \left[\frac{1}{2\lambda^2} - \frac{1}{\lambda^3} \frac{1 - e^{-\lambda\tau}}{\tau} + \frac{1}{4\lambda^3} \frac{1 - e^{-2\lambda\tau}}{\tau} \right] \\ &+ \sigma_3^2 \left[\frac{1}{2\lambda^2} + \frac{1}{\lambda^2} - \frac{1}{4\lambda} \tau e^{-2\lambda\tau} - \frac{3}{4\lambda^2} e^{-2\lambda\tau} - \frac{2}{\lambda^3} (1 - e^{-\lambda\tau}) + \frac{5}{8\lambda^2} \frac{1 - e^{-2\lambda\tau}}{\tau} \right] \\ &- \kappa_1^{\mathbb{Q}} \theta_1^{\mathbb{Q}} \left[\frac{2}{\tau \sigma_1^2} \log \left(\frac{\kappa_1^{\mathbb{Q}} + \eta + (\eta - \kappa_1^{\mathbb{Q}}) e^{-\eta\tau}}{2\eta} \right) + \frac{2}{\eta + \kappa_1^{\mathbb{Q}}} \right]. \end{aligned} \quad (3.3.1)$$

Unless mentioning, from now on we will speak of AFNS and MAFNS when referring to their independent-factor specifications. Since the conditional and unconditional expectation and variance of the state variables under \mathbb{P} -measure are useful for the estimation process, we state these in Appendix A.3 for easy reference.³

³For a more general affine process, see for example, Duan and Simonato (1999) on how to calculate conditional mean and variance.

4 Empirical analysis

4.1 Data

In this paper, we estimate the models using interest rates data of the US and Japan. For US data, we use monthly Treasury zero-coupon yield covering the period from January 1987 to December 2002. This data is constructed by Gurkayna, Sack and Wright (2006) based on the Svensson (1995)'s model, and is obtained from the Federal Reserve Board website.⁴ The zero-coupon yields to be included in our estimation are those at maturity spectrum of 3/12, 6/12, 9/12, 1, 2, 3, 4, 5, 7, 8, 9, 10, 15, 20, 30 years. The data contains zero-coupon yields with time to maturity ranging from 1 year to 30 year as well as the Svensson parameter estimates. Using these parameter estimates, we are able to calculate the 3-month, 6-month, and 9-month zero-coupon yields. The data is available in daily frequency. Choosing estimates on the last day of month as the monthly estimates, we obtain a total of 192 observations for monthly data sample. We also note that the sample period we use is the same as in the previous studies.

For Japan data, we use JGB zero-coupon yields estimated using the method proposed by Steeley (1991). Kikuchi and Shintani (2012) show that the Steeley's method is the best among popular yield-curve estimation methods based on some considered criteria.⁵ The daily data based on this method is downloaded from Bank of Japan (BoJ)'s website. We convert daily data to weekly data by choosing estimates on Thursday (if available, and on Wednesday if not) as the weekly estimates. We skip those weeks where both the estimates on Thursday and Wednesday are not available. The time-to-maturity spectrum included in our study are 0.5, 1, 1.5, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 15, 18, and 20 year, and the sample consists of 692 weekly observations covering the period from January 1999 to December 2012.

Figure 2 are plots of term structure of US treasury and JGB zero-coupon yields. Movement of the term structure of JGB yields is subject to zero-interest rate policy (ZIRP) started from February 1999. It was lifted temporarily during the ICT bubble in 2000, and restarted after the burst of the ICT bubble in 2001. Under this monetary easing, the long-term rate began to decrease to 0.43%, the lowest rate ever recorded in the history of the world, in 2003. The ZIRP was then lifted again in 2006 at the time of economic recovery, and restarted in December 2008 at the time it was introduced in the US, during the world financial crisis.

4.2 Estimation method

In this paper, the models are estimated using maximum likelihood estimation method based on the Kalman filter. The MAFNS model can be formulated in a state space form as follow.

System Equations:

$$\mathbf{X}_t = (\mathbf{I} - e^{-\mathbf{K}^p \Delta t}) \boldsymbol{\theta}^p + e^{-\mathbf{K}^p \Delta t} \mathbf{X}_{t-\Delta t} + \mathbf{u}_t. \quad (4.2.1)$$

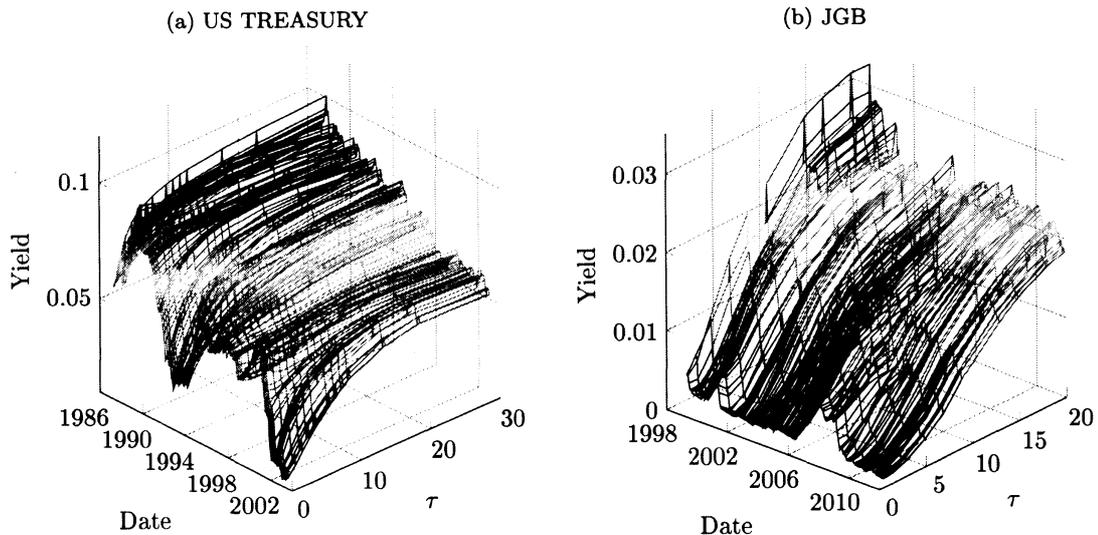
Measurement Equations:

$$\begin{bmatrix} y_t(\tau_1) \\ \cdot \\ \cdot \\ \cdot \\ y_t(\tau_N) \end{bmatrix} = \begin{bmatrix} -\frac{A(\tau_1)}{\tau_1} \\ \cdot \\ \cdot \\ \cdot \\ -\frac{A(\tau_N)}{\tau_N} \end{bmatrix} + \begin{bmatrix} -\frac{B_1(\tau_1)}{\tau_1} & -\frac{B_2(\tau_1)}{\tau_1} & -\frac{B_3(\tau_1)}{\tau_1} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ -\frac{B_1(\tau_N)}{\tau_N} & -\frac{B_2(\tau_N)}{\tau_N} & -\frac{B_3(\tau_N)}{\tau_N} \end{bmatrix} \mathbf{X}_t + \boldsymbol{\epsilon}_t. \quad (4.2.2)$$

⁴The data is updated periodically, and is available at <http://www.federalreserve.gov/pubs/feds/2006>.

⁵The criteria considered are 1) Non-negativity of interest rates 2) Ruling out abnormal values, 3) Goodness-of-fit to the market prices, and 4) Little unevenness in zero-coupon yield curves.

Figure 2: Term Structure of Zero-coupon Yields



where $B_1(\tau), B_2(\tau), B_3(\tau), A(\tau)$ are defined as in (3.1.7) - (3.1.10), Δt denotes infinitesimal time interval, and the variance of the measurement errors $Var(\epsilon_t) = \mathbf{H}$. In this paper, we assume that \mathbf{H} is a diagonal matrix, and the filtering algorithm is initialized with the unconditional expectation and variance. Note that the CIR process X_1 is approximated by a Gaussian process as in (4.2.1). This means that X_1 could turn out to be negative during the estimation process. When that happens, we put it equal to zero, and let the filtering algorithm continues.

4.3 Empirical results

We argue that our proposed model is not only closed related to DNS and AFNS, but can also be used as an alternative model. To justify this argument, we conduct empirical analysis by estimating these models with the term structure of US Treasury and JGB zero-coupon yield data, and then compare, in terms of performance and properties, our MAFNS model with the DNS and AFNS models. As mentioned earlier, we only focus on the independent-factor models. In-sample and out-of-sample performances are examined, and then to see whether modifying the first factor as in MAFNS would change the economic interpretation of the model's state variables, we calculate correlations of the estimates of these factors with long-term rates, empirical slopes, and empirical curvatures. First, the results for US treasury yields, and then the results for JGB yields are presented subsequently.

4.3.1 Results for US treasury yield data

The parameter estimates and log likelihood values are reported in Table 1 for the three models. The "Estimates" columns show parameter estimates, and the "Std. Err." columns show standard errors. The estimated value of the decay parameter λ is 0.4378 for DNS, and 0.361 for AFNS model. These values are lower than values estimated in previous studies, which range between 0.5 and 1; however, it is not surprising because in our study, Svesson's method is used as yield curve fitting method, whereas in previous studies Unsmoothed Fama Bliss method is used. We may expect that smaller value for the decay rate results in better fit for long-term yield than short-term yield. For MAFNS model, λ is estimated to be 0.6282.

Table 1: Parameter Estimates and Standard Errors (US TREASURY)

DNS			AFNS			MAFNS		
	Estimates	Std. Err.		Estimates	Std. Err.		Estimates	Std. Err.
μ_L	0.0758	0.0092	θ_1	0.0816	0.0077	θ_1	0.0517	0.0362
μ_S	-0.0306	0.0215	θ_2	-0.0353	0.0218	θ_2	-0.0110	0.0191
μ_C	-0.0093	0.0040	θ_3	-0.0148	0.0041	θ_3	0.0010	0.0054
a_1	0.9750	0.0155	κ_1	0.1201	0.0636	κ_1	0.0690	0.0482
a_2	0.9836	0.0136	κ_2	0.2458	0.1993	κ_2	0.2348	0.1812
a_3	0.8218	0.0444	κ_3	1.2519	0.3302	κ_3	1.3432	0.4096
q_1	0.0027	0.0002	σ_1	0.0054	0.0001	σ_1	0.0381	0.0021
q_2	0.0035	0.0002	σ_2	0.0142	0.0009	σ_2	0.0122	0.0006
q_3	0.0083	0.0005	σ_3	0.0233	0.0008	σ_3	0.0269	0.0012
λ	0.4378	0.0045	λ	0.3610	0.0037	λ	0.6282	0.0046
						κ_1^Q	0.0548	0.0030

Notes: The log likelihood value of each model is respectively 16778, 16576, and 17279, and the BIC is -33425, -33022, and -34421 respectively. Since lower value of BIC suggests better fit, based on this criterion, MAFNS has the best fit to the data.

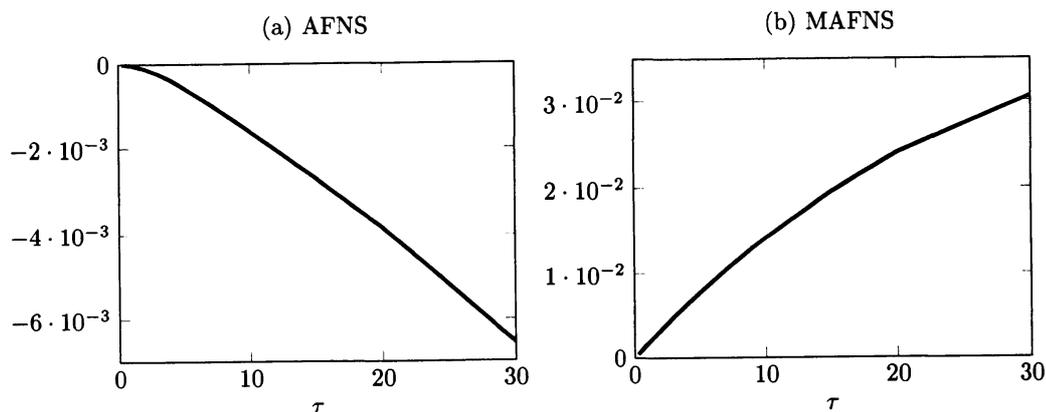
The estimate of the long run mean for DNS's level factor is 0.0758, which is slightly higher than the sample average of long-term yields: 0.069, 0.0715, 0.0723, and 0.0715 for the 10, 15, 20, and 30-year yields respectively. On the other hand, the estimate of the long-run mean for AFNS's X_1 factor is 0.0816, a magnitude higher than the sample average. This result may look skeptical at first; however, recall that unlike in DNS where the level factor X_1 alone represents the long-term yield, in AFNS, the long-term yield is represented by X_1 plus a deterministic adjustment term. This deterministic adjustment term is a downward function of time-to-maturity so that the average of the long-term yield estimated by the model is about the sample average. This is confirmed in Figure 3 (a) which shows a plot of AFNS's deterministic yield adjustment term as a downward function of time-to maturity. The same thing can be said about the result of MAFNS with the estimate of the long-run mean of X_1 of 0.0517. In MAFNS, the long-term yield is a linear transformation of the factor X_1 , with the loading as coefficient and the deterministic adjustment term as the intercept. Figure 3 (b) shows plot of MAFNS's deterministic adjustment term as upward function of time-to-maturity. In fact, the factor X_1 in both AFNS and MAFNS models can be considered as the long-term factor because as will be seen later, it captures the movement of long-term yields.

It is a stylized fact for yield curve dynamics that long-term yields are more persistent than short-term yields. To see how persistent our considered models' factors are, we examine the parameters $[a_1, a_2, a_3]$ for DNS, and the parameters $e^{-\kappa^P \Delta t}$ for AFNS and MAFNS. For DNS, these parameters are estimated to be $[0.975, 0.9836, 0.8218]$, suggesting that the most persistent is the slope factor, then the level factor, and the curvature is the least persistent. For AFNS, the estimates of $e^{-\kappa^P \Delta t}$ is $diag([0.9900, 0.9797, 0.9009])$,⁶ which means that the long-term factor is the most persistent, while the slope is less persistent. Finally, in MAFNS the long-term yield has the most persistent dynamic since estimates of $e^{-\kappa^P \Delta t}$ in our model is $diag([0.9943, 0.9806, 0.8941])$. From this result, we see that our model does capture this stylized fact of the term structure of interest rates.

In-sample fitting performance

⁶Where $diag(\nu)$ denotes a diagonal matrix with a diagonal vector ν .

Figure 3: Deterministic Adjustment Term (US TREASURY)



The log likelihood value of DNS is 16778 versus 16576 of AFNS. Since both models have the same number of parameters, this suggests that DNS better fitted to the data than AFNS. This is in line with Christensen et al.(2011). On the other hand, MAFNS has the highest likelihood value of all, but it has one parameter more than the other two models. However, the value of BIC, which places penalty on model complexity, suggests that MAFNS has the best fit to the data among the three models. That the overall performance of MAFNS is better than DNS, and AFNS can also be confirmed in Table 2, which reports value of RMSE in basis points for different time-to-maturities and the ratios of each model's RMSE to AFNS's RMSE, and in Figure 4, which shows plots of RMSE versus time-to-maturity for the three models. From this table, we also notice that MAFNS appears to fit better than AFNS for short- and medium-term yields. In addition, Figure 5 shows plots of sample average yield curve, and the models' fit average yield curves. As we can see, on average, these models appear to provide good fit especially for short and medium term. However, DNS does seem to have relatively hard time fitting the very long-term yield compared to AFNS and MAFNS. This may be because DNS does not have yield adjustment term. On the hand, MAFNS has the best fit for long-term yields on average.

Figure 4: RMSE (US TREASURY)

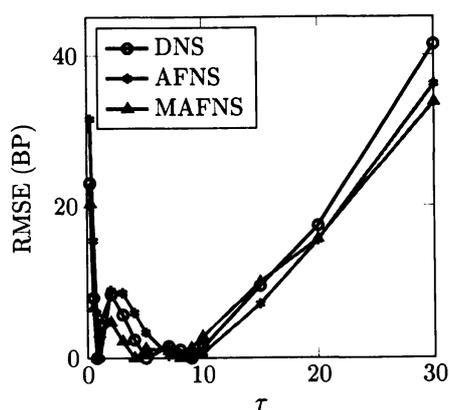


Figure 5: Average Yield Curve (US TREASURY)

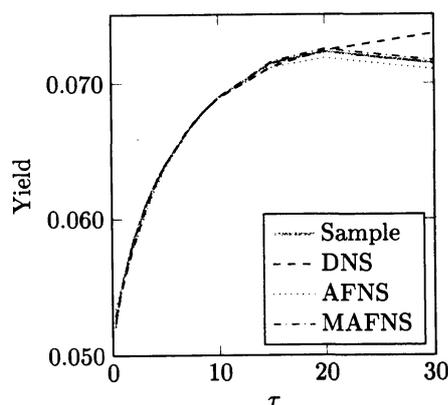


Table 2: RMSE (US TREASURY)

τ	DNS		AFNS		MAFNS	
	RMSE	DNS/AFNS	RMSE	AFNS/AFNS	RMSE	MAFNS/AFNS
3	23.06	0.73	31.67	1.00	20.29	0.64
6	7.99	0.52	15.48	1.00	6.63	0.43
9	0.00	0.00	6.01	1.00	0.00	0.00
12	4.45	194789.63	0.00	1.00	3.25	141926.90
24	8.45	0.95	8.93	1.00	4.74	0.53
36	5.74	0.67	8.60	1.00	2.21	0.26
48	2.42	0.40	6.00	1.00	0.00	0.00
60	0.00	0.00	3.38	1.00	1.22	0.36
84	1.54	4.93	0.31	1.00	1.01	3.21
96	1.04	5.31	0.20	1.00	0.00	0.00
108	0.00	0.27	0.00	1.00	1.33	5654.06
120	1.38	2.19	0.63	1.00	2.83	4.48
180	9.53	1.33	7.15	1.00	10.03	1.40
240	17.44	1.12	15.53	1.00	15.67	1.01
360	41.43	1.14	36.18	1.00	33.81	0.93

Notes: "RMSE" columns report root mean square errors (RMSE) in basis points for different time-to-maturity for each model. "DNS/AFNS", "AFNS/AFNS", and "MAFNS/AFNS" columns respectively report ratios of DNS's, AFNS's, and MAFNS's RMSE to AFNS's RMSE.

Out-of-sample forecasting performance

To examine the out-of-sample forecasting performance of our model as compared to the related models, we calculate 1-month-, 3-month-, 6-month-ahead forecasts. The forecasting window is the period from Sept. 1999 to Dec. 2002 having respectively 42, 40, and 37 observations for 1-month-, 3-month-, and 6-month-ahead forecasts. The estimation window starts from Jan. 1987, and we adopt the expanding estimation window approach, that is, to make forecasts at a time point, we use data starting from Jan. 1987 up to that time. Table 3 reports root mean squared forecast errors (RMSFE) of AFNS and ratios of DNS's and MAFNS's RMSFE to AFNS's RMSFE. Figure 6 are plots of RMSFE for different forecasting horizons. The result is obvious. For short and medium terms, the three models' performance is not much different from each other. However, it turns out that AFNS has the best performance among the three models for forecasting long-term yields. The result for DNS and AFNS is consistent with Christensen et al. (2011) who show that the independent-factor AFNS outperforms the independent-factor DNS for out-of-sample forecasts. That our model performs the worst for forecasting long-term yields might have been due to over-fitting as a result of introducing flexibility to the loading factor of X_1 and the deterministic adjustment term. As we can see in Figure 3, the magnitude of the MAFNS's deterministic adjustment term is so high compared to the AFNS's.

Economic interpretation of the latent variables

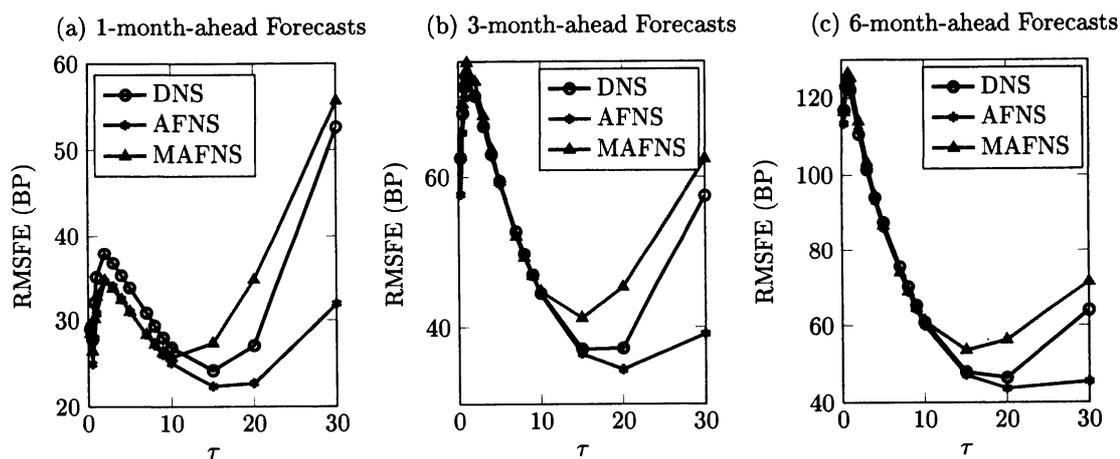
Finally, we want to know how modifying the loading factor of AFNS affects the economic interpretation of the state variables, namely level, slope, and curvature. To see this, we calculate correlations between the estimates of state variables and the empirical levels, slopes, and curvatures. Table 4 reports this result for the three models. Panel A reports correlations between

Table 3: RMSFE (US TREASURY)

τ	DNS/AFNS			AFNS			MANFS/AFNS		
	1-m	3-m	6-m	1-m	3-m	6-m	1-m	3-m	6-m
0.25	0.98	1.08	1.03	29.46	57.72	113.35	0.97	1.09	1.03
0.5	1.11	1.04	1.01	25.00	65.74	122.31	1.06	1.06	1.02
0.75	1.14	1.02	0.99	28.00	70.26	124.54	1.08	1.05	1.01
1	1.14	1.01	0.99	30.78	72.10	123.43	1.06	1.04	1.02
2	1.09	1.00	0.99	34.57	70.54	111.72	1.00	1.03	1.02
3	1.08	1.00	1.00	33.93	66.54	101.45	0.99	1.02	1.02
4	1.08	1.00	1.01	32.45	62.65	93.25	1.00	1.01	1.01
5	1.09	1.01	1.02	31.01	59.04	86.20	1.00	1.01	1.01
7	1.09	1.01	1.02	28.40	52.42	74.08	1.00	0.99	1.00
8	1.08	1.01	1.02	27.19	49.42	68.84	1.00	0.99	1.00
9	1.08	1.00	1.02	26.08	46.68	64.15	1.01	1.00	1.01
10	1.07	1.00	1.02	25.09	44.23	60.02	1.03	1.01	1.02
15	1.08	1.02	1.02	22.39	36.57	47.25	1.22	1.13	1.14
20	1.19	1.08	1.07	22.75	34.57	43.71	1.52	1.31	1.29
30	1.65	1.47	1.40	31.86	39.16	45.55	1.75	1.59	1.57

Notes: “ i -m” is short form for i -month. Column 5 to 7 are 1-month-, 3-month-, and 6-month-ahead root mean squared forecast errors (RMSFE) of AFNS in basis points. Column 2 to 4 are ratios of DNS’s RMSFE to AFNS’s RMSFE. Column 8 to 10 are ratios of MAFNS’s RMSFE to AFNS’s RMSFE.

Figure 6: Out-of-Sample Forecasting Performance (US TREASURY)



the estimates of X_1 of each model and long-term (10-year, 15-year, 20-year, and 30-year) yields. Panel B reports correlations between the estimates of X_2 of each model and negative empirical slopes defined as the difference between a short and a long-term (10-year, 15-year, 20-year, and 30-year) yield with 0.25-year yield chosen as the short-term yield. Finally, Panel C reports correlations between the estimates of X_3 and empirical curvatures defined as 2 times medium-term yield (4-year yield) minus a short-term (1-year) yield and minus a long-term (10-year, 15-year, 20-year, and 30-year) yield. As expected, for DNS and AFNS, X_1 , X_2 , and X_3 are respectively highly correlated with empirical levels, slopes, and curvatures. On the other hand, for MAFNS, X_1 and X_2 respectively have high correlations with long-term yields and empirical slopes, while X_3 appears to have weak correlations with empirical curvatures; the highest correlation is only about 73%. This suggests that the factor X_3 in the MAFNS may lose its interpretation as curvature.

To sum up, in this section we have examined the performance and properties of our proposed model relative to the related models when used as a term structure of US treasury yields. Our model appeared to have better fit to the data than the related models, but this was offset by poorer forecasting performance especially for very long-term yields as compared to the related models. However, our model to some extent inherits some good properties from its original AFNS model. This includes the fact that its long-term yield factor has strong persistent dynamic, and two of its factors can be interpreted as level, and slope. In Section 5 below, we will show by using an example that our model can be used in place of AFNS for modeling purpose. For now, we examine our model's performance and properties when used to describe the term structure of JGB yields.

Table 4: Correlations between the state variables and economic variables (US TREASURY)

Panel A: Correlations between X_1 and long-term yields				
	10y	15y	20y	30y
X_1^{DNS}	0.9020	0.9423	0.9512	0.9297
X_1^{AFNS}	0.8887	0.9344	0.9449	0.9155
X_1^{MAFNS}	0.9384	0.9669	0.9719	0.9557

Panel B: Correlations between X_2 and negative empirical slopes				
	[0.25y-10y]	[0.25y-15y]	[0.25y-20y]	[0.25y-30y]
X_2^{DNS}	0.9649	0.9792	0.9805	0.9657
X_2^{AFNS}	0.9438	0.9655	0.9699	0.9541
X_2^{MAFNS}	0.9858	0.9878	0.9839	0.9703

Panel C: Correlations between X_3 and empirical curvatures				
	$[2*(4y)-1y-10y]$	$[2*(4y)-1y-15y]$	$[2*(4y)-1y-20y]$	$[2*(4y)-1y-30y]$
X_3^{DNS}	0.8325	0.9688	0.9395	0.7750
X_3^{AFNS}	0.9279	0.9707	0.9151	0.6907
X_3^{MAFNS}	0.4505	0.7141	0.7364	0.6701

Notes: Panel A reports correlations between the estimated path of the factor X_1 of each model and long-term yields (10-year, 15-year, 20-year, and 30-year yields). Panel B reports correlations between the estimated path of X_2 of each model and negative empirical slopes defined as the difference between a short and a long-term (10-year, 15-year, 20-year, and 30-year) yield with 0.25-year yield chosen as the short-term yield. Finally, Panel C reports correlations between the estimates of X_3 and empirical curvatures defined as 2 times medium-term yield (4-year yield) minus a short-term (1-year) yield and a long-term yield.

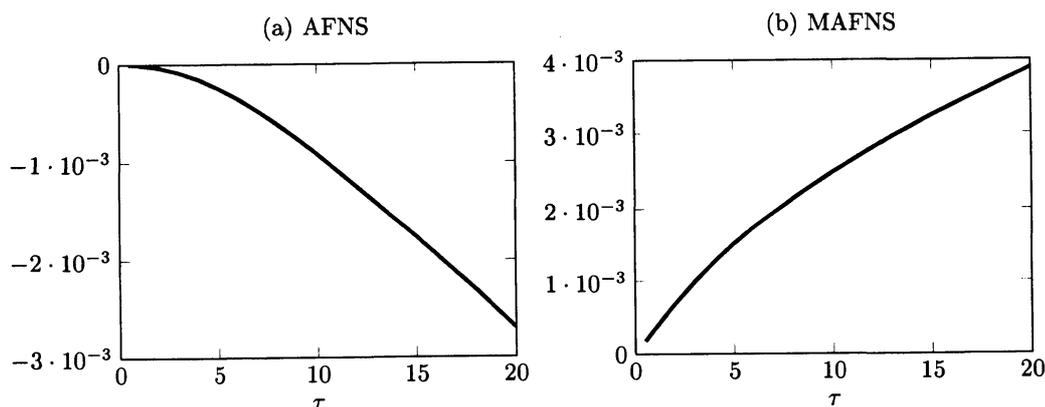
4.3.2 Results for JGB yield data

In this section, we examine the performance and properties of these models for the case of JGB zero-coupon yield data. It is important to note that, to our knowledge, our study is the first to estimate JGB zero-coupon yields using the AFNS model. Though DNS is sometimes seen used as a term structure model of JGB zero-coupon yields in previous literatures, this type of models is not popular among academic literatures because of at least two basic reasons. First, they are Gaussian models which do not rule out the negativity of the interest rates. This is of great concern especially for Japan whose standard short-term rate is near zero, as a result of which the estimates from these models are likely to be negative. The second reason may be the fact that these models do not incorporate regime switch which is often observed in JGB interest rates subject to ZIRP. Nevertheless, for the sake of reference, we notice the need for writing down a few pages in the literature regarding the results of using AFNS and the related models to describe the term structure of JGB zero-coupon yields.

The parameter estimates and standard errors are reported in Table 5. For DNS, the decay rate parameter λ is estimated to be equal to 0.3523, a magnitude that is often seen adopted in usual practice.⁷ Meanwhile, the estimates of λ for AFNS and MAFNS are 0.3072 and 0.2878 respectively. Furthermore, the estimates of the parameters (a_1, a_2, a_3) of DNS are equal to $[0.9912, 0.9882, 0.9828]$, and the estimates of $\exp(-\kappa^P \Delta t)$ of AFNS is $diag([0.9881, 0.9887, 0.9807])$ and of MAFNS is $diag([0.9995, 0.9914, 0.9825])$. This suggests that the dynamics of the three factors in the three models are highly persistent.

Figure 7 plots the deterministic adjustment term of AFNS and MAFNS. Just as in the case of US treasury data, the AFNS's deterministic adjustment term is a downward function of time-to-maturity, whereas the MAFNS's is an upward function of time-to-maturity. However, unlike in the case of US treasury data, the magnitude of the MAFNS's adjustment term is not much different from the AFNS's. This could have important consequence. As we have seen in the case US treasury, MAFNS have poor forecasting performance for long-term yields relative to AFNS. This may be due to large adjustment term estimated by MAFNS. With small adjustment just as in JGB case, we may expect that the MAFNS's performance and properties do not deviate much from the AFNS's. In fact, this is the case as we examine the results subsequently below.

Figure 7: Deterministic Adjustment Term (JGB)



In-sample fitting performance

⁷As usual practice, the decay rate parameter λ is sometimes assumed constant and fixed at about 0.36.

Table 5: Parameter Estimates and Standard Errors (JGB)

DNS			AFNS			MAFNS		
	Estimates	Std. Err.		Estimates	Std. Err.		Estimates	Std. Err.
μ_L	0.0303	0.0051	θ_1	0.0355	0.0049	θ_1	0.0310	0.0680
μ_S	-0.0270	0.0057	θ_2	-0.0325	0.0057	θ_2	-0.0309	0.0044
μ_C	-0.0315	0.0055	θ_3	-0.0359	0.0061	θ_3	-0.0361	0.0075
a_1	0.9912	0.0068	κ_1	0.6216	0.4412	κ_1	0.0237	0.0515
a_2	0.9882	0.0089	κ_2	0.5900	0.4295	κ_2	0.4484	0.2099
a_3	0.9829	0.0075	κ_3	1.0158	0.4719	κ_3	0.9159	0.4367
q_1	0.0006	0.0000	σ_1	0.0050	0.0002	σ_1	0.0273	0.0020
q_2	0.0008	0.0000	σ_2	0.0056	0.0002	σ_2	0.0057	0.0002
q_3	0.0020	0.0001	σ_3	0.0177	0.0006	σ_3	0.0181	0.0007
λ	0.3523	0.0014	λ	0.3072	0.0022	λ	0.2878	0.0026
						κ_1^Q	0.0578	0.0027

Notes: The log likelihood value of each model is respectively 67903, 68055, and 68133, and the BIC is -135636, -135940, and -136091 respectively. Since lower value of BIC suggests better fit, based on this criterion, MAFNS has the best fit to the data.

The MAFNS model has the smallest BIC, -136091 versus -135940 of the AFNS, and -135636 of the DNS. Thus, MAFNS fits the best to the data. This is consistent with RMSE values reported in Table 6, and plots of RMSE in Figure 8. The differences between the three though are not so obvious. Moreover, the three models fit well for medium and long-term yields, but happen to have hard time fitting 0.5-year yield. This may be because of the typical S-shape of JGB yield curves. The sample mean of 0.5-year yields is 0.32% larger than 0.28% and 0.31%, the sample mean of 1-year and 1.5-year yields respectively. Average yield curves of sample data, and of the models' fitted values are plotted in Figure 9. Although during the sample period, JGB yield curves may assume many different shapes, we can see the sample average yield curve takes an S-shape, and we see that the three models are not able to fit the very short end.

Figure 8: RMSE (JGB)

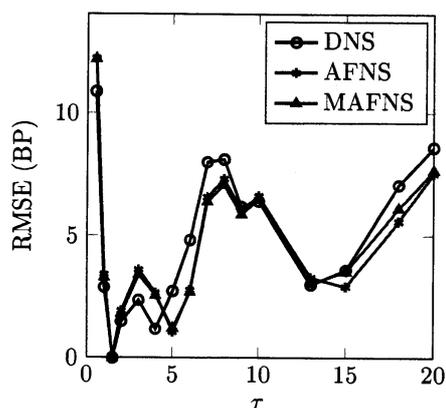
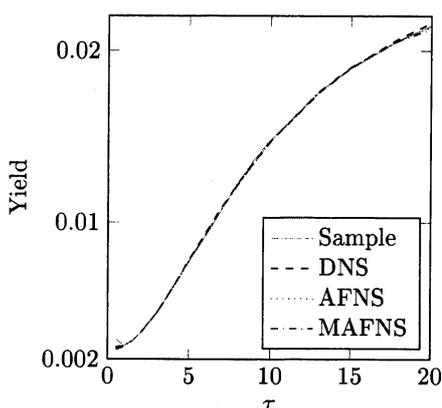


Figure 9: Average Yield Curve (JGB)



Out-of-sample forecasting performance

To make out-of-sample comparison, we calculate forecasts for 4-week, 12-week, and 26-

Table 6: RMSE (JGB)

τ	DNS		AFNS		MAFNS	
	RMSE	DNS/AFNS	RMSE	AFNS/AFNS	RMSE	MAFNS/AFNS
6	10.88	0.89	12.22	1.00	12.20	1.00
12	2.89	0.85	3.41	1.00	3.29	0.96
18	0.00	1.32	0.00	1.00	0.00	0.01
24	1.50	0.77	1.93	1.00	1.82	0.94
36	2.37	0.66	3.60	1.00	3.40	0.94
48	1.18	0.44	2.67	1.00	2.56	0.96
60	2.74	2.56	1.07	1.00	1.22	1.14
72	4.83	1.74	2.78	1.00	2.71	0.97
84	8.00	1.22	6.57	1.00	6.39	0.97
96	8.11	1.11	7.31	1.00	7.07	0.97
108	6.18	1.02	6.05	1.00	5.85	0.97
120	6.39	0.97	6.61	1.00	6.50	0.98
156	3.00	0.91	3.29	1.00	3.04	0.92
180	3.60	1.23	2.92	1.00	3.53	1.21
216	7.06	1.27	5.58	1.00	6.10	1.09
240	8.60	1.14	7.54	1.00	7.66	1.02

Notes: “RMSE” columns report root mean square errors (RMSE) in basis points for different time-to-maturity for each model. “DNS/AFNS”, “AFNS/AFNS”, and “MAFNS/AFNS” columns respectively report ratios of DNS’s, AFNS’s, and MAFNS’s RMSE to AFNS’s RMSE.

week forecasting horizons. The forecasting window is the period from week 3 of May 2010 to week 4 of Dec. 2011 having respectively 88, 80, and 66 observations for 4-week-, 12-week-, and 26-week-ahead forecasts. The estimation window, which starts from the 1st week of Jan. 1999, is expanded during the forecasting process. The RMSFE of AFNS and ratios of DNS’s and MAFNS’s RMSFE to AFNS’s RMSFE are reported in Table 7. Figure 10 are plots of RMSFE for different forecasting horizons. From Figure 10, it looks obvious that DNS has the best forecasting performance for 12-week, and 26-week. For the 4-week-ahead forecasts, the differences between the three models are not obvious. Also, comparing MAFNS with AFNS across different forecasting horizons, we see that the difference between them is not so obvious. This is in line with our expectation that MAFNS do not deviate much from AFNS because the magnitudes of the adjustment terms of both models are not much different.

Economic interpretation of the latent state variables

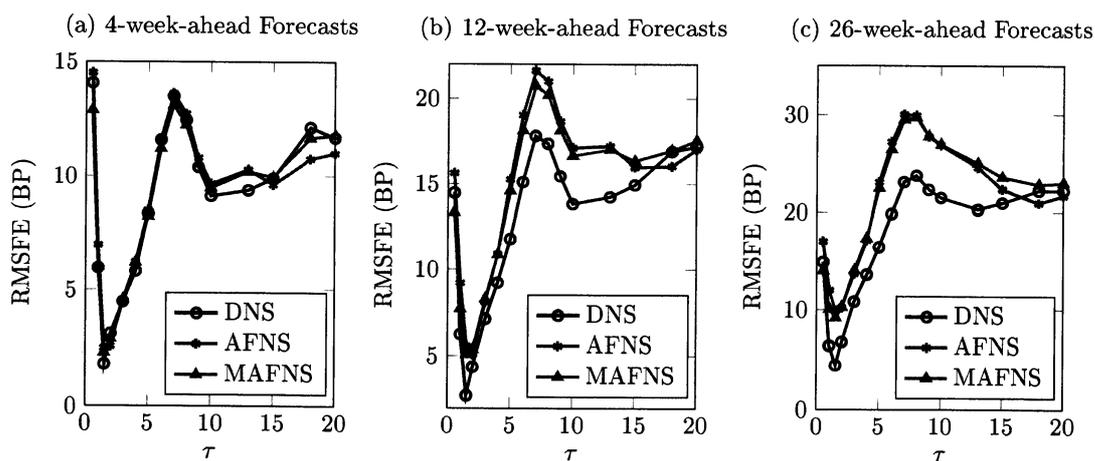
Same as above, to see whether the factors in our model can be interpreted as level, slope, and curvature, we calculate correlations between estimates of these factor and empirical levels, slopes, and curvatures. Table 8 Panel A reports correlations between the estimates of the factor X_1 of each model and long-term (13-year, 15-year, 18-year, and 20-year) yields. Panel B reports correlations between the estimates of X_2 of each model and negative empirical slopes defined as the difference between a short and a long-term (13-year, 15-year, 18-year, and 20-year) yield with 0.5-year yield chosen as the short-term yield. Finally, Panel C reports correlations between the estimates of X_3 and empirical curvatures defined as 2 times medium-term yield (5-year yield) minus a short-term (0.5-year) yield and a long-term (13-year, 15-year, 18-year, and 20-year) yield. As we can see, high correlations between X_2 and negative empirical slopes, and

Table 7: RMSFE (JGB)

τ	DNS/AFNS			AFNS			MANFS/AFNS		
	4-week	12-week	26-week	4-week	12-week	26-week	4-week	12-week	26-week
0.5	0.97	0.93	0.88	14.53	15.63	17.03	0.89	0.85	0.82
1	0.85	0.68	0.53	6.99	9.20	12.00	0.86	0.84	0.84
1.5	0.73	0.49	0.45	2.48	5.57	9.92	0.93	0.92	0.93
2	1.22	0.86	0.66	2.56	5.09	10.23	1.13	1.06	1.01
3	0.99	0.88	0.79	4.53	8.07	13.71	1.01	1.02	1.03
4	0.93	0.85	0.79	6.24	10.91	17.12	1.00	1.00	1.01
5	0.98	0.77	0.71	8.49	15.27	23.06	0.97	0.96	0.98
6	1.00	0.79	0.73	11.63	19.05	27.24	0.96	0.95	0.97
7	0.99	0.82	0.77	13.61	21.63	30.04	0.96	0.96	0.98
8	0.98	0.83	0.79	12.73	21.00	29.94	0.96	0.96	0.99
9	0.96	0.83	0.80	10.80	18.67	27.76	0.97	0.97	1.00
10	0.95	0.81	0.80	9.65	17.14	26.90	0.98	0.97	1.00
13	0.91	0.83	0.83	10.28	17.24	24.59	0.99	0.99	1.02
15	1.03	0.94	0.94	9.61	15.98	22.35	1.04	1.02	1.05
18	1.13	1.05	1.06	10.75	16.05	20.93	1.09	1.06	1.09
20	1.06	1.01	1.02	11.01	17.04	21.69	1.07	1.03	1.06

Notes: Column 5 to 7 are 4-week-, 12-week-, and 26-week-ahead RMSFE of AFNS in basis points. Column 2 to 4 are ratios of DNS's RMSFE to AFNS's RMSFE. Column 8 to 10 are ratios of MAFNS's RMSFE to AFNS's RMSFE.

Figure 10: Out-of-Sample Forecasting Performance (JGB)



between X_3 and empirical curvatures, are observed for the three models. However, X_1 of the three models has weaker correlation with long-term yields than expected, and X_1 of MAFNS has the weakest correlations.

Table 8: Correlations between the state variables and economic variables (JGB DATA)

Panel A: Correlations between X_1 and long-term yields				
	13y	15y	18y	20y
X_1^{DNS}	0.7597	0.8079	0.831	0.8424
X_1^{AFNS}	0.6416	0.7132	0.7581	0.7709
X_1^{MAFNS}	0.5939	0.6669	0.7132	0.7290

Panel B: Correlations between X_2 and negative empirical slopes				
	[0.5y-13y]	[0.5y-15y]	[0.5y-18y]	[0.5y-20y]
X_2^{DNS}	0.9238	0.9238	0.9118	0.9257
X_2^{AFNS}	0.8738	0.8952	0.9025	0.9193
X_2^{MAFNS}	0.8381	0.8702	0.8856	0.9031

Panel C: Correlations between X_3 and empirical curvatures				
	[2*(5y)-0.5y-13y]	[2*(5y)-0.5y-15y]	[2*(5y)-0.5y-18y]	[2*(5y)-0.5y-20y]
X_3^{DNS}	0.9505	0.9661	0.9579	0.9445
X_3^{AFNS}	0.9507	0.9766	0.9814	0.9653
X_3^{MAFNS}	0.9419	0.9681	0.9694	0.9533

Notes: Panel A reports correlations between the estimated path of the factor X_1 of each model and the long-term (13-year, 15-year, 18-year, and 20-year) yields. Panel B reports correlations between the estimated path of X_2 of each model and negative empirical slopes defined as the difference between a short and a long-term (13-year, 15-year, 18-year, and 20-year) yield with 0.5-year yield chosen as the short-term yield. Finally, Panel C reports correlations between the estimated path of X_3 and empirical curvatures defined as the 2 times medium-term yield (5-year yield) minus a short-term (0.5-year) yield and a long-term (13-year, 15-year, 18-year, and 20-year) yield.

5 An example of using MAFNS in credit risk pricing model

After checking the performance and properties of our model relative the existing related models, we can say that even though our model does not outperform the existing NS-type models, to some extent, it preserves the properties and performances of these models. We argue that for modeling purpose, our model can be used as an alternative model to AFNS model. In this section we illustrate this point by giving an example of using our model in credit risk modeling within the reduced-form framework.

In credit risk modeling, there are some stylized facts that are difficult to model. One of these stylized facts that most studies tend to ignore due to modeling and implementation difficulty is the fact that recovery rate is not constant, but time-varying and negatively correlated with default intensity. Most previous studies assume constant recovery rates. Studies that incorporate stochastic recovery rates include Christensen (2007), Hotch and Zagst (2010), Doshi (2011), Jaskowski (2011), among others. One other stylized fact is that default intensity is negatively correlated with interest rates. In most practices, interest rates are assumed constant, or independent of default intensity. Studies that attempt to capture this stylized fact are for instant, Doshi (2011), and Hotch and Zagst (2010).

Our attempt is to capture such difficult-to-model stylized facts using MAFNS as an interest

rate model. The reason for using NS-type model is two folds. First, since one factor of such models can be interpreted as slope, which is known to be an indicator of future economic growth, using NS-type models, we are able to incorporate such economic indicator into our credit pricing model. This is crucial because sub-prime loan crisis witnesses the need to incorporate future economic indicator into credit pricing models. Ignoring such important information may have led to underpricing of credit risk; as a consequence, the financial crisis may have occurred. Second, since the other two factors can be interpreted as level and curvature, we are able to explore the dynamic interaction between such economic factors and credit spread (and/or default probability and/or recovery rates).

In addition, the reason for preferring MAFNS to AFNS is that we are able to model correlation between interest rate factors and default intensity, and at the same time taking into account the positivity of default intensity. For example, we may consider modeling default intensity process h as follow:

$$\begin{aligned} h_t &= \rho X_{1,t} + X_{4,t}; -1 \leq \rho \leq 0, \\ dX_{4,t} &= \kappa_4^Q (\theta_4^Q - X_{4,t}) dt + \sigma_4 \sqrt{X_{4,t}} dW_{4,t}^Q. \end{aligned}$$

Since $X_{1,t}$ is a CIR process as in MAFNS, imposing the following parameter restrictions,

$$\kappa_4^Q = \kappa_1^Q, \quad \rho\theta_1^Q + \theta_4^Q > 0, \quad \text{and} \quad \sigma_4 = \sigma_1 \sqrt{|\rho|},$$

it can be easily shown that the default intensity is a positive process. The negative correlation between default intensity h and interest rate is captured through the negative parameter ρ . Also, we may consider stochastic recovery rate $\varphi(\cdot)$ model as follow:

$$\varphi(\mathbb{X}_t) = \exp(\beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \beta_3 X_{3,t} + X_{5,t}),$$

where $\mathbb{X}_t = (X_{1,t}, X_{2,t}, X_{3,t}, X_{4,t}, X_{5,t})$, and X_5 can be a Gaussian process independent of other factors, and can be considered as recovery rate idiosyncratic factor.

Under this model specification, first note that the default intensity and recovery rate share a common factor X_1 , so its correlation is captured by the parameter β_1 . Moreover, correlations between the recovery rate and interest rate factors depend on $\beta_i, i = 1, 2, 3, 4$. Imposing zero restriction on these parameters, the model reduces to a constant recovery rate model. Since, all factors are modeled as affine processes, explicit formulas for corporate bond, and CDS price can be obtained by applying the transform result of Duffie, Pan, and Singleton (2000). Using corporate bond price data and/or CDS spread data, we are able to estimate the joint model using Markov Chain Monte Carlo (MCMC) method or Monte Carlo Particle filtering method. Significances of these parameters can then be tested to see whether interest rate factors interact with default intensity as well as recovery rate processes. Furthermore, some stylized facts may also be drawn by investigating the sign of these parameters.

Nonetheless, it should be noted that in the above model, recovery rate does not take the value between 0 and 1. An alternative model as in Chen and Joslin (2012) for recovery rate may take the following form in order to ensure it takes the value in $[0, 1]$.

$$\varphi(\mathbb{X}_t) = \frac{1}{1 + e^{-\beta_0 + \beta_1 \cdot \mathbb{X}_t}}.$$

Since \mathbb{X} are affine processes, theoretical corporate bond price and CDS price may be calculated by applying the generalized transform result of Chen and Joslin (2012). However, explicit solution may not be guaranteed in this case.

Therefore, we have shown that using MAFNS model, we were able to model credit risk pricing model that captured some important stylized facts as mentioned above while guaranteeing the positivity of default intensity.

6 Conclusion

In this paper, we proposed a modified arbitrage-free Nelson-Siegel, a 3-factor arbitrage-free affine DTSM. We compared our model with the related models (DNS, and AFNS), in terms of empirical performance and properties using 2 sets of sample data; one is the US treasury zero-coupon yields, and the other one is the Japan government bond zero-coupon yields. For the US treasury data, we found that our model outperforms AFNS in terms of in-sample fit, but this is offset by poorer forecasting performance. However, our model still maintains some good properties, which include the fact that one of its factors can be interpreted as long-term yields, and one as slope, and that the long-term yield factor has the most persistent dynamic. For JGB data, we did not find significant difference between MAFNS and AFNS, which means that both models have similar properties and performance. Giving an example of using MAFNS in joint credit risk pricing model within the reduced-form framework, we conclude that our model can be used as an alternative to AFNS. Using MAFNS, we showed that not only is the joint model able to capture some important stylized facts, positivity of default intensity is also ensured. However, it should be noted that there are some important facts that are not addressed in our study. One is the positivity of interest rates, and another is the regime which is observed in JGB yields. Furthermore, even though our proposed joint model of credit risk seems to work fine, it still awaits empirical analysis to evaluate its practicality. Such problem is thus left for future research.

A Appendices

A.1 Proof of Proposition 1

Proof. With the represent as in (3.1.1) and (3.1.2), Duffie and Kan (1996) prove that the zero-coupon prices are given by

$$P(t, T) = \exp\{\mathbf{B}(T-t)' \mathbf{X}_t + A(T-t)\},$$

where $\mathbf{B}(t)$ and $A(t)$ satisfy the following ODEs

$$\dot{\mathbf{B}}(t) = -\rho_1 - \mathbf{K}^Q \mathbf{B}(t) + \frac{1}{2} \sum_{j=1}^n (\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{jj} \delta_1^j, \quad (\text{A.1.1})$$

$$\dot{A}(t) = -\rho_0 + (\mathbf{K}^Q \boldsymbol{\theta}^Q)' \mathbf{B}(t) + \frac{1}{2} \sum_{j=1}^n (\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{jj} \delta_0^j, \quad (\text{A.1.2})$$

with the boundary conditions: $\mathbf{B}(0) = \mathbf{0}, A(0) = 0$. The dot on \mathbf{B} and A denotes partial derivative w.r.t. t .

With the specification as in Proposition 1, we then have

$$\begin{aligned} \dot{\mathbf{B}}(t) &= - \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \begin{pmatrix} \kappa_1^Q & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \mathbf{B}(t) + \frac{1}{2} \begin{pmatrix} (\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{11} \\ 0 \\ 0 \end{pmatrix}, \\ \dot{A}(t) &= (\mathbf{K}^Q \boldsymbol{\theta}^Q)' \mathbf{B}(t) + \frac{1}{2} \sum_{j=2}^3 (\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{jj}. \end{aligned}$$

Since $(\boldsymbol{\Sigma}' \mathbf{B}(t) \mathbf{B}(t)' \boldsymbol{\Sigma})_{11} = \sigma_{11}^2 B_1^2(t)$, solving these ODEs is straightforward. \square

A.2 Parameter restrictions on the canonical representation of the $\mathbb{A}_1(3)$ class of affine term structure models

Dai and Singleton (2000) classify the n -factor affine term structure models into $n+1$ classes denoted by $\mathbb{A}_0(n), \mathbb{A}_1(n), \dots, \mathbb{A}_m(n), \dots, \mathbb{A}_n(n)$, where m in $\mathbb{A}_m(n)$ denotes the number of factors that enter into the volatility structure of the affine diffusion process. The canonical representation of the $\mathbb{A}_m(n)$ class is admissible (in the sense that the model is well-defined) and maximally flexible among the models within this class. In this appendix, using an invariant transformation we look for the parameter restrictions to be imposed on the canonical representation of $\mathbb{A}_1(3)$ class so as to arrive at the independent-factor MAFNS model.

The canonical representation of $\mathbb{A}_1(3)$:

The instantaneous short rate:

$$r_t = \delta_0 + \delta_1' \mathbf{Y}_t = \delta_0 + \delta_{11} Y_{1,t} + \delta_{12} Y_{2,t} + \delta_{13} Y_{3,t} \quad (\text{A.2.1})$$

The dynamic under \mathbb{Q} :

$$\begin{aligned} d \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \end{pmatrix} &= \begin{pmatrix} \kappa_{11}^{Q,Y} & 0 & 0 \\ \kappa_{21}^{Q,Y} & \kappa_{22}^{Q,Y} & \kappa_{23}^{Q,Y} \\ \kappa_{31}^{Q,Y} & \kappa_{32}^{Q,Y} & \kappa_{33}^{Q,Y} \end{pmatrix} \left[\begin{pmatrix} \theta_1^Q \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \end{pmatrix} \right] dt \\ &+ \begin{pmatrix} \sqrt{Y_{1,t}} & 0 & 0 \\ 0 & \sqrt{1 + \beta_2 Y_{1,t}} & 0 \\ 0 & 0 & \sqrt{1 + \beta_3 Y_{1,t}} \end{pmatrix} d \begin{pmatrix} W_{1,t}^Q \\ W_{2,t}^Q \\ W_{3,t}^Q \end{pmatrix} \end{aligned} \quad (\text{A.2.2})$$

The dynamic under \mathbb{P} :

$$d \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \end{pmatrix} = \begin{pmatrix} \kappa_{11}^{P,Y} & 0 & 0 \\ \kappa_{21}^{P,Y} & \kappa_{22}^{P,Y} & \kappa_{23}^{P,Y} \\ \kappa_{31}^{P,Y} & \kappa_{32}^{P,Y} & \kappa_{33}^{P,Y} \end{pmatrix} \left[\begin{pmatrix} \theta_1^P \\ \theta_2^P \\ \theta_3^P \end{pmatrix} - \begin{pmatrix} Y_{1,t} \\ Y_{2,t} \\ Y_{3,t} \end{pmatrix} \right] dt \\ + \begin{pmatrix} \sqrt{Y_{1,t}} & 0 & 0 \\ 0 & \sqrt{1 + \beta_2 Y_{1,t}} & 0 \\ 0 & 0 & \sqrt{1 + \beta_3 Y_{1,t}} \end{pmatrix} d \begin{pmatrix} W_{1,t}^P \\ W_{2,t}^P \\ W_{3,t}^P \end{pmatrix}. \quad (\text{A.2.3})$$

There are 24 parameters in this maximally flexible model, and there are only 12 parameters in the MAFNS model. Thus, 12 parameter restrictions need to be imposed in order to obtain the MAFNS model.

For the affine diffusion process $d\mathbf{Y}_t = \mathbf{K}^Y(\boldsymbol{\theta}^Y - \mathbf{Y}_t)dt + \boldsymbol{\Sigma}^Y \sqrt{\mathbf{S}_t^Y} d\mathbf{W}_t$, we consider the affine invariant transformation $\mathbf{X}_t = T_{\mathbf{A}}(\mathbf{Y}_t) = \mathbf{A}\mathbf{Y}_t + \boldsymbol{\zeta}$, where \mathbf{A} is a non-singular matrix, and $\boldsymbol{\zeta}$ is a vector having the same dimension as \mathbf{Y}_t .

From Itô formula,

$$d\mathbf{Y}_t = \mathbf{A}^{-1}d\mathbf{X}_t \\ = \mathbf{A}^{-1} \left[\mathbf{K}^X(\boldsymbol{\theta}^X - \mathbf{X}_t)dt + \boldsymbol{\Sigma}^X \sqrt{\mathbf{S}_t^X} d\mathbf{W}_t \right] \\ = \mathbf{A}^{-1} \mathbf{K}^X \mathbf{A} (\mathbf{A}^{-1} \boldsymbol{\theta}^X - \boldsymbol{\zeta} + \boldsymbol{\zeta} - \mathbf{A}^{-1} \mathbf{X}_t) dt + \mathbf{A}^{-1} \boldsymbol{\Sigma}^X \sqrt{\mathbf{S}_t^X} d\mathbf{W}_t \\ = \mathbf{K}^Y(\boldsymbol{\theta}^Y - \mathbf{Y}_t)dt + \boldsymbol{\Sigma}^Y \sqrt{\mathbf{S}_t^Y} d\mathbf{W}_t. \quad (\text{A.2.4})$$

Then, we must have $\mathbf{K}^Y = \mathbf{A}^{-1} \mathbf{K}^X \mathbf{A}$, $\boldsymbol{\theta}^Y = \mathbf{A}^{-1} \boldsymbol{\theta}^X - \boldsymbol{\zeta}$, and $\boldsymbol{\Sigma}^Y \sqrt{\mathbf{S}_t^Y} = \mathbf{A}^{-1} \boldsymbol{\Sigma}^X \sqrt{\mathbf{S}_t^X}$.

This is true both under \mathbb{P} and \mathbb{Q} . It can be verified that

$$\mathbf{A} = \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix}, \text{ and } \boldsymbol{\zeta} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

will transform the above canonical representation into the MAFNS model. Since

$$\mathbf{K}^{Y,Q} = \mathbf{A}^{-1} \mathbf{K}^{X,Q} \mathbf{A} = \begin{pmatrix} \sigma_1^{-2} & 0 & 0 \\ 0 & \sigma_2^{-1} & 0 \\ 0 & 0 & \sigma_3^{-1} \end{pmatrix} \begin{pmatrix} \kappa_1^Q & 0 & 0 \\ 0 & \lambda & -\lambda \\ 0 & 0 & \lambda \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \\ = \begin{pmatrix} \kappa_1^Q & 0 & 0 \\ 0 & \lambda & -\lambda \frac{\sigma_3}{\sigma_2} \\ 0 & 0 & \lambda \end{pmatrix},$$

four restrictions to be imposed on the mean reversion rate under \mathbb{Q} -measure are

$$\kappa_{21}^{Y,Q} = \kappa_{31}^{Y,Q} = \kappa_{32}^{Y,Q} = 0, \kappa_{22}^{Y,Q} = \kappa_{33}^{Y,Q}.$$

Since

$$\mathbf{K}^{Y,P} = \mathbf{A}^{-1} \mathbf{K}^{X,P} \mathbf{A} = \begin{pmatrix} \sigma_1^{-2} & 0 & 0 \\ 0 & \sigma_2^{-1} & 0 \\ 0 & 0 & \sigma_3^{-1} \end{pmatrix} \begin{pmatrix} \kappa_1^P & 0 & 0 \\ 0 & \kappa_2^P & 0 \\ 0 & 0 & \kappa_3^P \end{pmatrix} \begin{pmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \\ = \begin{pmatrix} \kappa_1^P & 0 & 0 \\ 0 & \kappa_2^P & 0 \\ 0 & 0 & \kappa_3^P \end{pmatrix},$$

four restrictions on the \mathbb{P} -dynamic mean reversion rate are

$$\kappa_{21}^{P,Y} = \kappa_{23}^{P,Y} = \kappa_{31}^{P,Y} = \kappa_{32}^{P,Y} = 0.$$

Also,

$$\begin{aligned} \Sigma^Y \sqrt{\mathbf{S}_t^Y} &= \mathbf{A}^{-1} \Sigma^X \sqrt{\mathbf{S}_t^X} = \begin{pmatrix} \sigma_1^{-2} & 0 & 0 \\ 0 & \sigma_2^{-1} & 0 \\ 0 & 0 & \sigma_3^{-1} \end{pmatrix} \begin{pmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{pmatrix} \begin{pmatrix} \sqrt{X_{1,t}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \sqrt{(1 \ 0 \ 0) \mathbf{A} \mathbf{Y}_t} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} \sqrt{Y_{1,t}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

We obtain another 2 restrictions: $\beta_2 = \beta_3 = 0$.

Finally, $r_t = X_{1,t} + X_{2,t} = (1 \ 1 \ 0) \mathbf{X}_t = (1 \ 1 \ 0) \mathbf{A} \mathbf{Y}_t = (\sigma_1^2 \ \sigma_2 \ 0) \mathbf{Y}_t$. Thus, the other 2 restrictions are $\delta_0 = 0, \delta_{13} = 0$.

A.3 Conditional and unconditional variance of state variables

Under the independent-factor assumption the conditional expectation and variance are explicitly given as follow.

The conditional expectation and variance:

$$\mathbb{E}_{t-\Delta t}^{\mathbb{P}}[\mathbf{X}_t] = (I - e^{-\mathbf{K}^{\mathbb{P}} \Delta t}) \boldsymbol{\theta}^{\mathbb{P}} + e^{-\mathbf{K}^{\mathbb{P}} \Delta t} \mathbf{X}_{r,t-\Delta t}, \quad (\text{A.3.1})$$

$$\text{Var}_{t-\Delta t}^{\mathbb{P}}[\mathbf{X}_t] = Q(X_{1,t-\Delta t}, \boldsymbol{\psi}), \quad (\text{A.3.2})$$

where $\mathbb{E}_t[\cdot] \equiv \mathbb{E}[\cdot | \mathcal{F}_t]$, and $\boldsymbol{\psi}$ is the vector containing all the parameters, and Q is a diagonal matrix with diagonal elements:

$$\begin{aligned} Q_{11} &= \frac{\sigma_1^2}{2\kappa_1^{\mathbb{P}}} \left[1 - e^{-\kappa_1^{\mathbb{P}} \Delta t} \right]^2 + \frac{\sigma_1^2}{\kappa_1^{\mathbb{P}}} X_{1,t} e^{-\kappa_1^{\mathbb{P}} \Delta t}, \\ Q_{ii} &= \frac{\sigma_i^2}{2\kappa_i^{\mathbb{P}}} \left[1 - e^{-2\kappa_i^{\mathbb{P}} \Delta t} \right], i = 2, 3. \end{aligned}$$

The unconditional expectation and variance:

$$\lim_{t \rightarrow \infty} \mathbb{E}^{\mathbb{P}}[\mathbf{X}_t] = \boldsymbol{\theta}^{\mathbb{P}}, \quad (\text{A.3.3})$$

$$\lim_{t \rightarrow \infty} \text{Var}^{\mathbb{P}}[\mathbf{X}_t] = \begin{pmatrix} \frac{\sigma_1^2}{2\kappa_1^{\mathbb{P}}} & 0 & 0 \\ 0 & \frac{\sigma_2^2}{2\kappa_2^{\mathbb{P}}} & 0 \\ 0 & 0 & \frac{\sigma_3^2}{2\kappa_3^{\mathbb{P}}} \end{pmatrix}. \quad (\text{A.3.4})$$

References

- [1] BIS (2005) *Zero-Coupon Yield Curves: Technical Documentation*, Bank of International Settlements, Basel
- [2] Björk T., and B.J. Christensen (1999). Interest Rate Dynamics and Consistent Forward Rate Curves. *Mathematical Finance* 9 (4): 323-348
- [3] Chen H., and S. Joslin (2012). Generalized Transform Analysis of Affine Processes and Applications in Finance. *Review of Financial Studies* 25 (7): 2225-2256
- [4] Christensen J.H.E. (2007). Joint Default and Recovery Risk Estimation: an Application to CDS Data. Working Paper, Federal Reserve Bank of San Francisco.
- [5] Christensen J.H.E., F.X. Diebold, and G. D. Rudebusch (2011). The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. *Journal of Econometrics* 164 (1): 4-20
- [6] Coroneo L., K. Nyholm, and R. Vidova-Koleva (2008). How Arbitrage-Free is the Nelson-Siegel Model? Working Paper Series, European Central Bank
- [7] Cox J.C., J.E. Ingersoll, and S.A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica* 53 (2): 385-407
- [8] Dai Q., and K.J. Singleton (2000). Specification Analysis of Affine Term Structure Models. *The Journal of Finance* 55 (5): 1943-1974
- [9] Diebold F.X., and C. Li (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics* 130 (2): 337-364
- [10] Doshi H. (2011). The Term Structure of Recovery Rates. Working paper, McGill University
- [11] Duan J.C., and J.G. Simonato (1999). Estimating and Testing Exponential-Affine Term Structure Models by Kalman Filter. *Review of Quantitative Finance and Accounting* 13 (2): 111-135
- [12] Duffee G.R. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *The Journal of Finance* 57 (1): 405-443
- [13] Duffie D., and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6 (4): 379-406
- [14] Duffie D., J. Pan, and K.J. Singleton (2000). Transform Analysis and Asset Pricing for Affine Jump-Diffusions. *Econometrica* 68 (6): 1343-1376
- [15] Gurkaynak R.S., B. Sack, and J.H. Wright (2006). The U.S. Treasury Yield Curve: 1961 to the Present. Finance and Economics Discussion Series, Federal Reserve Board
- [16] Heath D., R. Jarrow, and A. Morton (1992). Bond Pricing and the Term Structure of Interest Rates: A New Methodology for Contingent Claims Valuation. *Econometrica*, 60 (1):77-105
- [17] Hocht S., and R. Zagst (2010). Pricing Credit Derivatives Under Stochastic Recovery in a Hybrid Model. *Applied Stochastic Models in Business and Industry* 26 (3): 254-276
- [18] Hull J., and A. White (1990). Pricing Interest-Rate Derivative Securities. *The Review of Financial Studies* 3 (4): 573-592

- [19] Jaskowaki M. (2011). Estimation of Implied Recovery Rates. A Case Study of the CDS Spreads Market. Working paper, Vienna University of Economics and Business.
- [20] Kikuchi K., and K. Shintani (2012). Comparative Analysis of Zero Coupon Yield Curve Estimation Methods Using JGB Price Data. IMES Discussion Paper No. 2012-E-4, Bank of Japan
- [21] Lin H., and C. Wu (2010). Term Structure of Default-Free and Defaultable Securities: Theory and Empirical Evidence. in *Handbook of Quantitative Finance and Risk Management* (Lee C.-F., A.C. Lee, and J. Lee, Eds.), Springer, Chapter 63: 979-1005
- [22] Nelson C.R., and A.F. Siegel (1987). Parsimonious Modeling of Yield Curves. *Journal of Business* 60, 473-489
- [23] Steeley James M. (1991). Estimating the Gilt-edged Term Structure: Basis Splines and Confidence Intervals. *Journal of Business Finance and Accounting*, 18 (4): 513-529
- [24] Svensson Lars E.O. (1995). Estimating Forward Interest Rates with the Extended Nelson-Siegel Method. *Sveriges Riksbank Quarterly Review* 3, 13-26
- [25] Vasicek O. (1977). An Equilibrium Characterisation of the Term Structure. *Journal of Financial Economics* 5 (2): 177-188

Graduate School of Economics
Osaka University, Toyonaka 560-0043, Japan
E-mail address: simdara@gmail.com

大阪大学・経済学研究科 Dara Sim