



Australian  
National  
University

**CENTRE FOR APPLIED MACROECONOMIC ANALYSIS**

The Australian National University

---

**CAMA Working Paper Series**

**March, 2012**

---

A THEORETICAL FOUNDATION FOR THE NELSON AND SIEGEL  
CLASS OF YIELD CURVE MODELS

**Leo Krippner**

Reserve Bank of New Zealand

Centre for Applied Macroeconomic Analysis (CAMA), ANU

---

CAMA Working Paper 11/2012

<http://cama.anu.edu.au>

# A theoretical foundation for the Nelson and Siegel class of yield curve models

Leo Krippner\*

15 February 2012

## Abstract

Yield curve models within the popular Nelson and Siegel (hereafter NS) class are shown to arise from a formal low-order Taylor approximation to the generic Gaussian affine term structure model. That theoretical foundation provides an assurance that NS models correspond to a well-accepted framework for yield curve modeling. It further suggests that any yield curve from the GATSM class should be parsimoniously representable by an two-factor arbitrage-free NS model, which should prove useful for macrofinance applications. Such a model is derived and applied to provide evidence for changes in United States yield curve dynamics pre- and post-1988.

JEL: E43, G12, C58

Keywords: yield curve; term structure of interest rates; Nelson and Siegel model; affine term structure models.

## 1 Introduction

In this article, I establish a theoretical foundation for the popular Nelson and Siegel (1987, hereafter NS) class of yield curve models. In particular, I show how the Level, Slope, and Curvature components common to all models of the NS class correspond explicitly to the generic Gaussian affine term structure model (hereafter GATSM) outlined in Dai and Singleton (2002).

My primary motivation for establishing such a foundation is to “legitimize”, in a sense I discuss further below, the popularity of NS models. That popularity, from one perspective, is evident from the frequent and extensive applications of NS models to

---

\*Reserve Bank of New Zealand (P.O. Box 2498, Wellington, New Zealand. Phone: +64 4 471 3686. Email: leo.krippner@rbnz.govt.nz) and Centre for Applied Macroeconomic Analysis (CAMA). This working paper updates the related Reserve Bank of New Zealand discussion papers Krippner (2009, 2010). I thank Jens Christensen, Iris Claus, Darrell Duffie, Mardi Dungey, Manu De Veirman, Francis Diebold, Nicolas Groshenny, Campbell Harvey, Ben Jacobsen, John Knight, Miles Livingston, Charles Nelson, Ken Nyholm, Peter Phillips, Glenn Rudebusch, Christie Smith, participants at the New Zealand Association of Economists 2009 conference, the Society of Financial Econometrics 2010 conference, the New Zealand Econometrics Study Group 2011, the Australasian Macroeconomic Workshop 2011, and the Computational Financial Econometrics 2011 conference for helpful comments.

topics spanning finance, economics, and the emerging cross-over field of macrofinance.<sup>1</sup> A second perspective on the popularity of NS models is their pervasiveness beyond the academic patronage indicated by the references already footnoted. For example, Bank for International Settlements (2005) provides an overview of the routine use of NS models by many central banks, and NS models appear in standard textbooks aimed at financial market practitioners and university students; e.g. James and Webber (2000), Nyholm (2008), Filipović (2009) and Munk (2011). Moreover, a web search readily illustrates the ongoing inclusion of NS models in university courses and use by financial market practitioners, as also noted in Coroneo, Nyholm and Vidova-Koleva (2008).

Nevertheless, despite their apparent popularity, the Level, Slope, and Curvature components common to all models of the NS class were only given a heuristic basis for the original NS model. That basis is highlighted with a selection of quotations from the introduction and conclusion of NS:

“The purpose of this paper is to introduce a simple, parsimonious model that is flexible enough to represent the range of shapes generally associated with yield curves: monotonic, humped, and S shaped.” “A class of functions that readily generates the typical yield curve shapes is that associated with solutions to differential or difference equations. The expectations theory of the term structure provides heuristic motivation for investigating this class since, if spot rates are generated by a differential equation, then forward rates, being forecasts, will be the solution to the equations.” “A more parsimonious model that can generate the same range of shapes is given by the solution equation for the case of equal roots.” “Our objective in this paper has been to propose a class of models, motivated by but not dependent on the expectations theory of the term structure, that offers a parsimonious representation of the shapes traditionally associated with yield curves.”

Even the recent introduction of arbitrage-free NS models (e.g. Sharf and Filipović 2004, Krippner 2006a, and Christensen, Diebold and Rudebusch 2009, 2011), while at least imposing theoretical self-consistency by explicitly accounting for assumed Gaussian dynamics in the yield curve, still take the NS Level Slope and Curvature components as given latent factors. For example, from Christensen et al. (2011) footnote 4: “Our strategy is to find the affine AF [arbitrage-free] model with factor loadings that match Nelson-Siegel exactly.” Justification for the components themselves, if supplied, inevitably appeals to the practical benefits of NS models, such as their ease of estimation, close fit to the yield curve data, intuitive estimated components (including the

---

<sup>1</sup>A brief overview, with selected examples not otherwise referenced later in the article, includes: estimating zero-coupon interest rate data, e.g. Gürkaynak, Sack and Wright (2010); gauging monetary policy expectations, as originally advocated in Dahlquist and Svensson (1996); forecasting the yield curve, e.g. Diebold and Li (2006) and Vincente and Tabak (2008); modeling value at risk in fixed-interest portfolios, e.g. Diebold, Ji and Li (2006); general macrofinance, e.g. Diebold, Piazzesi and Rudebusch (2005); monitoring inflation compensation, e.g. Christensen, Lopez and Rudebusch (2010); monitoring the dynamics and liquidity conditions of non-government yields, e.g. Christensen, Lopez and Rudebusch (2009); investigating the uncovered interest parity puzzle, e.g. Krippner (2006b) and Chen and Tsang (2009); and investigating international yield curve linkages, e.g. Wong, Lucia, Price and Startz (2011).

similarity of the Level, Slope, and Curvature components to the first three principal components of the term structure), and their success in past applications.

So the more fundamental question of “why NS components in the first place?” therefore remains an open question. As such, it represents a point of vulnerability in the widespread application of NS models and, by extension, their empirical results.

Fortunately, I show in this article that NS models can be given an explicit foundation within the GATSM class of yield curve models, which is itself based on a well-accepted set of principles and assumptions for modeling the yield curve and its dynamics. To this end, section 2 specifies the generic GATSM from Dai and Singleton (2002) and then derives the associated forward rate curve. Section 3 shows how the original NS forward rate curve arises from the dynamic component of the generic GATSM forward rate curve using what I will call an “eigenvalue approximation”; specifically a low-order Taylor expansion around central measures of the eigenvalues associated with the generic GATSM. The NS Level component is shown to correspond to the persistent (i.e. slowly mean-reverting, or near-zero eigenvalue) components of the generic GATSM, and the NS Slope and Curvature components are shown to correspond to the non-persistent (i.e. non-zero eigenvalue) components of the generic GATSM. In light of this example, section 4 discusses how most models within the NS class, with one notable exception being the Svensson (1995)/NS model, can be classified as various eigenvalue approximations to the generic GATSM. That classification system therefore provides a useful guide to selecting an appropriate NS model for a given application.

A corollary from the theoretical foundation and classification system for the NS class of models is the evident absence of a fully flexible two-factor arbitrage-free NS model. Such a model is likely to prove particularly useful as a standard tool for yield curve modeling and analysis in macrofinance because the model will be very parsimonious (and so easy to apply empirically) while also being fully representative of any GATSM that might be generating the yield curve data. To highlight the latter point, regardless of the number of GATSM factors, their nature (e.g. economic, financial, etc.), and the potential complexity of their interactions (in terms of the mean-reversion and innovation matrices), the Level component of the NS model will represent the persistent components of the GATSM, and the Slope component will represent the non-persistent components. In addition, available empirical evidence to be discussed in section 5.1 already suggests that two factors should be appropriate for macrofinance applications.

I respectively derive constant and time-varying risk premium versions of the two-factor arbitrage-free NS model in sections 5.2 and 5.3. Section 6 contains my chosen empirical application of the model, in this case investigating changes in United States yield dynamics. While the results may simply be taken as a broad confirmation of the results from Rudebusch and Wu (2007) obtained with a two-factor GATSM (i.e. there was a material change in the yield curve data-generating process between the periods 1971 to 1987 and 1988 to 2002), the context of the two-factor arbitrage-free NS representing any GATSM model that could have been applied to the data makes the results and their interpretation much more general.

The remainder of the paper follows the outline already discussed above, and I conclude in section 7.

## 2 The generic Gaussian affine term structure model

The generic GATSM I specify in this section parallels the standard multifactor Gaussian dynamic term structure model as outlined in appendix A of Dai and Singleton (2002). It is the fully Gaussian subset of the affine framework outlined in Duffie and Kan (1996) with the essentially affine specification of market prices of risk from Duffie (2002).<sup>2</sup>

Three points of context for this article are worth noting up front. First, the state variables are completely generic, and may be interpreted as (potentially unobserved) economic and financial factors within the underlying economy. That interpretation follows the Duffie and Kan (1996) p. 321 comment that the state variables in an affine model can always, under standard assumptions, be related back to economic factors (e.g. preferences, technology, consumption, inflation, etc.) within a general equilibrium model.<sup>3</sup> For example, Cox, Ingersoll and Ross (1985a) originally provided a general equilibrium basis for term structure models (including the one-factor GATSM) based on a representative-agent economy, and that approach has since been extended by many authors to provide theoretical foundations for multifactor GATSMs; see, for example, Berardi and Esposito (1999) and Berardi (2009). More recently, Wu (2006) also shows how GATSMs may be given an explicit foundation within dynamic stochastic general equilibrium models. Regarding financial factors (e.g. default risk, liquidity risk, repurchase effects, etc.), Duffie and Singleton (1999) shows how they can be incorporated into the generic GATSM framework, and Singleton (2006) chapter 14 contains an extensive summary of that literature.

Second, to make the exposition more transparent from the perspective of the original NS model, this article derives and works with the forward rate curve associated with the generic GATSM. The affine term structure literature more commonly uses bond prices and/or interest rate curves, but all are within an elementary transformation of each other and are equivalent perspectives for representing the yield curve.<sup>4</sup>

Third, the approach for relating the generic GATSM to NS models does not extend to term structure models with full Cox, Ingersoll and Ross (1985b)/square-root dynamics.<sup>5</sup> NS models therefore inherit the same theoretical shortcomings of GATSMs, i.e. positive probabilities of negative interest rates and constant volatilities, and that perspective should be considered when deciding if it is appropriate to apply an NS model for the task at hand. That said, the assumption of Gaussian dynamics is standard in economics and macrofinance, whether explicitly or implicitly via the application of Gaussian-based econometrics,<sup>6</sup> and I presuppose from this point onward that the user has already assumed a Gaussian data-generating process for the yield curve.

---

<sup>2</sup>In the notation of Dai and Singleton (2000), the Dai and Singleton (2002) specification is  $A_0(N)$ . It may also be seen as an affine-invariant transformation (see Singleton (2006) pp. 319-321) of other GATSM specifications, such as the Joslin, Singleton and Zhu (2011) canonical GATSM.

<sup>3</sup>Of course, the state variables could simply be interpreted as points on the yield curve, without an explicit economic and/or financial context, which is the approach taken in Duffie and Kan (1996).

<sup>4</sup>See, for example, Filipović (2009) p. 7 for an overview of the standard relationships between forward rates, interest rates, and bond prices, which are used subsequently in the present article.

<sup>5</sup>Essentially because, anticipating section 3 of the present article, the functional form of Cox et al. (1985b) does not facilitate the factorization of terms like  $\exp(-\phi\tau)$  that characterize the NS class of models. Further details are available in appendix A of the working paper version of this article.

<sup>6</sup>For example, all of the macrofinance models summarized in Rudebusch (2010) are specified with Gaussian innovations.

Following Dai and Singleton (2002), I define the instantaneous short rate at time  $t$  as  $r(t) = \xi_0 + \xi_1' X(t)$ , where  $\xi_0$  is a constant,  $X(t)$  is an  $N \times 1$  vector of state variables, and  $\xi_1$  is a constant  $N \times 1$  vector. Under the physical  $\mathbb{P}$  measure, the state variables follow the process  $dX(t) = K_{\mathbb{P}}[\theta_{\mathbb{P}} - X(t)]dt + \Sigma dW_{\mathbb{P}}(t)$ , where  $K_{\mathbb{P}}$  is a constant  $N \times N$  mean-reversion matrix,  $\theta_{\mathbb{P}}$  is a constant steady-state  $N \times 1$  vector for  $X(t)$ ,  $\Sigma$  is a constant  $N \times N$  volatility matrix, and  $W_{\mathbb{P}}(t)$  is an  $N \times 1$  vector of independent Brownian motions. The market prices of risk are  $\Pi(t) = \Sigma^{-1}[\pi_0 + \pi_1 X(t)]$ , where  $\pi_0$  is a constant  $N \times 1$  vector and  $\pi_1$  is a constant  $N \times N$  matrix. Under the risk-neutral  $\mathbb{Q}$  measure, the state variables follow the process  $dX(t) = K_{\mathbb{Q}}[\theta_{\mathbb{Q}} - X(t)]dt + \Sigma dW_{\mathbb{Q}}(t)$ , where  $dW_{\mathbb{Q}}(t) = dW_{\mathbb{P}}(t) + \Pi(t)dt$ ,  $K_{\mathbb{Q}} = K_{\mathbb{P}} + \pi_1$ , and  $\theta_{\mathbb{Q}} = (K_{\mathbb{P}} + \pi_1)^{-1}(K_{\mathbb{P}}\theta_{\mathbb{P}} - \Sigma\pi_0)$ . Zero-coupon bond prices under measure  $\mathbb{Q}$  for the generic GATSM are  $P(t, \tau) = \exp[\check{A}(\tau) + B(\tau)' X(t)]$ , where  $\tau = T - t$  with  $T$  being the time of maturity ( $T \geq t$ , so  $\tau \geq 0$ ),  $B(\tau) = [\exp(-K'_{\mathbb{Q}}\tau) - I](K'_{\mathbb{Q}})^{-1}\xi_1$ , and  $I$  is the  $N \times N$  identity matrix. The full expression for  $\check{A}(\tau)$  is provided in Dai and Singleton (2002), but the present article requires only the summary results that  $\check{A}(\tau)$  is required for the system to be arbitrage free, and it can be expressed in the functional form  $-\xi_0\tau + A(\tau)$ .

The standard relationship between instantaneous forward rates and bond prices is  $f(t, \tau) = -\partial \log P(t, \tau) / \partial \tau$ . Therefore, under measure  $\mathbb{Q}$ , the generic GATSM forward rate curve is:

$$f(t, \tau) = \xi_0 + [\exp(-K'_{\mathbb{Q}}\tau)\xi_1]' X(t) - \frac{\partial}{\partial \tau} A(\tau) \quad (1)$$

Now express  $K'_{\mathbb{Q}}$  in eigensystem form; i.e.  $K'_{\mathbb{Q}} = Z\Psi Z^{-1}$ , where  $Z$  is the  $N \times N$  non-singular matrix of eigenvectors each normalized to 1, and  $\Psi$  is the  $N \times N$  diagonal matrix containing the  $N$  eigenvalues  $(\lambda_1, \dots, \lambda_n, \dots, \lambda_N)$ . The latter are assumed to be unique and positive, which follows the standard assumption in Duffie and Kan (1996) and Dai and Singleton (2002). Hence,  $\exp(-K'_{\mathbb{Q}}\tau) = \exp(-Z\Psi Z^{-1}\tau) = Z \exp(-\Psi\tau) Z^{-1} = Z \Lambda Z^{-1}$ ,<sup>7</sup> where  $\Lambda = \text{diag}[\exp(-\lambda_1\tau), \dots, \exp(-\lambda_n\tau), \dots, \exp(-\lambda_N\tau)]$ , an  $N \times N$  diagonal matrix. The forward rates in equation 1 are then  $f(t, \tau) = \xi_0 + \{Z \Lambda Z^{-1} \xi_1\}' X(t) - \frac{\partial}{\partial \tau} A(\tau)$ , which can be expressed equivalently as:

$$f(t, \tau) = \xi_0 + \sum_{n=1}^{n_0} q_n(t) \exp(-\lambda_n\tau) + \sum_{n=n_0+1}^N q_n(t) \exp(-\lambda_n\tau) - \frac{\partial}{\partial \tau} A(\tau) \quad (2)$$

where the coefficients  $q_n(t)$  associated with each unique  $\exp(-\lambda_n\tau)$  represent the collection of coefficients of the  $\exp(-\lambda_n\tau)$  terms that arise from the full matrix multiplication of  $\{Z \Lambda Z^{-1} \xi_1\}' X(t)$ .

For use in the example of the following section (but without loss of generality) I re-order the  $q_n(t) \exp(-\lambda_n\tau)$  components from the smallest to the largest eigenvalue, and then divide them into two groups. The first group contains the components with eigenvalues  $\lambda_1$  to  $\lambda_{n_0}$  that are close to zero (i.e. the persistent components, given

---

<sup>7</sup>For readers unfamiliar with the intermediate step, note that it is obtained via the polynomial expansion for the exponential of a matrix and the simplification  $Z^{-1}Z = I$ , i.e.:  $\exp(-Z\Psi Z^{-1}\tau) = I + \frac{1}{1!}(-Z\Psi Z^{-1}\tau) + \frac{1}{2!}(-Z\Psi Z^{-1}\tau)^2 + \dots = ZIZ^{-1} + \frac{1}{1!}Z(-\Psi\tau)Z^{-1} + \frac{1}{2!}Z(-\Psi\tau)Z^{-1}Z(-\Psi\tau)Z^{-1} + \dots = ZIZ^{-1} + \frac{1}{1!}Z(-\Psi\tau)Z^{-1} + \frac{1}{2!}Z(-\Psi\tau)^2Z^{-1} + \dots = Z\left[I + \frac{1}{1!}(-\Psi\tau) + \frac{1}{2!}(-\Psi\tau)^2 + \dots\right]Z^{-1} = Z \exp(-\Psi\tau) Z^{-1}$ .

they will have a slow exponential decay by time to maturity  $\tau$ ) and the second group contains the eigenvalues  $\lambda_{n_0+1}$  to  $\lambda_N$  that are not close to zero (i.e. the non-persistent components).

Also for use in the following section, I denote the first three components of equation 2 collectively as the dynamic component of the generic GATSM. That terminology reflects that the time-series properties of the generic GATSM forward rate curve are completely contained in the  $q_n(t) \exp(-\lambda_n \tau)$  components. That is, the coefficients  $q_n(t)$  are linear combinations of the original state variables  $X(t)$ , and the functions  $\exp(-\lambda_n \tau)$  determine the expected (as at time  $t$ ) evolution of forward rates from time  $t$  to time  $t + \tau$ . As time evolves, the stochastic term  $\Sigma dW_{\mathbb{P}}(t)$  for the generic GATSM imparts Gaussian innovations to the state variables  $X(t)$ , which are directly translated as Gaussian innovations to the coefficients  $q_n(t)$ . Conversely, the non-dynamic component  $\frac{\partial}{\partial \tau} A(\tau)$  is a time-invariant function of time to maturity.

### 3 The generic GATSM to the original NS model

I use formal Taylor approximation around the generic GATSM eigenvalues  $\lambda_n$ , which I hereafter denote an eigenvalue approximation, to reproduce the original NS model from the exact expression for the dynamic component of the generic GATSM forward rate curve in equation 2. The treatment of the time-invariant component  $\frac{\partial}{\partial \tau} A(\tau)$  is discussed further below.

For the first group of eigenvalues where  $\lambda_n \simeq 0$ , the first term of the Taylor expansion is  $\exp(-\lambda_n \tau) \simeq 1$ . For the second group of eigenvalues where  $\lambda_n \gg 0$ , express them relative to  $\phi = \text{mean}(\lambda_{n_0+1}, \dots, \lambda_N)$ , so that  $\lambda_n = \phi(1 - \delta_n)$  and  $\exp(-\lambda_n \tau) = \exp(-\phi \tau) \exp(\delta_n \phi \tau)$ . From the latter expression, taking two terms of the Taylor expansion around  $\delta_n$  gives  $\exp(-\phi \tau)(1 + \delta_n \phi \tau)$ , so  $q_n(t) \exp(-\lambda_n \tau) \simeq q_n(t) \exp(-\phi \tau) + q_n(t) \delta_n \phi \tau \exp(-\phi \tau)$ .<sup>8</sup> Substituting these results into the dynamic component of equation 2 gives:

$$f(t, \tau) \simeq \xi_0 + \sum_{n=1}^{n_0} q_n(t) + \left[ \sum_{n=n_0+1}^N q_n(t) \right] \exp(-\phi \tau) + \left[ \sum_{n=n_0+1}^N q_n(t) \delta_n \right] \phi \tau \exp(-\phi \tau) \quad (3)$$

The expression in equation 3 is precisely the functional form of the original NS model of the forward rate curve, i.e.:

$$f(t, \tau) \simeq f_{\text{NS}}(t, \tau) = L(t) + S(t) \exp(-\phi \tau) + C(t) \phi \tau \exp(-\phi \tau) \quad (4)$$

where  $f_{\text{NS}}(t, \tau)$  is the original NS forward rate curve, the functions 1,  $\exp(-\phi \tau)$ , and  $\phi \tau \exp(-\phi \tau)$  are the original NS forward rate factor loadings, and the expressions  $L(t) = \xi_0 + \sum_{n=1}^{n_0} q_n(t)$ ,  $S(t) = \sum_{n=n_0+1}^N q_n(t)$ , and  $C(t) = \sum_{n=n_0+1}^N q_n(t) \delta_n$  are the original NS forward rate coefficients expressed in terms of the generic GATSM state

---

<sup>8</sup>Any other central measure of  $(\lambda_{n_0+1}, \dots, \lambda_N)$  would suffice for the exposition in this article. In practice,  $\phi$  is an estimated parameter. Note also that there is a special case if the data were genuinely generated by a two-factor GATSM. That is,  $n_0 = 1$ ,  $N = 2$ ,  $\phi = \lambda_2$ ,  $\delta_2 = 0$ , and  $q_n(t) \exp(-\lambda_n \tau) = q_n(t) \exp(-\phi \tau)$ . The resulting eigenvalue approximation in equation 3 would then have just a Level and Slope component, i.e.:  $f(t, \tau) \simeq \xi_0 + q_1(t) + q_2(t) \exp(-\phi \tau)$ .

variables and parameters. The standard transformation  $R_{\text{NS}}(t, \tau) = \frac{1}{\tau} \int_0^\tau f_{\text{NS}}(t, \tau) d\tau$  then produces the original NS model for the interest rate curve.<sup>9</sup>

The exposition above shows explicitly how the Level component of the original NS model approximates the constant plus the persistent dynamic components of the generic GATSM, and how the NS Slope and Curvature components together approximate the non-persistent dynamic components. From the perspective of any particular GATSM, the eigenvalue approximations will always be “optimal” in the sense that each additional NS component will correspond precisely to each successive term in the associated eigenvalue expansion. Conversely, other functions (e.g. see James and Webber (2000) chapter 15) cannot provide a precise eigenvalue approximation. For example, a cubic spline (often applied to yield curve data) does not provide a precise third-order eigenvalue approximation (or even a precise zeroth-order eigenvalue approximation) to the non-persistent components of a GATSM.

The original NS model as derived above can be made arbitrage free (if required; see the discussion at the end of the following section) by adding appropriate terms to account for the effects that NS component innovations and their associated market prices of risk specifications have on the forward rate curve. It is first worth making the observation that the correspondence of the generic GATSM to the original NS model already discussed above means that the Gaussian innovations to the generic GATSM coefficients  $q_n(t)$  are translated directly into Gaussian innovations for the NS coefficients  $L(t)$ ,  $S(t)$ , and  $C(t)$ . That observation confirms the assumption of Gaussian innovations made in Krippner (2006a) p. 42, where the Heath, Jarrow and Morton (1992) framework is used to derive the appropriate arbitrage-free (hereafter AF) terms for the original NS model. Alternatively, Christensen, Diebold and Rudebusch (2011) directly derives the AF terms for  $R_{\text{NS}}(t, \tau)$  via a particular three-factor GATSM designed to reproduce the original NS factor loadings.<sup>10</sup>

## 4 A GATSM perspective for classifying, selecting, and applying NS models

By following the example in the previous section, it is possible to classify most NS models as a particular eigenvalue approximation of the generic GATSM. The key aspects are: (1) the number of groups of non-zero eigenvalues assumed for the non-persistent dynamic components of the generic GATSM, which determines how many mean eigenvalue parameters are required; (2) the number of terms in the eigenvalue approximation for each mean eigenvalue; and (3) whether the AF term is included, which determines

---

<sup>9</sup>That is,  $R_{\text{NS}}(t, \tau) = L(t) + S(t) \left( \frac{1 - \exp(-\phi\tau)}{\phi\tau} \right) + C(t) \left( \frac{1 - \exp(-\phi\tau)}{\phi\tau} - \exp(-\phi\tau) \right)$ . Interestingly, the original NS article, p. 475, also notes that “This model may also be derived as an approximation to the solution in the unequal roots case by expanding in a power series in the difference between the roots.” The connection to the present article is coincidental however; from a mathematical perspective, the second-order differential equation assumed in NS happens to produce the same general solution as a bivariate first-order differential equation, which would be a minimal multifactor GATSM without an allowance for stochastic dynamics. From an economic perspective, it would be theoretically untenable to propose an  $N^{\text{th}}$ -order differential equation as the basis for a generic  $N$ -factor interest rate model.

<sup>10</sup>As subsequently indicated in table 1 of the present article, Krippner (2006a) assumes independent innovations (and constant market prices of risk), while Christensen et al. (2011) allows for correlated innovations (and essentially-affine market prices of risk).



if the NS model is AF with respect to its dynamic components and the assumed specification for the market prices of risk.

From the perspective of these three aspects, table 1 summarizes the NS models already proposed in the literature. Permutations of the three aspects above can obviously generate an infinite variety of alternative NS models, but table 1 adds just three variants to cover evident absences in the range of proposed NS models to date, and then two further variants with six components to illustrate ongoing potential extensions.

Table 1:  
NS models from a generic GATSM perspective

NS model <sup>(5)</sup>	Components	Representation of GATSM		AF
		$\lambda_n \simeq 0^{(1)}$	$\lambda_n \gg 0^{(2)}$	
DLY2008	2	1 $[L]$	1 $[S]$	No
HJM1992	2	1 $[L]$	1 $[S]$	Yes <sup>(3)</sup>
Variant 1	2	1 $[L]$	1 $[S]$	Yes
NS1987	3	1 $[L]$	2 $[S, C]$	No
Krip.2006	3	1 $[L]$	2 $[S, C]$	Yes <sup>(3)</sup>
CDR2010	3	1 $[L]$	2 $[S, C]$	Yes
Variant 2	3	1 $[L]$	1 $[S_1]$ 1 $[S_2]$	Y/N
Bliss1997 <sup>(4)</sup>	3	1 $[L]$	1 $[S_1, \sim]$ 1 $[\sim, C_2]$	No
Variant 3	4	2 $[L_1, L_2]$	2 $[S_1, C_1]$	Y/N
Sven.1995 <sup>(4)</sup>	4	1 $[L]$	2 $[S_1, C_1]$ 1 $[\sim, C_2]$	No
SF2004 <sup>(4)</sup>	4	1 $[L]$	2 $[S_1, C_1]$ 1 $[\sim, C_2]$	Yes
CDR2009	5	1 $[L]$	2 $[S_1, C_1]$ 2 $[S_2, C_2]$	Yes
Variant 4	6	2 $[L_1, L_2]$	2 $[S_1, C_1]$ 2 $[S_2, C_2]$	Y/N
Variant 5	6	3 $[L_1, L_2, L_3]$	3 $[S, C, C^*]$	Y/N

Notes: (1) entry is the number of terms in the eigenvalue approximation around 0 followed by the related NS component/s; (2) as for (1), but approximation is around  $\phi = \text{mean}(\lambda_{n_0+1}, \dots, \lambda_N)$ , or  $\phi_1 = \text{mean}(\lambda_{n_0+1}, \dots, \lambda_{n_1})$  etc. depending on the number of groups assumed for the eigenvalues  $\lambda_n \gg 0$ ; (3) innovations are assumed to be independent; (4) without  $S_2$ , the model cannot be an eigenvalue approximation of the generic GATSM; (5) abbreviations are respectively Diebold et al. (2008), Heath et al. (1992), Nelson and Siegel (1987), Krippner (2006b), Christensen et al. (2011), Bliss (1997), Svensson (1995), Sharef and Filopovic (2004), Christensen et al. (2009), and NS model variants discussed in the present article.

As an example of interpreting table 1, the Diebold, Li and Yue (2008)/NS model with the forward rate form  $f(t) = L(t) + S(t) \exp(-\phi\tau)$  is the most parsimonious representation of the generic GATSM.<sup>11</sup> It represents both the persistent and non-

<sup>11</sup>This assumes one component to represent the near-zero group of eigenvalues and one component to represent the non-zero group of eigenvalues from the generic GATSM, which is arguably the minimal model one would want for empirical work. An “over-parsimonious” NS model would be to use just a Slope component, in which case the AF NS version would be the Vasicek (1977) model. The absolutely most parsimonious NS model would be to use just a Level component. That is the basis for the traditional “duration” calculations from Macauley (1938) which are often used to gauge the price-sensitivity of interest rate securities to a level shift in the yield curve. The AF NS version would be the Vasicek (1977) model with the limit of a zero mean-reversion parameter.

persistent components of the generic GATSM with a single term from the Taylor expansion and omits any AF adjustments. Interestingly, the two-factor model from Heath et al. (1992, pp. 91-92) turns out to be an AF version of the Diebold et al. (2008)/NS model, by coincidence of the assumed volatility functions for the forward rate curve being a constant and an exponential decay by time to maturity. However, the Heath et al. (1992) model assumes uncorrelated innovations; Variant 1 to be discussed and derived in section 5.2 effectively extends the Heath et al. (1992) model to allow for correlation (and then for time-varying market prices of risk in section 5.3).

At the other extreme, the Christensen, Diebold and Rudebusch (2009)/NS model is the most flexible model within the NS class to date. It has the forward rate form  $f(t, \tau) = L(t) + S_1(t) \exp(-\phi_1 \tau) + C_1(t) \phi_1 \tau \exp(-\phi_1 \tau) + S_2(t) \exp(-\phi_2 \tau) + C_2(t) \phi_2 \tau \exp(-\phi_2 \tau) + AF(\tau)$ , where  $AF(\tau)$  abbreviates the AF term from Christensen et al. (2009) in its forward rate form. In this model, the non-persistent components are represented with two groupings of non-zero eigenvalues (i.e.  $\phi_1 = \text{mean}(\lambda_{n_0+1}, \dots, \lambda_{n_1}) < \phi_2 = \text{mean}(\lambda_{n_1+1}, \dots, \lambda_N)$  each with two terms, and the appropriate AF terms is included to ensure the model is AF with respect to the five factor loadings. Variant 4 is a potential extension of the Christensen et al. (2009)/NS model that represents the persistent GATSM components with two terms of the Taylor expansion, thereby replacing  $L(t)$  with  $L_1(t) + L_2(t) \tau$ , where  $L_2(t) = -\sum_{n=1}^{n_0} q_n(t) \lambda_n$ .<sup>12</sup> Variant 5 is a six-component NS model that adds the third terms of the Taylor expansion for both the persistent and non-persistent components, i.e.  $L_3(t) \frac{1}{2} \tau^2$  with  $L_3(t) = \sum_{n=1}^{n_0} q_n(t) \lambda_n^2$ , and  $C^*(t) \frac{1}{2} (\phi \tau)^2 \exp(-\phi \tau)$  with  $C^*(t) = \sum_{n=n_0+1}^N q_n(t) \delta_n^2$ . Note that all of the variant NS models (and the two-component models) are “balanced” eigenvalue approximations of the generic GATSM, in the sense that the number of terms in the Taylor expansion is the same for the persistent and non-persistent components. That property has some theoretical appeal, because it guarantees the overall eigenvalue approximation of any GATSM to a given order.

The classification system suggests a systematic approach to selecting and applying an appropriate NS model for the application at hand rather than arbitrarily adding flexibility to seek a better fit to the yield curve data. Of particular note in this regard are the NS models of Bliss (1997), Svensson (1995), and Sharef and Filipović (2004) that add a second Curvature term  $C_2(t)$ . However, the latter by itself omits the first term of the eigenvalue approximation associated with the second group of non-zero eigenvalues for the generic GATSM. Hence, those models should be avoided if an explicit correspondence to the GATSM class is desired. Another aspect is more subtle: to maintain an explicit correspondence with the generic GATSM, NS models should be applied with a constant parameter  $\phi$  (or parameters  $\phi_1, \phi_2$ , etc.) because that parameter corresponds to the constant eigenvalues associated with the constant mean-reversion matrix  $K_{\mathbb{Q}}$  in the generic GATSM. Allowing the decay parameter/s to vary over time, which is often done with NS and Svensson (1995) models, would break the correspondence to the GATSM class of models.<sup>13</sup>

<sup>12</sup>That is,  $\exp(-\lambda_n \tau) \simeq 1 - \lambda_n \tau$ , and so  $\sum_{n=1}^{n_0} q_n(t) \exp(-\lambda_n \tau) \simeq \sum_{n=1}^{n_0} q_n(t) - \tau \cdot \sum_{n=1}^{n_0} q_n(t) \lambda_n$ .

<sup>13</sup>While it would be tempting to interpret time variation in  $\phi$  as representing time variation in the mean-reversion matrix  $K_{\mathbb{Q}}$ , a generic GATSM that formally allowed for such flexibility would not necessarily result in factor loadings reducible to the NS form using the eigenvalue approximation approach as in section 3. Of course, adding a second Curvature term  $C_2(t)$  and/or allowing time-varying parameter(s)  $\phi$  may be entirely appropriate if the purpose is simply to specify a flexible

Once the appropriate NS model has been chosen, one needs to consider whether the model should be made AF with respect to the NS components. Ideally, the AF NS term should be included in empirical applications to maintain theoretical consistency between the cross-sectional and time-series properties of the given NS model,<sup>14</sup> and correspondence back to the GATSM. Explicit estimates of the market prices of risk and the volatility parameters for the Gaussian innovations associated with the AF NS term may also provide useful information to the user (and are certainly essential when pricing instruments that are heavily influenced by interest rate volatility, such as options on fixed interest securities).

There are several points to note regarding the derivation of the AF NS term itself, whether via the Heath et al. (1992) framework as in Krippner (2006a) and section 5.2 of the present article or following the approach in Christensen et al. (2009, 2011). First, the AF NS term should be derived directly from the chosen base NS model to ensure the resulting model will be AF with respect to the NS components (which are themselves the eigenvalue approximation of the dynamic component of the GATSM). Conversely, a direct eigenvalue approximation of the generic GATSM including its AF terms would not necessarily guarantee a resulting AF model. Second, the component of the AF NS term purely associated with the volatility parameters for the Gaussian innovations in the NS model (denoted the “volatility effect” in the following section) will be a time-invariant function of time to maturity. That component corresponds to the time-invariant function of time to maturity component  $\frac{\partial}{\partial \tau} A(\tau)$  for the risk-neutral specification of the generic GATSM in equation 1. Third, the component of the AF NS term associated with the volatilities and the market prices of risk in the NS model represents a risk premium function (as denoted in the following section). That component explicitly accounts for the difference between the specifications of the AF NS model under the risk-neutral  $\mathbb{Q}$  measure and the physical  $\mathbb{P}$  measure. The risk premium can be time invariant (with constant market prices of risk) or time varying as a linear function of the NS coefficients (with an essentially-affine market price of risk specification), which are subsequently illustrated respectively in sections 5.2 and 5.3.

Notwithstanding the discussion above, non-AF NS models will sometimes be valid when applied to term structure data in some particular time-series contexts, for example to obtain constant time-to-maturity data that are subsequently used in a time-series regression. By way of explanation, in brief, the non-dynamic generic GATSM component  $\frac{\partial}{\partial \tau} A(\tau)$  and the volatility effect component of the AF NS term are both time-invariant functions of time to maturity, and the risk premiums for the respective models will either be time-invariant or a linear function of state variables/NS coefficients. Therefore, even though a non-AF NS model omits the AF NS term relative to the AF NS representation of the GATSM, that omission will typically be subsumed without loss in the constant of an OLS regression or a vector autoregressive model. However, applying a non-AF NS model will be invalid if simultaneous consistency between the cross-sectional and time-series perspectives of the yield curve is required or

---

function of time to maturity for obtaining zero-coupon interest rate data via atheoretical yield curve fitting (such as with the Gürkaynak, Sack and Wright (2008) data later used in section 6 of the present article). The point is that the results from such models cannot be interpreted from a GATSM perspective.

<sup>14</sup>The theoretical case for consistency was originally proposed in Björk and Christensen (1999) and further established in Filipović (1999, 2000).

implicitly assumed in the context of the application.

## 5 The two-factor arbitrage-free NS model

This section derives “Variant 1” mentioned in the previous section; i.e. the two-factor AF/NS model, hereafter denoted the AF/NS(2) model, based on just the Level and Slope components with an allowance for correlations between innovations to the Level and Slope. The motivation, from a macrofinance perspective, for developing the AF/NS(2) model is discussed in section 5.1. Section 5.2 derives the AF/NS(2) with constant market prices of risk (hence time-invariant risk premiums, resulting in an affine model), and is denoted the A/AF/NS(2) model. Section 5.3 extends the AF/NS(2) model with an essentially affine specification for the market prices of risk (hence resulting in time-varying risk premiums), and is denoted the EA/AF/NS(2) model.

### 5.1 A macrofinance motivation for the two-factor arbitrage-free NS model

The AF/NS(2) model should prove particularly useful as a standard tool for yield curve modeling and analysis in macrofinance for the following reasons: (1) it is very parsimonious, and so easy to apply empirically; (2) it is representative of any GATSM that might be generating the yield curve data; and (3) empirical evidence from related literature suggests that two-factors may be appropriate macrofinance applications.

To highlight the parsimony, the AF/NS(2) model with time-varying risk premiums will have just 10 parameters. By comparison, Christensen et al. (2011) p. 9 gives the number of parameters in the three-factor AF/NS model as 19, and the three-factor GATSM as 22.<sup>15</sup> With constant market prices of risk, which may prove adequate for some macrofinance applications, the number of parameters respectively become 6, 10, and 13. Of course, when yield curve data are modeled in conjunction with other financial and/or macroeconomic data, the proliferation of parameters will obviously be correspondingly less for the AF/NS(2) model than for three or more factors. Another welcome aspect related to the relative parsimony of the AF/NS(2) model is the ease of analytical manipulation. Specifically, the AF/NS(2) model derivations that follow in sections 5.2 and 5.3 result in very compact bivariate state-space representations. In addition, with time-varying risk premiums (i.e. for the EA/AF/NS(2) model) I am able to derive a closed-form analytic constraint for eigenvalues that ensures stability during estimation.

While being parsimonious, the eigenvalue approximation foundation for NS models established in sections 2 to 4 ensures that the AF/NS(2) model will provide a reliable representation of any GATSM that might be generating the yield curve data. Specifically, regardless of the number of factors that might be responsible for generating the yield curve data, their nature (e.g. economic, financial, etc.), and the potential complexity of their interactions (in terms of the mean-reversion and innovation matrices),

---

<sup>15</sup>The number of parameters in the three-factor GATSM is 28, but nine are lost with identifiability restrictions. A further three restrictions are imposed to obtain the three-factor AF/NS model. See Christensen et al. (2011) p. 9 for further details and discussion.

the Level component of the NS model will represent the persistent components of the GATSM, and the Slope component will represent the non-persistent components. The AF term allowing for correlated innovations to the Level and Slope components and a specification for the market prices of risk will enforce arbitrage-free self-consistency between the cross-sectional and time-series properties of the model.

Besides the case in principle discussed already for the AF/NS(2) model, available empirical evidence also suggests that two factors may be appropriate for macrofinance applications. For example, Diebold, Rudebusch and Aruoba (2006) empirically (and intuitively) associates the NS Level component with measures of inflation and the NS Slope component with measures of real activity and the central bank policy rate. Conversely, no apparent empirical macroeconomic interpretation (or intuition) for the Curvature component is found (from Diebold, Rudebusch and Aruoba (2006) p. 320: “Unfortunately ... we know of no reliable macroeconomic links to  $C_t$ .”). Another NS-related macrofinance application that succeeds admirably with just Level and Slope components is Diebold et al. (2008). That article investigates international yield curve linkages, finding very material common global Level and Slope components along with interesting regional differences.

At the same time, empirical evidence from Gürkaynak, Sack and Wright (2005a,b) suggests that more than one factor is required for macrofinance applications. For example, Gürkaynak, Sack and Swanson (2005a) provides evidence that the dynamics in long-term interest rates require an additional factor relative to innovations in short-term interest rates (consistent with the Diebold, Rudebusch and Aruoba (2006) association of the NS Level component with inflation measures, Gürkaynak et al. (2005a) also develops a model to explain long-term interest rates dynamics in terms of changing expectations of steady-state inflation). Even for relatively short times to maturity, Gürkaynak, Sack and Swanson (2005b) finds that a “current federal funds rate target” factor and a “future path of policy” factor are required to represent the impact of Federal Open Market Committee statements.

Of course, the AF/NS(2) model is unlikely to be appropriate for applications that require a close representation of yield curve data at all points in time, such as in financial markets. In such instances, one might still prefer the additional flexibility offered by a three-factor AF/NS model with the knowledge that it provides one better degree of eigenvalue approximation to the “true” GATSM generating the data (or more than three factors for an even closer eigenvalue approximation). In any case, applications purely in finance would likely prompt one to consider moving beyond the NS class of models to avoid the implicit constraint of constant volatility (and the negative interest rate issue) that NS models inherit from the GATSM class.

## 5.2 The A/AF/NS(2) model

### 5.2.1 Specifying and deriving the model

The NS model with two factors is  $\beta_1(t) + \beta_2(t) \cdot \exp(-\phi\tau)$ . I use that expression within the Heath et al. (1992) framework to derive the associated forward rate expression, including the modification from Tchuindjo (2008) that allows the factors to have

innovations  $\Omega^{1/2} [dW_1(t), \exp(-\phi\tau) \cdot dW_2(t)]'$  with a non-zero correlation, i.e.

$$\Omega = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \quad (5)$$

where  $\sigma_1$  and  $\sigma_2$  are the factor volatilities (annualized standard deviations of the factor innovations),  $\rho$  is the correlation of factor innovations, and  $dW_1(t)$  and  $dW_2(t)$  are independent Wiener increments.

The expression for the AF/NS(2) forward rate curve in the Tchuindjo (2008)/Heath et al. (1992) framework is therefore:

$$\begin{aligned} f(t, \tau) = & [1, \exp(-\phi\tau)] \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \\ & + \int_0^\tau [\sigma_1, \sigma_2 \exp(-\phi[\tau-s])] \begin{bmatrix} \gamma_{0,1} \\ \gamma_{0,2} \end{bmatrix} ds \\ & - \int_0^\tau \left\{ [\sigma_1, \sigma_2 \exp(-\phi[\tau-s])] \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right. \\ & \quad \left. \times \left( \int_s^\tau \begin{bmatrix} \sigma_1 \\ \sigma_2 \exp(-\phi[u-s]) \end{bmatrix} du \right) \right\} ds \end{aligned} \quad (6)$$

where the first line is  $\mathbb{E}_t[r(t+\tau)]$ , the expected value, conditional upon information available at time  $t$ , of the instantaneous short rate at time  $t+\tau$ ; the second line is the risk premium function involving the factor volatilities and the constant market prices of risk  $\gamma_{0,1}$  and  $\gamma_{0,2}$ ; the third line is the volatility effect involving the factor volatilities and their correlation  $\rho$ ; and  $u$  and  $s$  are dummy integration variables. Note that I define  $\gamma_{0,1}$  and  $\gamma_{0,2}$  as positive quantities in this article, so they (intuitively) add positive spreads to  $\mathbb{E}_t[r(t+\tau)]$  and therefore the observed yield curve (where, as usual, they impact negatively on security prices).

The solution to the forward rate expression in equation 6 is:<sup>16</sup>

$$\begin{aligned} f(t, \tau) = & \beta_1(t) + \beta_2(t) \cdot \exp(-\phi\tau) \\ & + \sigma_1\gamma_{0,1} \cdot \tau + \sigma_2\gamma_{0,2} \cdot F(\phi, \tau) \\ & - \sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2}[F(\phi, \tau)]^2 \\ & - \rho\sigma_1\sigma_2 \cdot \tau F(\phi, \tau) \end{aligned} \quad (7)$$

where  $F(\phi, \tau) = \frac{1}{\phi} [1 - \exp(-\phi\tau)]$ . The expression for  $f(t, \tau)$  may be more conveniently expressed as  $f(t, \tau) = a(\tau) + b(\tau)\beta(t)$ , where  $a(\tau) = \sigma_1\gamma_{0,1} \cdot \tau + \sigma_2\gamma_{0,2} \cdot F(\phi, \tau) - \sigma_1^2 \cdot \frac{1}{2}\tau^2 - \sigma_2^2 \cdot \frac{1}{2}[F(\phi, \tau)]^2 - \rho\sigma_1\sigma_2 \cdot \tau F(\phi, \tau)$ ,  $b(\tau) = [1, \exp(-\phi\tau)]$ , and  $\beta(t) = [\beta_1(t), \beta_2(t)]'$ .

Forward rates are not typically the observable quantities that define the yield curve in practice, so a conversion to interest rates  $R(t, \tau)$  is required. Hence, I use the standard relationship  $R(t, \tau) = \frac{1}{\tau} \int_0^\tau f(t, \tau) d\tau$  to derive:

---

<sup>16</sup> All the results in this and the following section follow from straightforward but lengthy calculus and algebra. Full workings of all the results are contained in appendix B of the working paper version of this article.

$$R(t, \tau) = \bar{a}(\tau) + \bar{b}(\tau) \beta(t) \quad (8)$$

where:

$$\bar{b}(\tau) = \left[ 1, \frac{1}{\tau} F(\phi, \tau) \right] \quad (9)$$

and:

$$\bar{a}(\tau) = \text{RP}(\tau) + \text{VE}(\tau)$$

with:

$$\text{RP}(\tau) = \sigma_1 \gamma_{0,1} \cdot \frac{1}{2} \tau + \sigma_2 \gamma_{0,2} \cdot \frac{1}{\phi} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) \right] \quad (10)$$

and:

$$\begin{aligned} \text{VE}(\tau) = & -\sigma_1^2 \cdot \frac{1}{6} \tau^2 - \sigma_2^2 \cdot \frac{1}{2\phi^2} \left( 1 - \frac{1}{\tau} F(\phi, \tau) - \frac{1}{2\tau} \phi [F(\phi, \tau)]^2 \right) \\ & - \rho \sigma_1 \sigma_2 \cdot \frac{1}{\phi^2} \left( 1 - \frac{1}{\tau} F(\phi, \tau) + \frac{1}{2} \phi \tau - \phi F(\phi, \tau) \right) \end{aligned} \quad (11)$$

The interest rate as a function of time to maturity  $\tau$  is therefore composed of: (1)  $\frac{1}{\tau} \int_0^\tau \mathbb{E}_t[r(t + \tau)] d\tau = \bar{b}(\tau) \beta(t)$ , i.e. the average of the expected path of the short rate up to  $\tau$ , which is obviously a quantity that varies over time due to innovations in the state variables  $\beta_1(t)$  and  $\beta_2(t)$ ; (2) the time-invariant risk premium function  $\text{RP}(\tau)$ , which represents the combined effect of the quantities and market prices of risk (i.e. innovation volatilities and the compensation required by the market to accept the unanticipated price effects associated with those innovations); and (3) the time-invariant volatility effect  $\text{VE}(\tau)$ , which represents the influence of volatility on expected returns due to Jensen's inequality (i.e. the expected compounded return from investing in a volatile short rate is less than the compounded return from investing in the expected [or mean] short rate).

### 5.2.2 State space representation

An observation of zero-coupon continuously compounding yield curve data at time  $t$  may be represented as:

$$R(t) = \bar{A} + \bar{B} \beta(t) + v(t) \quad (12)$$

where  $R(t)$  is the  $K \times 1$  vector of yield curve data (with  $K = 10$  or  $12$  in the empirical application below),  $\bar{A}$  is the  $K \times 1$  vector  $[\bar{a}(\tau_1), \dots, \bar{a}(\tau_k), \dots, \bar{a}(\tau_K)]'$ ,  $\bar{B}$  is the  $K \times 2$  matrix  $[\bar{b}(\tau_1), \dots, \bar{b}(\tau_k), \dots, \bar{b}(\tau_K)]$ , and  $\tau_1, \dots, \tau_k, \dots, \tau_K$  are the times to maturity of the yield curve data.

The evolution of  $\beta(t)$  over a finite time step  $\Delta t$  may be derived directly as:

$$\beta(t + \Delta t) = \Phi(\phi, \Delta t) \beta(t) + \varepsilon(t + \Delta t) \quad (13)$$

where  $\Phi(\phi, \Delta t) = \text{diag}[1, \exp(-\phi \Delta t)]$ , a  $2 \times 2$  diagonal matrix.

Equations 12 and 13 are a measurement and state equation that may be used in the Kalman filter to estimate the model. Regarding the additional elements required for the Kalman filter, the covariance matrix for the state equation may be evaluated

as:

$$\begin{aligned}
Q &= \int_0^{\Delta t} \Phi(\phi, s) \Omega[\Phi(\phi, s)]' ds \\
&= \begin{bmatrix} \sigma_1^2 \cdot \Delta t & \rho \sigma_1 \sigma_2 \cdot F(\phi, \Delta t) \\ \rho \sigma_1 \sigma_2 \cdot F(\phi, \Delta t) & \sigma_2^2 \cdot F(2\phi, \Delta t) \end{bmatrix}
\end{aligned} \tag{14}$$

and the covariance matrix for the measurement equation is assumed to take the form  $H = \text{diag}[\sigma_v^2(\tau_1), \dots, \sigma_v^2(\tau_K)]$ . That assumption is standard in the literature, as is the assumption that all other contemporaneous and intertemporal covariances are zero.

The starting values for the state variable vector and its covariance matrix are  $\beta_{1|0} = [0, 0]'$  and:

$$P_{1|0} = \begin{bmatrix} \sigma_\eta^2 / (1 - \psi^2) & \rho \sigma_1 \sigma_2 \cdot \frac{1}{\phi} \\ \rho \sigma_1 \sigma_2 \cdot \frac{1}{\phi} & \sigma_2^2 \cdot \frac{1}{2\phi} \end{bmatrix} \tag{15}$$

which are respectively the unconditional expectations for  $\beta(t)$  and its covariance with appropriate substitutions for the values purely associated with  $\beta_1(t)$ . The latter are technically undefined because  $\beta_1(t)$  is a unit-root process and so I substitute estimates based on a near-unit-root autoregression; i.e.  $\psi$  is the coefficient from the autoregression  $\beta_1^*(t + \Delta t) = \psi \cdot \beta_1^*(t) + \eta(t)$  and  $\beta_1^*(t)$  is the Level coefficient series obtained from a preliminary estimation of the non-AF two-factor NS model by OLS.<sup>17</sup>

### 5.2.3 Econometric discussion

It is worthwhile at this stage making several observations about the A/AF/NS(2) model regarding estimation and identification, given those issues have typically presented challenges for the estimation of latent-factor GATSMs in the past.<sup>18</sup> These observations apply equally to the EA/AF/NS model in the following section.

First, the A/AF/NS(2) model is globally identified. Following the discussion in Collin-Dufresne, Goldstein and Jones (2008) section I, that property is assured because the A/AF/NS(2) model effectively imposes the restriction that the mean-reversion parameter for the second factor (i.e.  $\phi > 0$ ) will always be greater than for the first factor (i.e. 0). That means the AF/NS(2) model estimation will have a unique maximum and estimated parameter set.

Second, the market prices of risk for both factors in the A/AF/NS(2) can be identified from an unrestricted estimation with a sample of zero-coupon data. That property arises because the Level component for the A/AF/NS(2) model subsumes the constant parameter for the mean level of the short rate in the generic GATSM. Conversely, Sin-

<sup>17</sup>Hamilton (1994) chapter 13 provides general background on the Kalman filter, and pp. 378-79 discusses the substitution of appropriate estimates for initial values and covariances when the Kalman filter contains unit roots. Note that the process may be iterated (using updated estimates of  $\beta_1^*(t)$  to update the parameter estimate  $\psi$ , etc.) which is the basis for the A/AF/NS(2) model estimates that I report subsequently in section 6.3. However, the additional iterations make an immaterial difference to the initial results. The estimates of  $\sigma_\eta^2 / (1 - \psi^2)$  are around 3 percentage points, which is quite diffuse (as would be expected from a near-unit-root process). Also note that the A/AF/NS(2) model and its Kalman filter set-up is equivalent to the bivariate latent-factor GATSM and set-up from Babbs and Nowman (1999), but with a limit of zero mean reversion for one of the factors.

<sup>18</sup>See the discussion in Hamilton and Wu (2010). That article and Joslin et al. (2011) have proposed techniques to alleviate identification and estimation issues associated with latent-factor GATSMs.



ingleton (2006) pp. 342-343 notes that the mean short rate parameter and the constant market prices of risk cannot all be identified for a latent-factor GATSM estimated with zero-coupon data.

Third, the implicit econometric identification of the AF/NS(2) model uses  $\theta_{\mathbb{P}} = 0$ , rather than  $\theta_{\mathbb{Q}} = 0$  as suggested in Singleton (2006) pp. 318-319 for the canonical GATSM. The practical implication is that the market prices of risk appear in the measurement equation rather than the state equation, which is convenient because it makes immediately transparent the risk premium component that underlies the yield curve data, i.e. equation 10. Conversely, the identification  $\theta_{\mathbb{Q}} = 0$ , which is used for the AF/NS model in Christensen et al. (2011), eliminates the constant market prices of risk from the measurement equation and embeds them in the constant for the state equation. Nevertheless, both identifications provide observationally equivalent representations of the data given they are within affine-invariant transformations of each other (see Singleton (2006) pp. 319-321).

Finally, from a practical perspective, standard reparametrizations are employed to ensure trouble-free numerical evaluation of the Kalman filter; i.e. a Cholesky specification to ensure the covariance matrices  $Q$  and  $P$  always remains positive definite, and  $\rho = \omega / (1 + |\omega|)$  to respect the  $\pm 1$  range for innovation correlations.<sup>19</sup>

### 5.3 The EA/AF/NS(2) model

The A/AF/NS(2) model may readily be extended to the EA/AF/NS(2) model using an essentially affine specification for the market prices of risk; i.e.  $\Gamma(t) = \gamma_0 + \gamma_1 \beta(t)$ , where  $\gamma_0 = [\gamma_{0,1}, \gamma_{0,2}]'$  and  $\gamma_1$  is a  $2 \times 2$  matrix of constants.

Following Dai and Singleton (2002) appendix A, the measurement equation for the EA/AF/NS(2) model has the same form as the A/AF/NS(2) model, i.e. equation 12, but the functions  $\bar{a}(\tau_k)$  in the  $K \times 1$  vector  $\bar{A} = [\bar{a}(\tau_1), \dots, \bar{a}(\tau_k), \dots, \bar{a}(\tau_K)]'$  now incorporate the following risk premium function:

$$\text{RP}(t, \tau) = \text{RP}(\tau) + \left( \left[ 1, \frac{1}{\tau} F(\phi, \tau) \right] - [1, 1] [\kappa \tau]^{-1} [I - \exp(-\kappa \tau)] \right) \beta(t) \quad (16)$$

where  $\text{RP}(\tau)$  is the expression in equation 10. Therefore, the essentially affine specification for the market prices of risk allows the risk premium function to vary as a linear function of the state variables.

Relative to equation 13 for the A/AF/NS(2) model, the state equation for the EA/AF/NS(2) model is modified by the matrix exponential  $\exp(-\gamma_1 \Delta t)$  to give:

$$\begin{aligned} \beta(t + \Delta t) &= \Phi(\phi, \Delta t) \exp(-\gamma_1 \Delta t) \beta(t) + \varepsilon(t + \Delta t) \\ &= \exp(-\kappa \Delta t) \beta(t) + \varepsilon(t + \Delta t) \end{aligned} \quad (17)$$

where  $\kappa = \text{diag}[0, \phi] + \gamma_1$ .

---

<sup>19</sup>See Hamilton (1994) for details and discussion.

The covariance matrix for the state equation may be evaluated as:

$$\begin{aligned} Q &= \int_0^{\Delta t} \exp(-\kappa s) \Omega \exp(-\kappa' s) ds \\ &= V \begin{bmatrix} u_{11} \cdot F(2d_1, \Delta t) & u_{12} \cdot F(d_1 + d_2, \Delta t) \\ u_{21} \cdot F(d_1 + d_2, \Delta t) & u_{22} \cdot F(2d_2, \Delta t) \end{bmatrix} V' \end{aligned} \quad (18)$$

where  $VDV^{-1}$  is the eigensystem decomposition of  $\kappa$ ,  $D = \text{diag}[d_1, d_2]$ , and the elements  $u_{ij}$  are those from  $U = V^{-1}\Omega(V^{-1})'$ . The covariance matrix for the measurement equations is again assumed to be  $H = \text{diag}[\sigma_v^2(\tau_1), \dots, \sigma_v^2(\tau_K)]$ .

The starting values for the state variables and their covariance are the unconditional expectations, respectively  $\beta_{1|0} = [0, 0]'$  and

$$\begin{aligned} P_{1|0} &= \int_0^\infty \exp(-\kappa s) \Omega \exp(-\kappa' s) ds \\ &= V \begin{bmatrix} u_{11} \frac{1}{2d_1} & u_{12} \frac{1}{d_1 + d_2} \\ u_{21} \frac{1}{d_1 + d_2} & u_{22} \frac{1}{2d_2} \end{bmatrix} V' \end{aligned} \quad (19)$$

### 5.3.1 EA/AF/NS(2) eigenvalue restriction

The system for EA/AF/NS(2) model requires restrictions to ensure that  $\kappa$  has positive eigenvalues. From a practical perspective, this can be viewed as an additional restriction to ensure trouble-free estimation, i.e. to ensure the covariance matrices  $Q$  and  $P$  always remains positive definite and to avoid “explosive” dynamics. From a theoretical perspective, the restriction also ensures consistency with the assumption from Dai and Singleton (2002) that the eigenvalues of the generic GATSM are strictly positive (which in turn means the eigenvalues of  $\exp(-\kappa\Delta t)$  will be less than 1 and so  $\beta(t)$  will be stationary).<sup>20</sup>

One advantage offered by the parsimony of the EA/AF/NS(2) model is that the required eigenvalue constraint may readily be manipulated into the closed-form analytic solution of a quadratic equation, as I summarize below. Expressing the eigenvalue constraint in such a form facilitates the reliable and timely empirical application of the model.

The eigenvalue constraint is enforced as follows: from Higgam (1996) p. 223 a non-symmetric matrix may be written as the sum of a symmetric matrix  $A_S$  and an antisymmetric matrix  $A_K$ , and  $A$  will be positive definite if  $A_S$  is positive definite. The latter can be generated with three additional parameters as  $A_S = LL' + \text{diag}[0, \phi]$ , where  $L$  is a  $2 \times 2$  lower-diagonal matrix. The antisymmetric matrix can be generated with one additional parameter  $d$  and then setting  $A_{12} = -A_{21} = d$  and  $A_{11} = A_{22} = 0$ . Directly calculating  $\kappa = A_S + A_K$  and its eigenvalues gives:

$$\text{eig} \begin{bmatrix} a & b+d \\ b-d & c \end{bmatrix} = \frac{1}{2}c + \frac{1}{2}a \pm \frac{1}{2}\sqrt{(a-c)^2 - 4d^2 + 4b^2} \quad (20)$$

---

<sup>20</sup>This restriction could easily be modified if one wanted to allow for the possibility of a pair of complex conjugate eigenvalues with positive real parts. From an economic perspective, that would correspond to an expectation that innovations to the state variables would follow the product of a sinusoidal cycle and an exponential decay (rather than just an exponential decay) when returning to equilibrium.

and a reparametrization:

$$d = \frac{e}{1 + |e|} \cdot \frac{1}{2} \sqrt{(a - c)^2 + 4b^2} \quad (21)$$

ensures that  $d$  will result in a positive value for the square root operand, therefore guaranteeing real positive eigenvalues for  $\kappa$ .

## 6 An application to U.S. yield curve dynamics

This section applies the AF/NS(2) models developed in the prior section to investigate changes in the data-generating process for the U.S. yield curve from 1971 to 2010. A similar application has already undertaken in Rudebusch and Wu (2007) over the period 1971 to 2002 using a two-factor GATSM with essentially affine market prices of risk, but section 6.1 discusses the distinctions between that application and the AF/NS(2) applications. Section 6.2 provides an overview of the yield curve data for the 1971 to 2010 period and the three subsamples I investigate. Section 6.3 applies the AF/NS(2) model to the full sample and three subsamples, and section 6.4 similarly applies the EA/AF/NS(2) model. I also compare the EA/AF/NS(2) results to the AF/NS(2) results to investigate whether time-invariant risk premiums offer an adequate representation of the data.

### 6.1 The AF/NS(2) models versus Rudebusch and Wu (2007)

The main difference between the Rudebusch and Wu (2007) and AF/NS(2) model applications in the following sections is conceptual; i.e. how the different models purport to represent the data. Put simply, Rudebusch and Wu (2007) implicitly takes the perspective that the data are generated by a two-factor GATSM. Conversely, applying the AF/NS(2) models implicitly acknowledges that the data could be generated by a GATSM with many factors, and the results represent the eigenvalue approximation of that model. Therefore, the AF/NS(2) results will be representative of the results from any GATSM that could (in principle) be applied to the data, but without explicitly needing to undertake such applications.

From an econometric perspective, the AF/NS(2) model offers more convenient identification and estimation than Rudebusch and Wu (2007), as discussed in section 5.2.3. However, there is no need to detail those relative advantages here because such issues for latent-factor GATSMs have been superseded by the estimation methods of Hamilton and Wu (2010) and Joslin et al. (2011).

From an entirely practical perspective, the maturity span and time span of the data differ in my application relative to Rudebusch and Wu (2007). Specifically, Rudebusch and Wu (2007) uses 1-month to 5-year data while, as detailed below, I use 3-month to 15-year or 30-year data (depending on availability). Regarding the time span, I also add a third subsample from 2002, simply to capture the data from where the Rudebusch and Wu (2007) sample ended. However, the results for sample C are reported only out of interest and for comparison to the estimates over samples A and B, and they should not necessarily be taken as an advocacy to apply NS models over this period. That caution is warranted because it is theoretically inconsistent to represent the nominal

yield curve in a near-zero interest rate environment with a yield curve model that cannot respect the zero bound for nominal interest rates.<sup>21</sup>

Regarding a comparison of the empirical results to be outlined subsequently below, the results for the AF/NS(2) models broadly confirm those from Rudebusch and Wu (2007). In brief, there was a material change in the yield curve data-generating process between the periods 1971 to 1987 and 1988 to 2002). The results for the AF/NS(2) models also confirm the Rudebusch and Wu (2007) finding that changes to average risk premiums made the main contribution to the change in the data-generating process.

## 6.2 Yield curve data

Figure 1 provides an overview of the U.S. yield curve and its dynamics by plotting the 3-month and 15-year government-risk interest rate data (as detailed further below), and also the spread between those two rates. The latter is defined opposite to the common convention of a long-maturity rate less a short-maturity rate so it coincides qualitatively with the inverted shape of the Slope function in the NS models. Therefore, troughs in the spread represent periods of easy monetary policy and values above zero represent an inverted yield curve.

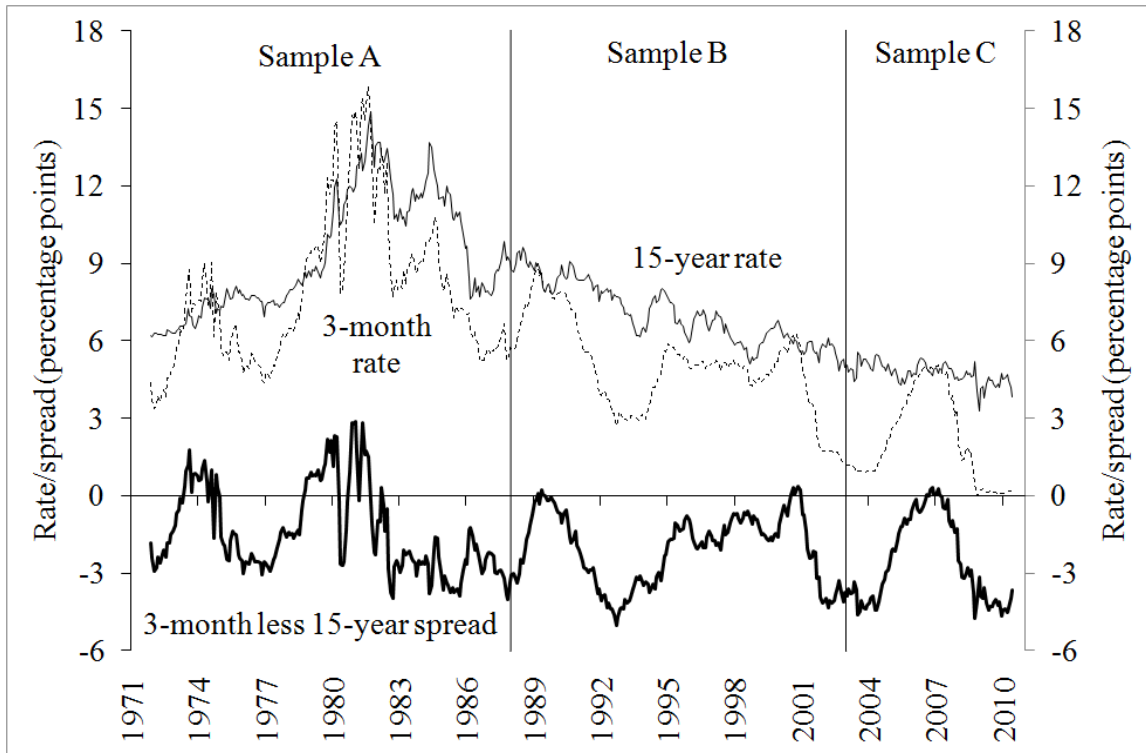


Figure 1: 3-month and 15-year interest rate data and the spread between those two rates.

The full sample period is divided into three samples. Sample A is from November 1971 (the beginning of the 15-year data) to December 1987, which matches the end

<sup>21</sup>Krippner (2012) p. 2 discusses that issue in greater detail and proposes a resolution using a shadow Gaussian term structure model in conjunction with bond options to impose the zero lower bound constraint.

of sample A from Rudebusch and Wu (2007) as determined by structural break tests. Sample B is from January 1988 to December 2002, which matches sample B from Rudebusch and Wu (2007). Sample C, from January 2003 to June 2010 (the latest data available at the time of the analysis), simply contains the additional data available relative to Rudebusch and Wu (2007). Sample C is arguably a unique period in its own right, because it conveniently begins amid the onset of U.S. deflation concerns in late-2002/early-2003 (e.g. see Billi, 2009 p. 83) and ends with the ultra-easy U.S. monetary policy following the onset of the Global Financial Crisis during 2007. The caveat, as already mentioned at the end of section 6.1, is that the data during period C is unlikely to be well-represented with a Gaussian model.

The maturity span of the available yield curve data changes over the full sample, reflecting the longest-maturity bond on issue at any point in time. Sample A uses 3- and 6-month Treasury bill rates (from the Federal Reserve Economic Database on the St. Louis Federal Reserve website, converted to a continuously compounding basis) and the 1-, 2-, 3-, 4-, 5-, 7-, 10-, and 15-year continuously compounding zero-coupon government interest rates from the data set described in Gürkaynak et al. (2008). Sample B is estimated with data of the same maturity span (hereafter denoted 3-m/15-y) to allow a direct comparison to sample A, and also with the addition of the Gürkaynak et al. (2008) 20- and 30-year data (which became available in July 1981 and November 1985 respectively). Sample C is estimated with data of the latter maturity span (hereafter denoted 3-m/30-y) to allow a direct comparison to sample B. All of the data are month-end rates taken from the original sets of daily data.

### 6.3 Estimation results for the A/AF/NS(2) model

The Kalman filter recursion is used to evaluate the log likelihood for the model, and the latter is maximized numerically using the Broyden-Fletcher-Goldfarb-Shanno algorithm as supplied in the “fminunc” function of the Matlab optimization toolbox. The asymptotic standard errors are calculated using the Hessian matrix evaluated at the parameter values that maximize the likelihood function.

Consistent with the discussion on global identification from section 5.2.3, convergence for the A/AF/NS(2) model estimation was timely and reliable with no apparent sensitivity in the end result to different starting values. Figure 2 illustrates the resulting state variables, i.e. the A/AF/NS(2) Level and Slope coefficients  $\beta_1(t)$  and  $\beta_2(t)$ , for the estimation using the 3-month to 15-year yield curve data over the entire sample. The A/AF/NS(2) Level and Slope coefficients respectively reflect the level and slope of the yield curve as represented by the 15-year rate and 3-month less 15-year spread in figure 1, and the model-implied short rate  $r(t) = \beta_1(t) + \beta_2(t)$  reflects the 3-month rate. Note that  $r(t)$  falls materially below zero in sample C, which further suggests that NS models (and so, by implication, GATSMs) are likely to be too simplistic for modeling the yield curve within a low to near-zero nominal interest rate environment (particularly if it were important to strictly respect the zero bound for a given application). The state variable estimates for the individual sample periods and using the 3-m/30-y data are very similar to figure 2, and so are not separately reported.

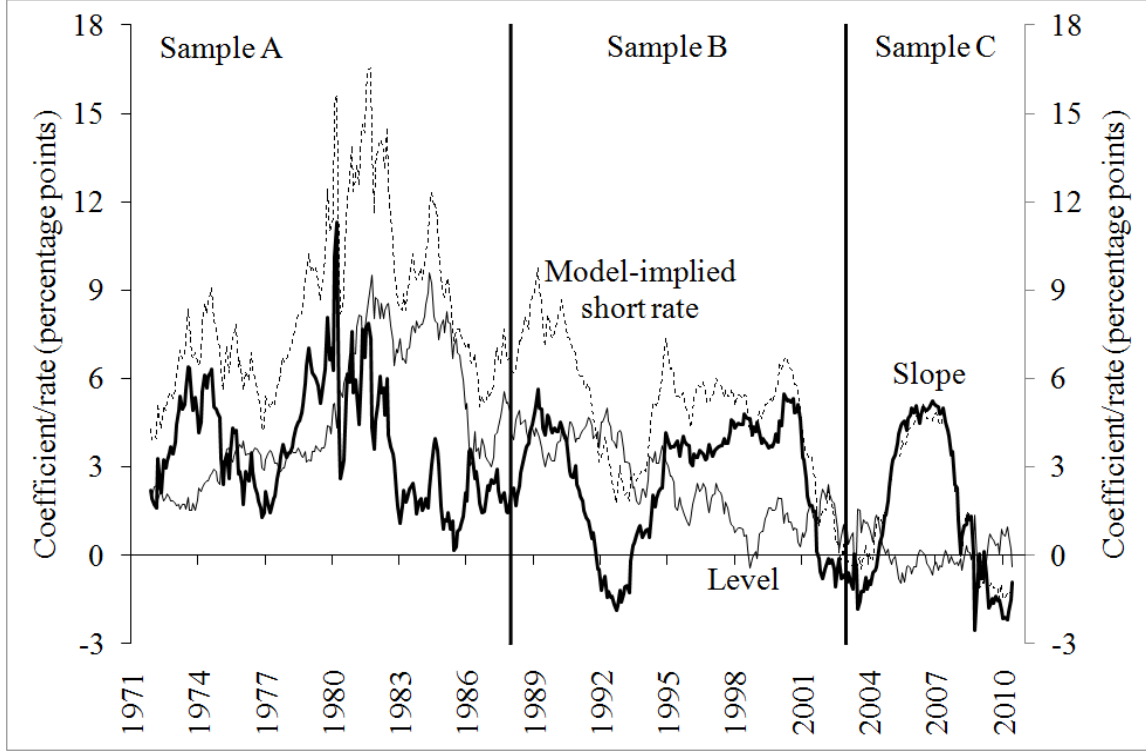


Figure 2: estimated Level and Slope coefficients and the model-implied short rate, i.e.  $\beta_1(t)$ ,  $\beta_2(t)$ , and  $r(t) = \beta_1(t) + \beta_2(t)$ , for the A/AF/NS(2) model estimated over the full sample using the 3-m/15-y data.

Table 2 contains the estimated parameter values for the A/AF/NS(2) model using the 3-m/15-y data over the joint sample A+B and the individual samples A and B.<sup>22</sup> Table 3 contains the estimated parameter values for the A/AF/NS(2) model using the 3-m/30-y data over the joint sample B+C and the individual samples B and C.

Table 2: Parameter estimates for the A/AF/NS(2) model with 3-m/15-y data						
Parameter	Sample A+B		Sample A		Sample B	
$\phi$	0.4994	(0.0083)	0.7028	(0.0140)	0.3931	(0.0082)
$\gamma_{0,1}$	0.1428	(0.0029)	0.1255	(0.0033)	0.1514	(0.0046)
$\gamma_{0,2}$	0.3079	(0.0130)	0.4659	(0.0223)	0.2223	(0.0169)
$\sigma_1$	0.0225	(0.0004)	0.0241	(0.0004)	0.0206	(0.0007)
$\sigma_2$	0.0339	(0.0014)	0.0367	(0.0018)	0.0295	(0.0021)
$\rho$	0.5729	(0.0307)	0.6728	(0.0427)	0.5257	(0.0403)
log-L	17424.7		9706.7		8803.3	
H <sub>0</sub> :A=B	910.5	[0.0000]				
Note: (standard errors), [probabilities]						

<sup>22</sup>The estimated parameters  $\sigma_v^2(\tau_k)$  are not reported here and for the EA/AF/NS(2) model in the following section. That is both to save space and because the model fit is not a focus of the analysis. As an indication, the typical values were respectively 0.10 and 0.16 percentage points for the 3-m/15-y and 3-m/30-y data.

Table 3: Parameter estimates for the A/AF/NS(2) model with 3-m/30-y data						
Parameter	Sample B+C		Sample B		Sample C	
$\phi$	0.3733	(0.0059)	0.3884	(0.0069)	0.3355	(0.0078)
$\gamma_{0,1}$	0.1410	(0.0026)	0.1435	(0.0031)	0.1192	(0.0043)
$\gamma_{0,2}$	0.2392	(0.0114)	0.2895	(0.0151)	0.0158	(0.0183)
$\sigma_1$	0.0172	(0.0002)	0.0172	(0.0002)	0.0191	(0.0003)
$\sigma_2$	0.0238	(0.0011)	0.0250	(0.0013)	0.0213	(0.0015)
$\rho$	0.3136	(0.0381)	0.4098	(0.0380)	-0.3324	(0.0621)
log-L	13621.6		10292.5		5095.4	
H <sub>0</sub> :B=C	3532.4	[0.0000]				
Note: (standard errors), [probabilities]						

The first aspect of note is that the hypotheses of no change in the yield curve data-generating process between samples A and B, and between samples B and C are soundly rejected (the log-likelihood ratio statistics are 910.5 and 3532.4 respectively). Figure 3 illustrates that one of the main points of difference between the two samples is that the (time invariant) risk premium function is smaller in sample B than in sample A. That in turn mainly reflects a lower market price of risk for the Slope component, but also somewhat lower volatilities for the Level and Slope components. The risk premium function falls again in sample C, mainly due to a further fall in the market price of risk for the Slope component.

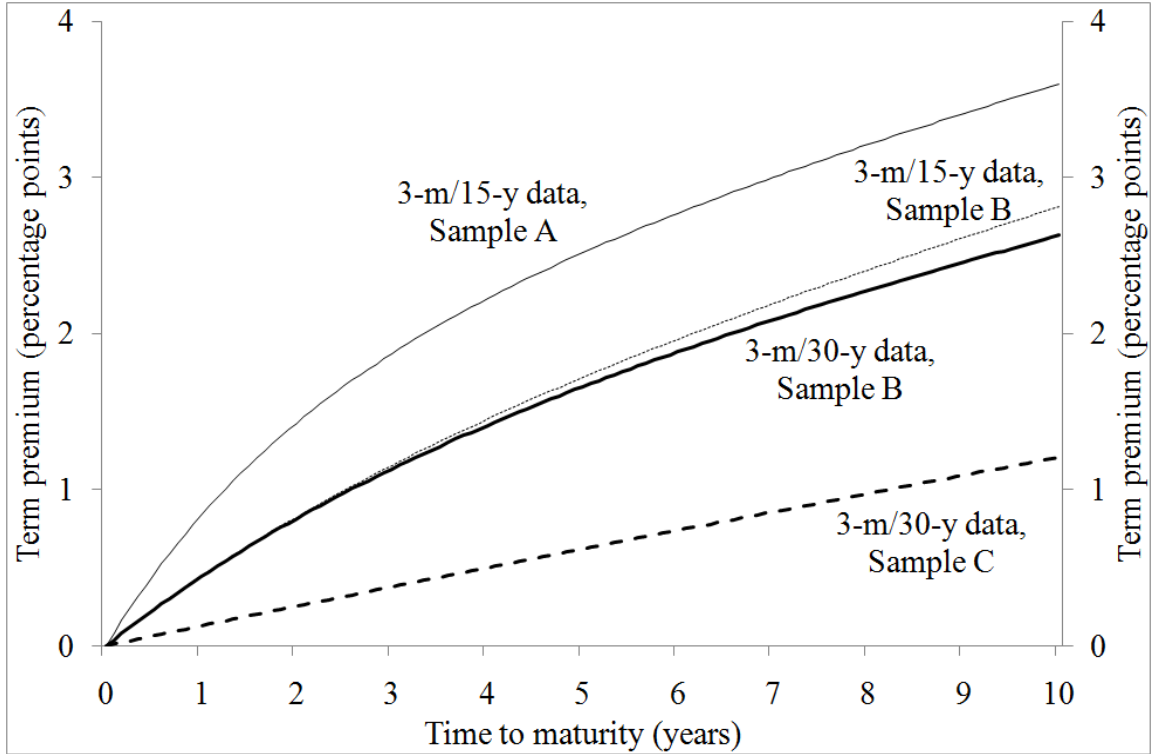


Figure 3: risk premium functions as implied by the point estimates of the A/AF/NS(2) model parameters in each sample, using the 3-m/15-y data and/or 3-m/30-y data as indicated.

Other points of note for the A/AF/NS(2) model estimates are the material decline in the mean-reversion parameter  $\phi$  for the Slope coefficient from sample A to B, and the sign reversal of the innovation correlation parameter  $\rho$  from sample B to C. In practical terms, the latter indicates that over sample C positive innovations to the Slope coefficient (e.g. policy tightenings) were on average associated with negative innovations to the Level coefficient (e.g. falls in long-maturity yields). That result is consistent with the so-called “bond-yield conundrum” during that period, where bond yields fell or stayed relatively steady despite short-term interest rates rising to reflect with increases in the policy interest rate; see, for example, the discussion in Rudebusch, Swanson and Wu (2006).

## 6.4 Estimation results for the EA/AF/NS(2) model

For the EA/AF/NS(2) model estimation, convergence via the Kalman filter was again timely and reliable with no apparent sensitivity in the end result to different starting values. All estimates of the Level and Slope coefficients for the EA/AF/NS(2) model are again very similar to figure 2 and so are not separately reported.

Tables 4 and 5 contain the estimated parameter values for the EA/AF/NS(2) model over the combined and individual samples. The hypotheses of no change in yield curve dynamics between the samples A and B, and the samples B and C are again soundly rejected by the likelihood ratio tests.

Figures 4 and 5 illustrate the risk premium estimates for the time to maturity of five years for each of the different sample estimated. These estimates are obtained using equation 16 with the relevant estimated parameters and state variables, and  $\tau = 5$  years.

Table 4: Parameter estimates for EA/AF/NS(2) model with 3-m/15-y data						
Parameters	Sample A+B		Sample A		Sample B	
$\phi$	0.5030	(0.0085)	0.7315	(0.0142)	0.3887	(0.0083)
$\gamma_{0,1}$	0.1355	(0.0123)	0.0914	(0.0122)	0.1923	(0.0056)
$\gamma_{0,2}$	0.2985	(0.0270)	0.5194	(0.0330)	0.1459	(0.0192)
$\sigma_1$	0.0227	(0.0006)	0.0193	(0.0006)	0.0276	(0.0007)
$\sigma_2$	0.0362	(0.0022)	0.0256	(0.0023)	0.0542	(0.0026)
$\rho$	0.3260	(0.2336)	0.1479	(0.5765)	0.4716	(0.1822)
$\gamma_{1,11}$	0.3081	(0.4400)	0.0002	(0.0326)	1.1247	(0.7635)
$\gamma_{1,12}$	-0.2610	(0.7073)	0.1892	(0.1555)	-0.0235	(0.6909)
$\gamma_{1,21}$	-0.3353	(0.4897)	-0.1925	(0.0673)	-0.0140	(3.7293)
$\gamma_{1,22}$	0.2920	(0.4770)	0.0152	(0.2650)	0.4468	(0.8513)
log-L	17429.3		9106.5		8835.7	
$H_0:A=B$	1025.7	[0.0000]				
$H_0:\gamma_1=0$	9.3	[0.0548]	59.6	[0.0000]	64.8	[0.0000]

Note: (standard errors), [probabilities]



Table 5: Parameter estimates for EA/AF/NS(2) model with 3-m/30-y data						
Parameters	Sample B+C		Sample B		Sample C	
$\phi$	0.3319	(0.0053)	0.3904	(0.0068)	0.3177	(0.0078)
$\gamma_{0,1}$	0.1385	(0.0112)	0.1332	(0.0090)	0.1232	(0.0566)
$\gamma_{0,2}$	0.2314	(0.0307)	0.3014	(0.0266)	0.2383	(0.0593)
$\sigma_1$	0.0182	(0.0003)	0.0169	(0.0003)	0.0219	(0.0008)
$\sigma_2$	0.0236	(0.0008)	0.0240	(0.0007)	0.0217	(0.0053)
$\rho$	0.0317	(0.2371)	0.1695	(0.2201)	-0.9920	(1.5593)
$\gamma_{1,11}$	0.2776	(0.1245)	0.1739	(0.1823)	6.8665	(6.0884)
$\gamma_{1,12}$	-0.0516	(0.2160)	-0.0565	(0.3121)	0.8017	(1.8018)
$\gamma_{1,21}$	-0.2274	(0.1165)	-0.2182	(0.0980)	1.7737	(1.4409)
$\gamma_{1,22}$	0.0701	(0.1153)	0.1092	(0.1693)	0.2415	(0.3182)
log-L	14982.3		10294.5		5180.2	
$H_0:B=C$	984.8	[0.0000]				
$H_0:\gamma_1=0$	2721.3	[0.0000]	4.0	[0.4066]	169.7	[0.0000]
Note: (standard errors), [probabilities]						

Figures 4 and 5 confirm that changes in the average level of the risk premium estimates are again a major point of difference between the individual samples, although the more complex specification for the risk premium function in the EA/AF/NS(2) model makes it harder to attribute differences to individual parameters. A more transparent alternative is simply to attribute the 5-year risk premium estimates to their time-invariant and time-varying components. Evaluating the point estimates of the time-invariant component  $RP(\tau)$  of the risk premium function for the EA/AF/NS(2) model confirms the pattern of results for the A/AF/NS(2) model; i.e. there is still a declining time-invariant risk premium component from sample A to sample C (except sample B based on 3-m/15-y data, which is arguably less reliable).<sup>23</sup> Nevertheless, the average risk premium during sample A is lower than sample B, meaning that the time-varying component of the 5-year risk premium was lower in sample A than sample B. Regarding variation in risk premium, sample B (based on 3-m/30-y data) shows the least variation in terms of the peak-to-trough range and the standard deviation of first differences. Sample C shows the most variation.

Other points of note for the EA/AF/NS(2) model estimates are the confirmation of the material decline in the mean-reversion parameter  $\phi$  and the sign reversal of the innovation correlation parameter  $\rho$  as already discussed for the AF/NS(2) model. Indeed, the model estimate suggests that the innovations become all but perfectly negative, albeit with an implausibly large confidence interval. That suggests some degree of overparametrization when applying the EA/AF/NS(2) model over the relatively short sample period and/or within a low to near-zero nominal interest rate environment.

<sup>23</sup>The point estimates are 1.77 percentage points (pps) for sample A 3-m/15-y, 1.60 pps for sample B 3-m/30-y, and 1.49 bps sample C 3-m/30-y. The sample B 3-m/15-y point estimate is 2.47 pps, but the results using the 3-m/30-y data should in principle be more reliable given they exploit the additional information from longer maturities.

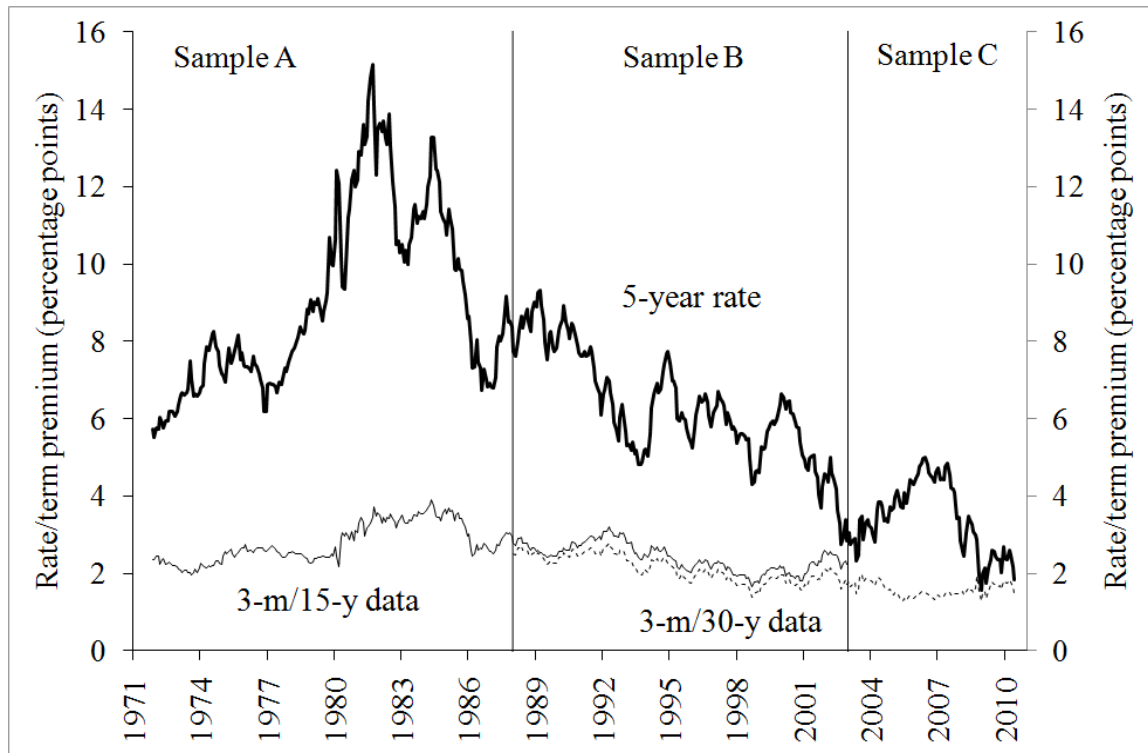


Figure 4: 5-year interest rate data and the 5-year risk premium implied by the point estimates of the EA/AF/NS(2) model parameters and state variables in each sample.

Sample A+B uses the 3-m/15-y data and sample B+C uses 3-m/30-y data, as indicated.

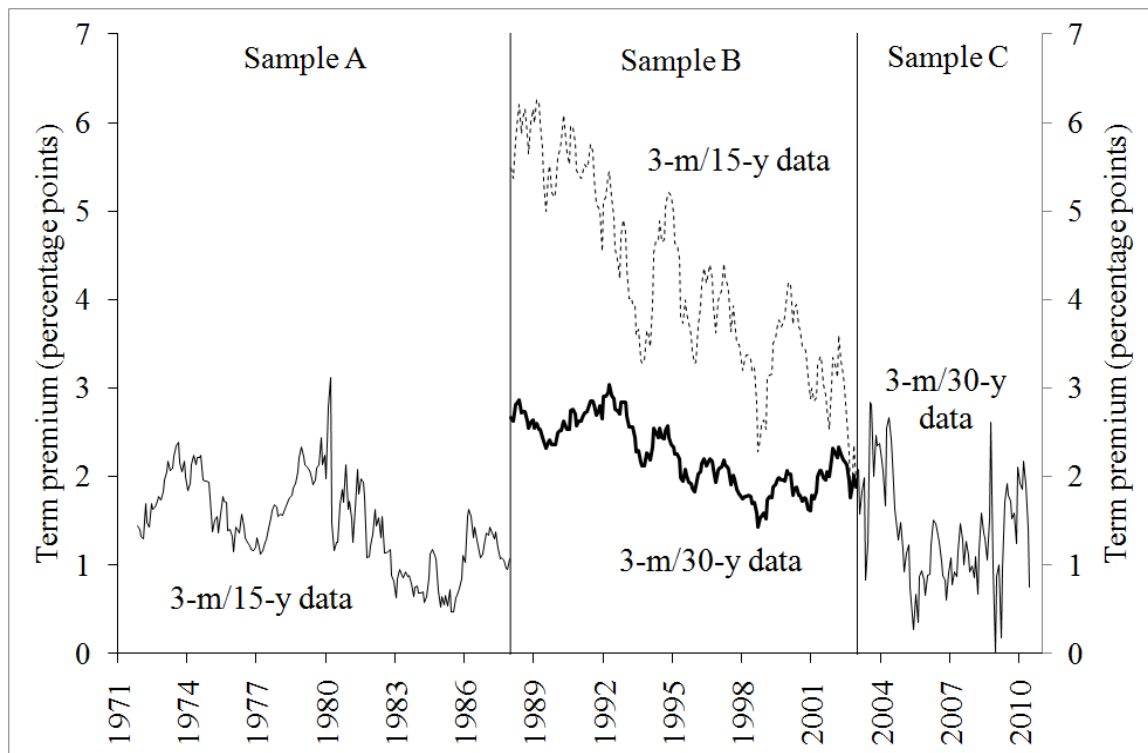


Figure 5: 5-year interest rate data and the 5-year risk premium implied by the point estimates of the EA/AF/NS(2) model parameters and state variables in each sample, using the 3-m/15-y data and/or 3-m/30-y data as indicated.

The significance of the time-varying component of the risk premium estimates can be assessed with a likelihood ratio test given that the EA/AF/NS(2) model nests the AF/NS(2) model; i.e. the latter is the EA/AF/NS(2) model with  $\gamma_1 = 0$  in  $\Gamma(t) = \gamma_0 + \gamma_1 \beta(t)$ . The likelihood ratio tests in tables 3 and 4 show that the estimates of  $\gamma_1$  are typically highly significant, with an exception being sample B using the 3-m/30-y data. These results imply that risk premiums are usually better modeled as time-varying rather than time-invariant, but assuming time-invariant risk premiums may sometimes provide an adequate representation of the data.

## 7 Conclusion

In this article, I have established that most NS models may be obtained as eigenvalue approximations to the dynamic component of the generic GATSM outlined in Dai and Singleton (2002). That theoretical foundation first provides an assurance that NS models correspond to a well-accepted framework for yield curve modeling. Second, the NS model classification arising from the theoretical foundation motivates the development and application of a two-factor arbitrage-free NS model as a standard model of the yield curve. That model is well-suited to macrofinance applications because it is very parsimonious while being representative of any GATSM that might underlie the yield curve data. A model with two factors is also consistent with available empirical evidence from related macrofinance literature.

As a practical illustration of applying the two-factor arbitrage-free NS model, I have investigated changes in U.S. yield dynamics. I find there was a very material change in the data-generating process for the U.S. yield curve between the samples from 1971 to 1987, 1988 to 2002, and 2003 to 2010, with the main contributor being a change to average risk premiums.

## References

- Babbs, S. and Nowman, K. (1999), ‘Kalman filtering of generalized Vasicek term structure models’, *Journal of Financial and Quantitative Analysis* **34(1)**, 115–130.
- Bank for International Settlements (2005), ‘Zero-coupon yield curves: technical documentation’, *BIS Papers* **25**.
- Berardi, A. (2009), ‘Term structure, inflation, and real activity’, *Journal of Financial and Quantitative Analysis* **44(4)**, 987–1011.
- Berardi, A. and Esposito, M. (1999), ‘A base model for multifactor specifications of the term structure’, *Economic Notes* **28(2)**, 145–170.
- Billi, R. (2009), ‘Was monetary policy optimal during past deflation scares?’, *Economic Review* **Third Quarter 2009**.
- Björk, T. and Christensen, B. (1999), ‘Interest rate dynamics and consistent forward rate curves’, *Mathematical Finance* **9(4)**, 323–348.

- Bliss, R. (1997), ‘Testing term structure estimation methods’, *Advances in Futures and Options Research* **9**, 197–231.
- Chen, Y. and Tsang, K. (2009), ‘What does the yield curve tell us about exchange rate predictability?’, *Working Paper, University of Washington and Virginia Tech*.
- Christensen, J., Diebold, F. and Rudebusch, G. (2009), ‘An arbitrage-free generalized Nelson-Siegel term structure model’, *Econometrics Journal* **12(3)**, 33–64.
- Christensen, J., Diebold, F. and Rudebusch, G. (2011), ‘The affine arbitrage-free class of Nelson-Siegel term structure models’, *Journal of Econometrics* **64**, 4–20.
- Christensen, J., Lopez, J. and Rudebusch, G. (2009), ‘Do central bank lending facilities affect interbank lending rates?’, *Working Paper, Federal Reserve Bank of San Francisco* **2009-13**.
- Christensen, J., Lopez, J. and Rudebusch, G. (2010), ‘Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields’, *Journal of Money, Credit, and Banking* **42**, 143–178.
- Collin-Dufresne, P., Goldstein, R. and Jones, C. (2008), ‘Identification of maximal affine term structure models’, *Journal of Finance* **63(2)**, 743–795.
- Coroneo, L., Nyholm, K. and Vidova-Koleva, R. (2008), ‘How arbitrage-free is the Nelson-Siegel model?’, *Working Paper, European Central Bank* **874**.
- Cox, J., Ingersoll, J. and Ross, S. (1985a), ‘An intertemporal general equilibrium model of asset prices’, *Econometrica* **53**, 363–384.
- Cox, J., Ingersoll, J. and Ross, S. (1985b), ‘A theory of the term structure of interest rates’, *Econometrica* **53**, 385–407.
- Dahlquist, M. and Svensson, L. (1996), ‘Estimating the term structure of interest rates for monetary policy analysis’, *Scandinavian Journal of Economics* **98(2)**, 163–183.
- Dai, Q. and Singleton, K. (2000), ‘Specification analysis of affine term structure models’, *Journal of Finance* **55(5)**, 1943–1978.
- Dai, Q. and Singleton, K. (2002), ‘Expectation puzzles, time-varying risk premia, and affine models of the term structure’, *Journal of Financial Economics* **63**, 415–441.
- Diebold, F., Ji, L. and Li, C. (2006), A three-factor yield curve model: non-affine structure, systematic risk sources and generalized duration, in L. Klein, ed., ‘Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando’, Cheltenham, U.K: Edward Elgar, pp. 240–274.
- Diebold, F. and Li, C. (2006), ‘Forecasting the term structure of government bond yields’, *Journal of Econometrics* **130(2)**, 337–364.
- Diebold, F., Li, C. and Yue, V. (2008), ‘Global yield curve dynamics and interactions: a dynamic Nelson-Siegel approach’, *Journal of Econometrics* **146**, 351–363.

- Diebold, F., Piazzesi, M. and Rudebusch, G. (2005), ‘Modelling bond yields in finance and macroeconomics’, *American Economic Review* **95**, 415–420.
- Diebold, F., Rudebusch, G. and Aruoba, S. (2006), ‘The macroeconomy and the yield curve: a dynamic latent factor approach’, *Journal of Econometrics* **131**, 309–338.
- Duffee, G. (2002), ‘Term premia and interest rate forecasts in affine models’, *Journal of Finance* **57**(1), 405–443.
- Duffie, D. and Kan, R. (1996), ‘A yield-factor model of interest rates’, *Mathematical Finance* **6**(4), 379–406.
- Duffie, D. and Singleton, K. (1999), ‘Modelling term structures of defaultable bonds’, *Review of Financial Studies* **12**(4), 687–720.
- Filipović, D. (1999), ‘A note on the Nelson-Siegel family’, *Mathematical Finance* **9**(4), 349–359.
- Filipović, D. (2000), ‘Exponential-polynomial families and the term structure of interest rates’, *Bernoulli Journal* **6**(6), 1081–1107.
- Filipović, D. (2009), *Term-Structure Models: A Graduate Course*, Springer.
- Gürkaynak, R., Sack, B. and Swanson, E. (2005a), ‘Do actions speak louder than words? The response of asset prices to monetary policy actions and statements’, *International Journal of Central Banking* **1**(1), 55–93.
- Gürkaynak, R., Sack, B. and Swanson, E. (2005b), ‘The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models’, *American Economic Review* **95**(1), 425–436.
- Gürkaynak, R., Sack, B. and Wright, J. (2008), ‘The U.S. Treasury yield curve: 1961 to the present’, *Journal of Monetary Economics* **54**, 2291–2304.
- Gürkaynak, R., Sack, B. and Wright, J. (2010), ‘The TIPS yield curve and inflation compensation’, *American Economic Journal: Macroeconomics* 2 pp. 70–92.
- Hamilton, J. (1994), *Time Series Analysis*, Princeton University Press.
- Hamilton, J. and Wu, J. (2010), ‘Identification and estimation of affine-term-structure models’, *Working paper, University of California, San Diego*.
- Heath, D., Jarrow, R. and Morton, A. (1992), ‘Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation’, *Econometrica* **60**(1), 77–106.
- Higam, N. (1996), *Accuracy and Stability of Numerical Algorithms*, Society for Industrial and Applied Mathematics.
- Hull, J. (2000), *Options, Futures and Other Derivatives, Fourth Edition*, Prentice Hall.
- James, J. and Webber, N. (2000), *Interest Rate Modelling*, Wiley and Sons.

- Joslin, S., Singleton, K. and Zhu, H. (2011), ‘A new perspective on Gaussian dynamic term structure models’, *Review of Financial Studies* **24(3)**, 926–970.
- Krippner, L. (2006a), ‘A theoretically consistent version of the Nelson and Siegel class of yield curve models’, *Applied Mathematical Finance* **13(1)**, 39–59.
- Krippner, L. (2006b), ‘A yield curve perspective on uncovered interest parity’, *Working Paper, University of Waikato* **06/10**.
- Krippner, L. (2009), ‘A theoretical foundation for the Nelson and Siegel class of yield curve models’, *Discussion paper, Reserve Bank of New Zealand* **DP2009/10**.
- Krippner, L. (2012), ‘Modifying Gaussian term structure models when interest rates are near the zero lower bound’, *Discussion paper, Centre for Applied Macroeconomic Analysis* **5/2012**.
- Macauley, F. (1938), *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the United States since 1865*, Columbia University Press.
- Munk, C. (2011), *Fixed Income Modelling*, Oxford University Press.
- Nelson, C. and Siegel, A. (1987), ‘Parsimonious modelling of yield curves’, *Journal of Business* **October**, 473–489.
- Nyholm, K. (2008), *Strategic Asset Allocation in Fixed-income Markets: a MATLAB-based User’s Guide*, Wiley and Sons.
- Rudebusch, G. (2010), ‘Macro-finance models of interest rates and the economy’, *The Manchester School Supplement* **2010**, 25–52.
- Rudebusch, G., Swanson, E. and Wu, T. (2006), ‘The bond yield ‘conundrum’ from a macro-finance perspective’, *Monetary and Economic Studies* **24(S-1)**, 83–128.
- Rudebusch, G. and Wu, T. (2007), ‘Accounting for a shift in the term structure behaviour with no-arbitrage and macro-finance models’, *Journal of Money, Credit and Banking* **39 (2-3)**, 395–422.
- Sharef, E. and Filipović, D. (2004), ‘Conditions for consistent exponential-polynomial forward rate processes with multiple nontrivial factors’, *International Journal of Theoretical and Applied Finance* **7(6)**, 685–700.
- Singleton, K. (2006), *Empirical Dynamic Asset Pricing: Model Specification and Econometric Assessment*, Princeton University Press.
- Svensson, L. (1995), ‘Estimating forward interest rates with the extended Nelson and Siegel model’, *Quarterly Review, Sveriges Riksbank* **1995(3)**, 13–26.
- Tchuindjo, L. (2008), ‘Factor’s correlation in the Heath-Jarrow-Morton interest rate model’, *Applied Stochastic Models in Business and Industry* **24**, 359–368.
- Vasicek, O. (1977), ‘An equilibrium characterisation of the term structure’, *Journal of Financial Economics* **5**, 177–188.

- Vincente, J. and Tabak, B. (2008), ‘Forecasting bond yields in the Brazilian fixed income market’, *International Journal of Forecasting* **24**, 490–497.
- Wong, E., Lucia, K., Price, S. and Startz, R. (2011), ‘The changing relationship between the Canadian and U.S. yield curves’, *Journal of International Money and Finance* **30**, 965–981.
- Wu, T. (2006), ‘Macro factors and the affine term structure of interest rates’, *Journal of Money, Credit and Banking* **38**(7), 1847–1875.

## A CIR dynamics

This appendix shows by example that dynamic term structure models with Cox et al. (1985b)/square-root innovations cannot be optimally approximated using NS factor loadings, in the sense of following a procedure analogous to the exposition in section 3.

Assume  $N$  independent factors each with the form  $dX_n(t) = \kappa_n [\theta_n - X_n(t)] dt + \sigma_n \sqrt{X_n(t)} dW(t)$  under the risk-neutral  $Q$  measure.

Then  $P(t, T) = \exp \left[ \sum_{n=1}^N A_n(t, T) + B_n(t, T) X_n(t) \right]$  where each  $B_n(t, T)$  has the standard Cox et al. (1985b) form:<sup>24</sup>

$$B_n(t, T) = \frac{2 [1 - \exp(\gamma_n \tau)]}{(\gamma_n + \kappa_n) [\exp(\gamma_n \tau) - 1] + 2\gamma_n} \quad (22)$$

with  $\gamma_n = \sqrt{\kappa_n^2 + 2\sigma_n^2}$ .

The associated forward rate curve is:

$$f(t, T) = a_0 + \sum_{n=1}^N \frac{4\gamma_n^2 \exp(\gamma_n \tau)}{[(\gamma_n + \kappa_n) [\exp(\gamma_n \tau) - 1] + 2\gamma_n]^2} X_n(t) - \frac{\partial}{\partial \tau} A_n(\tau) \quad (23)$$

The relative complexity of this functional form of maturity means that a central exponential decay term  $\exp(-\phi\tau)$  cannot be factored out of each factor loading as for the Gaussian case in section 3.

This incompatibility of the NS class of yield curve models with CIR/square-root dynamics is unfortunate, because CIR models have the well-known advantage over GATSMs of respecting the zero bound for interest rates. One resulting implication is that, in cases where the probability of zero interest rates from Gaussian dynamics is material, a non-Gaussian dynamic term structure model might be more appropriate than an NS model. That caveat applies in particular if the application requires the zero bound to be strictly respected (e.g. for financial market applications such as option pricing).

---

<sup>24</sup>See, for example, Hull (2000) p. 570. The associated  $A_n(t, T)$  terms have the form  $-a_0\tau + A^*(\tau)$ , and so have no influence on the factor loadings.

## B AF/NS(2) model derivations

### B.1 A/AF/NS(2) forward rate curve

#### B.1.1 A/AF/NS(2) forward rate expression

Equation 5 from the main text expresses the A/AF/NS(2) forward rate in terms of three components: (1) the expected path of the short rate component, the risk premium component, and the volatility effect component. These are evaluated respectively in the subsections below.

#### B.1.2 The expected path of the short rate component

$$\begin{aligned}\mathbb{E}_t[r(t + \tau)] &= \beta_1(t) + \beta_2(t) \cdot \exp(-\phi\tau) \\ &= [1, \exp(-\phi\tau)] \begin{bmatrix} \beta_1(t) \\ \beta_2(t) \end{bmatrix} \\ &= b(\tau) \beta(t)\end{aligned}$$

where  $b(\tau) = [1, \exp(-\phi\tau)]$  and  $\beta(t) = [\beta_1(t), \beta_2(t)]'$ .

#### B.1.3 Risk premium component

$$\int_0^\tau [\sigma_1, \sigma_2 \exp(-\phi[\tau - s])] \begin{bmatrix} \gamma_{0,1} \\ \gamma_{0,2} \end{bmatrix} ds = \sigma_1 \gamma_{0,1} \int_0^\tau ds + \sigma_2 \gamma_{0,2} \int_0^\tau \exp(-\phi[\tau - s]) ds$$

The Level and Slope risk premium components are evaluated respectively in the subsections below.

#### Risk premium component for Level

$$\begin{aligned}\sigma_1 \gamma_{0,1} \int_0^\tau ds &= \sigma_1 \gamma_{0,1} (s|_0^\tau) \\ &= \sigma_1 \gamma_{0,1} \cdot \tau\end{aligned}$$

#### Risk premium component for Slope

$$\begin{aligned}\int_0^\tau \sigma_2 \gamma_{0,2} \exp(-\phi[\tau - s]) ds &= \sigma_2 \gamma_{0,2} \frac{1}{\phi} \exp(-\phi[\tau - s]) \Big|_0^\tau \\ &= \sigma_2 \gamma_{0,2} \cdot \frac{1}{\phi} [1 - \exp(-\phi\tau)] \\ &= \sigma_2 \gamma_{0,2} \cdot F(\phi, \tau)\end{aligned}$$

where  $F(\phi, \tau) = \frac{1}{\phi} [1 - \exp(-\phi\tau)]$ .



#### B.1.4 Volatility effect component

$$\begin{aligned}
& \int_0^\tau [\sigma_1, \sigma_2 \exp(-\phi[\tau-s])] \left( \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \int_s^\tau \begin{bmatrix} \sigma_1 \\ \sigma_2 \exp(-\phi[u-s]) \end{bmatrix} du \right) ds \\
= & \int_0^\tau [\sigma_1, \sigma_2 \exp(-\phi[\tau-s])] \begin{bmatrix} \sigma_1 \int_s^\tau du + \rho \sigma_2 \int_s^\tau \exp(-\phi[u-s]) du \\ \rho \sigma_1 \int_s^\tau du + \sigma_2 \int_s^\tau \exp(-\phi[u-s]) du \end{bmatrix} ds \\
= & \sigma_1^2 \int_0^\tau \left( \int_s^\tau du \right) ds \\
& + \rho \sigma_1 \sigma_2 \int_0^\tau \left( \int_s^\tau \exp(-\phi[u-s]) du \right) ds \\
& + \rho \sigma_1 \sigma_2 \int_0^\tau \exp(-\phi[\tau-s]) \left( \int_s^\tau du \right) ds \\
& + \sigma_2^2 \int_0^\tau \exp(-\phi[\tau-s]) \left( \int_s^\tau \exp(-\phi[u-s]) du \right) ds
\end{aligned}$$

where the last four lines are, respectively, the Level/Level, Level/Slope, Slope/Level, and Slope/Slope components. These are evaluated in turn in the subsections below.

##### Volatility effect component for Level/Level

$$\begin{aligned}
\int_0^\tau \left( \int_s^\tau du \right) ds &= \int_0^\tau (u|_s^\tau) ds \\
&= \int_0^\tau (\tau - s) ds \\
&= \tau s - \frac{1}{2} s^2 \Big|_0^\tau \\
&= \frac{1}{2} \tau^2
\end{aligned}$$

##### Volatility effect component for Slope/Slope

$$\begin{aligned}
& \int_0^\tau \exp(-\phi[\tau-s]) \left( \int_s^\tau \exp(-\phi[u-s]) du \right) ds \\
= & \int_0^\tau \exp(-\phi s) \left( \frac{1}{\phi} (\exp(-\phi[\tau-s]) - 1) \right) ds \\
= & \frac{1}{\phi^2} \left( \exp(2\phi\tau) \left( \exp(-\phi[\tau-s]) - \frac{1}{2} \exp(-2\phi s) \right) \Big|_0^\tau \right) \\
= & \frac{1}{2\phi^2} [1 - \exp(-\phi\tau)]^2 \\
= & \frac{1}{2} [F(\phi, \tau)]^2
\end{aligned}$$

### Volatility effect component for Level/Slope

$$\begin{aligned}
\int_0^\tau \left( \int_s^\tau \exp(-\phi[u-s]) du \right) ds &= \int_0^\tau \left( -\frac{1}{\phi} \exp(-\phi[u-s]) \Big|_s^\tau \right) ds \\
&= \frac{1}{\phi} \int_0^\tau (1 - \exp(-\phi[\tau-s])) ds \\
&= \frac{1}{\phi} \left( s - \frac{1}{\phi} \exp(-\phi[\tau-s]) \Big|_0^\tau \right) \\
&= \frac{\tau}{\phi} - \frac{1}{\phi^2} + \frac{1}{\phi^2} \exp(-\phi\tau)
\end{aligned}$$

### Volatility effect component for Slope/Level

$$\begin{aligned}
\int_0^\tau \exp(-\phi[\tau-s]) \left( \int_s^\tau du \right) ds &= \int_0^\tau \exp(-\phi[\tau-s]) (\tau-s) ds \\
&= \left( \exp(-\phi[\tau-s]) \frac{1}{\phi^2} (1 + \tau\phi - s\phi) \Big|_0^\tau \right) \\
&= \frac{1}{\phi^2} - \frac{1}{\phi^2} \exp(-\phi\tau) - \frac{\tau}{\phi} \exp(-\phi\tau)
\end{aligned}$$

### Volatility effect component for Level/Slope + Slope/Level

$$\begin{aligned}
\int_0^\tau \left( \int_s^\tau \exp(-\phi[u-s]) du \right) ds + \int_0^\tau \exp(-\phi[\tau-s]) \left( \int_s^\tau du \right) ds &= \frac{\tau}{\phi} - \frac{\tau}{\phi} \exp(-\phi\tau) \\
&= \tau \cdot F(\phi, \tau)
\end{aligned}$$

#### B.1.5 A/AF/NS(2) forward rate curve

Substituting the results above into the A/AF/NS(2) forward rate expression gives equation 6 in the main text.

## B.2 A/AF/NS(2) interest rate curve

The A/AF/NS(2) interest rate curve may be evaluated using the standard relationship  $R(t, \tau) = \frac{1}{\tau} \int_0^\tau f(t, \tau) d\tau$  for each component of the A/AF/NS(2) forward rate curve. The calculations are undertaken respectively in the subsections below.

#### B.2.1 Short rate Level component

$$\begin{aligned}
\frac{1}{\tau} \int_0^\tau d\tau &= \frac{1}{\tau} (\tau|_0^\tau) \\
&= \frac{1}{\tau} (\tau) \\
&= 1
\end{aligned}$$

### B.2.2 Short rate Slope component

$$\begin{aligned}\frac{1}{\tau} \int_0^\tau \exp(-\phi\tau) d\tau &= \frac{1}{\tau} \left( -\frac{1}{\phi} \exp(-\phi\tau) \Big|_0^\tau \right) \\ &= \frac{1}{\tau} \left( \frac{1}{\phi} [1 - \exp(-\phi\tau)] \right) \\ &= \frac{1}{\tau} F(\phi, \tau)\end{aligned}$$

### B.2.3 Risk premium Level component

$$\begin{aligned}\frac{1}{\tau} \int_0^\tau \tau d\tau &= \frac{1}{\tau} \left( \frac{1}{2} \tau^2 \Big|_0^\tau \right) \\ &= \frac{1}{\tau} \left( \frac{1}{2} \tau^2 \right) \\ &= \frac{1}{2} \tau\end{aligned}$$

### B.2.4 Risk premium Slope component

$$\begin{aligned}\frac{1}{\tau} \int_0^\tau F(\phi, \tau) d\tau &= \frac{1}{\tau} \int_0^\tau \frac{1}{\phi} [1 - \exp(-\phi\tau)] d\tau \\ &= \frac{1}{\tau} \left( \frac{\tau}{\phi} + \frac{1}{\phi^2} \exp(-\phi\tau) \Big|_0^\tau \right) \\ &= \frac{1}{\tau} \left[ \frac{\tau}{\phi} - \frac{1}{\phi^2} + \frac{1}{\phi^2} \exp(-\phi\tau) \right] \\ &= \frac{1}{\phi} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) \right]\end{aligned}$$

### B.2.5 Volatility effect Level/Level component

$$\begin{aligned}\frac{1}{\tau} \int_0^\tau \frac{1}{2} \tau^2 d\tau &= \frac{1}{\tau} \left( \frac{1}{6} \tau^3 \Big|_0^\tau \right) \\ &= \frac{1}{\tau} \left( \frac{1}{6} \tau^3 \right) \\ &= \frac{1}{6} \tau^2\end{aligned}$$

### B.2.6 Volatility effect Slope/Slope component

$$\begin{aligned}
\frac{1}{\tau} \int_0^\tau \frac{1}{2} [F(\phi, \tau)]^2 d\tau &= \frac{1}{\tau} \int_0^\tau \frac{1}{2\phi^2} [1 - \exp(-\phi\tau)]^2 d\tau \\
&= \frac{1}{\tau} \left( \frac{1}{2\phi^2} \tau + \frac{1}{\phi^3} \exp(-\phi\tau) - \frac{1}{4\phi^3} \exp(-2\phi\tau) \right) \Big|_0^\tau \\
&= \frac{1}{\tau} \left[ \frac{\tau}{2\phi^2} - \frac{3}{4\phi^3} + \frac{1}{\phi^3} \exp(-\phi\tau) - \frac{1}{4\phi^3} \exp(-2\phi\tau) \right] \\
&= \frac{1}{2\phi^2} - \frac{1}{2\phi^3\tau} [1 - \exp(-\phi\tau)] - \frac{1}{4\phi^3\tau} [1 - \exp(-\phi\tau)]^2 \\
&= \frac{1}{2\phi^2} - \frac{1}{2\phi^2} \frac{1}{\tau} \frac{1}{\phi} [1 - \exp(-\phi\tau)] - \frac{1}{4\phi} \frac{1}{\tau} \left( \frac{1}{\phi} [1 - \exp(-\phi\tau)] \right)^2 \\
&= \frac{1}{2\phi^2} \left( 1 - \frac{1}{\tau} F(\phi, \tau) - \frac{\phi}{2\tau} [F(\phi, \tau)]^2 \right)
\end{aligned}$$

### B.2.7 Volatility effect Level/Slope + Slope/Level component

$$\begin{aligned}
\frac{1}{\tau} \int_0^\tau \tau F(\phi, \tau) d\tau &= \frac{1}{\tau} \int_0^\tau \frac{\tau}{\phi} [1 - \exp(-\phi\tau)] d\tau \\
&= \frac{1}{\tau} \left( \frac{1}{2\phi} \tau^2 + \frac{1}{\phi^3} \exp(-\phi\tau) + \frac{\tau}{\phi^2} \exp(-\phi\tau) \right) \Big|_0^\tau \\
&= \frac{1}{\tau} \left( \frac{1}{2\phi} \tau^2 - \frac{1}{\phi^3} + \frac{1}{\phi^3} \exp(-\phi\tau) + \frac{\tau}{\phi^2} \exp(-\phi\tau) \right) \\
&= \frac{1}{2\phi} \tau - \frac{1}{\phi^3\tau} + \frac{1}{\phi^2} \exp(-\phi\tau) + \frac{1}{\phi^3\tau} \exp(-\phi\tau) \\
&= \frac{1}{2\phi} \tau - \frac{1}{\phi^3\tau} - \frac{1}{\phi^2} [1 - \exp(-\phi\tau)] + \frac{1}{\phi^2} - \frac{1}{\phi^3\tau} [1 - \exp(-\phi\tau)] + \frac{1}{\phi^3\tau} \\
&= \frac{1}{2\phi} \tau + \frac{1}{\phi^2} - \frac{1}{\phi} F(\phi, \tau) - \frac{1}{\phi^2\tau} F(\phi, \tau)
\end{aligned}$$

### B.2.8 Final interest rate curve expression

$$\begin{aligned}
R(t, \tau) &= \beta_1(t) + \beta_2(t) \cdot \frac{1}{\tau} F(\phi, \tau) \\
&\quad + \sigma_1 \gamma_1 \cdot \frac{1}{2} \tau + \sigma_2 \gamma_2 \cdot \frac{1}{\phi} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) \right] \\
&\quad - \sigma_1^2 \cdot \frac{1}{6} \tau^2 - \sigma_2^2 \cdot \frac{1}{2\phi^2} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) - \frac{1}{2\tau} \phi [F(\phi, \tau)]^2 \right] \\
&\quad - \rho \sigma_1 \sigma_2 \cdot \frac{1}{\phi^2} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) + \frac{1}{2} \phi \tau - \frac{1}{\tau} F(\phi, \tau) \right]
\end{aligned}$$

which may be expressed in the form  $R(t, \tau) = \bar{a}(\tau) + \bar{b}(\tau) \beta(t)$  as in the main text.

## B.3 A/AF/NS(2) model Kalman filter calculations

### B.3.1 State equation

The A/AF/NS(2) state equation may be derived directly from the expected path of the short rate, i.e.:

$$\begin{aligned}
\mathbb{E}_t \{ \mathbb{E}_{t+\tau} [r(t+\tau+\Delta t)] \} &= \mathbb{E}_t [r(t+\tau+\Delta t)] \\
\mathbb{E}_t \{ [1, \exp(-\phi\tau)] \beta(t+\Delta t) \} &= [1, \exp(-\phi[\tau+\Delta t])] \beta(t) \\
[1, \exp(-\phi\tau)] \mathbb{E}_t \{ \beta(t+\Delta t) \} &= [1, \exp(-\phi\tau)] \begin{bmatrix} 1 & 0 \\ 0 & \exp(-\phi\Delta t) \end{bmatrix} \beta(t) \\
\mathbb{E}_t \{ \beta(t+\Delta t) \} &= \Phi(\phi, \Delta t) \beta(t)
\end{aligned}$$

Removing the expectations operator, the state equation is therefore:

$$\beta(t+\Delta t) = \Phi(\phi, \Delta t) \beta(t) + \varepsilon(t+\Delta t)$$

as in equation 8, where  $\mathbb{E}_t [\varepsilon(t+\Delta t)] = 0$ . Note that  $\varepsilon(t+\Delta t)$  has a correlated bivariate Gaussian distribution given the correlated innovations assumed to underlie equation 5, i.e.  $\Omega[dW_1(t), \exp(-\phi\tau) \cdot dW_2(t)]$ .

### B.3.2 State covariance matrix

$$\begin{aligned}
Q &= \int_0^{\Delta t} \Phi(\phi, s) \Omega [\Phi(\phi, s)]' ds \\
&= \int_0^{\Delta t} \begin{bmatrix} 1 & 0 \\ 0 & \exp(-\phi s) \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \exp(-\phi s) \end{bmatrix} ds \\
&= \int_0^{\Delta t} \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \exp(-\phi s) \\ \rho\sigma_1\sigma_2 \exp(-\phi s) & \sigma_2^2 \exp(-2\phi s) \end{bmatrix} ds \\
&= \begin{bmatrix} \sigma_1^2 s & -\rho\sigma_1\sigma_2 \frac{1}{\phi} \exp(-\phi s) \\ -\rho\sigma_1\sigma_2 \exp(-\phi s) & -\sigma_2^2 \frac{1}{2\phi} \exp(-2\phi s) \end{bmatrix} \Big|_0^{\Delta t} \\
&= \begin{bmatrix} \sigma_1^2 \Delta t & \rho\sigma_1\sigma_2 \left[ \frac{1}{\phi} - \frac{1}{\phi} \exp(-\phi \Delta t) \right] \\ \rho\sigma_1\sigma_2 \left[ \frac{1}{\phi} - \frac{1}{\phi} \exp(-\phi \Delta t) \right] & \sigma_2^2 \frac{1}{2\phi} [1 - \exp(-2\phi \Delta t)] \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 \cdot \Delta t & \rho\sigma_1\sigma_2 \cdot F(\phi, \Delta t) \\ \rho\sigma_1\sigma_2 \cdot F(\phi, \Delta t) & \sigma_2^2 \cdot F(2\phi, \Delta t) \end{bmatrix}
\end{aligned}$$

### B.3.3 Unconditional state covariance matrix

$$\begin{aligned}
P_{1|0} &= \int_0^{\infty} \Phi(\phi, s) \Omega [\Phi(\phi, s)]' ds \\
&= \begin{bmatrix} \sigma_1^2 s & -\rho\sigma_1\sigma_2 \frac{1}{\phi} \exp(-\phi s) \\ -\rho\sigma_1\sigma_2 \exp(-\phi s) & -\sigma_2^2 \frac{1}{2\phi} \exp(-2\phi s) \end{bmatrix} \Big|_0^{\infty} \\
&= \begin{bmatrix} \text{undef} & \rho\sigma_1\sigma_2 \frac{1}{\phi} \\ \rho\sigma_1\sigma_2 \frac{1}{\phi} & \sigma_2^2 \frac{1}{2\phi} \end{bmatrix}
\end{aligned}$$

where the indefinite integral evaluations use the results from the previous subsection. Note that  $P_{1|0}(1, 1) = \sigma_\eta^2 / (1 - \psi^2)$  replaces the undefined expression “undef” to give  $P_{1|0}$  as in equation 12 of the main text.

#### B.3.4 Cholesky specification for $\Omega$

$$\begin{aligned} LL' &= \Omega \\ \begin{bmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} \sigma_1 & \rho\sigma_2 \\ 0 & \sigma_2\sqrt{1-\rho^2} \end{bmatrix} &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \rho^2\sigma_2^2 + \sigma_2^2(-\rho^2 + 1) \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix} \end{aligned}$$

### B.4 A/AF/NS(2) risk premium functions

The AF/NS(2) forward risk premium function can be obtained by subtracting the A/AF/NS(2) forward rate expression excluding risk premiums (i.e. by setting  $\gamma_0 = 0$ ) from the A/AF/NS(2) forward rate expression with risk premiums, i.e.:

$$\begin{aligned} \text{FRP}(\tau) &= f(t, \tau) - [f(t, \tau) | \gamma_0 = 0] \\ &= \sigma_1\gamma_{0,1} \cdot \tau + \sigma_2\gamma_{0,2} \cdot F(\phi, \tau) \end{aligned}$$

The interest rate risk premium function is then:

$$\begin{aligned} \text{RP}(\tau) &= \frac{1}{\tau} \int_0^\tau \text{FRP}(\tau) d\tau \\ &= \sigma_1\gamma_1 \cdot \frac{1}{2}\tau + \sigma_2\gamma_2 \cdot \frac{1}{\phi} \left[ 1 - \frac{1}{\tau} F(\phi, \tau) \right] \end{aligned}$$

### B.5 The EA/AF/NS(2) model

There is almost certainly a straightforward way to establish the expressions for the EA/AF/NS(2) model using a suitable modification to the Heath et al. (1992) framework. But, assuming that route exists, it has unfortunately escaped the investigations of the author.

However, it turns out that the EA/AF/NS(2) model is precisely a special case of the Dai and Singleton (2002, hereafter DS) generic GATSM, as introduced in section 2, with two factors and limit of zero for one of the rates of mean reversion. The derivations in the following subsections are therefore based on the generic GATSM with the substitution of the EA/AF/NS(2) model parameters; i.e.  $X(t) = \beta(t)$ ,  $\xi_0 = 0$ ,  $\xi_1 = [1, 1]'$ ,  $K_{\mathbb{Q}} = \text{diag}[0, \phi]$ ,  $\pi_0 = -\gamma_0$ ,  $\pi_1 = -\gamma_1$ , and  $K_{\mathbb{P}} = \kappa = \text{diag}[0, \phi] + \gamma_1$ .

Regarding the EA/AF/NS(2) forward rate and interest rate curves, the time-invariant risk premiums associated with the constant market prices of risk  $\gamma_0$  remain via the identification  $\theta_{\mathbb{P}} = 0$ , as discussed in section 5.2.3. The time-varying risk premiums associated with matrix  $\gamma_1$  are captured in the state equation, as discussed in the following subsection. The measurement equation for the EA/AF/NS(2) model therefore remains identical to that of the AF/NS(2) model.

## B.6 EA/AF/NS(2) model Kalman filter calculations

### B.6.1 State equation

DS appendix A gives the following expression for the generic GATSM state equation:

$$\mathbb{E}_t[X(t + \Delta t)] = \exp(-K_{\mathbb{P}}\tau) X(t) + [I - \exp(-K_{\mathbb{P}}\tau)] \theta_{\mathbb{P}}$$

Substituting the EA/AF/NS(2) parameters from the previous subsection into the generic GATSM state equation expression gives  $\exp(-K_{\mathbb{Q}}\Delta t) = \text{diag}[1, \exp(-\phi\Delta t)] = \Phi(\phi, \Delta t)$ ,  $\exp(\pi_1\tau) = \exp(-\gamma_1\tau)$ , and  $\exp(-K_{\mathbb{P}}\Delta t) = \exp(-\kappa\Delta t)$ . Therefore the EA/AF/NS(2) state equation is:

$$\begin{aligned}\mathbb{E}_t[\beta(t + \Delta t)] &= \Phi(\phi, \Delta t) \exp(-\gamma_1\tau) \beta(t) \\ &= \exp(-\kappa\tau) \beta(t)\end{aligned}$$

as in equation 14. Note that  $\varepsilon(t + \Delta t)$  is Gaussian as for the AF/NS(2) model.

### B.6.2 State covariance matrix

$$\begin{aligned}Q &= \int_0^{\Delta t} \exp(-\kappa s) \Omega \exp(-\kappa' s) ds \\ &= \int_0^{\Delta t} \exp(-VDV^{-1}s) \Omega [\exp(-VDV^{-1}s)]' ds \\ &= \int_0^{\Delta t} V \exp(-Ds) V^{-1} \Omega [\exp(-VDV^{-1}s)]' ds \\ &= \int_0^{\Delta t} V \begin{bmatrix} \exp(-d_1 s) & 0 \\ 0 & \exp(-d_2 s) \end{bmatrix} V^{-1} \Omega (V^{-1})' \begin{bmatrix} \exp(-d_1 s) & 0 \\ 0 & \exp(-d_2 s) \end{bmatrix} V' ds \\ &= V \left( \int_0^{\Delta t} \begin{bmatrix} \exp(-d_1 s) & 0 \\ 0 & \exp(-d_2 s) \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{bmatrix} \begin{bmatrix} \exp(-d_1 s) & 0 \\ 0 & \exp(-d_2 s) \end{bmatrix} ds \right) V' \\ &= V \left( \int_0^{\Delta t} \begin{bmatrix} u_{11} \exp(-2d_1 s) & u_{12} \exp(-[d_1 + d_2] s) \\ u_{21} \exp(-[d_1 + d_2] s) & u_{22} \exp(-2d_2 s) \end{bmatrix} ds \right) V' \\ &= V \begin{bmatrix} u_{11} \cdot F(2d_1, \Delta t) & u_{12} \cdot F(d_1 + d_2, \Delta t) \\ u_{21} \cdot F(d_1 + d_2, \Delta t) & u_{22} \cdot F(2d_2, \Delta t) \end{bmatrix} V'\end{aligned}$$

### B.6.3 Unconditional state covariance matrix

$$\begin{aligned}P_{1|0} &= \int_0^{\infty} \exp(-\kappa s) \Omega \exp(-\kappa' s) ds \\ &= V \begin{bmatrix} u_{11} F_2(2d_1, \infty) & u_{12} F_2(d_1 + d_2, \infty) \\ u_{21} F_2(d_1 + d_2, \infty) & u_{22} F_2(2d_2, \infty) \end{bmatrix} V' \\ &= V \begin{bmatrix} u_{11} \frac{1}{2d_1} & u_{12} \frac{1}{d_1 + d_2} \\ u_{21} \frac{1}{d_1 + d_2} & u_{22} \frac{1}{2d_2} \end{bmatrix} V'\end{aligned}$$

where the integral evaluations use the results from the previous subsection.

## B.7 EA/AF/NS(2) risk premium functions

The time-varying component of the EA/AF/NS(2) forward risk premium function may be obtained using the time-varying component of the DS forward risk premium expression for the generic GATSM model.

### B.7.1 DS discrete-time expression for forward risk premium

DS appendix A obtains the following generic GATSM forward risk premium expression  $p_t^n$  using discrete increments of time  $\Delta t$ :

$$p_t^n = f_t^n - \mathbb{E}_t[r_{t+n\Delta t}]$$

where:

$$f_t^n = -\frac{1}{\Delta} \log \frac{P(t, [n+1]\Delta t)}{P(t, n\Delta t)}$$

and:

$$\begin{aligned} \mathbb{E}_t[r_{t+n\Delta t}] &= \mu_n + \nu'_n X(t) \\ \mu_n &= a_1 + \theta'_\mathbb{P} [I - \exp(-K'_\mathbb{P} n\Delta t)] b_1 \\ \nu_n &= \exp(-K'_\mathbb{P} n\Delta t) b_1 \end{aligned}$$

Note that  $a_1$  and  $b_1$  are constants associated with the one-period interest rate  $r_t = a_1 + b_1 X(t)$ . DS also notes that  $f_t^n$  is the one-period forward deliverable  $n$  periods forward, and  $\mathbb{E}_t[r_{t+n\Delta t}]$  is the conditional mean of the short rate.

### B.7.2 DS continuous-time expression for forward risk premium

In the reverse order of which they were introduced, take the limit of each quantity in the previous subsection as  $\Delta t \rightarrow 0$ . Hence,  $\lim_{\Delta t \rightarrow 0} r_t = r(t) = \xi_0 + \xi'_1 X(t)$  as defined in section 2, so  $\lim_{\Delta t \rightarrow 0} b_1 = \xi_1$  and  $\lim_{\Delta t \rightarrow 0} a_1 = \xi_0$ . Therefore:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} \nu_n &= \nu(\tau) = \exp(-K'_\mathbb{P} \tau) \xi_1 \\ \lim_{\Delta t \rightarrow 0} \mu_n &= \mu(\tau) = \xi_0 + \theta'_\mathbb{P} [I - \exp(-K'_\mathbb{P} \tau)] \xi_0 \\ \lim_{\Delta t \rightarrow 0} \mathbb{E}_t[r_{t+n\Delta t}] &= \xi_0 + \theta'_\mathbb{P} [I - \exp(-K'_\mathbb{P} \tau)] \xi_0 + \xi'_1 \exp(-K'_\mathbb{P} \tau) X(t) \end{aligned}$$

Regarding the forward rate:

$$\begin{aligned} \lim_{\Delta t \rightarrow 0} f_t^n &= -\frac{\partial}{\partial \tau} \log P(t, \tau) = f(t, \tau) \\ &= \xi_0 + [\exp(-K'_\mathbb{Q} \tau) \xi_1]' X(t) - \frac{\partial}{\partial \tau} A(\tau) \end{aligned}$$

where  $P(t, \tau)$  and  $f(t, \tau)$  are as outlined in section 2.



The forward risk premium expression therefore becomes:

$$\begin{aligned}
\lim_{\Delta t \rightarrow 0} p_t^n &= p(t, \tau) \\
&= f(t, \tau) - \mathbb{E}_t[r_{t+n\Delta t}] \\
&= \xi_0 + \xi'_1 \exp(-K_{\mathbb{Q}}\tau) X(t) - \frac{\partial}{\partial \tau} A(\tau) \\
&\quad - \xi_0 - \theta'_P [I - \exp(-K'_{\mathbb{P}}\tau)] \xi_0 - \xi'_1 \exp(-K'_{\mathbb{P}}\tau) X(t)
\end{aligned}$$

and the time-varying component of  $p(t, \tau)$  is:

$$p(t, \tau) - p(\tau) = \xi'_1 \exp(-K_{\mathbb{Q}}\tau) X(t) - \xi'_1 \exp(-K'_{\mathbb{P}}\tau) X(t)$$

### B.7.3 EA/AF/NS(2) forward risk premium function

Substituting the EA/AF/NS(2) parameters and expressions from section B.5 and B.6.1 into the time-varying component of  $p(t, \tau)$  from the previous section gives:

$$\begin{aligned}
p(t, \tau) - p(\tau) &= [1, 1] \exp(-K_{\mathbb{Q}}\tau) \beta(t) - [1, 1] \exp(-\kappa\tau) \beta(t) \\
&= \{[1, 1] \Phi(\phi, \tau)\} \beta(t) - [1, 1] \exp(-\kappa\tau) \beta(t) \\
\text{FRP}(t, \tau) - \text{FRP}(\tau) &= [1, \exp(-\phi\tau)] \beta(t) - [1, 1] \exp(-\kappa\tau) \beta(t)
\end{aligned}$$

### B.7.4 EA/AF/NS(2) interest rate risk premium function

$$\begin{aligned}
\text{RP}(t, \tau) &= \frac{1}{\tau} \int_0^\tau \text{FRP}(t, \tau) d\tau \\
&= \text{RP}(\tau) + \left( \frac{1}{\tau} \int_0^\tau [1, \exp(-\phi\tau)] d\tau \right) \beta(t) - [1, 1] \frac{1}{\tau} \left( \int_0^\tau \exp(-\kappa\tau) d\tau \right) \beta(t) \\
&= \text{RP}(\tau) + \left[ 1, \frac{1}{\tau} F(\phi, \tau) \right] \beta(t) - [1, 1] \left( \frac{1}{\tau} \left\{ -[\kappa]^{-1} \exp(-\kappa\tau) \Big|_0^\tau \right\} \right) \beta(t) \\
&= \text{RP}(\tau) + \left[ 1, \frac{1}{\tau} F(\phi, \tau) \right] \beta(t) - [1, 1] [\kappa\tau]^{-1} [I - \exp(-\kappa\tau)] \beta(t) \\
&= \text{RP}(\tau) + \left( \left[ 1, \frac{1}{\tau} F(\phi, \tau) \right] - [1, 1] [\kappa\tau]^{-1} [I - \exp(-\kappa\tau)] \right) \beta(t)
\end{aligned}$$