RESERVE
BANK
O F N E W Z E A L A N D

DP2009/10

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

Leo Krippner

September 2009

DP2009/10

A theoretical foundation for the Nelson and Siegel class of yield curve models [*]

Leo Krippner[†]

Abstract

This article establishes that most models within the popular and widely-used Nelson and Siegel (1987, hereafter NS) class, with one notable exception being the Svensson (1995) variant, are effectively reduced-form representations of the generic Gaussian affine term structure model outlined in Dai and Singleton (2002). That fundamental theoretical foundation provides a compelling case for applying certain NS models as standard tools for yield curve analysis in economics and finance: users get the well-established pragmatic benefits of NS models along with an assurance that they correspond to a well-accepted set of principles and assumptions for modelling the yield curve and its dynamics.

[†] Address: Economics Department, Reserve Bank of New Zealand, 2 The Terrace, PO Box 2498, Wellington, New Zealand. *email address:* leo.krippner@rbnz.govt.nz.

1 Introduction

This article establishes that most yield curve models within the Nelson and Siegel (1987, hereafter NS) class are effectively reduced-form representations of the generic Gaussian affine term structure model (hereafter GATSM) outlined in Dai and Singleton (2002). The primary motivation for producing this original result is to assure users of NS models that they correspond to a well-accepted set of principles and assumptions for modelling the yield curve and its dynamics. Prior to this fundamental theoretical foundation, the widespread application of NS models in an increasing variety of economic and financial fields has typically been justified from the perspective of their practical benefits and empirical successes.[1] However, that basis is less than satisfactory, given that the Level, Slope, and Curvature factor loadings at the core of all NS models had their origin in the somewhat arbitrary and atheoretical field of curve fitting.[2]

The corollary from establishing the theoretical foundation is a compelling case for applying certain NS models as standard tools for yield curve modelling and analysis in economics and finance, rather than GATSMs with low numbers of factors. That is, the user can bypass the relative complexity of specifying, identifying, estimating, and interpreting a particular GATSM by simply applying an NS model. The NS model will provide the same practical results in terms of summarising the shape of the yield curve and its evolution over time. Moreover, it will do so parsimoniously and reliably even when the yield curve may potentially be influenced by many factors, something that cannot be guaranteed within the practical limitations of GATSMs.

The exposition begins in section 2 by specifying the generic GATSM from Dai and Singleton (2002) within the context of this article. The forward rate curve associated with the generic GATSM is then derived, and section 3 explicitly shows how that forward rate curve may be re-expressed in the original NS representation. Specifically, the Level factor loading and its associated coefficient are shown to correspond to the highly-persistent (i.e very slowly mean-reverting) components of the generic GATSM, and the Slope and Curvature factor loadings with their associated coefficients are shown to correspond to the

---

[1]Bank for International Settlements (2005) provides an overview of routine central bank use of NS models, Coroneo, Nyholm and Vidova-Koleva (2008) notes their widespread use by financial market practitioners, Diebold, Piazzesi and Rudebusch (2005) summarises some recent time series applications, and more examples of NS model applications are referenced later in this article. Dahlquist and Svensson (1996) originally advocated the suitability of NS models for monetary policy purposes, noting their ease of application, close fit to the yield curve data, and ready interpretation of the estimated components.

[2]The forward rate curve of the original NS model was simply proposed as the general solution to a second-order differential equation with equal roots, and then integrated to obtain the NS interest rate curve.

non-persistent (i.e mean-reverting) 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 representations of the generic GATSM. This classification provides some guidance on selecting the appropriate NS model for the given application and for interpreting its output. Section 5 discusses the practical application of NS models versus GATSMs, highlighting that while both provide approximations to the "true" model, NS models have relative advantages from several perspectives. Section 6 concludes.

2　The generic Gaussian affine term structure model

The generic GATSM specified in this section parallels appendix A of Dai and Singleton (2002), and 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 Duffee (2002).[3] However, it is worth highlighting three points of context for this article.

First, while the state variables are completely generic, and so could represent points on the yield curve as in Duffie and Kan (1996), it is convenient for the subsequent discussion in section 5 to consider them as (potentially unobserved) economic and financial factors within the underlying economy. This follows the Duffie and Kan (1996) p. 321 interpretation that the state variables in an affine model can always, in principle and under standard assumptions, be related back to economic factors (e.g preferences, technology, consumption, inflation, etc.) within a general equilibrium model. Duffie and Singleton (1999) extends the latter result to financial factors.[4]

Second, to make the exposition more transparent from the perspective of the original NS model, this article works with the forward rate curve associated with the generic GATSM rather than the bond prices and interest rate curves typically used in affine term structure models.

Third, being fully Gaussian, the results for relating the generic GATSM to NS models do not extend to ATSMs with Cox, Ingersoll and Ross (1985)/square-root dynamics. Appendix A illustrates this by example, and briefly discusses the practical implications.

Define the instantaneous short rate at time t as $r(t) = \rho_0 + \rho_1' X(t)$, where ρ_0 is a constant, $X(t)$ is a $N \times 1$ vector of state variables, and ρ_1 is a constant

---

[3]In the notation of Dai and Singleton (2000) the specification is $A_0(N)$.

[4]While Duffie and Singleton (1999) focuses on default risk, pp. 193-94 notes that liquidity risk and repurchase effects, etc. may also be treated in a similar manner.

$N \times 1$ vector. Under the physical P measure, the state variables follow the process $dX(t) = K_P [\theta_P - X(t)] dt + \Sigma dW_P(t)$, where K_P is a constant $N \times N$ mean-reversion matrix, θ_P is a constant steady-state $N \times 1$ vector for $X(t)$, Σ is a constant $N \times N$ volatility matrix, and $W_P(t)$ is an $N \times 1$ vector of independent Brownian motions. Define the market prices of risk as $\Gamma(t) = \Sigma^{-1} [\gamma_0 + \gamma_1 X(t)]$, where γ_0 is a constant $N \times 1$ vector and γ_1 is an $N \times N$ matrix. Under the risk-neutral Q measure, the state variables follow the process $dX(t) = K_Q [\theta_Q - X(t)] dt + \Sigma dW_Q(t)$, where $dW_Q(t) = dW_P(t) + \Gamma(t) dt$, $K_Q = K_P + \gamma_1$, and $\theta_Q = (K_P + \gamma_1)^{-1} (K_P \theta_P - \gamma_0)$. Zero-coupon bond prices for the generic GATSM are $P(t, T) = \exp[A(t, T) + B(t, T)' X(t)]$, where $B(t, T) = [\exp(-K_Q' \tau) - I] (K_Q')^{-1} \rho_1$ with $\tau = T - t$ the time to maturity, and I the $N \times N$ identity matrix. The full expression for $A(t, T)$ is provided in Dai and Singleton (2002), but this article requires only the summary results that it has the functional form $-a_0 \tau + A(\tau)$ and is required for the system to be arbitrage free.

From Heath, Jarrow and Morton (1992), instantaneous forward rates are defined as $f(t, T) = -\partial \log P(t, T) / \partial T$, and so:

$$f(t, T) = a_0 + \left\{ \exp(-K_Q' \tau) (K_Q')^{-1} \rho_1 \right\}' X(t) - \frac{\partial}{\partial \tau} A(\tau) \qquad (1)$$

Now express K_Q' in eigensystem form; i.e $K_Q' = Z \Psi Z^{-1}$, where Z is the $N \times N$ non-singular matrix of eigenvectors each normalised to 1, and Ψ is the $N \times N$ diagonal matrix containing the N eigenvalues $(\lambda_1, \ldots, \lambda_n, \ldots, \lambda_N)$ that are assumed to be unique and positive.[5] Hence, $\exp(-K_Q' \tau) = \exp(-Z \Psi Z^{-1} \tau) = Z \exp(-\Psi \tau) Z^{-1} = Z \Lambda Z^{-1}$, where Λ is an $N \times N$ diagonal matrix containing the N elements $[\exp(\lambda_1 \tau), \ldots, \exp(\lambda_n \tau), \ldots, \exp(\lambda_N \tau)]$. The forward rates in equation 1 are then $f(t, T) = a_0 + \left\{ Z \Lambda Z^{-1} (K')^{-1} \rho_1 \right\}' X(t) - \frac{\partial}{\partial \tau} A(\tau)$, which can be expressed equivalently as:

$$f(t, T) = a_0 + \sum_{n=1}^{n_0} q_n(t) \exp(-\lambda_n \tau) + \sum_{n=n_0+1}^{N} q_n(t) \exp(-\lambda_n \tau) - \frac{\partial}{\partial \tau} A(\tau) \quad (2)$$

where the coefficients $q_n(t)$ represent the collection of coefficients associated with each unique $\exp(-\lambda_n \tau)$ term that arises from the full matrix multiplication of $\left\{ Z \Lambda Z^{-1} (K')^{-1} \rho_1 \right\}' X(t)$. For convenience in the example of the following section (but without loss of generality) it is assumed that the $q_n(t) \exp(-\lambda_n \tau)$ components have been re-ordered from the smallest to the largest eigenvalue, and then divided into two groups. The first group contains the components

---

[5] This follows the standard assumption in Duffie and Kan (1996) and Dai and Singleton (2002).

with eigenvalues λ_1 to λ_{n_0} that are close to zero (i.e persistent components, given their slow exponential decay by time to maturity τ) and the second group contains the eigenvalues λ_{n_0+1} to λ_N that are not close to zero (i.e non-persistent components).

3 From the generic GATSM to the original NS model

From the exact expression of the generic GATSM forward rate curve in equation 2, three approximations are required to reproduce the original NS model. First, drop the arbitrage-free (hereafter AF) term $\frac{\partial}{\partial \tau} A\left(\tau\right)$, while bearing in mind that an NS-consistent representation for $\frac{\partial}{\partial \tau} A\left(\tau\right)$ can be introduced later as discussed further below. Second, for the first group of eigenvalues where $\lambda_n \simeq 0$, the first term of the Taylor expansion is $\exp\left(-\lambda_n \tau\right) \simeq 1$. Third, for the second group of eigenvalues where $\lambda_n \gg 0$, express them relative to ϕ that is defined as the mean eigenvalue for the group; i.e $\phi =$ mean$(\lambda_{n_0+1}, \ldots, \lambda_N)$,[6] so that $\lambda_n = \phi\left(1 - \delta_n\right)$ and then $\exp\left(-\lambda_n \tau\right) = \exp\left(-\phi \tau\right) \exp\left(\delta_n \phi \tau\right)$. Now take the first-order Taylor approximation $\exp\left(\delta_n \phi \tau\right) \simeq 1 + \delta_n \phi \tau$, so $q_n\left(t\right) \exp\left(-\lambda_n \tau\right) \simeq q_n\left(t\right) \exp\left(-\phi \tau\right) + q_n\left(t\right) \delta_n \phi \tau \exp\left(-\phi \tau\right)$. Substituting these results into equation 2 gives:

$$f\left(t, T\right) \simeq a_0 + \sum_{n=1}^{n_0} q_n\left(t\right) + \exp\left(-\phi \tau\right) \sum_{n=n_0+1}^{N} q_n\left(t\right) + \phi \tau \exp\left(-\phi \tau\right) \sum_{n=n_0+1}^{N} q_n\left(t\right) \delta_n$$
(3)

This is precisely the functional form of the original NS model of the forward rate curve, i.e:

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

where 1, $\exp\left(-\phi \tau\right)$, and $\phi \tau \exp\left(-\phi \tau\right)$ are the forward rate factor loadings for the original NS model, and $L\left(t\right) = a_0 + \sum_{n=1}^{n_0} q_n\left(t\right)$, $S\left(t\right) = \sum_{n=n_0+1}^{N} q_n\left(t\right)$, and $C\left(t\right) = \sum_{n=n_0+1}^{N} q_n\left(t\right) \delta_n$ are the coefficients for the original NS model. The usual relationship $R_{\text{NS}}\left(t, T\right) = \frac{1}{\tau} \int_0^\tau f_{\text{NS}}\left(s\right) ds$ then produces the familiar form of the original NS model.[7]

This shows that the original NS model represents the generic GATSM with

---

[6] Any other central measure of $(\lambda_{n_0+1}, \ldots, \lambda_N)$ would suffice for the exposition in this article. In practice, ϕ is an estimated parameter.

[7] That is, the interest rate curve is $R_{\text{NS}}\left(t, T\right) = L\left(t\right) + S\left(t\right) \left(\frac{1 - \exp(-\phi \tau)}{\phi \tau}\right) + C\left(t\right) \left(\frac{1 - \exp(-\phi \tau)}{\phi \tau} - \exp\left(-\phi \tau\right)\right)$, where $L\left(t\right)$, $S\left(t\right)$, and $C\left(t\right)$ are already defined in the text.

no AF term (i.e $\frac{1}{\tau} \int_0^\tau \frac{\partial}{\partial s} A(s) \, ds = \frac{1}{\tau} A(\tau)$ that would have carried through from the generic GATSM forward rate curve), the persistent generic GATSM components "naturally" approximated to zeroth order by the Level component, and the non-persistent components "naturally" approximated to first order by the combination of Slope and Curvature components. The adverb "naturally" is used here in the sense that each additional NS component corresponds precisely to an additional term in the Taylor expansion of the generic GATSM.[8]

The AF version of the original NS model can be obtained by including the AF term for $R_{\text{NS}}(t, T)$ from Christensen, Diebold and Rudebusch (2007), which adds a further six parameters (the unique covariances for innovations in the NS coefficients) to the single parameter ϕ. Including that AF term creates a model that is AF with respect to the NS factor loadings, and the AF term is the approximation of $\frac{1}{\tau} A(\tau)$ from the generic GATSM.[9]

4   Classifying and applying NS models from a GATSM perspective

By following the example in the previous section, any specific NS model may be classified as a particular representation of the generic GATSM. The key aspects are: (1) the number of groups of non-zero eigenvalues assumed, which determines how many mean eigenvalue parameters (e.g ϕ for the original NS model) are required; (2) the degree of approximation chosen around each mean eigenvalue, which determines the number of components associated with each mean eigenvalue (e.g two for the original NS model); and (3) whether the AF term is included, which determines if the NS model is AF with respect to its factor loadings. The various permutations of each of those aspects can obviously generate a wide variety of NS models, but this section discusses just the range of NS models already in use and three parsimonious variants.

The Christensen, Diebold and Rudebusch (2008)/NS model, with the forward rate form $f(t) = L(t) + S_1(t) \exp(-\phi_1 \tau) + C_1(t) \phi_1 \tau \exp(-\phi_1 \tau) + S_2(t) \times \exp(-\phi_2 \tau) + C_2(t) \phi_2 \tau \exp(-\phi_2 \tau) + AF(\tau)$, is the most comprehensive model within the NS class to date. From the perspective of this article, it represents the persistent generic GATSM components to a zeroth-order approximation,

---

[8]Obviously, any other set of functions could be used to approximate the generic GATSM term structure, but the approximation would "unnatural" in the sense that the functions would not precisely represent a Taylor expansion of the generic GATSM.

[9]The Christensen et al. (2007)/NS model is actually a particular three-factor GATSM that by construction reproduces the three NS factor loadings with the NS coefficients as state variables. Hence, the NS AF term equals $\frac{1}{\tau} A(\tau)$ for that particular model, but is an approximation to $\frac{1}{\tau} A(\tau)$ when representing a generic GATSM from the more general perspective of this article.

the non-persistent components with two groupings of non-zero eigenvalues (i.e $\phi_1 = \text{mean}(\lambda_{n_0+1}, \ldots, \lambda_{n_1})$ and $\phi_2 = \text{mean}(\lambda_{n_1+1}, \ldots, \lambda_N)$) each to a first-order approximation, and includes the $AF(\tau)$ term to ensure the model is AF with respect to the five factor loadings.[10] The Svensson (1995)/NS model omits the AF term and $S_2(t) \exp(-\phi_2 \tau)$.

At the other extreme, the Diebold, Li and Yue (2008)/NS model with the forward rate form $f(t) = L(t) + S_1(t) \exp(-\phi \tau)$ is the most parsimonious representation of the generic GATSM. It ignores the AF adjustment and represents both the persistent and non-persistent components of the generic GATSM to zeroth order. The Diebold et al. (2008)/NS model is therefore "balanced", in the sense that the degree of approximation is the same for the persistent and non-persistent components.

Parsimonious balanced NS models can be obtained in three other ways: (1) approximate the persistent generic GATSM components to first order to match the first-order approximation inherent in the Slope and Curvature components,[11] resulting in $f(t) = L(t) + L_1(t)\tau + S(t)\exp(-\phi\tau) + C(t)\phi\tau\exp(-\phi\tau)$, where $L_1(t) = -\sum_{n=1}^{n_0} q_n(t)\lambda_n$; (2) add another Slope component to the Diebold et al. (2008)/NS model, i.e $f(t) = L(t) + S_1(t)\exp(-\phi_1\tau) + S_2(t) \times \exp(-\phi_2\tau)$, which would have two non-zero groups of eigenvalues (as for the Christensen et al. (2008)/NS model) with each approximated to zeroth order; (3) decompose the Level component of the Diebold et al. (2008)/NS model into the constant and the persistent exponential decay terms, i.e $f(t) = a_0 + L(t)\exp(-\phi_1\tau) + S_1(t)\exp(-\phi_2\tau)$ where $0 \simeq \phi_1 \ll \phi_2$ and $L(t) = \sum_{n=1}^{n_0} q_n(t)$. All of these NS models can be made AF with respect to their components as in Christensen et al. (2007, 2008), or directly via the Heath et al. (1992) framework as in the example of Krippner (2006).

Choosing the particular NS model to apply has in the past been largely an empirical matter; i.e trading off parsimony against goodness of fit to the yield curve data. However, the classification above suggests a systematic approach to introducing or omitting terms if maintaining a correspondence with the GATSM class is desired. For example, the Svensson (1995)/NS model would be avoided because the second Curvature term $C_2(t)$ cannot by itself represent a natural first-order approximation of a generic GATSM, given that the first term of the approximation to the second group of eigenvalues is omitted. Either the second Slope term $S_2(t)$ should also be added (creating the non-AF analogue of the Christensen, Diebold and Rudebusch (2008)/NS model), or the second

---

[10]Analogous to the Christensen et al. (2007)/NS model, the Christensen et al. (2008)/NS model is actually a particular five-factor GATSM, but is an approximation when representing the generic GATSM.

[11]That is, $\exp(-\lambda_n\tau) \simeq 1 - \lambda_n\tau$, and so $\sum_{n=1}^{n_0} q_n(t)\exp(-\lambda_n\tau) \simeq \sum_{n=1}^{n_0} q_n(t) - \sum_{n=1}^{n_0} q_n(t)\lambda_n\tau$.

Curvature term $C_2(t)$ dropped (recreating the original NS model). Another aspect is more subtle: from the strict perspective of maintaining a foundation within the generic GATSM, NS models should be applied with a constant decay parameter ϕ (or parameters ϕ_1 and ϕ_2) because that corresponds to a constant mean reversion matrix in the generic GATSM.[12]

The theoretical case for adding an AF term was originally established in Björk and Christensen (1999), and ideally it should be included in empirical applications to maintain consistency with the generic GATSM. The AF term also makes the cross-sectional and time series properties of the NS model consistent, which is important in applications where risk and volatility are key aspects (e.g assessing term premia or pricing options related to interest rates). That said, the practical relevance of the AF term has been questioned in Coroneo et al. (2008), and the empirical results in Christensen et al. (2007, 2008) using an NS model or the corresponding AF NS model are similar. Appendix B also shows that the typical estimation of NS or AFNS coefficients by OLS on constant maturity zero-coupon yields will yield time series of coefficients within a constant of each other. Hence, using either set of coefficients as data purely for time series analysis will yield equivalent econometric results.

5 NS models versus GATSMs in practice

The corollary to the exposition so far provides a compelling case for applying GATSM-corresponding NS (hereafter GCNS) models as standard tools for yield curve modelling and analysis in economics and finance, rather than applying GATSMs directly. For the purposes of illustration, assume that the "true" GATSM generating the yield curve data has N factors with arbitrary interactions in the mean-reversion, risk, and volatility matrices, the GATSM applied in practice has J factors with potentially restricted interactions, and $J < N$.

These assumptions are readily justified. For example, regarding N, Hördahl, Tristani and Vestin (2006) models the nominal government yield curve via the four economic factors of output growth, inflation, monetary policy, and an inflation target, but also notes (p. 408) that "The model is certainly too stylised - for example, in its ignoring foreign variables or the exchange rate - to provide a fully-satisfactory account of German macroeconomic dynamics.". Additional financial market factors influencing the government yield curve could include liquidity and repurchase effects (see Fleming (2003) and Fisher (2002) respec-

---

[12]While it would be tempting to interpret time variation in ϕ as representing time variation in the mean-reversion matrix $K_Q = K_P + \gamma_1$, a generic GATSM that formally allowed for such flexibility would result in more complex factor loadings that could no longer "naturally" be reduced to the NS factor loadings as in section 3.

tively), and corporate yield curves can contain many more factors in addition to the government yield curve, including default risk, corporate bond liquidity, and the three Fama and French (1993) factors (see Elton, Gruber, Agrawal and Mann (2001)).

Regarding J, three-factor GATSMs are commonly applied in practice.[13] And even two-factor and three-factor GATSMs are typically estimated with various assumptions to remove over-parametrisation (including prior restrictions on parameters and/or the "rather arbitrary procedure" of setting statistically-insignificant parameters to zero).[14]

Both GCNS models and the J-factor GATSM will provide approximations to the N-factor GATSM, but GCNS models have relative advantages from four perspectives.

The first perspective is the ease of application, where GCNS models have the property of reliable convergence (given their parsimony, and estimation via OLS or non-linear LS depending on whether the yield curve data are zero-coupon yields or bond prices) versus the problem of multiple local maxima for J-factor GATSMs.[15]

The second perspective is the interpretation of output. The mechanical output from a GCNS model is a time series of estimated coefficients that describe the evolution of the fitted yield curve in terms of the intuitive Level, Slope, and Curvature components. The economic and financial factors obviously cannot be identified from the reduced-form structure, but may nevertheless be suggested by the context of the application and the result that the Level and non-Level GCNS coefficients respectively reflect the persistent and non-persistent components of the N-factor GATSM. For example, one might anticipate the GCNS Level coefficient for a nominal government yield curve to covary with inflation (a persistent macroeconomic variable), and the non-Level components to covary with output growth (a non-persistent macroeconomic variable), and those results are established empirically in Diebold, Rudebusch and Aruoba (2006)

---

[13]Three factors are often justified on the grounds that three principal components explain the vast majority of US government yield curve movements, but it is also a practical limit for unrestricted or mildly-restricted GATSMs because the number of parameters is already high. For example Duffee (2002) notes on p. 418 that "Litterman and Sheinkman (1991) find that three factors explain the vast majority of Treasury bond price movements. This is fortunate, because general three-factor affine models are already computationally difficult to estimate owing to the number of parameters. Adding another factor would make this investigation impractical.". The number of parameters in the canonical form of GATSM models is $1 + 3J + 2J^2$, and so would rise from 28 to 45 parameters.

[14]Examples are Dai and Singleton (2002), Duffee (2002), Hördahl et al. (2006) and Rudebusch and Wu (2007). The quote is from Kim and Orphanides (2005) p. 10.

[15]For example, Rudebusch and Wu (2007) p. 406 mentions "it should be noted that estimation of the standard no-arbitrage latent factor model, which is highly nonlinear, often appears to be plagued by numerical problems."

using the original NS model.

The mechanical output from the J-factor GATSM is a time series of the J state variables $X(t)$ that describe the evolution of the fitted yield curve. However, the shape of the yield curve components associated with each state variable varies between applications, and so are less intuitive than for a GCNS model. In principle, the J-factor GATSM should have the advantage that estimates of individual parameters will reveal the structure underlying the yield curve, such as the interactions between the different factors in the mean reversion and volatility matrices. However, as already noted, the interpretation can be sensitive to the assumptions for reducing over-parametrisation even for a low number of factors; e.g Kim and Orphanides (2005) p. 11 notes that such assumptions risk "introducing significant biases in the resulting estimated model", and can potentially lead to "different conclusions with the same specification".

The third perspective is parsimony in the application to any yield curve. That is, the benefit of the reduced form nature of a GCNS model is that it will always parsimoniously represent the N-factor GATSM to a known and precise order of approximation regardless of (or even without knowledge of) the actual number of state variables, their nature, and their interactions. The J-factor GATSM cannot be extended arbitrarily due to problems with over-parametrisation, and so will always be limited to representing the N-factor GATSM as if it only had J factors. That imposes a prior structure on the data and introduces the issue of the $N - J$ omitted variables. Together with the parametrisation-reduction assumptions already mentioned earlier, it is not transparent how exactly the J-factor GATSM is representing the "true" N-factor GATSM.

The fourth perspective is the results in empirical applications. In cases where direct comparisons are available, GCNS models have generally given similar or superior results to GATSMs.[16] Examples are: (1) fitting the yield curve (the original NS model has lower RMSEs than the three-factor GATSM of Duffee (2002));[17] (2) forecasting the yield curve (e.g Vincente and Tabak (2008) obtains lower forecast RMSEs with the original NS model than a three-

---

[16]NS models have also applied successfully to other topics that have been investigated with GATSMs, but the empirical results are not directly comparable for various reasons. Examples include: (1) modelling non-government yields and liquidity conditions, e.g Christensen, Lopez and Rudebusch (2009); (2) the uncovered interest parity puzzle, e.g Chen and Tsang (2009); (3) portfolio management, e.g Diebold, Ji and Li (2006); and (4) investigating international yield curve transmissions.

[17]The result was obtained by fitted the original NS model to the Duffee (2002) US government yield curve data, giving an average RMSE of 53.2 basis points for the six yields. The three-factor GATSM estimated in Duffee (2002) had an average RMSE of 116.6 basis points. I thank Gregory Duffee for making the data available on his website.

factor GATSM);[18] (3) macrofinance (e.g in the Diebold et al. (2006) application mentioned earlier, the parsimony of the original NS model allows for bidirectionality between the yield curve and macroeconomic variables, while Diebold et al. (2005) notes that the complexity of the five-factor GATSM from Ang and Piazzesi (2003) requires prior restrictions that limit the directionality from macroeconomic factors to yields); and (4) monitoring inflation compensation (e.g using a modified AF NS model to jointly model the nominal and inflation-indexed yield curves, Christensen, Lopez and Rudebusch (2008) report better correlations between model-implied and surveyed 5-year inflation expectations than in D'Amico, Kim and Wei (2008) which uses three-factor GATSMs for both the nominal and inflation-indexed yield curves).

6   Conclusion

This article establishes most NS models as reduced-form representations of the generic GATSM. That result provides a compelling case for applying GATSM-consistent NS models as standard tools for yield curve analysis in economics and finance: users get the well-established pragmatic benefits of NS models along with an assurance that the model is consistent with a well-accepted set of principles and assumptions for modelling the yield curve and its dynamics. Moreover, GATSM-consistent models are guaranteed to parsimoniously and reliably represent the yield curve associated with any GATSM regardless of the "true" GATSM specification (in terms of the number of state variables, their nature, and their interactions within the underlying economic and financial system) that generated the data.

---

[18] Diebold and Li (2006) also obtains lower forecast RMSEs than Duffee (2002), but the sample is different.

References

Ang, A. and Piazzesi, M. (2003), 'A no-arbitrage vector autoregression of term structure dynamics with latent variables', *Journal of Monetary Economics* **50**, 745–787.

Bank for International Settlements (2005), 'Zero-coupon yield curves: technical documentation', *BIS Papers* **25**.

Björk, T. and Christensen, B. (1999), 'Interest rate dynamics and consistent forward rate curves', *Mathematical Finance* **9(4)**, 323–348.

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. (2007), 'The affine arbitrage-free class of Nelson-Siegel term structure models', *NBER Working Paper Series* **13611**.

Christensen, J., Diebold, F. and Rudebusch, G. (2008), 'An arbitrage-free generalized Nelson-Siegel term structure model', *Econometrics Journal (forthcoming)* .

Christensen, J., Lopez, J. and Rudebusch, G. (2008), 'Inflation expectations and risk premiums in an arbitrage-free model of nominal and real bond yields', *Working Paper, Federal Reserve Bank of San Francisco* **2008-34**.

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**.

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. (1985), '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.

D'Amico, S., Kim, D. and Wei, M. (2008), 'Tips from TIPS: the informational content of Treasury Inflation-Protected Securities', *Working Paper, Federal Reserve Board* **2008-30**.

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.

Elton, E., Gruber, M., Agrawal, D. and Mann, C. (2001), 'Explaining the rate spread on corporate bonds', *Journal of Finance* **1**, 247–277.

Fama, E. and French, K. (1993), 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics* **33**, 3–57.

Fisher, M. (2002), 'Special repo rates: an introduction', *Federal Reserve Bank of Atlanta Economic Review* **Second Quarter**.

Fleming, M. (2003), 'Measuring Treasury market liquidity', *Federal Reserve Bank of New York Bulletin* **9(3)**, 83–108.

Gorovoi, V. and Linetsky, V. (2004), 'Black's model of interest rates as options, eigenfunction expansions and Japanese interest rates', *Mathematical Finance* **14(1)**, 49–78.

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.

Hördahl, P., Tristani, O. and Vestin, D. (2006), 'A joint econometric model of macroeconomic and term structure dynamics', *Journal of Econometrics* **131**, 405–444.

Hull, J. (2000), *Options, Futures and Other Derivitives, Fourth Edition*, Prentice Hall.

Kim, D. and Orphanides, A. (2005), 'Term structure estimation with survey data of interest rate forecasts', *Federal Reserve Board Working Paper* **2005-48**.

Krippner, L. (2006), 'A theoretically consistent version of the Nelson and Siegel class of yield curve models', *Applied Mathematical Finance* **13(1)**, 39–59.

Litterman, R. and Sheinkman, J. (1991), 'Common factors affecting bond returns', *Journal of Fixed Income* **2**, 54–61.

Rudebusch, G. and Wu, T. (2007), 'A macro-finance model of the term structure, monetary policy, and the economy', *The Economic Journal* **118**, 906–926.

Svensson, L. (1995), 'Estimating forward interest rates with the extended Nelson and Siegel model', *Quarterly Review, Sveriges Riksbank* **1995(3)**, 13–26.

Vincente, J. and Tabak, B. (2008), 'Forecasting bond yields in the Brazilian fixed income market', *International Journal of Forecasting* **24**, 490–497.

A CIR dynamics

This appendix shows by example that ATSMs with Cox et al. (1985)/square-root dynamics cannot be "naturally" approximated by NS factor loadings. 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.

Hence, $P(t,T) = \exp\left[\sum_{n=1}^{N} A_n(t,T) + B_n(t,T)X_n(t)\right]$ where each $B_n(t,T)$ has the standard Cox et al. (1985) form:[19]

$$B_n(t,T) = \frac{2[1 - \exp(\gamma_n \tau)]}{(\gamma_n + \kappa_n)[\exp(\gamma_n \tau) - 1] + 2\gamma_n} \tag{5}$$

with $\gamma_n = \sqrt{\kappa_n^2 + 2\sigma_n^2}$. The associated forward rate curve is:

$$f(t,T) = a_0 + \sum_{n=1}^{N} \frac{4\gamma_n^2 \exp(\gamma_n \tau)}{[(\gamma_n + \kappa_n)[\exp(\gamma_n \tau) - 1] + 2\gamma_n]^2} X_n(t) - \frac{\partial}{\partial \tau} A_n(\tau) \tag{6}$$

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. Therefore, $f(t,T)$ cannot be "naturally" approximated by NS factor loadings following the procedure in section 3.

This result suggests that NS models should not be applied in situations where a non-Gaussian ATSM would be considered more appropriate than a GATSM; e.g when the probability of zero interest rates over the horizon of interest is material. Alternatively, analogous to Gorovoi and Linetsky (2004), a GCNS model could still be used to represent the "shadow yield curve" (i.e where interest rates are allowed to go negative) subject to an appropriate treatment being applied to derive an observed yield curve with non-zero interest rates.

---

[19]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   The time-series equivalence of NS and AF NS models

This appendix shows that when NS models are estimated with the commonly-used method of OLS on zero-coupon continuously-compounding yields of constant maturity, the estimated time series of coefficients $\beta_{\mathrm{NS}}(t)$ and the coefficients of the AF version of the NS model $\beta_{\mathrm{AF}}(t)$ will be identical to within a constant vector.

The AF NS model is the linear equation: $[Y(t) - D] = V\beta_{\mathrm{AF}}(t) + \varepsilon_{\mathrm{AF}}(t)$, where the yield curve data are $Y(t) = \{R_1(t, \tau_1), \ldots, R_K(t, \tau_I)\}'$, the AF terms underlying the yield curve data are $D = \left\{ \frac{1}{\tau_1} \int_0^{\tau_1} A(s)\, ds, \ldots, \frac{1}{\tau_K} \int_0^{\tau_K} A(s)\, ds \right\}'$, the NS factor loadings are $V_k(\tau_k) = \left\{ 1, \frac{1 - \exp(-\phi\tau_k)}{\phi\tau_k}, \frac{1 - \exp(-\phi\tau_k)}{\phi\tau_k} - \exp(-\phi\tau_k) \right\}$, so $V = \{V_1(\tau_1), \ldots, V_K(\tau_K)\}'$, and the AF NS coefficients are $\beta_{\mathrm{AF}}(t) = \{L(t), S(t), C(t)\}'$. The OLS estimate is $\beta_{\mathrm{AF}}(t) = [V'V]^{-1} V'[Y(t) - D] = [V'V]^{-1} V'Y(t) - [V'V]^{-1} V'D$.

The NS model will be represented by linear equation: $Y(t) = V\beta_{\mathrm{NS}}(t) + \varepsilon_{\mathrm{NS}}(t)$, which is the same as the AF NS system without the AF terms D. The OLS estimate is therefore $\beta_{\mathrm{NS}}(t) = [V'V]^{-1} V'Y(t)$.

Hence, $\beta_{\mathrm{AF}}(t) = \beta_{\mathrm{NS}}(t) + [V'V]^{-1} V'D$, where the latter term is a constant vector.