CENTRE FOR ECONOMETRIC ANALYSIS
CEA@BAYES

https://www.bayes.city.ac.uk/faculties-and-research/centres/cea

Bayes Business School (formerly Cass)
Faculty of Finance
106 Bunhill Row
London EC1Y 8TZ

Forecasting The Yield Curve: The Role of Additional and Time-Varying Decay Parameters, Conditional Heteroscedasticity, and Macro-Economic Factors

João F. Caldeira, Werley C. Cordeiro, Esther Ruiz, and André A.P. Santos

CEA@Bayes Working Paper Series

WP–CEA–07-2023

Forecasting the yield curve: the role of additional and time-varying decay parameters, conditional heteroscedasticity, and macro-economic factors

João F. Caldeira

Dept. Economics, Universidade Federal de Santa Catarina, Florianópolis (Brazil)

Werley C. Cordeiro

Dept. Economics, Universidade Federal de Santa Catarina, Florianópolis (Brazil)

Esther Ruiz*

Dept. Statistics, Universidad Carlos III de Madrid, Getafe (Spain)

André A.P. Santos

Dept. Quantitative Methods, CUNEF Universidad, Madrid (Spain)

July 26, 2023

Abstract

In this paper, we analyze the forecasting performance of several parametric extensions of the popular Dynamic Nelson-Siegel (DNS) model for the yield curve. Our focus is on the role of additional and time-varying decay parameters, conditional heteroscedasticity, and macroeconomic variables. We also consider the role of several popular restrictions on the dynamics of the factors. Using a novel dataset of end-of-month continuously compounded Treasury yields on US zero-coupon bonds and frequentist estimation based on the extended Kalman filter, we show that a second decay parameter does not contribute to better forecasts. In concordance with the preferred habitat theory, we also show that the best forecasting model depends on the maturity. For short maturities, the best performance is obtained in a heteroscedastic model with a time-varying decay parameter. However, for long maturities, neither the time-varying decay nor the heteroscedasticity plays any role, and the best fit is obtained in the basic DNS model with the shape of the yield curve depending on macroeconomic activity. Finally, we find that assuming non-stationary factors is helpful in forecasting at long horizons.

Keywords: Dynamic Nelson-Siegel-Svensson model, Time-varying decay parameter, Term Structure, Extended Kalman filter

MSC classification: 62P20; 91B84.
JEL classification: C53; E43; G17.

*Corresponding author. E-mail: ortega@est-econ.uc3m.es

1 Introduction

The term structure of interest rates, also referred to as the yield curve of government bonds, is a crucial component of financial markets due to its high liquidity and low credit risk. Accurate predictions of this yield curve are valuable for investors for the pricing of interest rate-contingent assets, the construction of bond portfolios, and the management of financial risk, as outlined by Hodges and Schaefer (1977) and Ronn (1987). Yield curve forecasting holds significance in macroeconomic and monetary policy analysis. Studies such as Ang et al. (2006), Estrella and Hardouvelis (1991), and Giacomini and Rossi (2006) explore the yield curve's predictive capability for economic growth, while Ang et al. (2006) and Estrella and Trubin (2006) analyze its function as a leading recession indicator. Additional contributions include Wu and Xia (2016)'s and Carriero et al. (2018)'s examination of the macroeconomic impact of monetary policy at the zero lower bound on inflation. Amisano and Tristani (2023) model the interplay between monetary policy and long-term yields across the business cycle. Term structure forecasts are also instrumental in developing unconventional monetary policy tools such as yield curve control and forward guidance, as discussed by Gambacorta et al. (2014), Kuttner (2018), Bernanke (2020), Inoue and Rossi (2021), and Rossi (2021), among others.

During the last two decades, a large number of works have been devoted to developing alternative methodologies for modelling and forecasting the term structure. These methodologies can be classified into three main alternative approaches. First, several models focus on fitting the term structure at a point in time to ensure that no-arbitrage possibilities exist. These models are usually estimated using regression-based procedures; see the recent work by Golinski and Spencer (2021) and the references therein. Many empirical studies suggest that imposing no-arbitrage conditions does not generally lead to more accurate predictions; see, for example, Joslin et al. (2011).

Second, several authors estimate the term structure using affine models, originally characterised by Duffie and Kan (1996). Affine models specify bond yields as linear functions of multiple state variables; see Duffee and Stanton (2012) for a comparison of alternative estimators of affine models, including Kalman filter-based estimators. Recent estimation techniques of affine models for the term structure rely on Bayesian procedures, which are computationally demanding; see, for example, Carriero et al. (2021) for a recent Bayesian estimator of the canonical affine term structure model, in its equivalent but computationally more stable representation of Joslin et al. (2011).

Finally, a large number of works represent the term structure of interest rates using the dynamic Nelson-Siegel (DNS) model. Motivated by the rational expectation theory, Nelson and Siegel (1987) express spot interest rates in terms of forward rates and propose a three factor model for the yield curve, with the factors representing its level, slope and curvature. As a result of this interpretation of the factors, the factor loadings are heavily parameterised depending on a single exponential decay rate parameter; see Coroneo et al. (2011) and Krippner (2012) for the connection between the DNS model and affine and arbitrage-free models, respectively. DNS models were popularised by Diebold and Li (2006), who propose modelling the dynamic evolution of the yield curve allowing for time-varying factors. The resulting DNS model is a major workhorse among academics and it is also widely used in the financial community as well as by Central Banks to decide about its monetary policy; see, for example, Diebold et al. (2008), Yu and Salyards (2009), Christensen et al. (2011), Laurini and Hotta (2014) and Caldeira et al. (2016), for some few selected empirical applications, and BIS (2005), Almeida and Vicente (2008) and ECB (2018) for its implementation by practitioners. In spite of its popularity, the Achilles heal of the DNS model lies in its poor forecasting performance, with forecasts that hardly beat those obtained by a random walk model; see Matsumura et al. (2011), De Rezende and Ferreira (2013) and Choi and Kang (2023), among many others. As far as we are concerned, Christensen et al. (2011) and Coroneo et al. (2016) are among the few that conclude that the DNS has a good forecasting performance improving the random walk forecasts. Two main reasons have been put forward to explain this poor performance. First, the specification of the model and, in particular, the restrictions imposed on the factor loadings, may not hold in practice; see Jungbacker et al. (2014) and Carriero et al. (2021). Alternatively, there is also a literature suggesting that the performance of DNS type models has deteriorated in the post global financial crisis due to the low variability of interest rates during the zero-lower-bound interest rate constraints period from 2008 to 2012; see, for example, Diebold and Rudebusch (2013) and Altavilla et al. (2017).

The main contribution of this paper is the analysis of the forecasting performance of a very general and flexible specification of the DNS model, which encapsulates four popular extensions of the original specification often considered in separate empirical applications. We investigate whether these extensions could mitigate the adverse effects of potential misspecification. The first extension considered is due to Svensson (1994), who proposes a four-factor version of the DNS model; see Almeida et al. (2018) and Swanson et al. (2020) for applications with four factors. Other two important extensions of the DNS model are due to Koopman et al. (2010). First, they propose allowing the decay parameter that controls the shape of the yield curve to evolve over time; see Laurini and Hotta (2010), Hevia et al. (2015) and Swanson et al. (2020), for other proposals in which the decay parameter is not constant over time. Second, Koopman et al. (2010) propose representing the overall volatility by a conditionally heteroscedastic GARCH model. Allowing for the evolution of volatility seems to be an important characteristic of the yield curve. Caldeira et al. (2010) and Laurini and Caldeira (2016) also allow for time-variation in both the decay parameter and volatilities. Finally, the fourth extension of the DNS model considered in this paper, is the inclusion of macroeconomic variables to explain the shape of the yield curve as proposed by Diebold et al. (2006); see Gürkaynak and Wright (2012) and Morley (2016) for surveys on the relationship between the yield curve and the macroeconomic activity and Bianchi et al. (2009), Favero et al. (2012), Exterkate et al. (2013), Pedersen and Swanson (2019), and Swanson et al. (2020), for some selected applications.

In order to disentangle the role played by the four extensions described above in forecasting the yield curve, the extended DNS model with two time-varying decay parameters, macroeconomic variables, and conditional heteroscedasticity, is used to fit and forecast a novel data set of end-of-month continuously compounded Treasury yields on US zero-coupon bonds. The data set is divided into an in-sample period from January 1972 to December 1993, and an out-of-sample period, from January 1994 to December 2019. Estimation of the model parameters is carried out by maximum likelihood (ML) using the extended Kalman filter (EKF) to obtain the prediction error decomposition form of the likelihood and pseudo out-of-sample forecasts of yields are obtained using a rolling-window scheme. The out-of-sample performance is analysed with forecasts obtained up to 2008 and with forecasts obtained over the full out-of-sample period up to 2019. Several important conclusions are obtained from this analysis. First, we show that the second decay parameter does not have any role in obtaining a better forecasting performance of the factor model. Second, when looking at results over the full out-of-sample period, we show that the best forecasts are obtained with different specifications depending on the maturity. For short maturities, the best performance is obtained in a heteroscedastic DNS model with time-varying decay. However, for long maturities, the simplest homoscedastic model with constant decay performs better if the shape of the yield curve depends on economic activity. In any case, in concordance with results in the related literature, the performance of the best DNS specifications hardly beats the RW forecasts, being relatively worse as the forecast horizon increases.

We provide two robustness checks that complement the results and provide new insights about the performance of extended DNS specifications. First, we analyse the role played by restrictions on i) the dynamics of the factors, ii) the dynamics of the decay parameter, and iii) the specification of the covariance matrix of residuals. The analysis reveals that assuming non-stationary factors and a decay parameter with a random-walk dynamics enhance forecast performance relative to their baseline versions. The results of the fluctuation test of Giacomini and Rossi (2010) show that some of the restricted specifications outperform the RW in different points in time for different maturities and forecast horizons. Second, we consider a shorter out-of-sample window that excluded the zero-lower-bound interest rates period. In this case, the forecasting performance improves and the extended DNS model is able to beat the RW for very short maturities.

The outline of the paper is as follows. Section 2 describes the DNS model and its extensions considered in this paper. In Section 3, the out-of-sample performance of the extended DNS model is analysed empirically by obtaining forecasts of end-of-month continuously compounded Treasury yields on US zero-coupon bonds. Section 4 deals with the analysis of the forecasting performance when the specification of the factors and the decay parameters are assumed to be random walks and/or when the zero lower bound interest rates period is excluded. Finally, Section 5 concludes.

2 Extensions of the Dynamic Nelson-Siegel model

In this section, we describe the DNS model for the yield curve as well as some of its more popular extensions. We also describe estimation procedures of the model parameters.

2.1 Dynamic Nelson-Siegel model

The DNS model, originally proposed by Diebold and Li (2006) to represent the term structure of interest rates, is given by

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}\right) + \varepsilon_{it}, \tag{1}$$

$$\beta_{t+1} = \mu + \Phi\left(\beta_t - \mu\right) + \eta_t, \tag{2}$$

where $y_t(\tau_i)$ is the yield of a security with maturity τ_i, $i = 1, \ldots, N$, observed at time t, for $t = 1, \ldots, T$, $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t})'$ is the vector of factors, which represent the level (β_1), slope (β_2), and curvature (β_3) of the yield curve; see Hännikäinen (2017) for their high correlations with empirical counterparts of the actual level, slope and curvature of the yield curve. The noise $\varepsilon_t = (\varepsilon_{1t}, ..., \varepsilon_{Nt})'$ is an $N \times 1$ Gaussian white noise vector with diagonal covariance matrix Σ_ε; see, for example, Diebold and Li (2006), Diebold et al. (2006), Koopman et al. (2010), Exterkate et al. (2013) and Jungbacker et al. (2014) for the diagonality of Σ_ε. η_t is a Gaussian white noise 3×1 vector with full covariance matrix Σ_η and assumed to be independent of ε_t for all lags and leads. The λ parameter is a strictly positive decay parameter that governs the decay of interest rates when maturity increases. Small (large) values of λ produce a slow (fast) decay. The decay parameter, λ, also governs where the loading on the curvature, β_{3t}, achieves its maximum. Finally, μ is a 3×1 vector of constant parameters and Φ is a 3×3 matrix of autoregressive parameters that govern the dynamics of the factors. The matrix Φ satisfies the stationarity conditions.[1]

Two main characteristics explain the popularity of the DNS model in equations (1)-(2) to represent the dynamic evolution of the yield curve. First, the parameters of the DNS model can be easily estimated using standard frequentist estimation techniques, avoiding the heavy computations often associated to Bayesian estimation procedures, which are often used to estimate, for example, affine models of the yield curve. Early on Diebold and Li (2006) propose estimating the DNS parameters by a simple two-step estimation procedure after fixing the exponential decay rate, λ, to a constant chosen by the researcher. In particular, they propose fixing $\lambda = 0.0609$, with the curvature having its largest impact on the 30-month maturity bond; see Bianchi et al. (2009), Swanson and Williams (2014), van Dijk et al. (2014), Byrne et al. (2017), Almeida et al. (2018) and Freire and Riva (2023) for implementations using this value.[2] In the first step, if there are sufficient interest rates with different maturities, the factors, β_t, are estimated by Ordinary Least Squares (OLS) at each period of time t, while, in the second step, assuming that Φ is diagonal, univariate AR(1) models are fitted to each estimated factor; see, for example, Matsumura et al. (2011), Diebold and Rudebusch (2013), Swanson and Williams (2014), van Dijk et al. (2014), Byrne et al. (2017), Hännikäinen (2017), Almeida et al. (2018), Inoue and Rossi (2021) and Freire and Riva (2023) for implementations of this two-step procedure.

Alternatively, following Diebold et al. (2006), the DNS model can be viewed as a dynamic factor model (DFM) with restricted factor loadings as follows

$$y_t = \Lambda\beta_t + \varepsilon_t, \tag{3}$$

where $y_t = (y_t(\tau_1), ..., y_t(\tau_N))'$ is the $N \times 1$ vector of yields at time t, and, $\forall i$, the (i, j) element of the $N \times 3$ matrix Λ is given by

$$\Lambda_{ij} = \begin{cases} 1, & j = 1 \\ \dfrac{1 - \exp(-\lambda\tau_i)}{\lambda\tau_i}, & j = 2 \\ \dfrac{1 - \exp(-\lambda\tau_i)}{\lambda\tau_i} - \exp(-\lambda\tau_i), & j = 3. \end{cases} \tag{4}$$

The representation of the DNS model in equations (3)-(2) is an state space model and, consequently, if the model parameters were known, the Kalman filter and smoothing (KFS) algorithms can be implemented to extract the factors. The filter can be initialised using the unconditional mean and covariance matrix of β_t, namely, $b_{1|0} = E[\beta_t] = \mu$ and $P_{1|0} = E[\beta_t\beta_t'] = \Sigma_\beta$, with the latter being the solution of $\Sigma_\beta - \Phi\Sigma_\beta\Phi = \Sigma_\eta$, which can be solved using the properties of the vectorization operator; see Christensen and van der Wel (2019) for details. In practice, the parameters, including the decay parameter, λ, can be estimated by quasi maximum likelihood (QML), using the prediction error decomposition of the Gaussian likelihood. The start-up values of the parameters in Φ can be obtained using the two-step OLS estimator described above, while all variances can be initialised at 1.0 and $\lambda = 0.0609$; see Exterkate et al. (2013) and Joslin et al. (2013) for implementations. Once the parameters are estimated, the Kalman filter can be used to obtain h-step-ahead forecasts of the underlying level, slope and curvature, and, consequently, of the yields, by running the prediction equations of the filter alone with the model parameters substituted by their corresponding QML estimates.

On top of being easily estimated using frequentist procedures, the second reason why DNS models are attractive is because the factors have meaningful economic interpretations and can embody different aspects of monetary policy. In particular, the level, β_{1t}, is the limit of the yield when maturity increases, embodying any effects of monetary policy that simultaneously shift all interest rates. Moreover, the slope, β_{2t}, is related with conventional monetary policy, which typically affects short rates more than long ones, thereby changing the so-called term spread of the yield curve. Finally, an increase in the curvature, β_{3t}, increases medium-term yields and has little effect on short and long interest rates. It is related with unconventional monetary shocks, such as forward guidance or monetary policy announcements. Also note that some linear combinations of the factors are of interest. For example, $\beta_{1t} + \beta_{2t}$ represents the instantaneous yield while $\beta_{3t} - \beta_{1t}$ represents changes in long-run expectations or risk premium that do not result in parallel shifts in the term structure.

Next, we describe several popular extensions of the DNS model and how the Kalman filter methodology for factor extraction can be updated to each of them. In particular, we consider additional and time-varying decay parameters, the inclusion of macroeconomic variables and of conditional heteroscedasticity.

2.2 Additional and time-varying decay parameters

Svensson (1994) proposes extending the DNS model in (1) and (2) by including an additional decay parameter, which allows the yield curve to have more flexible shapes with two humps.

The so-called dynamic Nelson-Siegel-Svensson (DNSS) model is given by

$$y_t(\tau_i) = \beta_{1t} + \beta_{2t}\left(\frac{1-e^{-\lambda_1\tau_i}}{\lambda_1\tau_i}\right) + \beta_{3t}\left(\frac{1-e^{-\lambda_1\tau_i}}{\lambda_1\tau_i} - e^{-\lambda_1\tau_i}\right) + \beta_{4t}\left(\frac{1-e^{-\lambda_2\tau_i}}{\lambda_2\tau_i} - e^{-\lambda_2\tau_i}\right) + \varepsilon_{it}, \tag{5}$$

where λ_1 and λ_2 are both strictly positive and distinct to avoid multicollinearity. Note that $\beta_t = (\beta_{1t}, \beta_{2t}, \beta_{3t}, \beta_{4t})'$ can be defined as in (2) with the dimensions of Φ and Σ_η adapted accordingly. Gürkaynak et al. (2007) show that the yield curve often needs two humps, one at short maturities associated with monetary policy expectations and another at long maturities to capture convexity effects; see also the results by Almeida et al. (2018) and Swanson et al. (2020). The second hump in the DNSS model is difficult to identify without imposing additional restrictions; see, for example, the empirical results in Gürkaynak et al. (2007). Consequently, several authors propose restrictions to guarantee that the two humps are far apart; see, for example, Ferstl and Hayden (2010), Pedersen and Swanson (2019), Sasongko et al. (2019) and Walstrøm et al. (2022). However, it is important to note that Walstrøm et al. (2022) conclude that the restrictions can be disadvantageous when using the yield curve for monetary policy decisions.

The estimation procedures described above for the DNS model can be easily adapted to estimate the DNSS model; see, for example, De Rezende and Ferreira (2013), who propose estimating the parameters of a DNS model with five factors using the two-step procedure.

Alternatively, Koopman et al. (2010) propose making the yield curve more flexible by allowing the decay parameter in the DNS model in equations (1)-(2), λ, to be time-varying. To guarantee that $\lambda > 0$, they specify $\log(\lambda_t)$ and included it into the vector β_t. The resulting state space representation is given by

$$y_t = \Lambda(\lambda_t)B_t + \varepsilon_t, \tag{6}$$

$$B_t = \Phi B_{t-1} + \eta_t, \tag{7}$$

where $B_t = (\beta_{1t} - \mu_1, \beta_{2t} - \mu_2, \beta_{3t} - \mu_3, \log(\lambda_t) - \mu_4)$ follows a VAR process and $\Lambda(\lambda_t)$ is a $N \times 4$ matrix with the elements of its first three columns defined as in (3) with λ substituted by λ_t and the fourth column being a column of zeros. The model with time-varying decay parameter is denoted as DNS-TVL; see also Zantedeschi et al. (2011) for a DNS model with a time-varying decay parameter. Note that the DNSS model with two decay parameters can also be extended to allow for both parameters to be time-varying; see, for example, Choi and Kang (2023) for a two-varying factor-specific decay parameters that govern the slope and curvature. In this case, the model will be referred to as DNSS-TVL while if only the first (second) decaying parameter is time-varying, the model is called DNSS-TVL1 (DNSS-TVL2).

Since the measurement equation of the state space representation of the DNS-TVL model is non-linear, estimation of its parameters can no longer rely on the Kalman filter. To overcome this problem, estimation can be carried out using the extended Kalman filter (EKF), which accounts for non-linearities by using, at each time t, a first-order Taylor expansion around the current state estimate; see Jazwinski (1970), who demonstrates that the EKF is particularly effective in dealing with this type of non-linearity. Denote by $H(B_t) = \Lambda_1(\lambda_t)\beta_{1t} + \Lambda_2(\lambda_t)\beta_{2t} + \Lambda_3(\lambda_t)\beta_{3t}$ with $\Lambda_k(\lambda_t)$ being the k-th column of $\Lambda(\lambda_t)$, and by $b_{t|t-1} = \left(b_{1,t|t-1}, b_{2,t|t-1}, b_{3,t|t-1}, l_{t|t-1}\right)'$, the KF one-step-ahead predictions of B_t. The measurement equation in (6) can be linearised around $b_{t|t-1}$ by approximating $H(B_t)$ as follows

$$H(B_t) \approx H\left(b_{t|t-1}\right) + \dot{H}\left(b_{t|t-1}\right)\left(B_t - b_{t|t-1}\right), \tag{8}$$

where $\dot{H}(B_t) = \frac{\partial H(B_t)}{\partial B_t} = [\iota_N, \Lambda_2(\lambda_t), \Lambda_3(\lambda_t), A_t]$, with ι_N being an $N \times 1$ vector of ones and $A_t = \lambda_t \frac{\partial H(B_t)}{\partial \lambda_t}$ with its i-th element given by $\frac{\exp(-\lambda_t\tau_i)}{\lambda_t\tau_i}[(\beta_{2t} + \beta_{3t})(\lambda_t\tau_i + 1 - \exp(\lambda_t\tau_i)) + \lambda_t^2\tau_i^2\beta_{3t}]$. The Kalman filter is then run with the following measurement equation[3]

$$y_t = \beta_{1t} + \Lambda_2\left(l_{t|t-1}\right)\beta_{2t} + \Lambda_3\left(l_{t|t-1}\right)\beta_{3t} + A_t\left(log\lambda_t - l_{t|t-1}\right) + \varepsilon_t. \tag{9}$$

2.3 Adding macroeconomic variables

There is a generalised consensus about macroeconomic variables being significant for explaining bond yield dynamics, which could be due to the fact that Central Banks around the world use interest rates as their main monetary policy instrument responding to macroeconomic variables such as inflation or output; see, for example, Diebold et al. (2006), who find strong in-sample evidence in the US in favour of causal linkages between manufacturing capacity utilisation, monthly average of the federal funds rate, and 12-month percent change in the price deflator for personal consumption expenditures, and future yield curve dynamics. Byrne et al. (2017) also find evidence in the US of a better forecasting performance when including macroeconomic variables in the DNS model, among them the Federal Fund Rate, CPI inflation and Industrial Production while Bianchi et al. (2009) use detrended output, annualised monthly inflation and the policy interest rate to explain the term structure of interest rates in the UK.

Consequently, the DNS model in (1) and (2) has also been extended by assuming that the level, slope and curvature of the yield curve, depend on the macroeconomic and financial activity, which could be represented by the inclusion of key macro-finance indicators in the equation that governs the dynamic evolution of the three factors in the DNS model. Instead of using a set of specific macroeconomic variables to explain the factors of the yield curve, the macroeconomic information is often summarised by extracting diffusion indexes from large sets of economic variables using principal components (PC); see Pedersen and Swanson (2019) for a survey on recent empirical findings regarding the out-of-sample forecast usefulness of including diffusion indexes in DNS type models.[4] Favero et al. (2012) favour the forecast performance of factor-augmented DNS (DNS-Macro) models when compared with a large number of alternatives. They conclude that macroeconomic information is more useful at longer forecast horizons and longer maturities. Exterkate et al. (2013) conclude that DNS-Macro models perform well in relatively volatile periods with reductions of 20%-30% in mean square forecast errors (MSFEs) when compared with the simplest DNS model. Swanson and Williams (2014) also observe decreasing sensitivity, beginning in late 2011, of medium-term interest rates to macroeconomic news. It is important to note that Swanson and Xiong (2018) and Pedersen and Swanson (2019) point out that the usefulness of diffusion indexes is crucially dependent upon whether real-time-data are used or not. When real-time data are used, pure DNS models based only on historical information on interest rates, deliver forecasts with smaller MSFEs. However, when data are not real-time, diffusion indexes always have marginal forecasting content for interest rates.

Denote by f_t the $r \times 1$ vector of PC factors at time t extracted from a large set of macroeconomic variables. The vector β_t in (2) is substituted by

$$\begin{bmatrix} \beta_t - \mu \\ f_t \end{bmatrix} = \Phi \begin{bmatrix} \beta_{t-1} - \mu \\ f_{t-1} \end{bmatrix} + \eta_t, \tag{10}$$

where Φ is a $(3+r) \times (3+r)$ matrix of parameters allowing interrelations between the shape of the yield curve and the macro-financial factors. The structure of the matrix Φ has information about the possibility of different characteristics of the yield curve being related to different macroeconomic aspects. Note that the link between the yield curve and macroeconomic aggregates may also exist in the reverse direction due to economic agents responding to changes in interest rates; see, for example, Estrella and Hardouvelis (1991) and Giacomini and Rossi (2006), who analyse the performance of the yield curve as a predictor of growth. Finally, η_t is a $(3+r) \times 1$ Gaussian white noise vector with covariance matrix Σ_η.

2.4 Conditional heteroscedasticity

In the DNS model in (1)-(2), the variances in the diagonal covariance matrix of ε_t are assumed to be constant over time. To allow for conditional heteroscedasticity in the yields, Koopman et al. (2010) propose modelling ε_t with a common pattern of evolving variances as follows

$$\varepsilon_t = \Gamma \varepsilon_t^* + \varepsilon_t^\dagger, \tag{11}$$

where Γ is an $N \times 1$ vector of constants, $\varepsilon_t^\dagger$ is an $N \times 1$ Gaussian white noise vector with diagonal covariance matrix $\Sigma_\varepsilon^\dagger$, and ε_t^* is given by a conditionally normal GARCH(1,1) model with its conditional variance given by[5]

$$E(\varepsilon_t^* | \varepsilon_{t-1}^*, \varepsilon_{t-2}^*, ...) = h_t = \gamma_0 + \gamma_1 \varepsilon_{t-1}^{*2} + \gamma_2 h_{t-1}, \tag{12}$$

where, following Koopman et al. (2010), identification is achieved by fixing $\gamma_0 = 10^{-4}$ and the other parameters satisfy the usual positivity and stationarity conditions, namely, $\gamma_1, \gamma_2 \geq 0$ and $\gamma_1 + \gamma_2 < 1$, respectively.[6] Furthermore, the initial conditional variance is given by the marginal variance, $h_1 = \gamma_0(1 - \gamma_1 - \gamma_2)^{-1}$.

The DNS model with conditionally heteroscedastic errors is denoted as DNS-GARCH. The volatility of each yield is related to a common conditional variance in (12) that can be interpreted as the volatility of an underlying bond market portfolio; see Engle and Ng (1993).

The DNS-GARCH model can be rewritten as a state-space model as follows

$$y_t = \begin{bmatrix} \Lambda & \Gamma \end{bmatrix} \begin{bmatrix} \beta_t \\ \varepsilon_t^* \end{bmatrix} + \varepsilon_t^\dagger, \tag{13}$$

$$\begin{bmatrix} \beta_{t+1} - \mu \\ \varepsilon_{t+1}^* \end{bmatrix} = \begin{bmatrix} \Phi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_t - \mu \\ \varepsilon_t^* \end{bmatrix} + \omega_t, \tag{14}$$

where $\omega_t = \begin{bmatrix} \eta_t, \varepsilon_{t+1}^* \end{bmatrix}$ has covariance matrix $\Sigma_\omega = \begin{bmatrix} \Sigma_\eta & 0 \\ 0 & \gamma_0 + \gamma_1 \varepsilon_t^{*2} + \gamma_2 h_t \end{bmatrix}$. In order to run the Kalman filter in model (13)-(14), Harvey et al. (1992) propose to substitute the last term in the diagonal of Σ_ω by

$$\hat{h}_{t+1|t} = \gamma_0 + \gamma_1 \left[\varepsilon_{t|t}^{*2} + P_{t|t}^{\varepsilon} \right] + \gamma_2 \hat{h}_{t|t-1}, \tag{15}$$

where $\hat{\varepsilon}_{t|t}^*$ is the last element of the filtered state in (14) and $P_{t|t}^{\varepsilon^*}$ is its variance, both given by the Kalman filter; see Hansen (2022), who implements the same methodology in the context of a no-arbitrage yield curve with time-varying conditional variation.

3 Empirical forecasts of yields

In this section, the extended versions of the DNS models are used to obtain out-of-sample forecasts of end-of-month continuously compounded yields on US zero-coupon bonds. Table 1 lists the several admissible specifications considered along with the acronyms used for each of them.

3.1 Data

We analyse the novel zero-coupon Treasury yield curve data set constructed by Liu and Wu (2021).[7] Yield curves, plotted in Figure 1, are available at the end-of-month from January

Table 1: **DNS specifications of the yield curve**

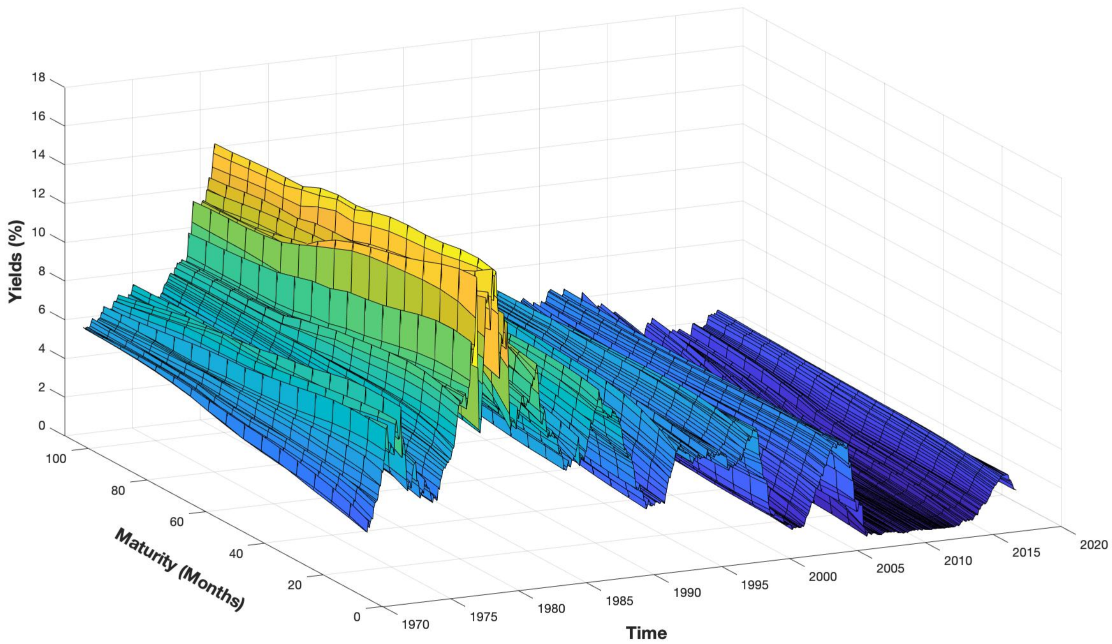
Acronym	Number of factors	Volatility	Decay parameter(s)	Macroeconomic variables
DNS	3	constant	constant	No
DNS-Macro	3	constant	constant	Yes
DNS-GARCH	3	time-varying	constant	No
DNS-GARCH-Macro	3	time-varying	constant	Yes
DNS-TVL	3	constant	time-varying	No
DNS-TVL-Macro	3	constant	time-varying	Yes
DNS-GARCH-TVL	3	time-varying	time-varying	No
DNS-GARCH-TVL-Macro	3	time-varying	time-varying	Yes
DNSS	4	constant	constant	No
DNSS-GARCH	4	time-varying	constant	No
DNSS-TVL1	4	constant	time-varying λ_{1t}	No
DNSS-TVL2	4	constant	time-varying λ_{2t}	No
DNSS-GARCH-TVL1	4	time-varying	time-varying λ_{1t}	No
DNSS-GARCH-TVL2	4	time-varying	time-varying λ_{2t}	No
DNSS-GARCH-TVL-Macro	4	time-varying	time-varying λ_{1t} and λ_{2t}	Yes

1972 through December 2019 for US securities with maturities of 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108, and 120 months, with $N = 17$ maturities and $T = 576$ monthly observations; see Freire and Riva (2023) for an implementation using the same data set.

Table 2 reports some descriptive statistics of the time series of annual yields for each maturity. As expected, the mean of interest rates increases with the maturity, while the standard deviation decreases. This latter reduction is due to the fact that while the minimum yield increases with maturity, the maximum decreases. Table 2 also shows that, for each maturity, yields can be approximated by a normal distribution with the skewness coefficient being close to zero while the kurtosis is close to 3. When looking at the dynamic dependence of yields, we can observe that there is some evidence of non-stationarity; see also Figure 1. Table 2 also reports descriptive statistics for time series of some proxies of the level, slope and curvature. In particular, as proposed by Diebold and Li (2006), the proxy for the level is the highest maturity bond, i.e. the bond for 120 months, while that for the slope is the difference between the bond of 120 months and the bond of 3 months. The proxy for curvature is two times the bond of 24 months minus the sum of 3 months and 120 months bonds. The sample autocorrelations reported in Table 2 show that the dependence of the level, slope and curvature is characterised by large persistence.

Finally, the macroeconomic diffusion indexes used to explain the shape of the yield curve are extracted using PC from the FRED-MD data base, which contain real-time data observed

Figure 1: **US Treasury yield curves observed monthly from January 1972 to December 2019**



monthly over 130 variables, covering output and income, labour market, prices, and interest rates variables; see McCracken and Ng (2016) for a description. The number of macroeconomic diffusion indexes extracted is three; see Exterkate et al. (2013), Pedersen and Swanson (2019), Swanson et al. (2020), and Freire and Riva (2023), who also consider three factors extracted from the same data base.

3.2 Out-of-sample forecasts of interest rates

The data set of US yields described above is divided into an initial in-sample period from January 1972 to December 1993, with $R = 264$ observations, used to estimate the parameters of the specifications described in Table 1, and an out-of-sample period, from January 1994 to December 2019, with $P = 312 - h$ observations, with h being the forecast horizon and $T = R + P + h$. Note that the out-of-sample period includes the zero-lower-bound interest rates constraints period from 2008 to 2012. For each maturity, τ_i, and forecast horizon, $h = 1, 3, 6, 12$, pseudo-real-time h-step-ahead forecasts of the interest rates are obtained at time $R + p$, for $p = 0, ..., P$, denoted as $\hat{y}_{R+p+h|R+p}(\tau_i)$. These forecasts are obtained using a rolling window estimation scheme; see Inoue et al. (2017) for a description of the advantages of the rolling

Table 2: **Descriptive statistics of time series of yields with different maturities and proxies of level, slope and curvature of the yield curve**: Mean, standard deviation (Std.dev.), minimum (Min.), maximum (Max), three autocorrelation coefficients, 1 lag ($\rho(1)$), 6 lags ($\rho(6)$), and 12 lags ($\rho(12)$), skewness (Skew.) and Kurtosis (Kurt.)

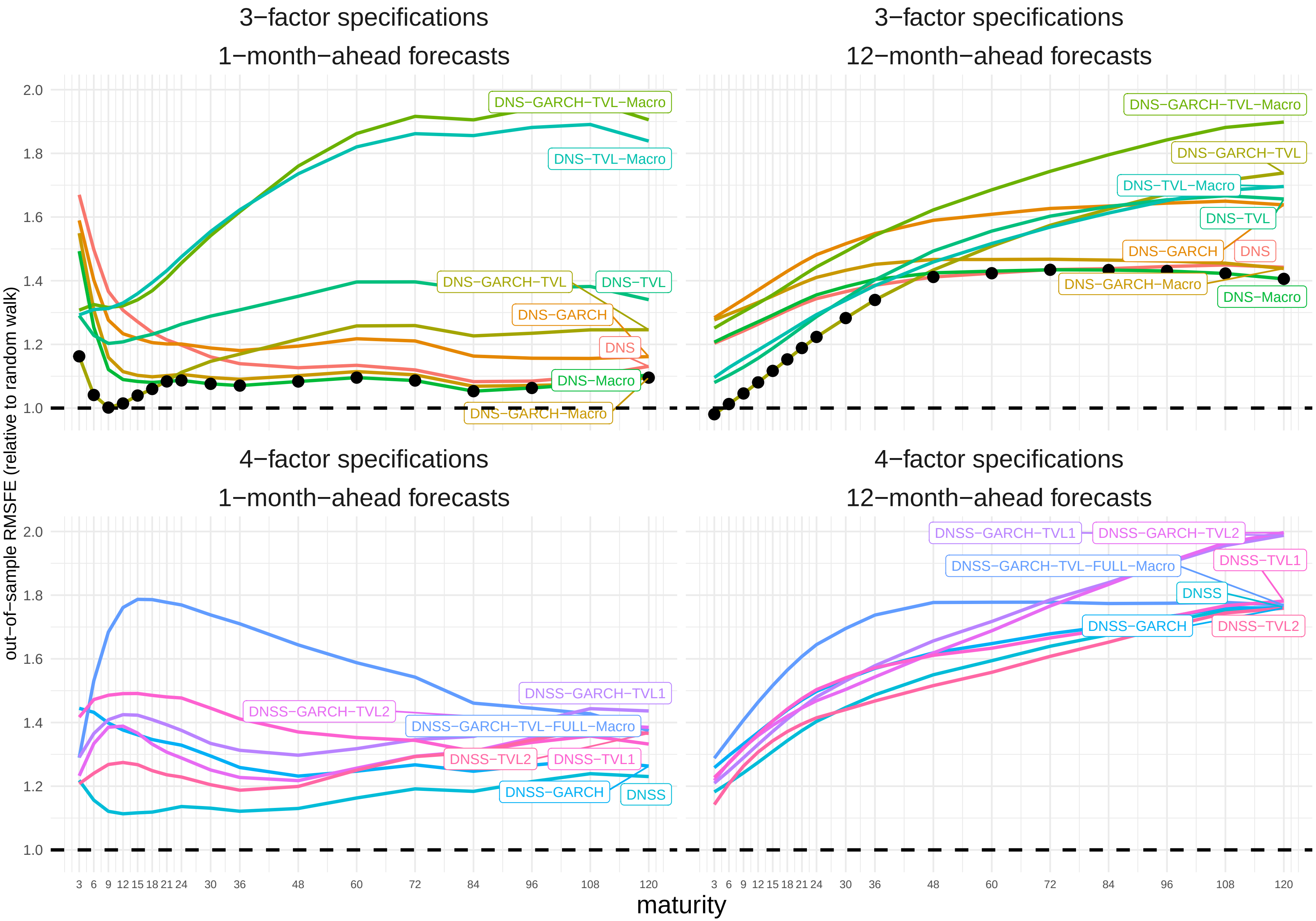
Maturities	Mean	Std.dev.	Min.	Max.	$\rho(1)$	$\rho(6)$	$\rho(12)$	Skew.	Kurt.
3	4.722	3.522	0.020	16.170	0.988	0.928	0.865	0.610	3.210
6	4.873	3.561	0.040	16.210	0.989	0.934	0.874	0.569	3.073
9	4.985	3.567	0.070	16.180	0.990	0.937	0.880	0.524	2.946
12	5.072	3.559	0.100	16.030	0.990	0.940	0.886	0.483	2.844
15	5.147	3.550	0.130	15.950	0.991	0.943	0.890	0.455	2.781
18	5.216	3.544	0.160	15.960	0.991	0.945	0.895	0.441	2.754
21	5.274	3.530	0.180	15.900	0.991	0.947	0.898	0.429	2.729
24	5.321	3.501	0.200	15.660	0.991	0.948	0.900	0.410	2.685
30	5.415	3.448	0.240	15.510	0.991	0.950	0.905	0.383	2.629
36	5.518	3.411	0.320	15.550	0.992	0.952	0.907	0.387	2.643
48	5.699	3.329	0.470	15.420	0.992	0.953	0.910	0.388	2.627
60	5.834	3.237	0.640	15.010	0.992	0.953	0.912	0.384	2.600
72	5.971	3.183	0.820	14.990	0.992	0.955	0.913	0.413	2.619
84	6.070	3.116	1.000	14.960	0.992	0.954	0.911	0.433	2.663
96	6.157	3.061	1.210	14.900	0.992	0.955	0.913	0.445	2.677
108	6.229	3.008	1.410	14.810	0.993	0.955	0.913	0.459	2.705
120 (Level)	6.285	2.932	1.500	14.780	0.992	0.952	0.908	0.444	2.722
Slope	1.564	1.417	-4.280	4.340	0.942	0.713	0.476	-0.632	3.445
Curvature	-0.365	0.969	-2.680	3.080	0.921	0.746	0.631	-0.250	2.931

window scheme in the context of potential parameter instability. The h-step-ahead out-of-sample forecasts of the yield $y_{T^*+p+h}(\tau_i)$ obtained at time $T^* + p$ are assessed by their root mean square forecast errors (RMSFE), which are calculated as follows

$$\text{RMSFE}(h, \tau_i) = \sqrt{\frac{1}{P+1}\sum_{p=0}^{P}[\hat{y}_{T^*+p+h|T^*+p}(\tau_i) - y_{T^*+p+h}(\tau_i)]^2}. \quad (16)$$

Figure 2 plots the RMSFEs of the interest rate forecasts obtained by each of the models described in Table 1, relative to those of the forecasts obtained by the benchmark RW model, for $h = 1$ (first column) and 12 (second column), as functions of the maturities. A relative-to-RW RMSFE greater (less) than 1 indicates that the model has a higher (lower) RMSFE in comparison to the RW model. The first row of Figure 2 plots the relative RMSFE for DNS models with 3 factors while the second row deals with DNSS models with four factors. For every individual maturity and forecast horizon, the model yielding the best performance in terms of relative RMSFE is marked by a black circle.

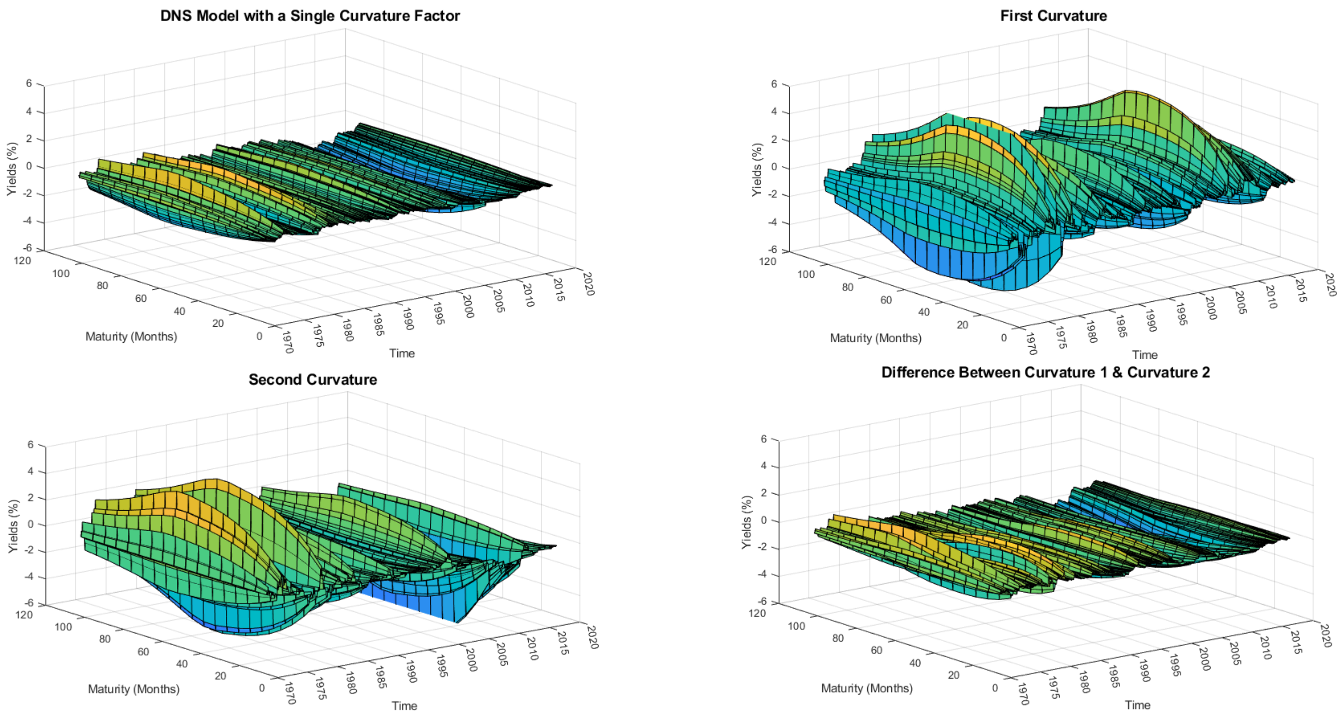
Figure 2: **One-step-ahead (left column) and twelve-step-ahead (right column) relative RMSFEs of out-of-sample yield forecasts obtained with 3-factor DNS models (first row) and with 4-factor DNSS models (second row)**. The RMSFEs are relative to forecasts obtained by the random walk (RW) model. For every individual maturity and forecast horizon, the model yielding the best performance in terms of relative RMSFE is marked by a black circle. Models and acronyms are described in Table 1.



Two main important conclusions emerge from Table 2. First, we can observe that, regardless of h, the RMSFEs of yield forecasts obtained by the DNSS models are larger than those obtained by the DNS models with just one decay factor. In particular, none of plots of the four-factor DNSS specifications has a black circle that indicates the best performance for a given maturity and forecast horizon. Note that the lack of forecasting power of the second decay is even smaller when looking at twelve-step-ahead forecasts than for one-step-ahead forecasts, when the DNSS models generate forecasts with slightly larger RMSFEs than those of their best competitors. To understand the role of the second decay in forecasting the yield curve, Figure 3 plots the curvature estimated by the simplest DNS model (top-left plot), when only one decay is included, as well as the two curvatures estimated by the DNSS model, with two-decays, together with the difference between these two latter curvatures (bottom-right plot). Comparing the curvature estimated by the DNS model with the difference between the curvatures estimated by the DNSS model, we can observe that both are very similar, explaining why, in practice, the forecasting power of the second decay is very mild. Our results are in concordance with those by Walstrøm et al. (2022), who analyse daily market prices of Treasury instruments with maturities up to 30 years, and also conclude that one of the two curvatures of the DNSS model is superfluous due to confounding effects. Diebold and Li (2006) also fit an extended model with four and five factors and conclude that this extension provides negligible improvement in model fit; see also Dahlquist and Svensson (1996) and Almeida et al. (2018), who fit more complicate shapes to the yields and also conclude that the second decay parameter does not contribute to the predictive power of the DNS model.

The second conclusion from Figure 2 is that the DNS model with best forecasting performance depends on the maturity. For short maturities, the forecasts with lowest RMSFEs are obtained by the model proposed by Koopman et al. (2010), DNS-GARCH-TVL, in which the decay parameter is allowed to change over time and there is conditional heteroscedasticity. This conclusion about conditional heteroscedasticity and time-varying decay parameter being related to short maturities is closely related to the DSGE model proposed by Amisano and Tristani (2023), who show that regime switches in the variance of shocks lead to changes in the demand of households for precautionary savings, mainly related to short maturities; see also Jungbacker et al. (2014), who observe that volatility tends to be lower for the yields of bonds with a longer

Figure 3: **Estimated curvatures of the term structure.** DNS model (top left panel) and DNSS model (top right and bottom left panels) along with their difference (bottom right panel)



time to maturity. However, for long maturities, Figure 2 shows that the best performance is obtained when the decay parameter is constant and there is conditional homoscedasticity but allowing the shape of the yield curve to depend on macroeconomic conditions as in the DNS-Macro model; see Favero et al. (2012), who also conclude that macroeconomic variables have forecasting power at longer maturities.

Consequently, according to the RMSFEs plotted in Figure 2, it seems that, in order to represent adequately the yield curve, one should use different models depending on whether short or long maturities are being forecast. This finding supports the existence of a segmented yield curve with yields of different maturities being affected by different risk factors. In particular, volatility is a risk factor for short maturities while macroeconomic variables are a risk factor for yields with long maturities. This segmentation is postulated by the preferred habitat theory of the term structure of interest rates, according to which each investor may demand bonds of specific maturities. For instance, pension funds may prefer long-term bonds while speculators may chose short-term bonds. Arbitrageurs may also participate in the market aiming to maximise a mean-variance utility function and, consequently, choosing bonds with any maturity. By doing so, arbitrageurs guarantee some smoothness among yields with different maturities; see Modigliani and Sutch (1966) for the preferred habitat theory. Almeida et al. (2018) also conclude that introducing segmentation in term structure models consistently improves long-horizon forecasts. However, in their model the segmentation is not able to identify different risk factors for yields with short and long maturities.

Finally, Figure 2 also shows that, for all maturities, the RMSFEs of the one-step-ahead forecasts of the preferred segmented specification are slightly larger than those of the corresponding RW. Consequently, next we test whether the RMSFEs of the two preferred specifications, namely the DNS-GARCH-TVL and DNS-Macro models, are statisitically different from those of the RW model. Table 3 reports the RMSFEs of these models, together with those of the basic DNS model and of the benchmark RW model. The RMSFEs reported in bold in Table 3 are those which are significantly different from the RMSFEs of the RW specification according to the Conditional Predictive Ability (CPA) test proposed by Giacomini and White (2006), which introduces estimation error under the null hypothesis and is valid for nested as well as non-nested models. Note that the asymptotic distribution of the test is derived for $R < P \to \infty$ and, in our case, $R = 264$ and $P - h = 312$.

Consider first the RMSFEs of the benchmark RW forecasts when $h = 1$, which are generally larger the larger the maturity (more than four times larger) until $\tau_i = 30$ months, decreasing afterwards with maturity. A similar pattern is observed when forecasting $h = 6$ steps ahead. with RMSFEs more than doubling those obtained when $h = 3$, and increasing until $\tau_i = 15$ months. Finally, when $h = 12$, the RMSFEs decrease with maturity. Looking at the RMSFEs reported in Table 3, it seems that the RW hypothesis for the yields could be compatible with forecasting yields with long maturities but not those with short maturities, opening the door to the possibility of different models for different maturities. Table 3 also shows that, except when forecasting in the long-run for maturities larger than 18 months, the RMSFEs are always significantly larger than those of the RW. Finally, we can observe that, although the minimum RMSFEs are obtained with the DNS-GARCH-TVL model for short maturities and with the DNS-Macro model for long maturities, the RMSFEs of this latter model are always significantly larger than those of the RW. It is important to note that the CPA test is oversized, rejecting the null of the differences between RMSFEs being a martingale difference more often than it

Table 3: **RMSFE of selected models**. Significantly different RMSFE relative to the random walk (RW) model according to the CPA test are indicated in bold. Model descriptions and acronyms are provided in Table 1.

	3	6	9	12	15	18	21	24	30	36	48	60	72	84	96	108	120
1-step-ahead forecasts																	
RW	0.043	0.037	0.039	0.043	0.047	0.051	0.054	0.057	0.062	0.065	0.070	0.071	0.072	0.071	0.069	0.067	0.067
DNS	**0.072**	**0.055**	**0.053**	**0.056**	**0.060**	**0.063**	**0.066**	**0.068**	**0.072**	**0.074**	**0.078**	**0.080**	**0.080**	**0.077**	**0.075**	**0.073**	**0.076**
DNS-Macro	**0.064**	**0.046**	**0.043**	**0.047**	**0.051**	**0.055**	**0.059**	**0.062**	**0.067**	**0.069**	**0.075**	**0.077**	**0.078**	**0.075**	**0.073**	**0.072**	**0.074**
DNS-GARCH-TVL	**0.050**	0.038	0.039	0.044	0.049	**0.054**	**0.059**	**0.063**	**0.071**	**0.076**	**0.085**	**0.089**	**0.090**	**0.087**	**0.085**	**0.083**	**0.084**
3-step-ahead forecasts																	
RW	0.179	0.183	0.195	0.206	0.214	0.220	0.226	0.231	0.240	0.239	0.239	0.231	0.217	0.204	0.196	0.190	0.183
DNS	**0.260**	**0.256**	**0.262**	**0.271**	**0.279**	**0.282**	**0.287**	**0.290**	**0.293**	**0.288**	**0.284**	**0.276**	**0.260**	**0.242**	**0.232**	**0.223**	**1.177**
DNS-Macro	**0.231**	0.226	0.234	0.245	0.254	0.260	0.267	0.273	0.279	**0.276**	**0.274**	**0.266**	**0.250**	**0.233**	**0.223**	**0.214**	**1.123**
DNS-GARCH-TVL	0.180	0.183	0.196	0.212	0.228	**0.240**	**0.252**	**0.264**	**0.281**	**0.288**	**0.302**	**0.307**	**0.299**	**0.287**	**0.281**	**0.275**	**1.461**
6-step-ahead forecasts																	
RW	0.521	0.520	0.526	0.531	0.533	0.530	0.528	0.525	0.521	0.506	0.491	0.473	0.442	0.411	0.394	0.377	0.358
DNS	0.686	0.685	0.691	0.698	0.700	**0.696**	**0.692**	**0.688**	**0.677**	**0.654**	**0.629**	**0.607**	**0.569**	**0.530**	**0.505**	**0.482**	**0.458**
DNS-Macro	0.636	0.638	0.649	0.658	0.662	0.662	0.662	0.661	**0.654**	**0.634**	**0.610**	**0.585**	**0.545**	**0.505**	**0.480**	**0.455**	**0.428**
DNS-GARCH-TVL	0.509	0.522	0.543	0.564	0.583	**0.595**	**0.607**	**0.618**	**0.633**	**0.634**	**0.646**	**0.651**	**0.633**	**0.608**	**0.594**	**0.580**	**0.560**
12-step-ahead forecasts																	
RW	1.523	1.497	1.461	1.416	1.363	1.304	1.246	1.192	1.103	1.014	0.896	0.824	0.756	0.700	0.655	0.618	0.588
DNS	1.832	1.831	1.816	1.789	1.752	**1.703**	**1.652**	**1.602**	**1.506**	**1.405**	**1.266**	**1.173**	**1.085**	**1.006**	**0.946**	**0.896**	**0.848**
DNS-Macro	1.838	1.841	1.827	1.800	1.762	1.713	1.664	1.616	**1.523**	**1.423**	**1.277**	**1.179**	**1.085**	**1.003**	**0.937**	**0.879**	**0.827**
DNS-GARCH-TVL	1.493	1.516	1.528	1.531	1.523	**1.503**	**1.481**	**1.458**	**1.414**	**1.358**	**1.285**	**1.243**	**1.191**	**1.137**	**1.095**	**1.060**	**1.022**

should; see Zhu and Timmermann (2022).

Finally, we compare the performance of each model with respect to the RW at each point in time using the fluctuation test of Giacomini and Rossi (2010). The test is designed to detect any changes in the relative performance of two competing forecasting models, and it is particularly useful in situations where the superiority of one model over another may vary across different time periods. The results of the test are plotted in Figure 4. Test statistics above (below) the critical value indicate that the model outperforms (underperforms) the RW model. Interestingly, the Figure shows that the lack of forecasting power of DNS models could be attributed to short periods of time as, for example, the period before 2005, when the Fed was in the midst of a steady round of monetary tightening and it raised short-term interest-rate target by a quarter of a percentage at each policy meeting trying to cool the housing market.

Nevertheless, we also observe that, for some individual maturities and forecast horizons, the DNS-Macro specification significantly *outperformed* the RW model. In particular, in the period immediately after 2005 the DNS-Macro model performed better than the RW in delivering one-month-ahead forecasts for the 120-month maturity yields. This corroborates the findings obtained from Figure 2 which indicate that macroeconomic variables are helpful in forecasting long-term maturities.

4 Robustness checks

In the previous section, two models emerge as having the best forecasting performance, namely, the DNS-GARCH-TVL model for short maturities and the DNS-Macro model for long maturities. To avoid problems of overfitting and numerical issues associated with estimating the large number of parameters in Φ, the specification of these two models can be simplified imposing restrictions on the dynamics of the factors and/or decay parameter. Furthermore, some authors have suggested that the forecasting performance of DNS models is enhanced when the zero coupon lower bound period is not included in the analysis. In this section, we consider three restrictions often assumed in empirical analysis of the yield curve based on the DNS framework. We also consider an alternative out-of-sample period up to 2008 without including the zero lower bound period.

4.1 Restricted specifications: Non-stationary factors and decay parameter

Consider first the specification of the factors in equation (2). Several authors observe that the factors display little cross-correlation, so that Φ can be assumed to be diagonal; see, for example, Exterkate et al. (2013) and van Dijk et al. (2014). Second, the results in Diebold and Li (2006) and Diebold et al. (2006) suggest that β_{1t} and β_{2t} may have a unit root while β_{3t} does not; see also Bowsher and Meeks (2008), Christensen and Rudebush (2012), and Almeida et al. (2018) for non-stationary models for the factors and Jungbacker et al. (2014) for large persistence of the factors and diagonality of the autoregressive matrix.

Figure 4: **Giacomini and Rossi (2010) fluctuation test: Extended DNS models**

Non-stationarity of the factors of the yield curve may be the consequence of non-stationarity of interest rates; see Hall et al. (1992) and Bauer and Rudebush (2020). Christensen and Rudebush (2012) point out that due to this persistence, stationary AR models fitted to the factors may suffer from substantial small-sample bias with estimates implying much less persistence than the true process. Therefore, we estimate the DNS, DNS-GARCH-TVL and DNS-Macro models assuming that $\Phi = I$. Furthermore, the matrix Σ_η is also assumed to be diagonal. Finally, several authors suggest that the persistence of the log-decay parameter can be large and propose modelling it using a random walk; see, for example, Koopman et al. (2010). Consequently, we consider the specifications of the two preferred models, DNS-GARCH-TVL and DNS-Macro, with the restrictions described above; see Table 4 for the acronyms of the corresponding restricted versions.

Table 4: **Acronyms and description of restricted DNS-TVL-GARCH and DNS-Macro models**

Acronym	Factor dynamics (Φ)	Covariance matrix Σ_η	Decay parameter λ_1
Panel A: Restricted DNS-GARCH-TVL models			
DNS-GARCH-TVL-RW	Stationary VAR	Unrestricted	RW
DNS-RW-GARCH-Q-DIAG	$\Phi = I$	Diagonal	constant
DNS-RW-GARCH-TVL-RW	$\Phi = I$	Unrestricted	RW
DNS-RW-GARCH-TVL-RW-Q-DIAG	$\Phi = I$	Diagonal	RW
Panel B: Restricted DNS-Macro models			
DNS-Macro-RW	$\Phi = I$	Unrestricted	constant
DNS-Macro-Q-DIAG	Stationary VAR	Diagonal	constant
DNS-Macro-RW-Q-DIAG	$\Phi = I$	Diagonal	constant

Figures 5 and 6 plot the relative RMSFEs of the restricted DNS-GARCH-TVL and DNS-Macro models with respect to the RW model. As in Exterkate et al. (2013), we observe the forecast performance of the several restricted models is improved relative to their baseline versions and is on par with the RW model. Figure 5 shows that two restricted specifications (DNS-RW-GARCH-TVL-RW and DNS-RW-GARCH-TVL-RW-Q-DIAG) perform well relative to the baseline DNS-GARCH-TVL across all forecast horizons and for maturities grater than 6 months. This results indicates that two restrictions are helpful to improve forecasts: assuming non-stationary factors and random-walk dynamics for the decay parameter. Figure 6 confirms that assuming non-stationary factors is indeed helpful to improve the forecasting performance of the DNS-Macro model. Nevertheless, Figures 5 and 6 are inconclusive as to whether the restricted specifications outperform the RW in terms of RMSFEs. To address this question, we plot in Figure 8 the results of the fluctuation test for two restricted specifications: DNS-RW-GARCH-TVL-RW-Q-DIAG and DNS-Macro-RW-Q-Diag; see Table 4 for a description of the restricted models. The Figure shows that the restricted models outperformed the RW model in different points in time specially for the 12-month maturity at both 1-month and 12-month forecast horizons.

4.2 Reduced out-of-sample period

We reduce the out-of-sample window leaving out the zero-lower bound period. Consequently, we end the out-of-sample period in December 2008. The corresponding relative RMSFEs are plotted in Figure 7, which shows that in this out-of-sample period, the short maturities are best forecast with a DNS-GARCH-TVL model without restrictions. With respect to longer maturities, it seems that the DNS-GARCH-TVL or DNS-Macro models with restrictions have similar performances.

5 Conclusions

We implement a flexible specification of the popular DNS model for the yield curve which encapsulates four popular extensions that could mitigate the adverse effects of potential misspecification. The extensions include a four-factor version of the DNS model, a time-varying decay parameter, a conditionally heteroscedastic GARCH model, and the inclusion of macroeconomic variables. First, we observe that the yield curve could be segmented with different factors affecting short and long maturities. We also show that two restrictions increase forecast accuracy: assuming non-stationary factors and a decay parameter with random-walk dynamics. It is also important to consider the presence of structural breaks in the dynamics of yields dealing mainly with the zero lower bound period. All in all, we show the DNS model can be carefully specified to obtain more accurate yield forecasts relative to the RW, at least for some periods of time.

Figure 5: **Relative RMSFEs of restricted versions of the DNS-TVL-GARCH model**

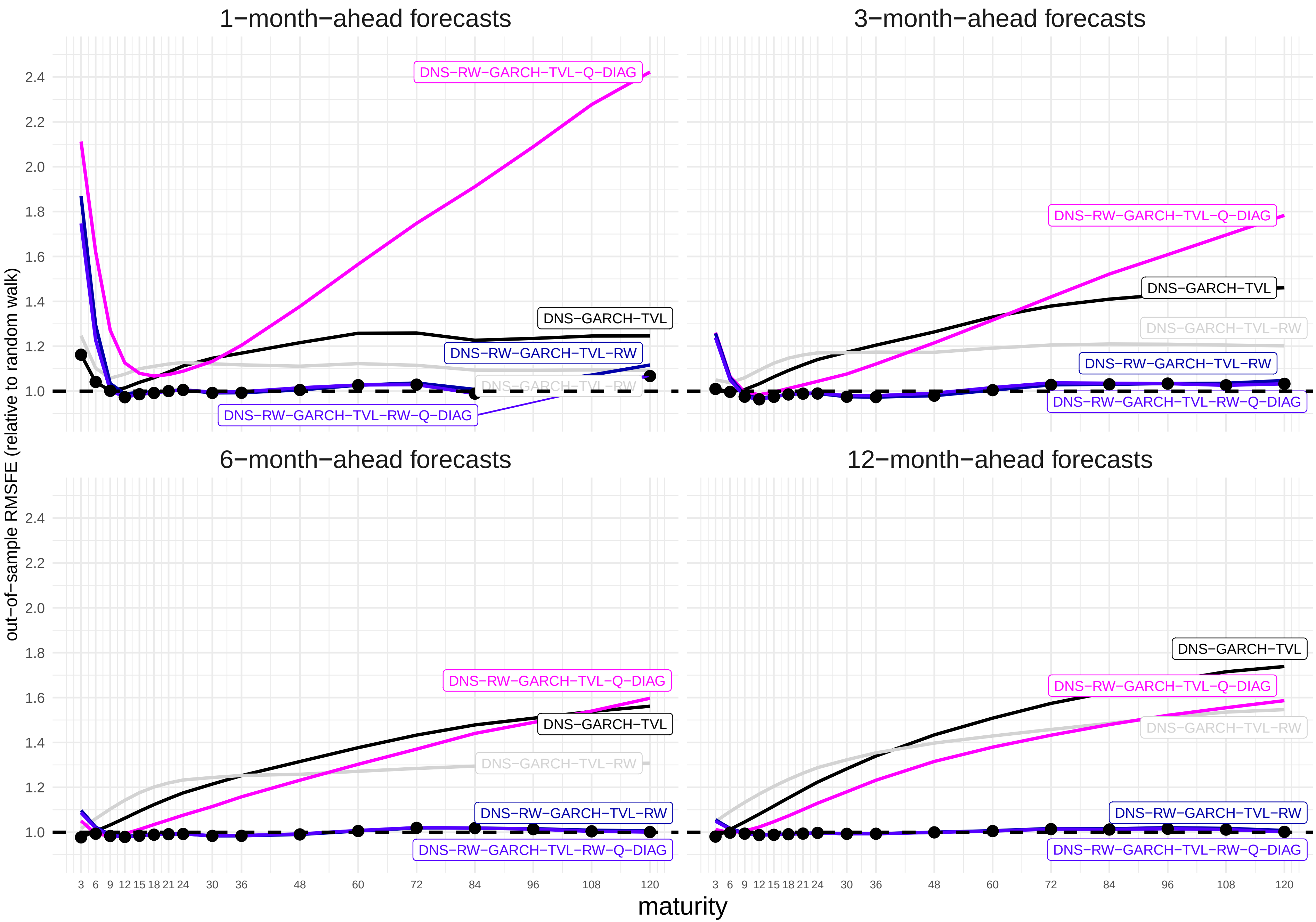


Figure 6: **Relative RMSFEs of restricted versions of the DNS-Macro model**

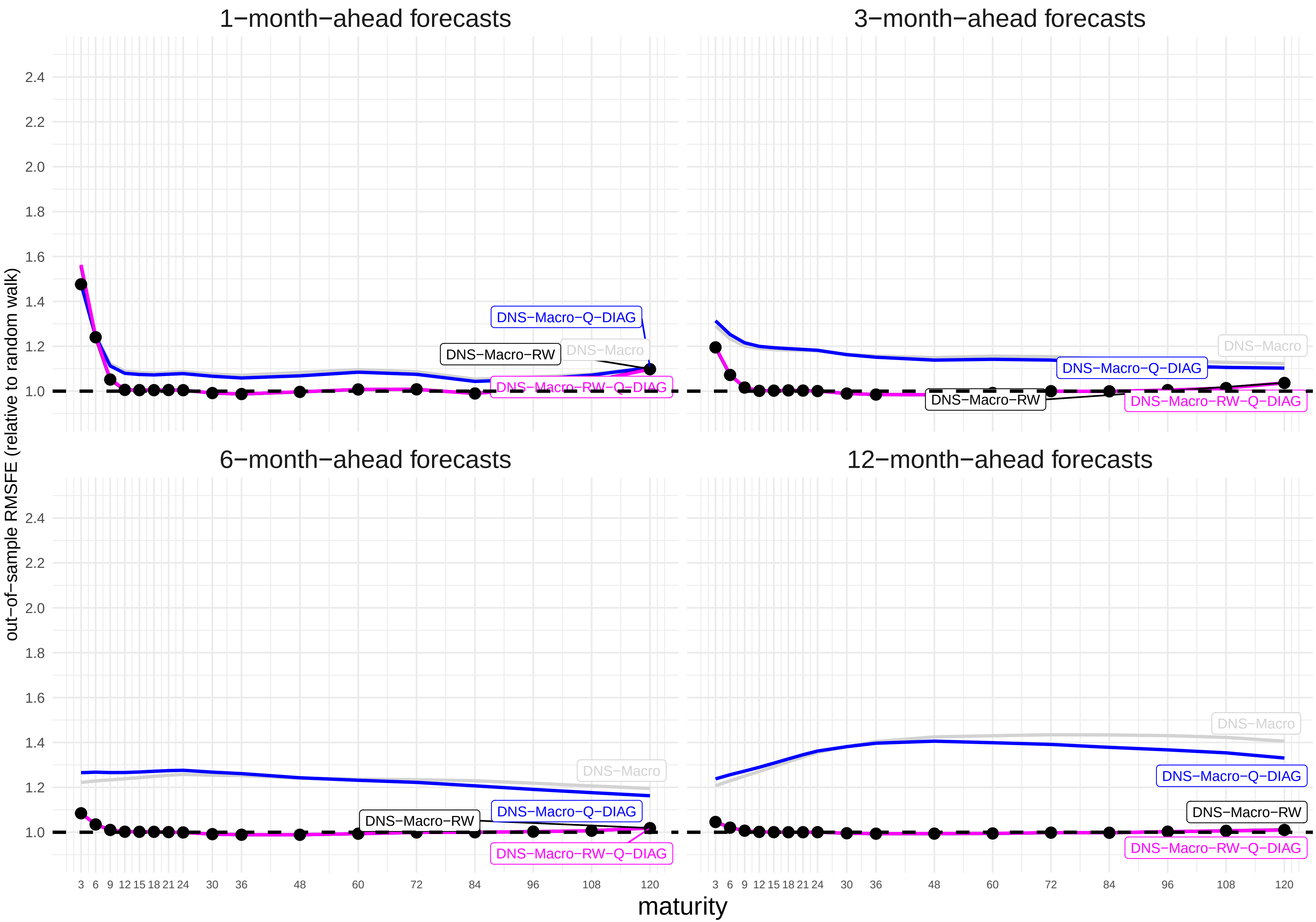


Figure 7: **Relative RMSFEs of models over the out-of-sample period up to 2008**

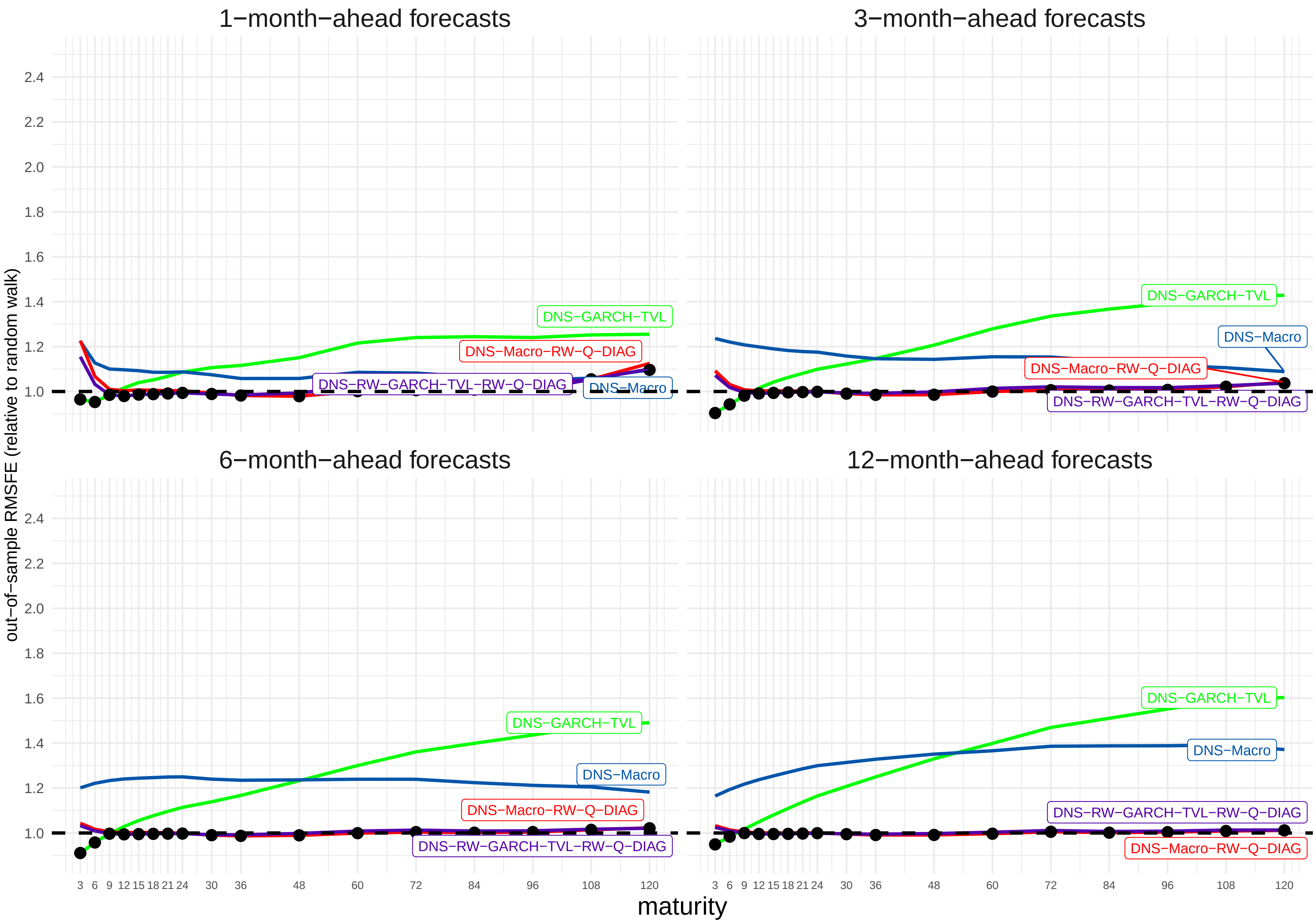
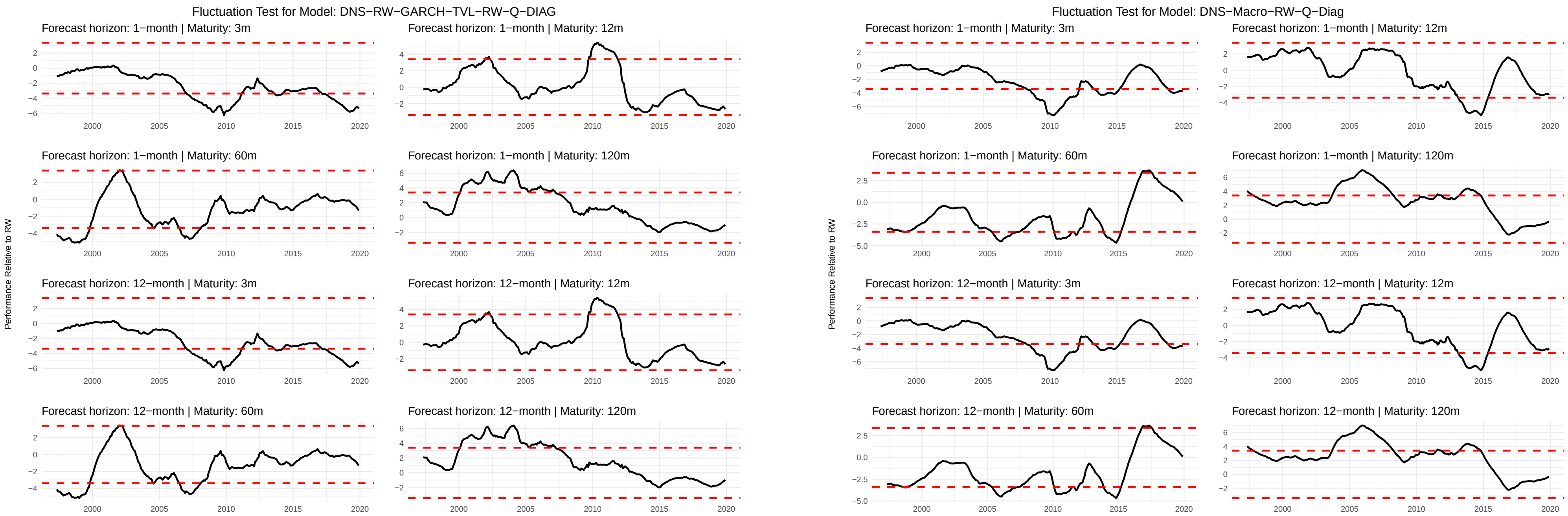


Figure 8: **Giacomini and Rossi (2010) fluctuation test: Restricted DNS models**



Notes

[1]Alternatively, several authors propose modelling the term structure with the loadings represented by polynomial splines instead of exponential functions; see, for example, Bowsher and Meeks (2008), Koopman and van der Wel (2013) and Almeida et al. (2018). Also, Jungbacker et al. (2014) do not impose a fixed structure of the loadings further than some smoothing conditions.

[2]Alternatively, Almeida and Vicente (2008) and De Rezende and Ferreira (2013) propose searching for λ by optimising the model fit.

[3]Alternatively, Gimeno and Nave (2018) propose a genetic algorithm for the estimation of the parameters of the DNSS-TVL model and Choi and Kang (2023) propose an MCMC estimator.

[4]Exterkate et al. (2013) consider alternative procedures for the extraction of the macroeconomic factors and rank PC diffusion indexes second best after Partial Least Squares factor extraction. Pedersen and Swanson (2019) also survey procedures using targeted prediction, in which the variables used in the construction of the diffusion indexes are pre-selected using methods based on Machine Learning (ML). They note that, in periods when interest rates are more volatile, ML techniques may have much to offer. Swanson et al. (2020) also propose factor extraction based on the ML and Elastic net procedures proposed by Bai and Ng (2008). Alternatively, Coroneo et al. (2016) propose fitting a DFM treating macroeconomic factors as unobservable components that are extracted simultaneously with the traditional yield curve factors.

[5]Alternatively, Bianchi et al. (2009), Hautsch and Yang (2012) and Byrne et al. (2017) allow for stochastic volatility in the DNS model and Choi and Kang (2023) assume a covariance matrix of the noises ε_t, which follows a Wishart process.

[6]Note that $E\left(\varepsilon_t \varepsilon_t'\right) = \Gamma\left(\frac{\gamma_0}{1-\gamma_1-\gamma_2}\right)\Gamma' + \Sigma_{\varepsilon}^{\dagger}$. Therefore, identification is usually achieved by assuming either that $\Gamma\Gamma' = I$ or by fixing γ_0 to any known constant.

[7]The data is publicly available in the Journal of Financial Economics Data Archive, as part of their supplementary material.

Statements and declarations

The first gratefully acknowledges support provided by CNPq under Grant 309448/2022-0. The third author acknowledges financial support from the Spanish Government grant PID2019-108079GBC21/AIE/10.13039/501100011033 (MINECO/FEDER). The fourth acknowledges the financial support from the Comunidad de Madrid Government through project 2022-T1/SOC-24167 and Spanish Agencia Estatal de Investigación through project PID2022-138289NB-I00. The authors have no further competing interests to declare that are relevant to the content of this article. The datasets analysed during the current study are available from

the corresponding author on reasonable request.

References

Almeida, C., K. Ardison, D. Kubudi, A. Simonsen, and J. Vicente (2018). Forecasting bond yields with segmented term structure models. *Journal of Financial Econometrics 16*(1), 1–33.

Almeida, C. and J. Vicente (2008). The role of no-arbitrage on forecasting: Lessons from a parametric term structure model. *Journal of Banking & Finance 32*(12), 2695–2705.

Altavilla, C., R. Giacomini, and G. Ragusa (2017). Anchoring the yield curve using survey expectations. *Journal of Applied Econometrics 32*, 1055–1068.

Amisano, G. and O. Tristani (2023). Monetary policy and long-term interest rates. *Quantitative Economics*.

Ang, A., M. Piazzesi, and M. Wei (2006). What does the yield curve tell us about GDP growth? *Journal of Eonometrics 131*(1-2), 359–403.

Bai, J. and S. Ng (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics 146(2)*, 304–317.

Bauer, M. and G. Rudebush (2020). Interest rates under falling stars. *American Economic Review 110(5)*, 1316–1354.

Bernanke, B. (2020). The new tools of monetary policy. *American Economic Review 110*(4), 943–983.

Bianchi, F., H. Mumtaz, and P. Surico (2009). The great moderation of the term structure of uk interest rates. *Journal of Monetary Economics 56*(6), 856–871.

BIS (2005). Zero-coupon yield curves: Technical documentation. *BIS Papers 25*.

Bowsher, C. and R. Meeks (2008). The dynamics of economic functions: Modelling and forecasting the yield curve. *Journal of the American Statistical Association 101*, 1419–1437.

Byrne, J., S. Cao, and D. Korobilis (2017). Forecasting the term structure of government bond yields in unstable environments. *Journal of Empirical Finance 44*, 209–225.

Caldeira, J., M. Laurini, and M. Portugal (2010). Bayesian inference applied to dynamic nelson-siegel model with stochastic volatility. *Brazilian Review of Econometrics 30*(1), 123–161.

Caldeira, J., G. Moura, and A. Santos (2016). Bond portfolio optimization using dynamic factor models. *Journal of Empirical Finance 37*, 128–158.

Carriero, A., T. Clark, and M. Marcellino (2021). No-arbitrage priors, drifting volatilities, and the term structure of interest rates. *Journal of Applied Econometrics 36*, 495–516.

Carriero, A., S. Mouabbi, and E. Vangelista (2018). Uk term structure decompositions at the zero lower bound. *Journal of Applied Econometrics 33*(5), 643–661.

Choi, A. and K. Kang (2023). Modeling the time-varying dynamic term structure of interest rates. *Journal of Banking and Finance 153*, 106908.

Christensen, B. J. and M. van der Wel (2019). An asset pricing approach to testing general term structure models. *Journal of Financial Economics 134*(1), 165–191.

Christensen, J., F. Diebold, and G. Rudebusch (2011). The affine arbitrage-free class of nelson–siegel term structure models. *Journal of Econometrics 164*(1), 4–20.

Christensen, J. and G. Rudebush (2012). The response of interest rates to us and uk quantitative easing. *The Economic Journal 122*, F385–F414.

Coroneo, L., D. Giannone, and M. Modugno (2016). Unspanned macroeconomic factors in the yield curve. *Journal of Business & Economic Statistics 34(3)*, 472–485.

Coroneo, L., K. Nyholm, and R. Vidova-Koleva (2011). How arbitrage-free is the nelson-siegel model? *Journal of Empirical Finance 18(3)*, 393–407.

Dahlquist, M. and L. Svensson (1996). Estimating the term structure of interest rates for monetary analysis. *Scandinavian Journal of Economics 98*, 163–183.

De Rezende, R. and M. Ferreira (2013). Modeling and forecasting the yield curve by an extended nelson-siegel class of models: A quantile autoregression approach. *Journal of Forecasting 32*, 111–123.

Diebold, F. and C. Li (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics 130*(2), 337–364.

Diebold, F., C. Li, and V. Yue (2008). Global yield curve dynamics and interactions: a dynamic nelson–siegel approach. *Journal of Econometrics 146*(2), 351–363.

Diebold, F. and G. Rudebusch (2013). *Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach*. Princeton University Press.

Diebold, F., G. Rudebusch, and S. Aruoba (2006). The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of Econometrics 131*(1-2), 309–338.

Duffee, G. and R. Stanton (2012). Estimation of dynamic term structure models. *Quarterly Journal of Finance 2*(2), 1250008.

Duffie, D. and R. Kan (1996). A yield-factor model of interest rates. *Mathematical Finance 6*(4), 379–406.

ECB (2018). Yield curve modelling and a conceptual framework for estimating yield curves: evidence from the european central bank's yield curves. Technical report, European Central Bank.

Engle, R. and V. K. Ng (1993). Time-varying volatility and the dynamic behavior of the term structure. *Journal of Money, Credit and Banking 25*(3), 336–49.

Estrella, A. and G. Hardouvelis (1991). The term structure as a predictor of real economic activity. *Journal of Finance Finance 46*, 555–576.

Estrella, A. and M. Trubin (2006). The yield curve as a leading indicator: Some practica issues. *Current Issues in Economics and Finance 12*(5), 1–7.

Exterkate, P., D. van Dijk, C. Heij, and P. Groenen (2013). Forecasting the yield curve in a data-rich environment using the factor-augmented nelson-siegel model. *Journal of Forecasting 32*, 193–214.

Favero, C., L. Niu, and L. Sala (2012). Term structure forecasting: no arbitrage restrictions versus large information. *Journal of Forecasting 31*, 124–156.

Ferstl, R. and J. Hayden (2010). Zero-coupon yield curve estimation with the package termtrc. *Journal of Statistical Software 36*, 1–34.

Freire, G. and R. Riva (2023). Asymmetric violations of the spanning hypothesis. *Manuscript*.

Gambacorta, L., B. Hofmann, and G. Peersman (2014). The effectiveness of unconventional monetary policy at the zero lower bound: A cross-country analysis. *Journal of Money, Credit and banking 46*(4), 615–644.

Giacomini, R. and B. Rossi (2006). How stable is the forecasting performance of the yield curve for output growth. *Oxford Bulletin of Economics and Statistics 68*(s1), 783–795.

Giacomini, R. and B. Rossi (2010). Forecast comparisons in unstable environments. *Journal of Applied Econometrics 25*(4), 595–620.

Giacomini, R. and A. White (2006). Tests of conditional predictive ability. *Econometrica 74*(6), 1545–1578.

Gimeno, R. and J. Nave (2018). A genetic algorithm estimation of the term structure of interest rates. *Computational Statistics & Data Analysis 53*, 2236–2250.

Golinski, A. and P. Spencer (2021). Estimating the term structure with linear regressions: Getting to the roots of the problem. *Journal of Financial Econometrics 19*(5), 960–984.

Gürkaynak, R., B. Sack, and J. Wright (2007). U.s. treasury yield curve: 1961 to present. *Journal of Monetary Economics 54*, 2291–2304.

Gürkaynak, R. and J. Wright (2012). Macroeconomics and the term structure. *Journal of Economic Literature 50*(2), 331–367.

Hall, A., H. Anderson, and C. Granger (1992). A cointegration analysis of treasury bill yields. *Review of Economics and Statistics 74*, 116–126.

Hännikäinen, J. (2017). When does the yield curve contain predictive power? *International Journal of Forecasting 33*(4), 1044–1064.

Hansen, A. (2022). A joint model for the term structure of interest rates and realized volatility. *Journal of Financial Econometrics forthcoming.*

Harvey, A., E. Ruiz, and E. Sentana (1992). Unobserved component time series models with arch disturbances. *Journal of Econometrics 52*(1/2), 129–158.

Hautsch, N. and F. Yang (2012). Bayesian inference in a stochastic volatility nelson–siegel model. *Computational Statistics & Data Analysis 56*(11), 3774–3792.

Hevia, C., M. Gonzalez-Rozada, M. Sola, and F. Spagnolo (2015). Estimating and forecasting the yield curve using a markov switching dynamic nelson and siegel model. *Journal of Applied Econometrics 30*(6), 987–1009.

Hodges, S. and S. Schaefer (1977). A model for bond portfolio improvement. *Journal of Financial and Quantitative Analysis 120*(2), 0243.

Inoue, A., L. Jin, and B. Rossi (2017). Rolling window selection for out-of-sample forecasting with time-varying parameters. *Journal of Econometrics 196*, 55–67.

Inoue, A. and B. Rossi (2021). A new approach to measuring economic policy shocks, with an application to conventional and unconventional monetary policy. *Quantitative Economics 12*, 1085–1138.

Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory.* Mathematics in Science and Engineering. Academic Press.

Joslin, S., A. Le, and K. Singleton (2013). Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. *Journal of Financial Economics 109*(3), 604–622.

Joslin, S., K. Singleton, and H. Zhu (2011). A new perspective on gaussian dynamic term structure models. *Journal of Financial and Quantitative Analysis 24*(3), 926–970.

Jungbacker, B., S. Koopman, and M. van der Wel (2014). Smooth dynamic factor analysis with application to the us term structure of interest rates. *Journal of Applied Econometrics 29*(1), 65–90.

Koopman, S., M. Mallee, and M. Van der Wel (2010). Analyzing the term structure of interest rates using the dynamic nelson–siegel model with time-varying parameters. *Journal of Business & Economic Statistics 28*(3), 329–343.

Koopman, S. and van der Wel (2013). Forecasting the u.s. term structure of interest rates using a macroeconomic smooth dynamic factor model. *International Journal of Forecasting 29*, 676–694.

Krippner, L. (2012). A theoretical foundation for the nelson and siegel class of yield curve models. Technical report, CAMA Working Paper, 11/2012.

Kuttner, K. (2018). Unconventional monetary policy in the great recession and beyond. *Journal of Economic Perspectives 32*(4), 121–146.

Laurini, M. and J. Caldeira (2016). A macro-finance term structure model with multivariate stochastic volatility. *International Review of Economics & Finance 44*, 68–90.

Laurini, M. and L. Hotta (2010). Bayesian extensions to diebold-li term structure model. *International Review of Financial Analysis 19*(5), 342–350.

Laurini, M. and L. Hotta (2014). Forecasting the term structure of interest rates using integrated nested laplace approximations. *Journal of Forecasting 33*(3), 214–230.

Liu, Y. and J. Wu (2021). Reconstructing the yield curve. *Journal of Financial Economics 142(3)*, 1395–1425.

Matsumura, M., A. Moreira, and V. Vicente (2011). Forecasting the yield curve with linear factor models. *International Review of Financial Analysis 20*, 237–243.

McCracken, M. and S. Ng (2016). Fred-md: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics 34*(4), 574–589.

Modigliani, F. and R. Sutch (1966). Innovations and interest-rate policy. *American Economic review 56*, 178–197.

Morley, J. (2016). Macro-finance linkages. *Journal of Economic Surveys 30*(4), 698–711.

Nelson, C. and A. Siegel (1987). Parsimonious modeling of yield curves. *Journal of Business 60*(4), 473–489.

Pedersen, H. and N. Swanson (2019). A survey of dynamic nelson-siegel models, difussion indexes and big data methods for predicting interest rates. *Quantitative Finance and Economics 3*(1), 22–45.

Ronn, E. (1987). A new linear programming approach to bond portfolio management. *Journal of Financial and Quantitative Analysis 220*(4), 439–466.

Rossi, B. (2021). Identifying and estimating the effects of unconventional monetary policy in the data: How to do it and what have we learned? *The Econometrics Journal 24*(1), C1–C32.

Sasongko, A., C. Utama, B. Wibowo, and Z. Husodo (2019). Modifying hybrid optimisation algorithms to construct spot term structure of interest rates and proposing standardised assessment. *Computational Economics 54*, 957–1003.

Svensson, L. (1994). Estimating and interpreting forward interest rates: Sweden 1992-1994. Technical report, National Bureau of Economic Research.

Swanson, E. and J. Williams (2014). Measuring the effect of the zero lower bound on medium- and long-term interest rates. *American Economic Review 104*(10), 3154–3185.

Swanson, N. and W. Xiong (2018). Big data analytics in economics: What have we learned so far, and where should we go from here? *Canadian Journal of Economics 3*, 695–746.

Swanson, N., W. Xiong, and X. Yang (2020). Predicting interest rates using shrinkage methods, real time diffusion indexes, and model combinations. *Journal of Applied Econometrics 35(5)*, 587–613.

van Dijk, D., S. Koopman, M. van der Wel, and J. Wright (2014). Forecasting interest rates with shifting endpoints. *Journal of Applied Econometrics 29*(5), 693–712.

Walstrøm, R., F. Paraschiv, and M. Schürle (2022). A comparative analysis of parsimonious yield curve models with focus on the Nelson-Siegel, Svensson and Bliss versions. *Computational Economics 59*, 967–1004.

Wu, J. and F. Xia (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Money, Credit and Banking 48*(2-3), 253–291.

Yu, W.-C. and D. Salyards (2009). Parsimonious modeling and forecasting of corporate yield curve. *Journal of Forecasting 28*(1), 73–88.

Zantedeschi, D., P. Damien, and N. Polson (2011). Predictive macro-finance with dynamic partition models. *Journal of the American Statistical Association 106*(494), 427–439.

Zhu, Y. and A. Timmermann (2022). Can two forecasts have the same conditional expected accuracy? *Manuscript*.