UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SOCIOECONÔMICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Werley da Costa Cordeiro

Essays on the Yield Curve

Florianópolis
2023

Werley da Costa Cordeiro

Essays on the Yield Curve

Tese submetida ao Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina para a obtenção do título de doutor em Economia.
Orientador: Prof. João Frois Caldeira, Dr.

Florianópolis
2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Cordeiro, Werley da Costa
Essays on the Yield Curve / Werley da Costa Cordeiro ; orientador, João Frois Caldeira, 2023.
122 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Socioeconômico, Programa de Pós-Graduação em Economia, Florianópolis, 2023.

Inclui referências.

1. Economia. 2. Modelo DNS. 3. Filtro de Kalman. 4. Variáveis Macroeconômicas. 5. Prêmio a Termo. I. Caldeira, João Frois. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Economia. III. Título.

Werley da Costa Cordeiro

Essays on the Yield Curve

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof Marcelo Fernandes, Dr.
FGV

Prof.Márcio Poletti Laurini, Dr.
FEA-RP/USP

Prof.Guilherme Valle Moura, Dr.
UFSC

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de doutor em Economia.

ICP-Edu
Documento assinado digitalmente
Guilherme Valle Moura
Data: 19/07/2023 10:08:49-0300
CPF: ***.234.686-**
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de
Pós-Graduação

ICP-Edu
Documento assinado digitalmente
Joao Frois Caldeira
Data: 19/07/2023 09:53:52-0300
CPF: ***.518.836-**
Verifique as assinaturas em https://v.ufsc.br

Prof. João Frois Caldeira, Dr.
Orientador

Florianópolis, 2023.

“Why then didn’t you put my money on deposit,
so that when I came back,
I could have collected it with interest?”
(St. Luke XIX. 23)

ABSTRACT

This dissertation investigates models to predict the USA and Brazilian yield curves with forward-looking and backward-looking macroeconomic factors and time-varying parameters. Although the literature focuses on forecasting the yield curves' level, which is a difficult task, we also propose forecasts of the direction-of-change returns of the yield curves. Lastly, we use real and nominal yield curves to forecast inflation. In all four papers, the results suggest that the proposed models can outperform traditional benchmarks in the literature.
Keywords: DNS model. Affine Term-Structure Models. Macroeconomic factors.

RESUMEN

Esta disertación investiga modelos para predecir las Curva cupón cero de EE. UU. y Brasil con factores macroeconómicos prospectivos y retrospectivos y parámetros variables en el tiempo. Si bien la literatura se enfoca en pronosticar el nivel de las curvas de rendimiento, lo cual es una tarea difícil, también proponemos pronósticos de los retornos de la dirección del cambio de las curvas de rendimiento. Por último, utilizamos curvas de rendimiento real y nominal para pronosticar la inflación. En los cuatro documentos, los resultados sugieren que los modelos propuestos pueden superar los puntos de referencia tradicionales en la literatura.
Palabras clave: Modelo DNS. Modelos de estructura de términos afines. Factores macroeconómicos. Prima a término

RESUMO EXPANDIDO

Introdução

A importância das previsões de indicadores econômicos e financeiros, como taxas de juros, expectativas de inflação, atividade econômica e retornos de ativos, têm norteado a tomada de decisão dos governos na condução da política monetária e dos investidores na alocação de carteiras. Uma conhecida série temporal estudada por acadêmicos, banqueiros centrais e investidores são as curvas de juros ou a estrutura a termo das taxas de juros. A curva de juros é uma representação gráfica das taxas de juros negociadas no mercado com diferentes vencimentos, ela indica quais seriam as expectativas dos agentes de mercado sobre as taxas de juros de curto e longo prazo. Esta tese utiliza extensões de modelos populares na literatura para prever as curvas de juros dos EUA e do Brasil com fatores macroeconômicos observados e expectativas, e parâmetros variantes no tempo. Embora a literatura se concentre em prever o nível das curvas de juros, o que é uma tarefa difícil, também propomos previsões da direção de retornos das curvas de juros. Por fim, usamos curvas de juros reais e nominais para prever a inflação. Em todos os quatro capítulos, os resultados sugerem que os modelos propostos podem superar os benchmarks tradicionais da literatura.

Objetivos

Dado o avanço dos modelos Nelson-Sigel (DNS) e Modelos Afim da Estrutura a Termo AFTSM′s, suas interpretações, avanços e literatura macrofinanceira, nossa pesquisa visa preencher lacunas no campo da previsibilidade. O primeiro objetivo é prever as taxas de juros usando o modelo DNS e extensões com parâmetros variando no tempo usando variáveis macroeconômicos, além de combinações de previsões. O segundo é prever a direção das taxas de juros com o modelo DNS com volatilidade estocástica e o último é prever inflação utilizando as curvas de juros nominais e reais e prêmio pelo risco com o modelo ATSM.

Metodologia

Nossos avanços estão sustentados em quatro literaturas que chamamos de “bases": Koopman, Mallee, and Van der Wel (2010), Peter Christoffersen et al. (2006), Peter F Christoffersen and Francis X Diebold (2006) e Michael Abrahams et al. (2016). Na primeira base, adicionamos novas caracteríticas ao modelo DNS com parâmetros variando no tempo: segunda curvatura (modelo de Svensson), variáveis macroeconômicas observadas e expectativas de mercado, e combinações de previsões, além de avaliar as previsões do ponto de vista econômico. Utilizamos os modelos de volatidade estocástica da primeira base com os modelos de previsão de direção da segunda e terceira base. Por fim, utilizamos o modelo da quarta base para fazer estimação dos premios a termo e previsão de inflação.

Resultados e Discussão

No primeiro artigo, analisamos o desempenho de previsão de vários modelos de fatores para a curva de juros, com foco no papel do fator de decaimento variando no tempo, da heterocedasticidade e do efeito de variáveis macroeconômicas. Usando novos dados do Tesouro compostos continuamente no final do mês sobre títulos de cupom zero dos EUA e estimativa freqüentista com base no filtro de Kalman estendido, mostramos que a segunda curvatura não tem papel na obtenção de um melhor desempenho de ajuste e previsão do modelo de fator. Além disso, mostramos que a melhor especificação depende da maturidade e do horizonte de previsão quando se olha para as previsões pontuais. Para vencimentos curtos, o melhor desempenho é obtido em um modelo heterocedástico com fator de decaimento variando no tempo. No entanto, se as variáveis macro adicionam informações, depende do horizonte de previsão. No entanto, o modelo homocedástico mais simples com fator de decaimento constante apresenta melhor desempenho para vencimentos longos. Neste caso, as variáveis macro podem ter um papel em horizontes curtos de previsão. Esses resultados sugerem que o modelo de fator deve incorporar algum tipo de não linearidade dependendo da maturidade. No segundo artigo, investigamos a previsibilidade da taxa de juros usando modelos de fatores dinâmicos com expectativas macroeconômicas e volatilidade variável no tempo. Os resultados sugerem que a introdução de expectativas macroeconômicas da pesquisa Focus dos participantes do mercado supera diferentes modelos de referência em previsões fora da amostra, principalmente em vencimentos curtos. Além disso, volatilidade variando no tempo melhora a relevância econômica dessas previsões. No terceiro artigo, exploramos o desempenho de previsão bem documentado do modelo dinâmico de Nelson-Siegel (DNS) e mostramos como empregar as previsões para seus dois primeiros momentos condicionais para obter previsões de direção de mudança para a curva de juros, uma vez que tomadas de decisões na política monetária ou na seleção de carteiras de renda fixa e/ou cobertura geralmente requer estimativas da direção futura das taxas de juros. A incorporação de informações sobre a assimetria e a curtose da curva de juros leva a previsões direcionais que superam o modelo de referência, principalmente para maturidades longas e previsões de curto prazo. Por fim, no último artigo, usamos os modelos afins dinâmicos e livres de arbitragem para a estrutura a termo das taxas de juros AFTSM′s para modelar as taxas de juros nominais e reais conjuntamente. A abordagem permite decompor as taxas de juros em expectativas de taxas de juros futuras e o prêmio de risco que os investidores compensam ao comprar títulos de longo prazo. Além disso, analisamos sua capacidade de capturar expectativas de inflação ajustadas ao risco usando-o para previsão de inflação. Os resultados sugerem que os prêmios a prazo reais e nominais são variantes no tempo e aumentam ao longo dos vencimentos. Além disso, as expectativas de inflação ajustadas ao risco superam a pesquisa FOCUS em horizontes de previsão longos.

Considerações Finais

Em todos os artigos, observamos que os melhores resultados dependem do horizonte de previsão e dos vértices das curvas de juros. Para a curva brasileira, os modelos com volatilidade estocástica apresentam resultados melhores que o Random Walk, mas a maioria não são estatisticamente significativos. Além disso, usar variáveis macroeconômicas observadas ou expectativas de mercado pondem melhorar pontualmente as previsões, mas os modelos ficam com muitos parâmentros. De todo modo, a aplicação econômica sugere que os modelos com volatilidade estocástica tem utilidade para um investidor de média-variância. Na previsão de ditreção, os modelos conseguem ter bons desempenhos em maturidades curtas, isto é, conseguem apontar de forma mais acertiva a probabilidade de retorno positivo das curvas de juros. Por fim, os resultados do quarto artigo sugere que o prêmio de inflação é positivo na maior parte do tempo para a economia brasileira. Além disso, pode-se extrair previsões competitivas para a

inflação usando a expectativa de inflação do mercado.

Palavras-chave: Modelo DNS. Fatores Macroeconômicos. Filtro de Kalman. Modelos de estrutura de termo afim. Prêmio a termo.

LIST OF FIGURES

LIST OF TABLES

CONTENTS

1 INTRODUCTION

Can we sharpen our ability to predict future yields and deliver value to investors? To answer this question, let us start with the credit market. It comprises lenders, borrowers, intermediaries, and remuneration for borrowed money. There are loans with different rates according to the terms and risks involved. Families, companies, and mainly governments participate in this market to anticipate future consumption, investments, and spending. The government, one of the principal borrowers, finances spending by issuing sovereign bonds. These securities are known for being low-risk, highly liquid, and have benchmark yields.

Fixed-income securities have several characteristics, such as different maturities, fixed or inflation-protected securities, and some with periodic interest payments (coupons). The relationship between a zero-coupon bond's yield and the time to maturity of its cash flow is known as the term structure of the interest rate or (nominal) yield curve. Also, real yield curves can be derived from inflation-protected securities with some intermediate steps. The difference between nominal and real yields for a given maturity, known as Breakeven inflation, reflects inflation expectations. Therefore, the yield curve is a reference for pricing other fixed-income instruments and as input for various models, for example, risk management, monetary policy, derivative pricing, and portfolio allocation. It also sheds light on expected economic growth Fisher (1907), on forecasting real economic activity Campbell R Harvey (1989), on predicting interest rates, inflation, and real returns Fama (1990), and on information about future inflation Mishkin (1990), Mishkin (1991). Therefore, practitioners and academics are interested in interpreting its behavior and implications, modeling, and generating good yield curve forecasts.

However, modeling the yield curve is a challenging task. It presents well-known stylized facts such as various shapes, including upward-sloping, downward-sloping ("inverted"), humped, and inverted-humped. In general, the average yield curve is increasing and concave along maturities. The literature presents theories to explain the relation among the returns on bonds of various maturities, and we highlight the three traditional ones. According to the Liquidity Preference theory, investors prefer most liquidity bonds, that is, short-term bonds. Then they charge to buy long-term bonds, a liquidity premium, see Keynes (1936). The Expectations Hypothesis theory, as formulated by Fisher (1896) and restated by Hicks (1939) and Lutz (1940), suggests that the long-term yields are the average of the expected short-term so that the risk premium would be zero or constant over time. However, its weaker version states that there is a risk premium, and it is time-varying. Lastly, the Market Segmentation theory suggests that individual investors have a preferred range of bond maturity lengths. They are only willing to buy bonds outside their maturity preference if a higher yield for other maturity ranges is available, see Modigliani and Sutch (1966).

With theoretical groundwork, many models were developed to model the yield curve. They are theoretically rigorous or empirically successful, but not both, as said by Francis X Diebold, Glenn D Rudebusch, et al. (2012). For instance, there are One-Factor Short-Rate Models, Two-Factor Short-Rate Models, The Heath-Jarrow-Morton Framework, Non-parametric Estimation Methods, and Parametric Estimation Methods. See Brigo, Mercurio, et al. (2001) and Filipovic (2009) for a mathematically straightforward but rigorous development.

In this work, we focus on two classes of models well-documented in the literature: Nelson and Siegel (1987) and Svensson (1994) families and affine term-structures models. The Nelson-Siegel and Svensson models, which belong to the exponential-polynomial functions inside the Parametric Estimation Methods, are models that most central banks use for term-structure estimation according to BIS (1999). These models and their dynamic version of Francis X Diebold and Li (2006) (DNS) have attractiveness because of their parsimony and good empirical performance. Also, they can capture most of the behavior of the term structure of interest rate using only three and four factors, respectively. These frameworks help us to include additional factors in the model and forecast the yields. See Francis X Diebold, Glenn D Rudebusch, et al. (2012) for a practical approach. The second one is the affine term-structures models (ATSM), from One-Factor Short-Rate Models, popularized by Duffie and Kan (1996), which encompass the equilibrium models of Vasicek (1977), Cox, Ingersoll Jr, and Ross (1985), Longstaff and Schwartz (1992), and others. This class is constructed by assuming that the bond price function is a linear function of the underlying state variables. It can be derived under several hypotheses about the risk premium, which helps us to investigate whether investors are being compensated for bearing risk on bonds and generate inflation forecasts based on market expected inflation.

Therefore, our next focus is to evaluate the Nelson-Siegel and affine term-structures extensions on forecasting performance. To answer the question at the beginning of this introduction, we investigate whether observed and expected macroeconomic variables improve forecasting performance, whether time-varying volatility forecastings point in the right direction, and whether breakeven decomposition predicts inflation. In the first essay, we assess the forecastability of DNS with time-varying loading parameter and time-varying volatility in the US economy. We also use backward-looking macroeconomic factors such as interest rate, economic activity, and inflation rate. Besides, we assess the DNS-Svensson (DNSS) model's extensions with time-varying volatility and time-varying loading parameter. Lastly, we include four combinations of forecasts. In the second essay, we evaluate the forecast performance of the DNS with time-varying volatility with forward- and backward-looking macroeconomic variables to the Brazilian economy. We include observed interest rate, economic activity, and inflation rate and its respectively expected series from a survey of market practitioners. Since generating good magnitude yield forecasts is a tough job, we evaluate the direction of change of returns in the third essay. Thereunto, we use yields and volatility returns as input to calculate the probability of positive return in three models: the Logit function, the Non-Parametric model, and the Extended model. Lastly, in the fourth essay, we use nominal and real yield curves and ATSM to decompose the breakeven inflation in expected inflation and inflation risk premium for the Brazilian economy. Also, we use this decomposition to predict the Extended National Consumer Price Index (IPCA).

In the first two assays, economic evaluation suggest that the forecasts can have worth to an mean-variance investor. In the first, the extensions outperform the more simple DNS model. In the second, the extensions outperform the Random Walk model (RW). In the third assay, the results suggest that using time-varying volatility forecasting using DNS model can outperform a naive strategy of direction-of-change returns forecasts. Lastly, the results of decomposing the breakeven inflation suggest that its possible to have better forecast to inflation that the FOCUS survey, mostly to long horizons.

2 FORECASTING THE YIELD CURVE: THE ROLE OF ADDITIONAL AND TIME-VARYING DISCOUNTING, HETEROSCEDASTICITY, AND MACRO-ECONOMIC FACTORS

Accurate forecasts of the term structure of interest rates are relevant for bond portfolios, risk management, and pricing derivatives; see Hodges and Schaefer (1977) and Ronn (1987) for early references. More recently, they have also become important in the context of new tools of unconventional monetary policy as yield curve control or forward guidance; see, for example, Kuttner (2018) and Bernanke (2020). Consequently, many works have been devoted to developing alternative modeling and forecasting methodologies for the term structure during the last two decades.

There are three main alternative approaches to modeling the term structure. First, the models based on a no-arbitrage 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 Goliński and Spencer (2021) and the references therein. Recently, Bauer (2018) also proposed a Bayesian estimator. Many empirical studies suggest that imposing no-arbitrage conditions do not generally lead to more accurate forecasts; for example, Joslin, Kenneth J Singleton, and Zhu (2011). Alternatively, many authors estimate the term structure using affine models that Duffie and Kan (1996) characterized originally. Affine models allow multiple state variables to drive interest rates, with bond yields being linear functions of these variables; see Duffee and Stanton (2012) for a comparison of alternative estimators of affine models, including Kalman filter based estimation. Recently, estimations of affine models for the term structure rely on Bayesian procedures, which are computationally demanding; see, for example, Carriero, Clark, and Marcellino (2021) for a recent Bayesian estimator of the canonical affine term structure model of Duffie and Kan (1996), in its equivalent but computationally more stable representation of Joslin, Kenneth J Singleton, and Zhu (2011).

Lastly, Francis X Diebold and Li (2006) propose modeling the dynamic evolution of the yield curve using the three factor model proposed by Nelson and Siegel (1987). These three factors represent the curve's level, slope, and curvature; consequently, the factor loadings are heavily parametrized depending on a single exponential decay rate parameter; see Krippner et al. (2010) for arguments on the connection between affine models and the factor model. Francis X Diebold and Li (2006) propose a simple two-step estimation procedure assuming that the exponential decay rate is known and constant over time. In the first step, the factors are estimated by Ordinary Least Squares (OLS), while in the second step, AR(1) models are fitted to the estimated factors. This approach has become a major workhorse among academics and practitioners and is widely used in the financial community; see empirical applications by De Pooter (2007), Francis X Diebold, Li, and Yue (2008), Yu and Salyards (2009), De Pooter, Ravazzolo, and Dick JC Van Dijk (2010), Jens HE Christensen, Francis X Diebold, and Glenn D Rudebusch (2011), MÁrcio Poletti Laurini and Hotta (2014), João F. Caldeira, Guilherme V. Moura, and André A. P. Santos (2016), and BIS (2005) and ECB (2018). However, the Achilles' heel of this methodology is that forecasts hardly beat those obtained by a random walk model. Carriero, Clark, and Marcellino (2021) argue that one reason could be that the model specification may not hold in the data, and Jungbacker, Koopman, and Van Der Wel (2014), show that the Likelihood ratio test rejects the restrictions imposed in the factor loadings.

In practice, there is a debate about forecasting the yield curve using either the Kalman filter for the factor model or the affine model estimated by the Bayesian methods. The first is computationally simpler but can hardly beat the random walk specification. And the second is computationally intensive however has been shown to improve over the random walk. Within this debate, the contribution of this paper is to analyze the empirical performance of a very general and flexible specification of the three factor model when fitted to forecast the yield curve based on a novel data set of end-of-month continuously compounded Treasury yields on US zero-coupon bonds. In particular, we consider three extensions of the original specification of the DNS in Francis X Diebold and Li (2006), which would flexibility the model and mitigate the adverse effects of potential misspecification.

First, early on Svensson (1994) propose a four-factor version of the Nelson and Siegel (1987) model. As far as we know, the role of the fourth factor has yet to be analyzed in the related literature so our results will be novel in this respect. Other important extensions of the DNS model are due to Koopman, Mallee, and Van der Wel (2010), who extend it in two directions. First, they propose allowing the discount parameter time-varying; see Márcio Laurini and Hotta (2010) and Hevia et al. (2015) for another proposal in which the discount parameter is allowed to change over time. Second, the overall volatility is represented by a GARCH model[1], allowing for the evolution of volatility seems to be an essential characteristic of the yield curve. However, it is important to note that this extension could be more relevant for density forecasts of the yield curve than when obtaining point forecasts; see, for example, Carriero, Clark, and Marcellino (2021) and Shin and Zhong (2017). Density forecasting of interest rates is important for derivatives pricing and risk management. João F Caldeira, Márcio P Laurini, and Portugal (2010) and Márcio P Laurini and João F Caldeira (2016) also allow for time variation in the discounting parameter and volatilities. A final helpful extension of the DNS model considered in this paper is the inclusion of macroeconomic variables to explain the yield curve; see Gürkaynak and Wright (2012) and Morley (2016) for surveys on the relationships between the yield curve and the macroeconomy.

[1] There are many other extensions, which could also be of interest in empirical applications. For example, Jungbacker, Koopman, and Van Der Wel (2014) propose a DFM with smooth loadings with the smoothness conditions introduced via spline functions.

Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) propose augmenting the yield curve specification by adding macroeconomic variables to explain the evolution of the level, slope, and curvature factors. In particular, they suggest writing the model for the yield curve as a state space model in which the three elements (level, slope, and curvature) depend on macroeconomic variables. The state-space model (SSM) representation allows the estimation of all the parameters in the model (including the decay rate) in just one step as compared with the two-step estimator proposed by Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006)[2]. They find strong evidence of the effects of macroeconomic variables on the yield curve and evidence for a reverse influence.

Our main objective is to assess the role of the four elements described above in forecasting the yield curve, namely, the role of macroeconomic variables, time-varying discounting, conditional heteroscedasticity, and the fourth factor. We will fit an encompassing model with all these four characteristics to the above data set. Estimation of the general model is carried out using the Extended Kalman filter. We show that the second discount rate has no role in obtaining a better fitting and forecasting performance of the factor model. Furthermore, we show that the best specification depends on the maturity and forecast horizon when looking at point forecasts. For short maturities, the best performance is obtained in a heteroscedastic model with a time-varying discount. However, whether macro variables add information depends on the large forecast horizon. However, the simplest homoscedastic model with constant discount performs better for large maturities. In this case, macro variables may have a role at short forecast horizons. These results suggest that the factor model should incorporate some sort of non-linearity depending on the maturity.

The outline of the paper is as follows. Section 2 describes the general model and its estimation. In section 3, the model is fitted to a data set of end-of-month continuously compounded Treasury yields on US zero-coupon bonds. Section 4 deals with out-of-sample point and density forecasting. Section 5 proposes further specifications designed to enhance the performance of the DNS model beating the random walk specification. Section 6 concludes.

2.1 DYNAMIC NELSON-SIEGEL-SVENSSON MODEL

The Nelson-Siegel model was extended to a more flexible form by Svensson (1994) by the inclusion of an additional term in Equation (40) which allows the curve to assume a double-humped shape. This so-called Dynamic-Nelson-Siegel-Svensson

2 Note that the SSM representation allows interpreting the model proposed by Francis X Diebold and Li (2006) as a Dynamic factor model (DFM) with restrictions on the loadings.

model has the form

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

The second curvature is governed by $\beta_{4,t}$ with decay parameter λ_2. Both the Nelson and Siegel and the Svensson models have been extensively adopted for pricing, hedging, and monetary policy purposes Francis X. Diebold and Glenn D. Rudebusch (2013).

The DNS and DNSS framework can also be represented as a state space model by treating $\boldsymbol{\beta_t} = \beta_{j,t}$, for $j = 1, \ldots, 4$, as a latent vector. For these purpose, the general specification of the dynamic factor model is given by:

$$y_t = \Lambda(\lambda_t)\boldsymbol{\beta_t} + \varepsilon_{\boldsymbol{t}}, \qquad \varepsilon_{\boldsymbol{t}} \sim \mathcal{N}(0,\Sigma_\varepsilon), \tag{2}$$

where Λ is a $N \times K$ matrix of factor loadings, $\boldsymbol{\beta_t}$ is a K-dimensional stochastic process, and $\varepsilon_{\boldsymbol{t}}$ is the $N \times 1$ vector of measurement errors, whose covariance matrix given by Σ_ε. For any given, strictly positive and distinct, λ_1 and λ_2, the $N \times K$ factor loading matrix $\Lambda(\lambda_t)$ is given by:

$$\Lambda_{ij}(\lambda_k) = \begin{cases} 1, & j = 1 \\ \psi_{i2} = \dfrac{1-z_{1i}}{\lambda_1\tau_i}, & j = 2 \\ \psi_{i3} = \dfrac{1-z_{1i}}{\lambda_1\tau_i} - z_{1i}, & j = 3 \\ \psi_{i4} = \dfrac{1-z_{2i}}{\lambda_2\tau_i} - z_{2i}, & j = 4, \end{cases}$$

where $z_{1,i} = \exp(-\lambda_1\tau_i)$ and $\psi_{2,i} = \exp(-\lambda_2\tau_i)$. The DNS model is obtained by eliminating the second curvature ψ_{i4} of the DNSS specification.

The transition equation increase appropriately to include the DNSS extension. Also details on factor-augmented DNS model, time-varying volatility, and estimation procedure with Kalman filter are available in subsection 3.1.3, subsubsection 3.1.4.2, and subsection 3.1.4 respectively.

2.1.1 Time-Varying Loading Parameter

In previous models, the decay parameters are time-invariant over the full sample period, despite being freely estimated alongside the other parameters with the state space approach. However, Koopman, Mallee, and Van der Wel (2010) argued that the characteristics of the yield curve could vary over time, such that the factor loadings associated with the slope and curvature also change over time. The introduction of time-varying loading parameters results in non-linear models, which makes estimation slightly more complicated.

We followed Koopman, Mallee, and Van der Wel (2010) and included the time-varying loading parameters (λ_{1t} and λ_{2t}, the first for DNS and both DNSS models, respectively) in the set of factors, so they are considered latent factors. To prevent the loading parameters from assuming negative values we use $\log(\lambda_t)$ in the set of factors. The new state vector is $B_t = \left[\beta_{1t}, \beta_{2t}, \beta_{3t}, \beta_{4t}, \boldsymbol{\lambda_t}\right]'$, where $\boldsymbol{\lambda_t} = \left[\log(\lambda_{1t}), \log(\lambda_{2t})\right]$. Making $\beta_{4t} = \log(\lambda_{2t}) = 0$ we have the DNS model. The dynamic of the new state vector B_t is again modeled by using VAR(1) processes described by Equation (16). In this case, the new observation equation becomes non-linear and the resulting state space representation is given by

$$y_t = \Lambda\left(\exp(\boldsymbol{\lambda}_t)\right)\boldsymbol{\beta}_t + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}\left(0, \Sigma_\varepsilon\right) \tag{3}$$

$$B_{t+1} = \boldsymbol{\mu}^{\mathrm{D}} + \Phi^{\mathrm{D}}\left(B_t - \boldsymbol{\mu}^{\mathrm{D}}\right) + \boldsymbol{\eta}_t^{\mathrm{D}}, \quad \boldsymbol{\eta}_t^{\mathrm{D}} \sim \mathcal{N}\left(0, \Sigma_\eta^{\mathrm{D}}\right), \tag{4}$$

where the dimension of $\Phi^{\mathrm{D}}, \boldsymbol{\mu}^{\mathrm{D}}, \boldsymbol{\eta}_t^{\mathrm{D}}$, e Σ_η^{D} are adjusted as appropriate. We refer to this model as the DNSS-TVL, and DNS-TVL when $\beta_{4t} = \log(\lambda_{2t}) = 0$. Since the loading matrix are time-varying, due to the time-varying $\boldsymbol{\lambda_t}$, the measurement equation is non-linear and the estimation can no longer rely on the Kalman filter. To overcome this problem, we perform the estimations using the extended Kalman filter (EKF), which we describe in the next section.

2.1.2 Estimation of Nonlinear State Space Models

As earlier mentioned, with the introduction of the time-varying factor loading, the observation equation becomes non-linear in the state vector. Because of the nonlinearity in Equations Equation (3) and Equation (4), the linear Kalman filter can no longer used. Hence, to circumvent this problem, we can proceed with the extended Kalman filter. The extended Kalman filter is much the same as the Kalman filter discussed in subsubsection 3.1.4.1, with the addition of first-order Taylor series terms to the transition and observation equations to account for non-linearities. For each time t, the non-linear dynamic and measurement equations are linearized locally around the current state estimates, $b_{t|t-1}$, using first-order Taylor series expansion. The measurement equation is then rewritten as follow

$$y_t = H_t\left(B_t\right) + \varepsilon_t, \tag{5}$$

where,

$$\begin{aligned} H_t\left(B_t\right) &= \Lambda\left(\exp\left(\boldsymbol{\lambda}_t\right)\right)\left(\beta_{1t}, \beta_{2t}, \beta_{3t}, \beta_{4t}\right)', \\ &= \Lambda_1\left(\lambda_{1t}\right)\beta_{1t} + \Lambda_2\left(\lambda_{1t}\right)\beta_{2t} + \Lambda_3\left(\lambda_{1t}\right)\beta_{3t} + \Lambda_4\left(\lambda_{2t}\right)\beta_{4t}, \end{aligned} \tag{6}$$

and $\Lambda_i\left(\exp\left(\boldsymbol{\lambda}_t\right)\right)$ is the i-th column of $\Lambda\left(\boldsymbol{\lambda}_t\right)$. Note that, in this case, the non-linearity of the state space model is limited to the measurement equation. Jazwinski (1970) demonstrated that EKF method is particularly effective in dealing with this type of non-linearity. The measurement equation are then linearized around the state estimates by substituting Jacobian matrix H_{it} in place of Λ_t in Equation Equation (21) and Equation (22), with

$$H_{it} = \left. \frac{\partial H_t\left(B_t\right)}{\partial B_t} \right|_{B_t = \hat{b}_{t|t-1}}, \tag{7}$$

where the (i,j) entry of H_{it} carries the partial derivative of the i-th measurement equation with respect to the j-th latent variable, evaluated at $\hat{b}_{t+1|t}$. The resulting approximation is given by

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

where $\hat{b}_{t|t-1}$ is an estimate of the states B_t in time t based on the past observations up to time $t-1$, with Jacobian matrix of the non-linear equation as follows

$$\dot{H}_t = \left. \frac{\partial H_t\left(B_t\right)}{\partial B_t} \right|_{\boldsymbol{\beta}=\boldsymbol{\beta}_{t|t-1}} = \left[\left. \frac{\partial H_t}{\partial \beta_{1t}} \right|_{\beta_{1t}=b_{1t|t-1}}, \dots, \left. \frac{\partial H_t}{\partial \beta_{jt}} \right|_{\beta_{jt}=b_{jt|t-1}}, \left. \frac{\partial H_t}{\partial \log(\lambda_{kt})} \right|_{\log(\lambda_{kt})=\log(\lambda_{kt|t-1})} \right], \tag{9}$$

with referring to $j=3$ and $k=1$ for DNS-TVL, and $j=4$ and $k=2$ for DNSS-TVL1 and DNSS-TVL2 models. The first j columns of $\dot{H}_t$ follow straightforward from differentiating Equation (6) with respect to $\beta_{1t},\dots,\beta_{jt}$, with $j=3$ or 4, for DNS and DNSS models, respectively. For the columns $j+1$ and $j+2$ we need to make use of the chain-rule as follows

$$\begin{aligned} \frac{\partial H(B_t)}{\partial \lambda_{kt}} &= \frac{\partial H(B_t)}{\partial \log(\lambda_{kt})} \cdot \frac{\partial \log(\lambda_{1t})}{\partial \lambda_{kt}} = \frac{\partial H(B_t)}{\partial \log(\lambda_{kt})} \cdot \frac{1}{\lambda_{kt}}, \\ \frac{\partial H(B_t)}{\partial \log(\lambda_{kt})} &= \frac{\partial H(B_t)}{\partial \lambda_{kt}} \cdot \lambda_{kt}. \end{aligned} \tag{10}$$

Applying the chain rule in Equation (10) results

$$\begin{aligned} \frac{\partial H(B_t)}{\partial \lambda_{1t}} &= \frac{\beta_{2t} \cdot \exp(-\tau_i\lambda_{1t})(\tau_i\lambda_{1t} - \exp(\tau_i\lambda_{1t}) + 1)}{\tau_i\lambda_{1t}^2} + \dots \\ &+ \frac{\beta_{3t} \cdot \exp(-\tau_i\lambda_{1t})(\tau_i^2\lambda_{1t}^2) + \tau_i\lambda_{1t} - \exp(\tau_i\lambda_{1t}) + 1)}{\tau_i\lambda_{1t}^2}, \qquad \text{if } k=1, \\ \frac{\partial H(B_t)}{\partial \lambda_{2t}} &= \frac{\beta_{4t} \cdot \exp(-\tau_i\lambda_{2t})(\tau_i^2\lambda_{2t}^2) + \tau_i\lambda_{2t} - \exp(\tau_i\lambda_{2t}) + 1)}{\tau_i\lambda_{2,t}^2}, \qquad \text{if } k=2. \end{aligned}$$

Therefore, the decay factors λ_{kt} are filtered in EKF recursions, where $k=1$ implies that we are filtering λ_{1t} in DNS-TVL, DNS-GARCH-TVL, DNSS-TVL1, and DNSS-GARCH-TVL1 specifications, and $k=2$ implies that we are filtering λ_{2t} in the DNSS-TVL2 and DNSS-GARCH-TVL2 models. In general, this yields the following result

$$\dot{H}_t = \left[\iota_{N\times 1} \quad \Lambda_2\left(\exp(\lambda_{kt})\right) \ \dots \ \Lambda_j\left(\exp(\lambda_{p,t})\right) \quad \frac{\partial H(B_t)}{\partial \lambda_{kt}} \cdot \lambda_{kt} \right], \tag{11}$$

where $\imath_{N\times 1}$ is a vector of ones of length N.

Last but not least, we consider the model specification with time-varying factor loadings and time-varying volatility, named DNS-GARCH-TVL. The measurement equation of the DNS-GARCH-TVL models is given by

$$\begin{aligned} y_t &= H_t\left(B_t\right) + \varepsilon_t^{\dagger}, \\ &= \Lambda\left(\exp(\boldsymbol{\lambda} t)\right)\left(\beta_{1,t},\ldots,\beta_{jt}\right)' + \Gamma_{\varepsilon}\varepsilon_t^{*} + \varepsilon_t^{\dagger}, \quad \varepsilon_t^{\dagger} \sim \mathcal{N}(0,\Sigma_{\varepsilon}^{\dagger}), \end{aligned} \tag{12}$$

where $B_t = (\beta_{1t},\beta_{2t},\beta_{3t},\log(\lambda_{1t}))'$ for the DNS-GARCH-TVL model, and where $B_t = (\beta_{1t},\beta_{2t},\beta_{3t},\beta_{4t},\ldots \log(\lambda_{1t}),\log(\lambda_{2t}))'$ the DNSS-GARCH-TVL1 or DNSS-GARCH-TVL2 models. This equation has the additional component $\Gamma_{\varepsilon}\varepsilon_t^{*}$ and the state vector gets additional variables, from witch only ε_t^{*} is represented in the measurement equation. The gradient matrix for Equation (12) is given by

$$\dot{H}_t = \begin{bmatrix} \imath_{N\times 1} & \Lambda_2\left(\exp(\lambda_{kt})\right) \;\ldots\; \Lambda_j\left(\exp(\lambda_{kt})\right) & \dfrac{\partial H(B_t)}{\partial \lambda_{kt}}\cdot\lambda_{kt} & \Gamma_{\varepsilon} \end{bmatrix}. \tag{13}$$

If we replace Equation (11) or Equation (13) in Equation (7) and then substitute Equation (7) in the in the measurement equation we obtain the linearized model. Since we have at this point again a linear system, we can apply the usual recursions of the Kalman filter. In the following Section, we explain the algorithm used in combined forecasts models.

2.1.3 Combined Forecasts

An important result from the methodological literature on forecasting is that a linear combination of two or more forecasts may yield more accurate predictions than using only a single forecast, see Granger (1989), Newbold and David I Harvey (2002), and Aiolfi and Timmermann (2006). Adaptive strategies for combining forecasts might also mitigate structural breaks and model misspecification and thus lead to more accurate forecasts, see Pesaran and Timmermann (2007) and Newbold and David I Harvey (2002). In particular, there is evidence that combining forecasts of nested models can significantly improve forecasting precision upon forecasts obtained from single model specifications, see Clark and McCracken (2009).

Assuming we are combining forecasts from $\mathcal{M}$ different forecast models, a combined forecast for a h-month horizon for the yield with maturity τ_i given by

$$\hat{y}_{t+h|t}\left(\tau_i\right) = \sum_{m=1}^{\mathcal{M}} w_{t+h|t,m}\left(\tau_i\right)\hat{y}_{t+h|t,m}\left(\tau_i\right),$$

where $w_{t+h|t,m}(\tau_i)$ denotes the weight assigned to the time-t forecast from the m^{th} model, $\hat{y}_{t+h|t,m}(\tau_i)$. Most of the forecast combination schemes considered are adaptive, meaning that the forecasts included in $\mathcal{M}$: $\hat{y}_{t+h|t,m}(\tau_i)$ and/or corresponding weights m^{th} are based on alternative selection criteria within a sub-sample of realized observations.

Note that since a forecaster would only have information available up to the forecast origin ω, the sub-sample for forecast selection and computation of weights must contain data on or before that period. Thus, we start by setting equal weights to all forecasts until the selection of forecasts and weighting schemes could be based on the evaluation of realized forecast errors. This procedure guarantees that we use only information available up to a particular period ω to set weights of forecasts for period $\omega + h$. The following 5 alternative combination strategies $\mathcal{M}$ ={FC-EW, FC-OLS, FC-RANK, FC-MSE, FC-RMSE}={1,2,. . . ,5} are considered:

1. *Equally weighted forecasts (FC-EW)*: Various studies have demonstrated that simple averaging of a multitude of forecasts works well in relation to more sophisticated weighting schemes. Therefore, the first forecast combination method we consider assigns equal weights to the forecasts from all individual models, i.e. $w_{t+h|t,m}(\tau_i) = \frac{1}{\mathcal{M}}$ for $m = 1, \ldots, \mathcal{M}$. We denote the resulting combined forecast as Forecast Combination - Equally Weighted (FC-EW). As explained in Timmermann (2006), this approach is likely to work well if forecast errors from different models have similar variances and are highly correlated.

2. *Thick modeling approach with OLS weights (FC-OLS)*: A study by Granger and Jeon (2004) proposes the so-called thick modeling approach (TMA), which consists of selecting the z-percent of the best forecasting models in the sub-sample period for model evaluation, according to the root mean square error (RMSE) criterion. We use the selection process of Granger and Jeon and subsequently compute weights by means of OLS regressions along with the constraint that the weights are all positive and sum up to one. The z-percent of top forecasts selected is set to 2 (i.e., about z = 40%).

3. *Rank-weighted combinations (FC-RANK)*: The FC-RANK scheme, suggested by Aiolfi and Timmermann (2006), consists of first computing the RMSE of all models in the sub-sample period for evaluation. Defining $RANK^{-1}_{t+h|t,m}$ as the rank of the m^{th} model based on its historical RMSE performance up to time t for horizon h, the weight for the m^{th} forecast is then calculated as

$$\widehat{w}_{t+h|t,m}(\tau_i) = RANK^{-1}_{t+h|t,m} / \sum_{m=1}^{\mathcal{M}} RANK^{-1}_{t+h|t,m}.$$

4. *Thick modeling approach with MSE-Frequency weights (FC-MSE)*: This scheme consists of selecting models employing the thick modeling approach and assigning to each m^{th} forecast a weight equal to a model's empirical frequency of minimizing the squared forecast error over-realized forecasts. The weight for model m is computed as

$$\widehat{w}_{l+h|t,m}(\tau_i) = \frac{1/MSE_{t+h|t,m}(\tau_i)}{\sum_{m=1}^{\mathcal{M}} 1/MSE_{t+h|t,m}(\tau_i)}.$$

5. *Thick modeling approach with RMSE-weights (FC-RMSE)*: This scheme consists of selecting models by means of the thick modelling approach, then computing the RMSE of all selected models m and setting:

$$\widehat{w}_{l+h|t,m}(\tau_i) = \frac{1/RMSE_{t+h|t,m}(\tau_i)}{\sum_{m=1}^{\mathcal{M}} 1/RMSE_{t+h|t,m}(\tau_i)}.$$

In the next section, we present the in-sample fit and out-of-sample results.

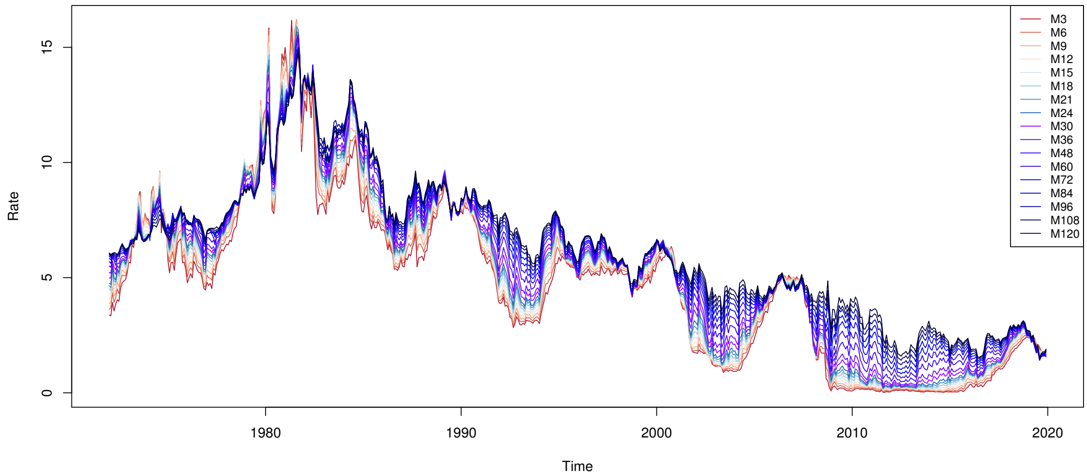
2.2 EMPIRICAL RESULTS

The empirical analysis is structured in the data summary, in-sample and out-of-sample analysis. For in-sample examination, we used a database with monthly observations from January 1972 through December 2019, and for out-of-sample analysis, we made forecasts from January 1994 through December 2019. In the following subsection, we presented the dataset and the respective in-sample discussion as well for out-of-sample.

2.2.1 Data

The data set consists of end-of-month continuously compounded yields on U.S. zero-coupon bonds. Specifically, we use the novel zero-coupon Treasury yield curve data set constructed by Liu and Jing Cynthia Wu (2021) and is publicly available on on the Journal of Financial Economics Data Archive, as part of their supplementary material. The dataset covers the period from January 1972 through December 2019 for maturities of $3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108$, and 120 months of US securities, with 576 monthly observations. Also, we use backward-looking macroeconomic variables such as manufacturing capacity utilization (CU_t), the monthly average of the federal funds rate (FFR_t), and 12-month percent change in the price deflator for personal consumption expenditures ($INFL_t$). As Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) suggest, these three variables represent, respectively, the level of real economic activity relative to potential, the monetary policy instrument, and the inflation rate, which are widely considered to be the minimum set of fundamentals needed to capture basic macroeconomic dynamics.

Figure 1 – U.S. Treasury Bonds per Maturities.



Source: Elaborated by the author.

Legend: U.S. Treasury Bonds per Maturities. The figure shows all seventeen time series maturities for τ = 3, 6, 9, 12, 15, 18, 21, 24, 30, 36, 48, 60, 72, 84, 96, 108 and 120 months. The figure also shows periods associated with the yield curve inversion, where short-term rates are higher than long-term rates. The sample contains 576 monthly observations from January 1972 through December 2019.

Table 1 – Summary Statistics.

Maturities	Mean	Std.dev.	Min.	Max.	ρ_1	ρ_6	ρ_{12}	Skewness	Kurtosis
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

Source: Elaborated by the author.

Legend: The table reports summary statistics for U.S. treasury yields from January 1972 through December 2019. Maturity is measured in months. For each maturity we show mean, standard deviation (Std. dev.), minimum, maximum, and three autocorrelation coefficients, 1 month [$\hat{\rho}(1)$], 6 months [$\hat{\rho}(6)$], and 12 months [$\hat{\rho}(12)$]. The proxies for level is the highest maturity bond (120 months), for slope, the difference between the bond of 120 months and the bond of 3 months, and for curvature, two times the bond of 24 months minus the sum of bond of 3 months and bond of 120 months.

The yield curve shape is analyzed since it provides indications of future interest rates and economic activity. Figure 1 shows the term structure of the interest rate on a three-dimensional surface as a function of maturity over time. The Figure reveals some common stylized facts for the yield curve as the dynamics of levels and slopes.

In Table 1, we presented the dataset in some descriptive statistics for each maturity, such as mean, standard deviation, skewness, and kurtosis. One can confirm some well-known stylized facts of the yield curve, as autocorrelations close to one and increase with maturity, that is, regardless of maturity, are persistent processes. Also, volatility decreases through maturities, skewness, and excess kurtosis, suggesting no normality in the presented sample.

The following section discusses the in-sample extension results regarding DNS and DNSS reference models and and also analyzes the dynamics of Svensson extensions with the inclusion of time-varying loading parameters time-varying volatility.

2.2.2 In-Sample Results

Our first set of extensions is the factor augmented Nelson-Siegel models with time-varying loading parameter and volatility. We summarize in Panel (A) of Table 2 the in-sample results as the root mean squared errors (RMSE), log-likelihood, and AIC of Nelson-Siegel extensions. The result of DNS-Macro in terms of RMSE is similar to the DNS model, which could be expected since Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) findings suggest the same conclusions, see Table 1 - summary statistics for measurement errors of yields - in Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006). For most maturities, the results in factor augmented model extensions are approximated to the extant models. Otherwise, the results of 3-month and 6-month maturities for all factor augmented extensions outperform the DNS benchmark model, especially the DNS-GARCH-TVL-Macro model.

Figure Figure 2 presents the filtered level, slope, and first and second curvature of all DNSS extension models. All filtered factors are quite similar throughout all extensions, so we could infer the results in-sample should be similar. In fact, prior assumption is verified in Table 2. The log-likelihood values in Table 2 also suggest the models with time-varying volatility outperform the benchmark model. These values are used as a basis of performance comparison for the other models, besides the Akaike Information Criterion (AIC) and the Likelihood Ratio test (LR statistic).

We consider the DNS (DNSS) model a benchmark to DNS (DNSS) extensions throughout our in-sample analysis. We start proposing extensions to the DNSS model. We aim to improve the flexibility of curves and adjustments by allowing time-varying volatility, so we have the DNSS-GARCH with 68 parameters. The RMSE results, in Table 2, suggest the DNSS-GARCH and DNS-GARCH outperform their respective benchmark model for almost all maturities, mainly for the first two and the last two maturities, as well as the analog extension in the Nelson-Siegel. Overall the DNS(S)-

GARCH model is better in terms of RMSE than the DNS (DNSS) model, however the DNS(S)-GARCH model is more parameterized than the latter two.

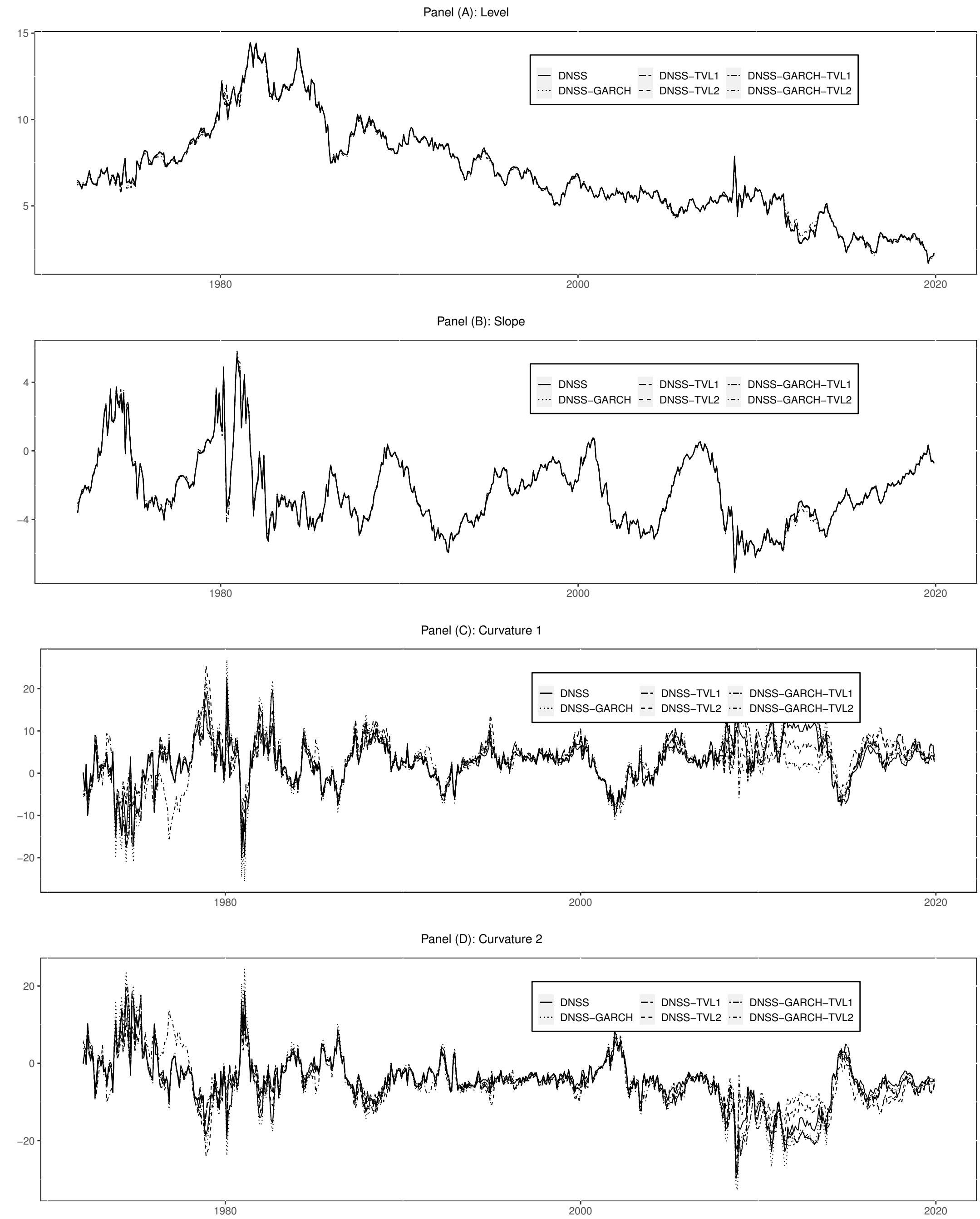
The second and third extensions of the DNSS model refer to the models with time-varying factor loadings, i.e., DNSS-TVL1 and DNSS-TVL2. In the DNS model, Francis X Diebold and Li (2006) suggest setting λ_t at 0.0609 for every t, instead of treating it as an unknown parameter for a dataset from January 1972 through December 2000. Otherwise, Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) is treated as a parameter to be estimated, whose value after the estimation is 0.077, a result similar to that found by Koopman, Mallee, and Van der Wel (2010), 0.0778. However, Koopman, Mallee, and Van der Wel (2010) and Márcio Laurini and Hotta (2010) suggest that λ is time-varying for each observed yield curve.

Our proposed models to λ is twofold, that is, the DNSS-TVL1 (DNSS-TVL2) model with λ_{1t} (λ_{2t}) time-varying while the parameter λ_2 (λ_1) is estimated. Figure 3 (Figure 4) shows the evolution of the factor loading in the DNSS-TVL1 (DNSS-TVL2) model. The results suggest the US yield curve presented the highest value and volatility of λ_{1t} in the 1980s. On the other hand, λ_{2t} values laying around below 0.06. In terms of RMSE, these two models outperform the DNSS model only in maturity of 3 months, see Table 2. Also, the log-likelihood gain for the DNSS-TVL1 and DNSS-TVL2 models are not expressive against the DNSS model as the gain of the DNSS-GARCH model against the DNS model. Despite gains in terms of RMSE, almost all Svensson extensions have more estimated parameters, except for the DNSS-GARCH model so far.

The evolution of the h_t over time is quite similar in the DNSS-GARCH and the DNSS-GARCH-TVL1 models, see Figure 5. Koopman, Mallee, and Van der Wel (2010) suggests the publication of the Nelson and Siegel (1987) paper, after the mid-1980s, reduced the volatility in US bond rates, except around the financial crisis of 2008. The previous period has more peaks and greater amplitude in the GARCH component. In general, there are no major differences among the DNS-GARCH-TVL1, DNS-GARCH-TVL2, and the DNSS-GARCH model, for instance, is in the amplitude of the peaks of the h_t.

The Svensson extension model with $\lambda_{1,t}$ ($\lambda_{2,t}$) and time-varying obtained the lowest RMSE results in the first and second maturities against the DNSS model. Figure 5 presents the filtered estimates of the common volatility for the DNS-GARCH-TVL1 and DNS-GARCH-TVL2 models. In general, they have the same path over time. As in Koopman, Mallee, and Van der Wel (2010), in all the analyzed models, the model with the best performance in the analyzed criteria is the model with time-varying volatility. In the following Section, the out-of-sample results of the developed models are presented.

Figure 2 – Level, Slope, Curvature 1, and Curvature 2.



Source: Elaborated by the author.

Legend: This figure reports the level, slope, curvature 1, and curvature 2 obtained from the baseline dynamic Svensson (DNSS) model, models with time-varying factor loadings (DNSS-TVL1 and DNSS-TVL2), model with time-varying volatility (DNSS-GARCH), models with both time-varying factor loadings and volatility (DNSS-GARCH-TVL1 and DNSS-GARCH-TVL2). Panels (A), (B), (C) and (D) present the level, slope, curvature 1, and curvature 2 estimates, respectively.

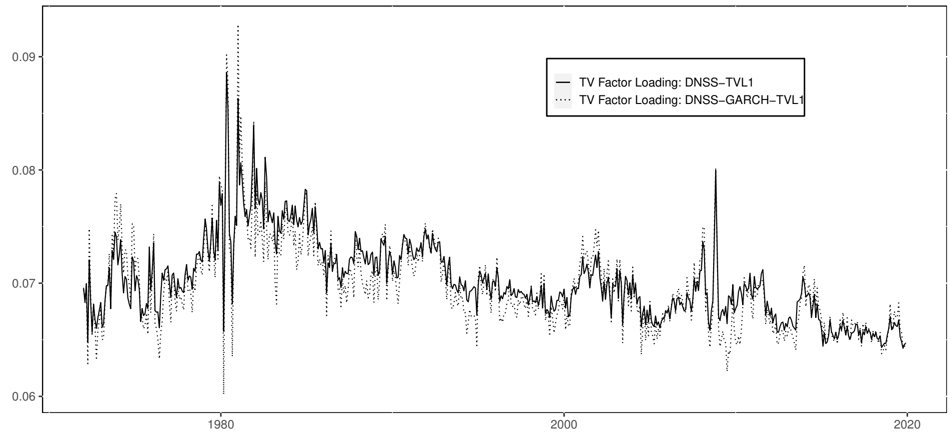
Table 2 – RMSE In-sample, Log-likelihood, and AIC of Model Extensions.

	Panel (A): Nelson-Siegel Extensions								Panel (B): Svensson Extensions					
Maturities	DNS	DNS -Macro	DNS -TVL	DNS-TVL -Macro	DNS -GARCH	DNS-GARCH -Macro	DNS-GARCH -TVL	DNS-GARCH -TVL-Macro	DNSS	DNSS -TVL1	DNSS -TVL2	DNSS -GARCH	DNSS-GARCH -TVL1	DNSS-GARCH -TVL2
3	32.19	32.21	16.01	16.63	28.20	28.19	22.32	9.68	18.75	18.45	18.42	0.17	4.25	2.60
6	16.66	16.68	8.77	8.78	17.11	17.36	11.01	9.36	6.27	8.05	6.43	2.04	7.06	4.30
9	7.52	7.53	8.49	7.88	4.74	4.83	6.42	10.52	0.01	5.80	3.02	0.02	7.08	4.28
12	2.80	2.80	9.18	8.33	0.05	0.04	5.89	11.16	2.40	5.90	3.91	2.57	7.14	4.98
15	1.22	1.22	9.15	8.41	1.49	1.57	6.33	11.82	1.96	5.20	3.35	2.14	6.28	4.14
18	2.45	2.45	8.89	8.59	1.02	1.18	6.44	12.24	1.24	4.47	2.66	1.47	5.56	3.31
21	2.34	2.36	8.66	8.73	0.88	0.83	6.49	12.07	1.69	4.11	2.58	1.68	5.06	2.98
24	1.32	1.35	8.77	8.96	1.99	1.97	6.74	11.86	2.44	4.28	2.89	2.29	4.88	2.97
30	3.87	3.85	9.58	10.00	3.59	3.72	7.40	12.80	3.99	5.57	4.17	3.92	5.71	4.23
36	5.46	5.43	9.85	10.35	2.98	3.22	7.17	13.81	3.49	5.16	3.55	3.52	5.27	3.85
48	7.65	7.61	10.87	11.59	3.63	3.46	7.72	15.62	3.37	6.75	4.09	3.22	6.91	5.28
60	8.75	8.72	11.70	12.16	4.76	4.35	8.22	16.62	4.65	8.51	5.76	4.63	9.26	7.07
72	7.35	7.32	10.48	10.80	5.75	5.82	8.29	16.28	5.58	8.57	6.76	5.50	9.45	7.84
84	4.65	4.65	8.89	8.99	4.48	4.67	7.05	15.39	4.51	7.49	5.49	4.55	8.62	6.43
96	2.06	2.05	7.81	7.87	1.98	0.81	5.67	15.15	2.25	6.28	4.34	2.77	7.70	5.82
108	5.95	5.96	8.56	8.55	4.52	5.33	6.89	15.45	4.12	7.31	5.60	3.61	7.87	6.40
120	10.96	10.97	11.89	12.14	12.54	13.64	11.66	17.87	10.19	11.74	10.54	9.53	11.85	10.56
Log-likelihood	10188.9	10222.9	11557.4	11717.0	11953.6	11640.6	12175.4	12898.9	12254.8	12583.9	12580.2	13215.8	13215.8	13353.2
Parameters	36	81	47	101	55	100	66	120	49	63	63	68	82	82
AIC	-20301.9	-20279.9	-23016.8	-23228.0	-23793.2	-23077.2	-24214.8	-25553.7	-24407.6	-25037.9	-25030.4	-26291.7	-26263.7	-26538.5
LR-statictic		-22.0	2714.9	2926.1	3491.3	2775.3	3912.9	5251.8		630.3	622.8	1884.1	1856.1	2130.9

Source: Elaborated by the author.

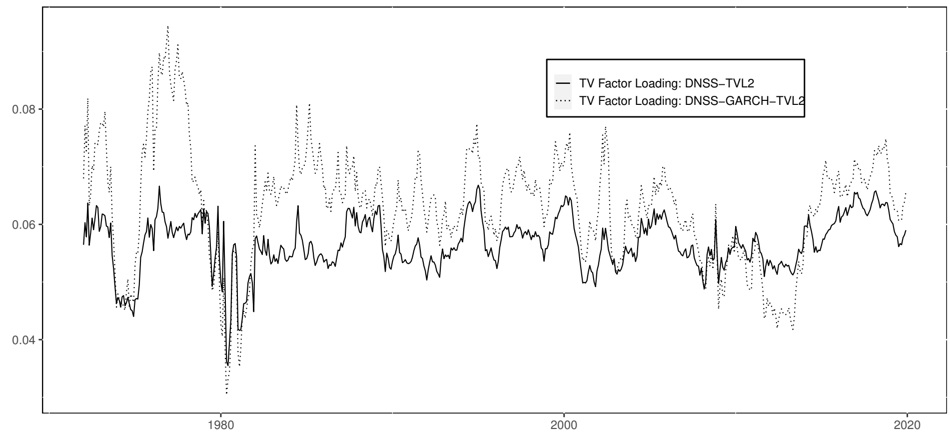
Legend: We present the RMSE in-sample. Benchmark model is located in the first column and maturities by rows. The table presents the log-likelihood results, the Akaike Information Criterion (AIC), the number of parameters for the different models, and the LR test. The DNS model in Panel (A) and DNSS in Panel (B) correspond to Nelson-Siegel and Svensson latent factor models with constant loading parameter (λ) and constant volatility. The models DNS-TVL, DNSS-TVL1 and DNSS-TLV2 correspond to models with time-varying λ_t. The DNS-GARCH and DNSS-GARCH models correspond to models with a common GARCH component for time-varying volatility. The models DNS-TVL-GARCH, DNSS-GARCH-TVL1 and DNSS-GARCH-TVL2 correspond to models with a factor loading and the common GARCH component for volatility time-varying. The models with macroeconomic factors are followed by “-Macro". We present in total seven Nelson-Siegel extensions in Panel A and the five Svensson extensions in Panel B.

Figure 3 – Factor Loading $\lambda_{1,t}$ Svensson.



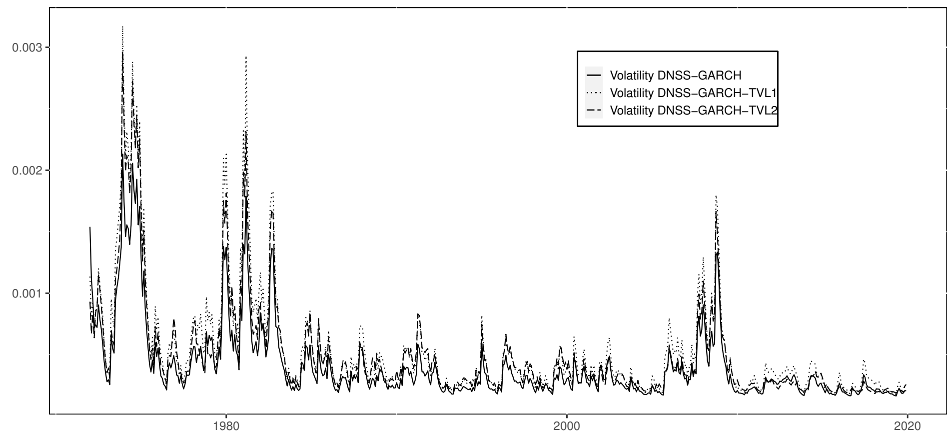
Source: Elaborated by the author.
Legend: This figure shows the DNSS-TVL1 λ estimates (dots) with those obtained from DNSS-GARCH-TVL1 (solid line) in the Svensson (cross-section) model.

Figure 4 – Factor Loading $\lambda_{2,t}$ Svensson.



Source: Elaborated by the author.
Legend: This figure shows the DNSS-TVL2 λ estimates (dots) with those obtained from DNSS-GARCH-TVL2 (solid line) in the Svensson (cross-section) model.

Figure 5 – Common Component $h_t^{0.5}$ Svensson.



Source: Elaborated by the author.
Legend: We compare the time-varying common GARCH volatility estimates of DNSS-GARCH (solid line) with the DNSS-GARCH-TVL1 (dots) and DNSS-GARCH-TVL2 (dashed)

2.2.3 Out-of-Sample Results

The forecasting exercise is performed in pseudo real time, i.e., we never use information which is not available at the time the forecast is made. The forecasting experiment uses a rolling estimation window and relies on information up to period t to compute interest rate forecasts for period $t + h$, where $h = 1,3,6,12$ is the forecast horizon. We reestimate a given model at each time t and produce out-of-sample forecasts for 1-month, 3-month, 6-month, and 12-month ahead. More specifically, we divide our full sample into two parts (in- and out-of-sample): the first one over the period 1972:1 - 1993:12 with 264 monthly observations and the second over the period 1994:1 - 2019:12 with 312 monthly observations, and make predictions recursively extending the sample by 1 month every period. We constructed a moving window with 264 observations so that month after month a data line was added to the end of the sample and a line of data at the beginning of the sample was deleted to re-estimate parameters. Thus through the (extended) Kalman filter prediction step the forward predicted states were used to construct forecasts of yield curves for h-months ahead.

The period under analysis covers two recession periods 2000-2001 and 2007-2009. In these periods, there were reversal of the term structure of interest rate. First, DNS confirmed to be a very competitive benchmark in forecasting the term structure of bond yields, mainly for twelve-step-ahead forecast and longest maturities. The hypothesis is the competing models would obtain expressive results against DNS model, in periods with more volatility and different shapes in yield curve. In fact, the DNS-GARCH and DNSS-GARCH had more statistically significant results in 1-month ahead than the benchmark model, with lowest RMSPE in the set of best models according to Model Confidence Set. One can note, however, all competing models outperform the DNS model in maturity of 3 months in 1-month forecast horizon, see Table 3. We observed the same pattern in 3-month forecast horizon related to forecast to maturity of 3 months see Table 4. In this horizon, the DNS-GARCH model had statistical significance over all maturity, whereas the DNSS-GARCH-TVL2 model had more results inside the MCS (17 maturities).

The results for 6-months and 12-months ahead models with time-varying volatility outperform the benchmark in all maturities see Table 5 and see Table 6. Overall models with time-varying loading factor had RMSPE lower than DNS model, nevertheless models with both time-varying parameters had statistical significance over most maturities and only the DNSS-GARCH-TVL2 was considered the best model according to MCS. We explore this last result in detail in Figure 8 and Figure 9, which show the cumulative squared forecast errors (CSFE), relative to the DNS model, of selected maturities for 6 and 12 months forecast horizon. The model DNSS-GARCH-TVL2 have outperform other models over forecasts and mainly outperform DNS model. One can also note this consistency over recessions periods for instance 2000-2001 and 2007-2009.

In general, the DNS-GARCH model (dashed red line) and DNSS-GARCH-TVL2 outperform the DNS model and others in out-of-sample analysis, whereas the models with the $\lambda_{1,t}$ were not so successful, for instance DNS-TVL and DNSS-TVL1. We also note models with only time-varying loading factor had fall in performance in recession periods, this can suggest it is difficult forecast λ over time. Overall recession periods clearly sift the wheat from the chaff, for instance the DNS model outperform the DNSS model, mainly after 2000's.

Table 3 – Relative Root Mean Squared Forecast Errors: 1-month ahead

Models	3-Month	6-Month	9-Month	12-Month	18-Month	24-Month	30-Month	36-Month	48-Month	60 Month	72-Month	84-Month	96-Month	108-Month	120-Month
DNS	34.59	28.66	26.9	27.12	27.99	28.54	28.88	29	29.72	30.16	29.95	28.87	28.49	28.42	29.33
DNS – Macro	0.661**	**0.680****	**0.751****	**0.810****	**0.887**	0.955	1.015	1.059	1.115	1.144	1.150	1.163	1.154	1.128	1.095
DNS – GARCH	0.952**	0.936**	0.934	**0.944**	**0.974**	**1.003**	**1.024**	**1.036**	**1.060**	**1.074**	1.081	1.074	1.066	1.051	**1.027**
DNS – GARCH – Macro	0.661**	**0.680****	**0.751****	**0.810**	**0.887**	0.956	1.016	1.060	1.116	1.146	1.153	1.166	1.157	1.132	1.100
DNS – TVL	0.773	0.820	0.881	0.924	0.995	1.055	1.110	1.148	1.199	1.231	1.247	1.267	1.270	1.257	1.185
DNS – TVL – Macro	**0.630****	**0.720****	**0.809****	0.859**	0.938	1.014	1.085	1.134	1.210	1.280	1.334	1.397	1.434	1.450	1.398
DNS – GARCH – TVL	0.696**	**0.695****	**0.733****	**0.776****	**0.857****	0.929	0.988	1.026	1.079	1.109	1.124	1.133	1.138	1.133	1.102
DNS – GARCH – TVL – Macro	**0.604****	**0.658****	**0.740****	**0.801****	**0.888****	0.963	1.024	1.061	1.111	1.153	1.188	1.238	1.26	1.261	1.226
DNSS	0.730**	0.772**	0.821**	0.852	**0.904**	**0.949**	**0.974**	**0.984**	**1.003**	**1.025**	1.064	1.093	1.120	1.127	**1.088**
DNSS – GARCH	0.865**	0.956	1.024	1.053	1.088	1.110	1.116	1.105	1.093	1.100	1.132	1.151	1.167	1.166	1.117
DNSS – TVL1	0.849**	0.983	1.088	1.141	1.200	1.233	1.245	1.239	1.216	1.193	1.200	1.209	1.233	1.235	1.178
DNSS – TVL2	0.723**	0.828	0.928	0.975	1.009	1.026	1.038	1.042	**1.064**	1.102	1.155	1.206	1.240	1.254	1.208
DNSS – GARCH – TVL1	0.773**	0.911	1.031	1.090	1.139	1.148	1.150	1.152	1.152	1.162	1.202	1.252	1.292	1.313	1.270
DNSS – GARCH – TVL2	0.738**	0.890	1.014	1.062	1.077	1.076	1.078	1.077	**1.081**	**1.108**	1.156	1.210	1.246	1.268	1.224
FC – EW	0.770**	0.832**	0.901	0.943	0.994	1.033	1.064	1.082	1.109	1.133	1.159	1.186	1.203	1.203	1.163
FC – RANK	0.804	0.838	0.908	0.933	0.980	1.02	1.042	1.065	1.090	1.127	1.130	1.155	1.188	1.174	1.157
FC – MSE	0.780	0.828	0.893	0.934	0.989	1.028	1.057	1.073	1.097	1.119	1.145	1.173	1.191	1.193	1.156
FC – RMSE	0.784	0.837	0.901	0.941	0.992	1.029	1.057	1.073	1.097	1.12	1.146	1.174	1.191	1.194	1.156
FC – OLS	**0.645****	**0.660****	**0.716****	**0.763**	**0.832**	**0.889**	**0.932**	**0.962**	**1.006**	**1.030**	1.050	1.069	1.086	1.090	**1.061**

Source: Elaborated by the author.

Legend: The table reports relative root mean squared forecast errors (RMSFE) relative to the DNS model obtained by using individual yield models and different forecast combination methods, for the 1-month, 3-month, 6-month, and 12-month forecast horizons. The evaluation sample is 1994:1 to 2019:12 (312 out-of-sample forecasts). The first line in each panel of the table reports the value of RMSFE (expressed in basis points) for the DNS model, while all other lines reports statistics relative to the DNS. The following model abbreviations are used in the table: DNS for the dynamic Nelson-Siegel model and DNSS for the dynamic Svensson model. Macro refers to macroeconomic factors, GARCH for time varying volatility, and TVL refers to time varying factor loading. FC-EW, FC-OLS and FC-RANK stand for forecast combinations based on equal weights, OLS-based weights, and rank-weighted combinations, respectively. FC-RMSE and FC-MRMSE refer to forecast combinations base on the thick modeling approach with RMSE-weights and MSE-Frequency weights, respectively. Numbers smaller than one indicate that models outperform the DNS, whereas numbers larger than one indicate underperformance. Numbers in **bold** indicate that model is on the model confident set at 25% level. Stars indicate the level at which the Giacomini and White (2006) test rejects the null of equal forecasting accuracy (*, and ** mean respectively rejection at 5%, and 1% level).

Table 4 – Relative Root Mean Squared Forecast Errors: 3-month ahead

Models	3-Month	6-Month	9-Month	12-Month	18-Month	24-Month	30-Month	36-Month	48-Month	60 Month	72-Month	84-Month	96-Month	108-Month	120-Month
DNS	61.47	59.69	59.48	59.86	60.25	60.35	59.92	58.98	58.06	57.4	55.81	53.67	52.3	51.14	50.34
DNS – Macro	0.706*	0.741	0.782	0.820	0.887	0.950	1.007	1.054	1.127	1.183	1.224	1.268	1.292	1.306	1.308
DNS – GARCH	0.958	0.972	0.989	1.006	1.038	1.063	1.082	1.096	1.120	1.133	1.142	1.144	1.143	1.140	1.127
DNS – GARCH – Macro	0.706*	0.741*	0.782	0.820	0.887	0.951	1.007	1.055	1.128	1.185	1.226	1.270	1.294	1.308	1.311
DNS – TVL	0.805	0.828	0.853	0.881	0.941	1.000	1.057	1.106	1.184	1.237	1.274	1.303	1.320	1.327	1.300
DNS – TVL – Macro	0.688*	0.724	0.762	0.797	0.864	0.924	0.974	1.012	1.075	1.129	1.178	1.235	1.277	1.312	1.329
DNS – GARCH – TVL	0.694*	0.715*	0.747	0.782	0.850	0.909	0.959	0.998	1.064	1.113	1.151	1.186	1.212	1.235	1.241
DNS – GARCH – TVL – Macro	0.668*	0.708*	0.749	0.786	0.850	0.907	0.955	0.993	1.054	1.107	1.154	1.209	1.248	1.278	1.296
DNSS	0.819	0.834	0.856	0.880	0.927	0.965	0.993	1.014	1.051	1.085	1.125	1.160	1.197	1.228	1.232
DNSS – GARCH	0.955	0.983	1.006	1.025	1.063	1.088	1.104	1.113	1.134	1.156	1.189	1.214	1.244	1.268	1.264
DNSS – TVL1	0.928	0.982	1.027	1.061	1.120	1.162	1.188	1.203	1.216	1.219	1.237	1.258	1.285	1.310	1.310
DNSS – TVL2	0.793	0.858	0.918	0.960	1.003	1.019	1.029	1.038	1.075	1.124	1.187	1.246	1.300	1.345	1.358
DNSS – GARCH – TVL1	0.884	0.933	0.980	1.016	1.056	1.069	1.080	1.092	1.114	1.144	1.200	1.262	1.324	1.385	1.419
DNSS – GARCH – TVL2	0.851	0.934	0.997	1.034	1.066	1.076	1.082	1.090	1.122	1.173	1.243	1.316	1.384	1.446	1.473
FC – EW	0.826	0.861	0.895	0.925	0.972	1.009	1.039	1.063	1.106	1.143	1.182	1.222	1.255	1.282	1.288
FC – RANK	0.869	0.886	0.914	0.922	0.955	0.986	1.011	1.044	1.076	1.108	1.156	1.195	1.214	1.237	1.236
FC – MSE	0.832	0.865	0.900	0.929	0.974	1.007	1.033	1.054	1.092	1.125	1.162	1.197	1.227	1.251	1.256
FC – RMSE	0.838	0.871	0.904	0.932	0.976	1.009	1.035	1.056	1.093	1.127	1.164	1.199	1.230	1.255	1.260
FC – OLS	**0.617****	**0.621***	**0.639***	**0.661***	**0.706***	**0.749***	**0.784***	**0.811***	**0.863***	**0.904**	**0.937**	**0.967**	**0.991**	**1.009**	**1.015**

Source: Elaborated by the author.

Legend: The table reports relative root mean squared forecast errors (RMSFE) relative to the DNS model obtained by using individual yield models and different forecast combination methods, for the 1-month, 3-month, 6-month, and 12-month forecast horizons. The evaluation sample is 1994:1 to 2019:12 (312 out-of-sample forecasts). The first line in each panel of the table reports the value of RMSFE (expressed in basis points) for the DNS model, while all other lines reports statistics relative to the DNS. The following model abbreviations are used in the table: DNS for the dynamic Nelson-Siegel model and DNSS for the dynamic Svensson model. Macro refers to macroeconomic factors, GARCH for time varying volatility, and TVL refers to time varying factor loading. FC-EW, FC-OLS and FC-RANK stand for forecast combinations based on equal weights, OLS-based weights, and rank-weighted combinations, respectively. FC-RMSE and FC-MRMSE refer to forecast combinations base on the thick modeling approach with RMSE-weights and MSE-Frequency weights, respectively. Numbers smaller than one indicate that models outperform the DNS, whereas numbers larger than one indicate underperformance. Numbers in **bold** indicate that model is on the model confident set at 25% level. Stars indicate the level at which the Giacomini and White (2006) test rejects the null of equal forecasting accuracy (*, and ** mean respectively rejection at 5%, and 1% level).

Table 5 – Relative Root Mean Squared Forecast Errors: 6-month ahead

Models	3-Month	6-Month	9-Month	12-Month	18-Month	24-Month	30-Month	36-Month	48-Month	60 Month	72-Month	84-Month	96-Month	108-Month	120-Month
DNS	95.11	94.97	95.33	95.75	95.62	94.95	93.69	91.88	89.71	88.22	85.65	82.61	80.49	78.48	76.44
DNS – Macro	0.785	0.815	0.842	0.868	0.918	0.971	1.020	1.065	1.135	1.191	1.237	1.283	1.312	1.334	1.352
DNS – GARCH	1.038	1.054	1.067	1.079	1.100	1.116	1.128	1.138	1.152	1.158	1.161	1.163	1.164	1.163	1.157
DNS – GARCH – Macro	0.785	0.815	0.842	0.867	0.917	0.97	1.020	1.064	1.134	1.191	1.237	1.283	1.312	1.335	1.353
DNS – TVL	0.853	0.871	0.890	0.909	0.952	0.992	1.032	1.067	1.126	1.162	1.189	1.209	1.221	1.229	1.22
DNS – TVL – Macro	**0.764**	0.791	0.815	0.840	0.89	0.937	0.978	1.013	1.072	1.12	1.162	1.207	1.245	1.277	1.304
DNS – GARCH – TVL	**0.742**	0.763	0.785	0.809	0.855	0.897	0.936	0.970	1.027	1.073	1.112	1.147	1.176	1.203	1.223
DNS – GARCH – TVL – Macro	**0.75**	0.778	0.803	0.826	0.872	0.916	0.956	0.991	1.051	1.102	1.146	1.194	1.23	1.261	1.289
DNSS	0.910	0.917	0.928	0.941	0.972	0.998	1.020	1.037	1.070	1.100	1.134	1.164	1.197	1.228	1.242
DNSS – GARCH	1.015	1.033	1.049	1.065	1.096	1.117	1.131	1.140	1.158	1.173	1.195	1.212	1.233	1.254	1.258
DNSS – TVL1	0.951	0.976	1.003	1.027	1.070	1.099	1.120	1.135	1.153	1.166	1.187	1.211	1.240	1.269	1.286
DNSS – TVL2	0.878	0.934	0.980	1.009	1.040	1.051	1.058	1.066	1.095	1.128	1.171	1.212	1.252	1.291	1.310
DNSS – GARCH – TVL1	0.949	0.979	1.011	1.036	1.073	1.092	1.108	1.122	1.147	1.178	1.226	1.277	1.328	1.380	1.420
DNSS – GARCH – TVL2	0.942	0.999	1.037	1.060	1.079	1.084	1.089	1.097	1.133	1.179	1.24	1.301	1.359	1.414	1.450
FC – EW	0.889	0.914	0.937	0.957	0.992	1.020	1.044	1.066	1.105	1.138	1.173	1.207	1.236	1.264	1.281
FC – RANK	0.910	0.921	0.938	0.950	0.979	0.996	1.016	1.037	1.073	1.114	1.143	1.172	1.180	1.214	1.232
FC – MSE	0.897	0.922	0.944	0.962	0.994	1.019	1.040	1.058	1.092	1.121	1.152	1.182	1.209	1.234	1.248
FC – RMSE	0.900	0.924	0.946	0.964	0.995	1.020	1.042	1.06	1.094	1.124	1.155	1.185	1.213	1.238	1.253
FC – OLS	**0.673***	**0.685***	**0.700***	**0.716***	**0.745***	**0.772***	**0.796***	**0.815***	**0.851***	**0.883***	**0.911**	**0.935**	**0.955**	**0.972**	**0.982**

Source: Elaborated by the author.

Legend: The table reports relative root mean squared forecast errors (RMSFE) relative to the DNS model obtained by using individual yield models and different forecast combination methods, for the 1-month, 3-month, 6-month, and 12-month forecast horizons. The evaluation sample is 1994:1 to 2019:12 (312 out-of-sample forecasts). The first line in each panel of the table reports the value of RMSFE (expressed in basis points) for the DNS model, while all other lines reports statistics relative to the DNS. The following model abbreviations are used in the table: DNS for the dynamic Nelson-Siegel model and DNSS for the dynamic Svensson model. Macro refers to macroeconomic factors, GARCH for time varying volatility, and TVL refers to time varying factor loading. FC-EW, FC-OLS and FC-RANK stand for forecast combinations based on equal weights, OLS-based weights, and rank-weighted combinations, respectively. FC-RMSE and FC-MRMSE refer to forecast combinations base on the thick modeling approach with RMSE-weights and MSE-Frequency weights, respectively. Numbers smaller than one indicate that models outperform the DNS, whereas numbers larger than one indicate underperformance. Numbers in **bold** indicate that model is on the model confident set at 25% level. Stars indicate the level at which the Giacomini and White (2006) test rejects the null of equal forecasting accuracy (*, and ** mean respectively rejection at 5%, and 1% level).

Table 6 – Relative Root Mean Squared Forecast Errors: 12-month ahead

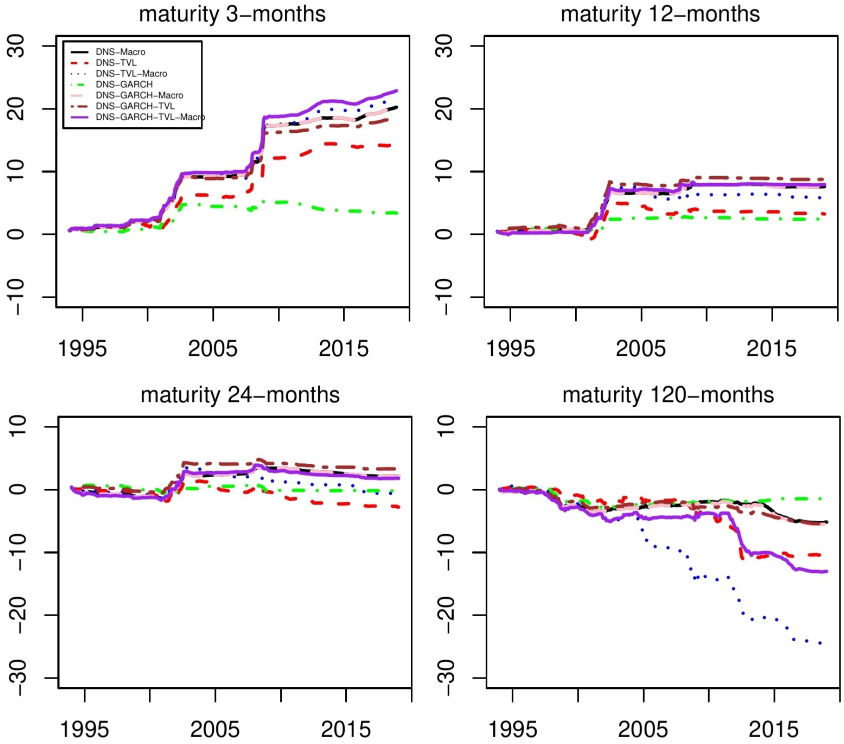
Models	3-Month	6-Month	9-Month	12-Month	18-Month	24-Month	30-Month	36-Month	48-Month	60 Month	72-Month	84-Month	96-Month	108-Month	120-Month
DNS	148.5	149.62	150.18	150.36	149.17	146.71	143.39	139.52	133.67	129.24	124.8	120.3	116.89	113.93	110.59
DNS – Macro	0.910	0.922	0.933	0.945	0.977	1.016	1.058	1.098	1.165	1.224	1.270	1.313	1.341	1.364	1.385
DNS – GARCH	1.067	1.074	1.080	1.085	1.094	1.103	1.110	1.117	1.126	1.130	1.134	1.137	1.138	1.138	1.136
DNS – GARCH – Macro	0.908	0.919	0.931	0.943	0.975	1.015	1.057	1.097	1.164	1.223	1.269	1.312	1.341	1.363	1.386
DNS – TVL	0.898	0.902	0.908	0.915	0.934	0.958	0.985	1.012	1.058	1.093	1.117	1.136	1.145	1.150	1.149
DNS – TVL – Macro	**0.887**	0.896	0.906	0.917	0.943	0.973	1.004	1.034	1.084	1.131	1.172	1.214	1.246	1.274	1.301
DNS – GARCH – TVL	**0.815**	0.828	0.842	0.855	0.882	0.911	0.939	0.967	1.016	1.060	1.097	1.130	1.158	1.183	1.205
DNS – GARCH – TVL – Macro	**0.877**	0.886	0.895	0.905	0.929	0.959	0.990	1.021	1.073	1.123	1.165	1.207	1.238	1.265	1.293
DNSS	0.982	0.990	1.000	1.009	1.028	1.045	1.060	1.073	1.098	1.120	1.143	1.166	1.188	1.209	1.222
DNSS – GARCH	1.046	1.060	1.073	1.083	1.102	1.115	1.125	1.133	1.147	1.158	1.170	1.184	1.199	1.215	1.220
DNSS – TVL1	1.020	1.041	1.062	1.079	1.104	1.119	1.128	1.134	1.141	1.147	1.161	1.178	1.198	1.219	1.235
DNSS – TVL2	0.949	0.988	1.017	1.035	1.050	1.053	1.055	1.059	1.074	1.094	1.121	1.149	1.177	1.203	1.220
DNSS – GARCH – TVL1	1.005	1.021	1.038	1.054	1.080	1.102	1.121	1.139	1.173	1.206	1.244	1.279	1.315	1.349	1.378
DNSS – GARCH – TVL2	1.012	1.043	1.064	1.076	1.087	1.093	1.101	1.114	1.147	1.186	1.231	1.275	1.318	1.355	1.384
FC – EW	0.958	0.972	0.985	0.996	1.016	1.035	1.054	1.073	1.106	1.137	1.166	1.194	1.218	1.239	1.256
FC – RANK	0.952	0.974	0.980	0.994	1.001	1.003	1.018	1.040	1.082	1.112	1.130	1.154	1.178	1.190	1.207
FC – MSE	0.963	0.974	0.984	0.994	1.011	1.028	1.045	1.062	1.091	1.118	1.144	1.169	1.191	1.210	1.225
FC – RMSE	0.965	0.976	0.987	0.997	1.015	1.031	1.048	1.064	1.093	1.120	1.146	1.172	1.194	1.213	1.229
FC – OLS	$\mathbf{0.717}^{**}$	$\mathbf{0.725}^{**}$	$\mathbf{0.733}^{**}$	$\mathbf{0.741}^{**}$	$\mathbf{0.755}^{**}$	$\mathbf{0.771}^{**}$	$\mathbf{0.787}^{*}$	$\mathbf{0.803}^{*}$	$\mathbf{0.837}^{*}$	$\mathbf{0.870}^{*}$	**0.900**	**0.928**	**0.950**	**0.968**	**0.983**

Source: Elaborated by the author.

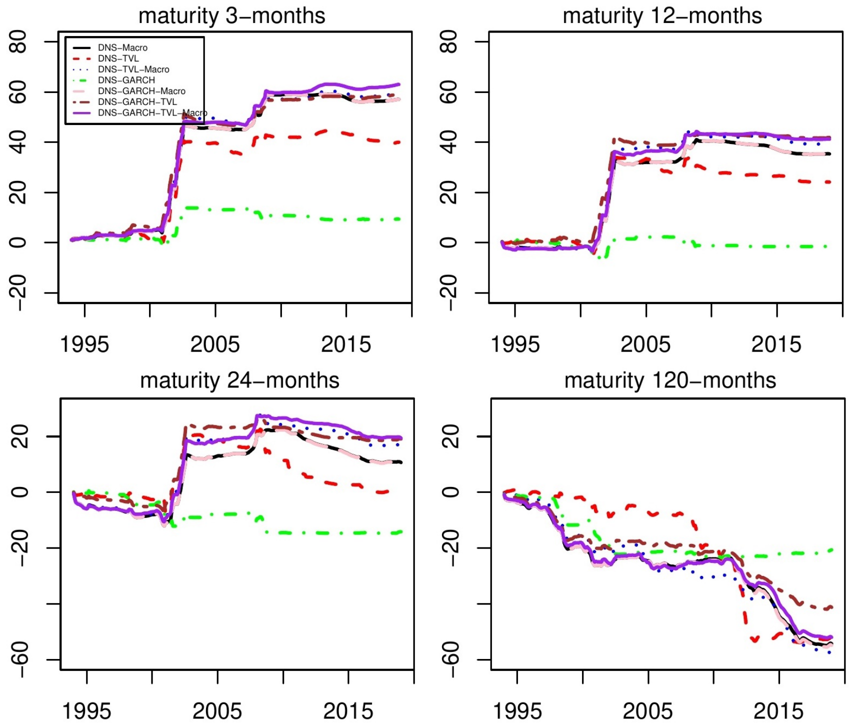
Legend: The table reports relative root mean squared forecast errors (RMSFE) relative to the DNS model obtained by using individual yield models and different forecast combination methods, for the 1-month, 3-month, 6-month, and 12-month forecast horizons. The evaluation sample is 1994:1 to 2019:12 (312 out-of-sample forecasts). The first line in each panel of the table reports the value of RMSFE (expressed in basis points) for the DNS model, while all other lines reports statistics relative to the DNS. The following model abbreviations are used in the table: DNS for the dynamic Nelson-Siegel model and DNSS for the dynamic Svensson model. Macro refers to macroeconomic factors, GARCH for time varying volatility, and TVL refers to time varying factor loading. FC-EW, FC-OLS and FC-RANK stand for forecast combinations based on equal weights, OLS-based weights, and rank-weighted combinations, respectively. FC-RMSE and FC-MRMSE refer to forecast combinations base on the thick modeling approach with RMSE-weights and MSE-Frequency weights, respectively. Numbers smaller than one indicate that models outperform the DNS, whereas numbers larger than one indicate underperformance. Numbers in **bold** indicate that model is on the model confident set at 25% level. Stars indicate the level at which the Giacomini and White (2006) test rejects the null of equal forecasting accuracy (*, and ** mean respectively rejection at 5%, and 1% level).

Figure 6 – Cumulative squared forecast errors (Nelson-Siegel Extensions)

(a) One-month ahead



(b) Three-month ahead

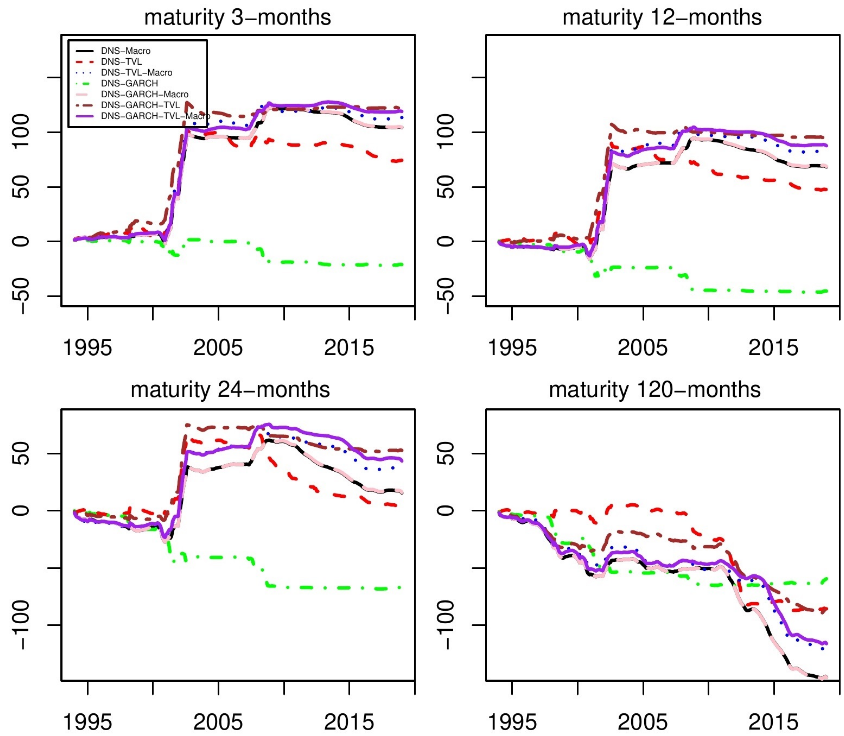


Source: Elaborated by the author.

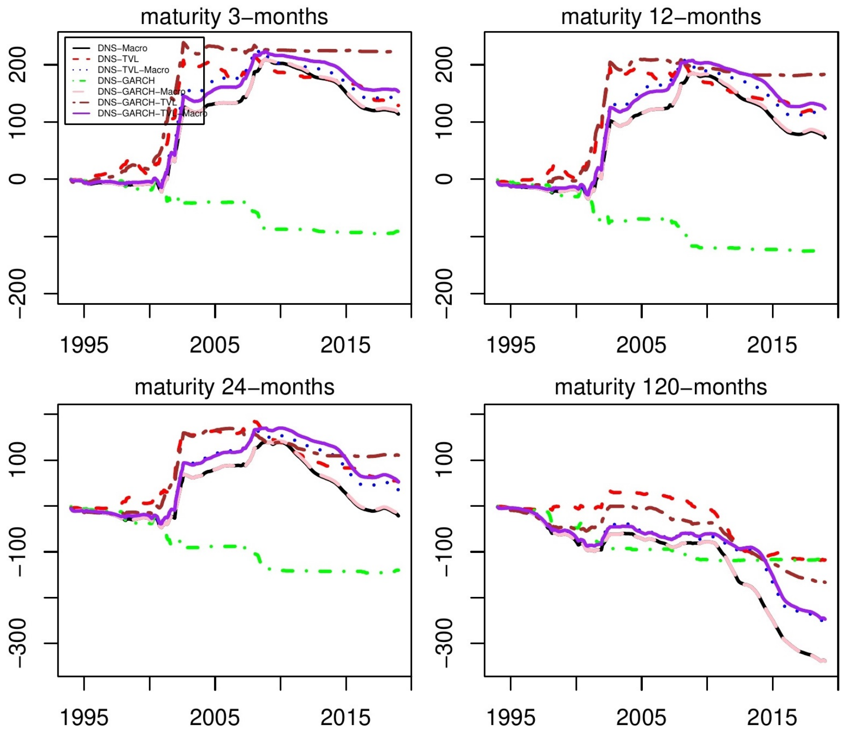
Legend: Figures show the cumulative squared forecast errors (CSFE) of Nelson-Siegel Extensions relative to the DNS baseline model. Figure shows CSFEs for a 1- and 3-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (312 out-of-sample forecasts).

Figure 7 – Cumulative squared forecast errors (Nelson-Siegel Extensions)

(a) Six-month ahead



(b) Twelve-month ahead

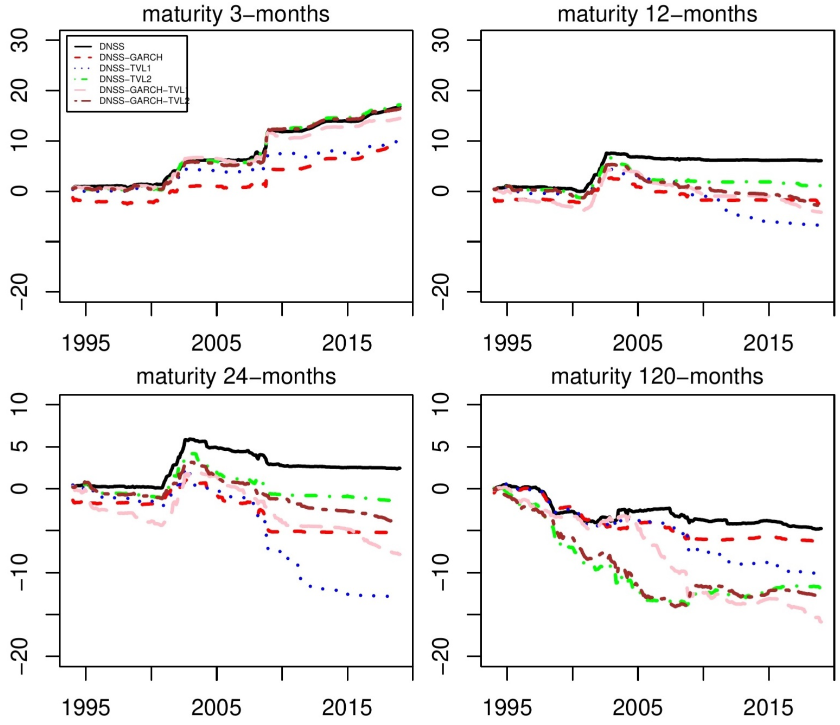


Source: Elaborated by the author.

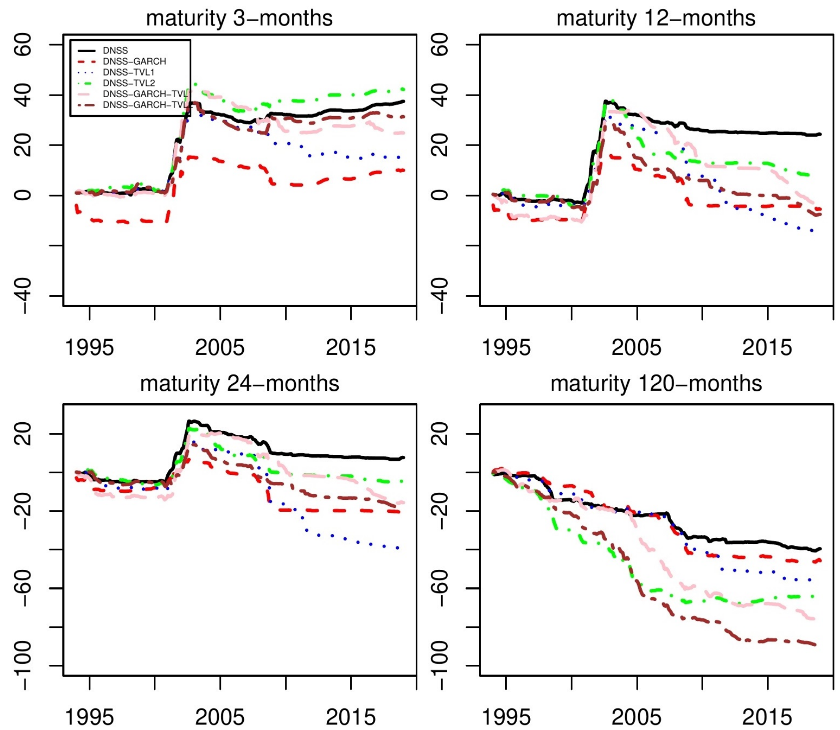
Legend: Figures show the cumulative squared forecast errors (CSFE) of Nelson-Siegel Extensions relative to the DNS baseline model. Figure shows CSFEs for a 6- and 12-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (312 out-of-sample forecasts).

Figure 8 – Cumulative squared forecast errors (Svensson Extensions)

(a) One-month ahead



(b) Three-month ahead

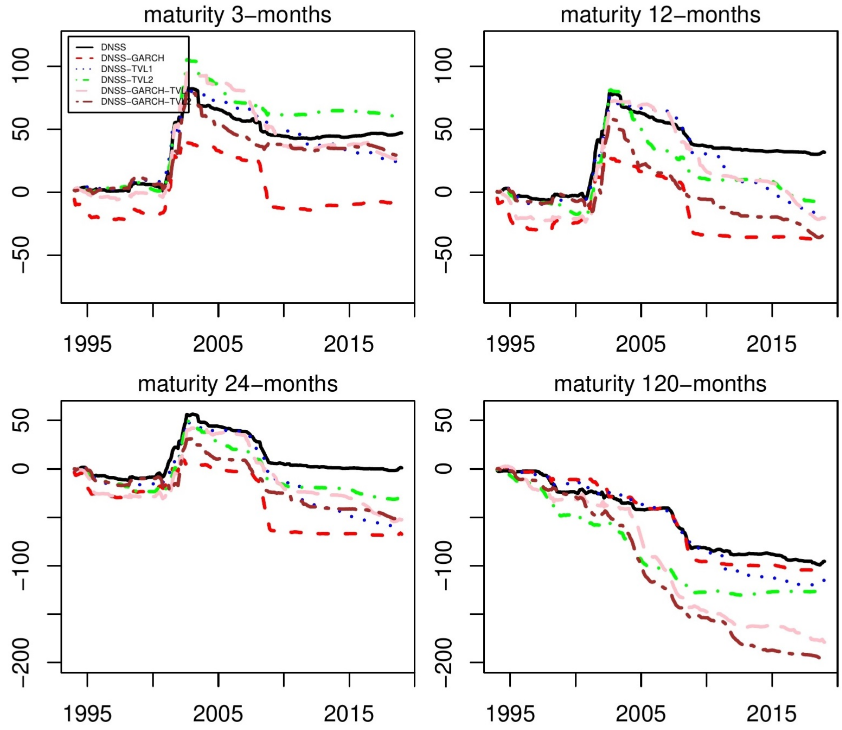


Source: Elaborated by the author.

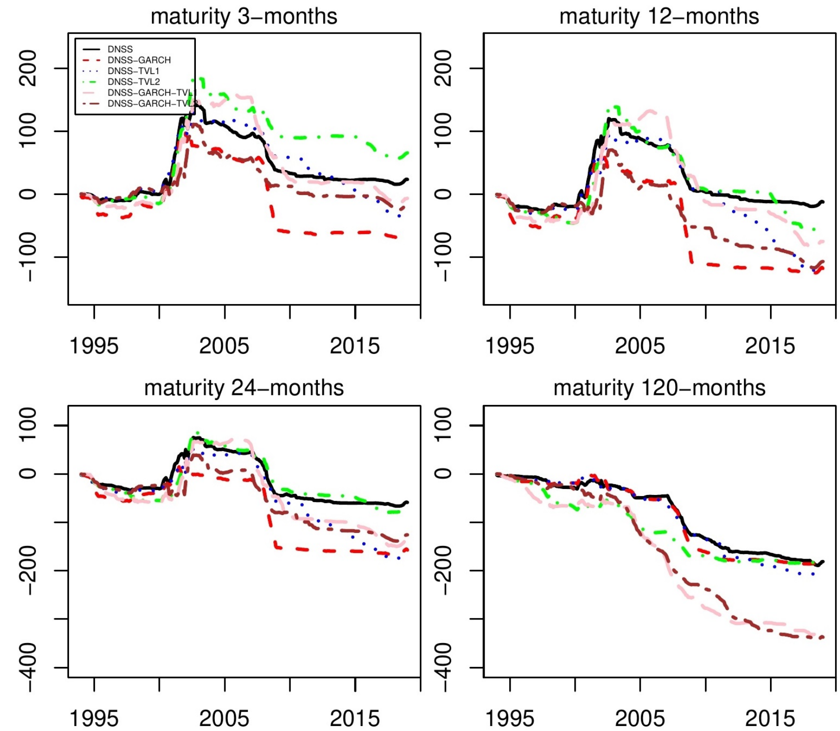
Legend: Figures show the cumulative squared forecast errors (CSFE) of Svensson Extensions relative to the DNS baseline model. Figure shows CSFEs for a 1- and 3-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (312 out-of-sample forecasts).

Figure 9 – Cumulative squared forecast errors (Svensson Extensions)

(a) Six-month ahead



(b) Twelve-month ahead

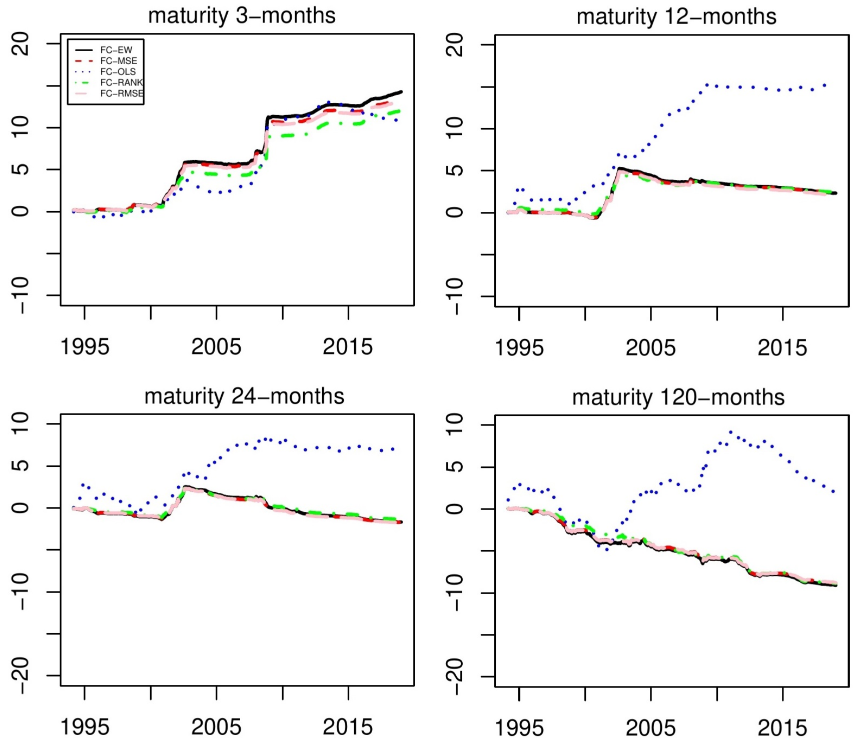


Source: Elaborated by the author.

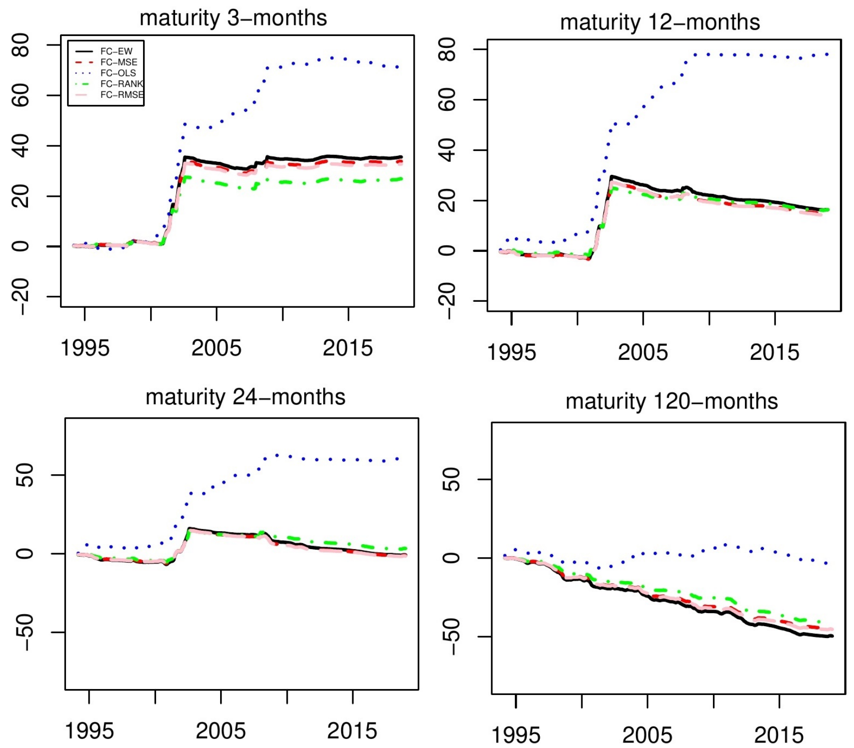
Legend: Figures show the cumulative squared forecast errors (CSFE) of Svensson Extensions relative to the DNS baseline model. Figure shows CSFEs for a 1- and 3-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (312 out-of-sample forecasts).

Figure 10 – Cumulative squared forecast errors (Forecasting Combination)

(a) One-month ahead



(b) Three-month ahead

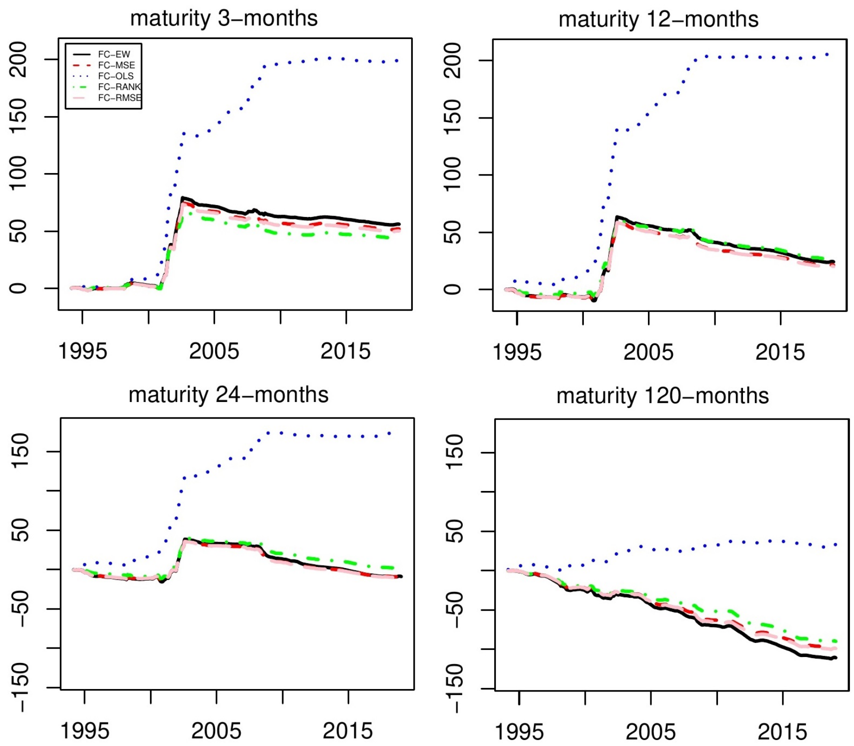


Source: Elaborated by the author.

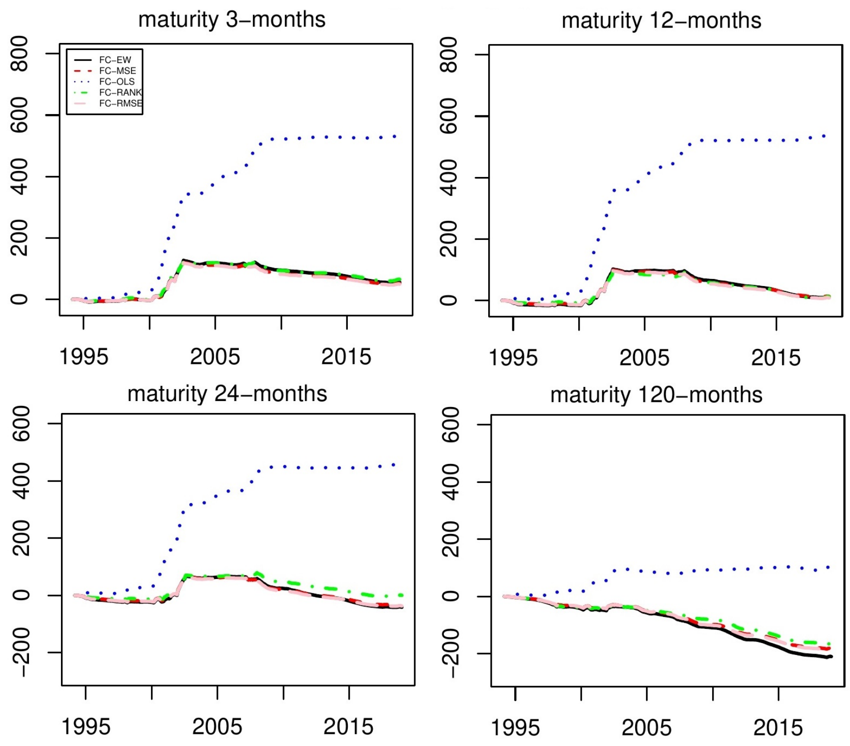
Legend: Figures show the cumulative squared forecast errors (CSFE) of Svensson Extensions relative to the DNS baseline model. Figure shows CSFEs for a 1- and 3-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (300 out-of-sample forecasts).

Figure 11 – Cumulative squared forecast errors (Forecasting Combination)

(a) Six-month ahead



(b) Twelve-month ahead



Source: Elaborated by the author.

Legend: Figures show the cumulative squared forecast errors (CSFE) of Svensson Extensions relative to the DNS baseline model. Figure shows CSFEs for a 1- and 3-month forecast horizon. The evaluation sample is from January 1994:01 through December 2019:12 (300 out-of-sample forecasts).

2.2.3.1 Economic Evaluation Results

In the previous subsection, we showed that alternative specifications of individual prediction models and forecast combination schemes could deliver more accurate forecasts concerning the benchmark when considering statistical criteria. We observe, however, that in some instances, the improvement in forecasting performance (as indicated by lower forecasting errors) is small in magnitude. Therefore, a question that remains unanswered is whether or not this statistical gain is also economically meaningful.

Table 7 to Table 9 reports annualized certainty equivalent (CER) for a mean-variance investor with $\delta = 0.1, 1,$ and 5, who allocates among τ periods to maturity risky bonds and one-month T-bill that pays the risky free rate using forecasts based on competitors models in place of dynamic Nelson-Siegel baseline model forecasts. Therefore, positive entries indicate that the alternative models perform better than DNS baseline model. For analyzed scenarios, the CER values generally decrease with the bond maturity. The highest CER values are generally obtained for the DNS-GARCH-Macro and DNSS-TLV1 models. These models are the only individual ones that generate positive CER values for all risk levels and maturities considered for time forecast horizons of 1- and 3-months. With $\delta = 5$, the DNSS-TVL1 model is best for roughly half the models for longest maturities ($\tau = 3, 5, 10$ years), while the DNS-GARCH-Macro model accomplishes a similar level of performance for $\tau = 1$. In summary, when we consider individual models, the DNS-GARCH-Macro performs better 21 out of 48 times, while the DNSS-TVL1 outperforms all other competitors by 22 out of 48 times.

There is a vast empirical literature with strong evidence of statistical predictability of the yield curve, but there is so scarce evidence that the predictability of the interest rates could be translated into economic utility gains in real time by investors. Comparing our results with other studies, our finds are quite different from these reported by Thornton and Valente (2012) and Sarno, Paul Schneider, and Wagner (2016). Specifically, these studies test for out of sample forecasting of bond excess returns and find that predictive models based on forward rates do not generate significant economic value to investors. In other words, the statistical evidence of out-of-sample predictability fails to translate into an ability for investors to use bond excess returns forecasts in a way that generates higher out-of-sample certainty equivalent than forecasts from the benchmark model. More recently, Gargano, Pettenuzzo, and Timmermann (2019) and Daniele Bianchi et al. (2021) report that models having bond return predictors based on macro variables, forward rates or yields can generate significant economic value to investors. As a matter of fact, our setup is different from these as they consider excess returns, not yields.

We conclude that there is strong statistical and economic evidence that the returns on 1-10 years bonds can be exploited using dynamic factor models proposed in the literature. Moreover, the best performing models allow for time-varying decay parameter volatility dynamics.

Table 7-Table 9 also reports annualized certainty equivalent CER for 5 forecast combination schemes. We observe that in the vast majority of the instances, the mean-variance allocations obtained with forecast combinations achieve statistically higher CER in comparison to the mean-variance allocations obtained with the individual models. Moreover, this result is robust to the value of the risk aversion coefficient and to the maturity spectrum. Interestingly, the forecast combination best performance seems to be concentrated in shortest forecast horizon, with positive CER values for all considered risk aversion level and maturities. Notably, whereas we find that CER values decrease with maturity for one-month ahead forecasting, when we look at the longer forecasting horizons the CER values increase for the longest bond maturity.

One of the most important lessons from the economic evaluation of forecasts is that the differences in performance in terms of CER are much more pronounced that those based on statistical measures. In this sense, it seems to be much easier to distinguish between “good” and “bad” predictions when looking at economic criteria. As we noted earlier, differences in statistical performance measures are usually small in magnitude. In contrast, differences in CER’s tend to be much more evident. In particular, we observe that in several cases the CER’s obtained by some forecast combination schemes is twice as the ones obtained by some individual models.

Table 7 – Performance evaluation for an Investor with Mean Variance Utility.

Models	1-month ahead				3-month ahead			
	1-Year	3-Year	5-Year	10-Year	1-Year	3-Year	5-Year	10-Year
Individual Models								
DNS – Macro	45.17	2.799	–0.592	–0.402	6.744	–0.658	–1.037	0.485
DNS – GARCH	91.32	11.40	3.882	1.055	10.24	0.463	0.436	0.634
DNS – GARCH – Macro	146.59	24.94	9.585	4.547	9.488	1.448	0.210	0.844
DNS – TVL	–133.50	–3.87	–1.511	–2.841	–40.35	–2.393	–0.719	–0.711
DNS – TVL – Macro	–133.49	–3.916	–1.603	–2.933	–40.36	–2.487	–0.937	–0.874
DNS – GARCH – TVL	–133.46	–3.882	–1.541	–2.886	–40.34	–2.414	–0.821	–0.920
DNS – GARCH – TVL – Macro	–133.40	–3.916	–1.603	–2.933	–40.35	–2.487	–0.937	–0.874
DNSS	25.286	6.171	1.265	–1.434	5.485	2.280	0.894	0.290
DNSS – GARCH	–133.41	–3.868	–1.508	–2.889	–40.35	–2.394	–0.742	–0.788
DNSS – TVL1	122.33	25.96	7.013	–1.031	7.821	6.329	3.626	1.714
DNSS – TVL2	34.69	9.278	1.895	–2.012	6.503	3.734	1.077	–0.016
DNSS – GARCH – TVL1	–133.43	–3.872	–1.529	–2.949	–40.35	–2.409	–0.819	–0.964
DNSS – GARCH – TVL2	–133.48	–3.863	–1.51	–2.931	–40.34	–2.388	–0.764	–0.874
Forecast Combination								
FC – EW	445.26	67.55	34.06	17.04	–1.498	2.822	4.308	3.82
FC – RANK	447.73	66.88	34.02	16.56	–1.705	2.684	3.660	3.397
FC – MSE	443.96	67.94	34.51	17.00	–1.487	3.16	4.524	3.767
FC – RMSE	443.42	67.93	34.49	17.07	–1.636	3.163	4.535	3.828
FC – OLS	443.69	56.75	26.37	7.827	0.889	1.207	2.984	2.920
	6-month ahead				12-month ahead			
Individual Models								
DNS – Macro	1.034	–1.083	–0.901	1.577	–0.141	–1.416	–0.249	4.534
DNS – GARCH	1.664	–1.262	0.156	1.082	0.064	–1.440	–0.831	1.250
DNS – GARCH – Macro	0.435	0.258	0.882	2.214	–0.136	0.275	4.407	9.118
DNS – TVL	–17.24	–1.432	–0.285	–0.214	–7.344	0.179	0.700	0.452
DNS – TVL – Macro	–17.23	–1.557	–0.602	–0.281	–7.320	0.107	0.433	1.142
DNS – GARCH – TVL	–17.22	–1.459	–0.484	–0.741	–7.292	0.216	0.494	–0.34
DNS – GARCH – TVL – Macro	–17.25	–1.557	–0.602	–0.281	–7.320	0.107	0.433	1.142
DNSS	1.197	0.653	0.695	1.281	0.034	0.274	0.772	1.859
DNSS – GARCH	–17.22	–1.426	–0.334	–0.347	–7.297	0.264	0.724	0.248
DNSS – TVL1	1.220	2.889	1.96	2.097	0.101	1.224	1.453	1.838
DNSS – TVL2	1.615	1.781	1.005	0.827	0.018	0.514	0.588	0.809
DNSS – GARCH – TVL1	–17.23	–1.451	–0.44	–0.548	–7.315	0.230	0.637	0.177
DNSS – GARCH – TVL2	–17.22	–1.435	–0.394	–0.514	–7.317	0.211	0.617	0.084
Forecast Combination								
FC – EW	–4.404	6.807	18.15	13.36	–0.852	10.95	20.68	27.80
FC – RANK	–3.717	5.632	19.38	10.75	–0.904	10.30	17.32	23.87
FC – MSE	–4.312	7.489	18.61	12.15	–0.859	11.41	21.08	26.25
FC – RMSE	–4.359	7.565	18.72	12.60	–0.858	11.56	21.40	26.51
FC – OLS	–4.262	7.106	16.68	5.970	–0.823	10.29	19.30	20.02

Source: Elaborated by the author.

Legend: This table presents the average utility gain (Δ) in the portfolio management fee (in annualized percent return) that an investor with mean-variance preferences and risk aversion coefficient of $\delta = 0.1$ would be willing to pay to have access to the DNS forecasts.

Table 8 – Performance evaluation for an Investor with Mean Variance Utility.

Models	1-month ahead				3-month ahead			
	1-Year	3-Year	5-Year	10-Year	1-Year	3-Year	5-Year	10-Year
Individual Models								
DNS – Macro	4.517	0.280	–0.059	–0.04	0.674	–0.066	–0.104	0.048
DNS – GARCH	9.132	1.139	0.388	0.105	1.024	0.046	0.044	0.063
DNS – GARCH – Macro	14.66	2.494	0.958	0.455	0.949	0.145	0.021	0.084
DNS – TVL	–13.35	–0.387	–0.151	–0.284	–4.035	–0.239	–0.072	–0.071
DNS – TVL – Macro	–13.35	–0.392	–0.16	–0.293	–4.036	–0.249	–0.094	–0.087
DNS – GARCH – TVL	–13.34	–0.388	–0.154	–0.289	–4.034	–0.241	–0.082	–0.092
DNS – GARCH – TVL – Macro	–13.35	–0.392	–0.160	–0.293	–4.036	–0.249	–0.094	–0.087
DNSS	2.529	0.617	0.126	–0.143	0.548	0.228	0.089	0.029
DNSS – GARCH	–13.35	–0.387	–0.151	–0.289	–4.035	–0.239	–0.074	–0.079
DNSS – TVL1	12.23	2.596	0.701	–0.103	0.782	0.633	0.363	0.171
DNSS – TVL2	3.469	0.928	0.189	–0.201	0.650	0.373	0.108	–0.002
DNSS – GARCH – TVL1	–13.35	–0.387	–0.153	–0.295	–4.035	–0.241	–0.082	–0.096
DNSS – GARCH – TVL2	–13.34	–0.386	–0.151	–0.293	–4.034	–0.239	–0.076	–0.087
Foreacst Combination								
FC – EW	44.53	6.755	3.406	1.704	–0.150	0.282	0.431	0.382
FC – RANK	44.77	6.688	3.402	1.656	–0.171	0.268	0.366	0.34
FC – MSE	44.40	6.794	3.451	1.700	–0.149	0.316	0.452	0.377
FC – RMSE	44.34	6.793	3.450	1.707	–0.164	0.316	0.453	0.383
FC – OLS	44.37	5.675	2.637	0.783	0.089	0.121	0.298	0.292
	6-month ahead				12-month ahead			
Individual Models								
DNS – Macro	0.103	–0.108	–0.09	0.158	–0.014	–0.142	–0.025	0.453
DNS – GARCH	0.166	–0.126	0.016	0.108	0.006	–0.144	–0.083	0.125
DNS – GARCH – Macro	0.044	0.026	0.088	0.221	–0.014	0.027	0.441	0.912
DNS – TVL	–1.724	–0.143	–0.028	–0.021	–0.734	0.018	0.070	0.045
DNS – TVL – Macro	–1.723	–0.156	–0.06	–0.028	–0.732	0.011	0.043	0.114
DNS – GARCH – TVL	–1.721	–0.146	–0.048	–0.074	–0.729	0.022	0.049	–0.034
DNS – GARCH – TVL – Macro	–1.723	–0.156	–0.06	–0.028	–0.732	0.011	0.043	0.114
DNSS	0.120	0.065	0.069	0.128	0.003	0.027	0.077	0.186
DNSS – GARCH	–1.723	–0.143	–0.033	–0.035	–0.73	0.026	0.072	0.025
DNSS – TVL1	0.122	0.289	0.196	0.210	0.010	0.122	0.145	0.184
DNSS – TVL2	0.161	0.178	0.100	0.083	0.002	0.051	0.059	0.081
DNSS – GARCH – TVL1	–1.723	–0.145	–0.044	–0.055	–0.731	0.023	0.064	0.018
DNSS – GARCH – TVL2	–1.722	–0.143	–0.039	–0.051	–0.732	0.021	0.062	0.008
Foreacst Combination								
FC – EW	–0.44	0.681	1.815	1.336	–0.085	1.095	2.067	2.780
FC – RANK	–0.372	0.563	1.938	1.075	–0.090	1.03	1.732	2.387
FC – MSE	–0.431	0.749	1.861	1.215	–0.086	1.141	2.108	2.625
FC – RMSE	–0.436	0.757	1.872	1.260	–0.086	1.156	2.140	2.651
FC – OLS	–0.426	0.711	1.668	0.597	–0.082	1.029	1.930	2.002

Source: Elaborated by the author.

Legend: This table presents the average utility gain (Δ) in the portfolio management fee (in annualized percent return) that an investor with mean-variance preferences and risk aversion coefficient of δ = 1 would be willing to pay to have access to the DNS forecasts.

Table 9 – Performance evaluation for an Investor with Mean Variance Utility.

Models	1-month ahead				3-month ahead			
	1-Year	3-Year	5-Year	10-Year	1-Year	3-Year	5-Year	10-Year
Individual Models								
DNS – Macro	0.903	0.056	–0.012	–0.008	0.135	–0.013	–0.021	0.01
DNS – GARCH	1.826	0.228	0.078	0.021	0.205	0.009	0.009	0.013
DNS – GARCH – Macro	2.932	0.499	0.192	0.091	0.19	0.029	0.004	0.017
DNS – TVL	–2.670	–0.077	–0.03	–0.057	–0.807	–0.048	–0.014	–0.014
DNS – TVL – Macro	–2.670	–0.078	–0.032	–0.059	–0.807	–0.05	–0.019	–0.017
DNS – GARCH – TVL	–2.670	–0.078	–0.031	–0.058	–0.807	–0.048	–0.016	–0.018
DNS – GARCH – TVL – Macro	–2.670	–0.078	–0.032	–0.059	–0.807	–0.05	–0.019	–0.017
DNSS	0.506	0.123	0.025	–0.029	0.110	0.046	0.018	0.006
DNSS – GARCH	–2.670	–0.077	–0.03	–0.058	–0.807	–0.048	–0.015	–0.016
DNSS – TVL1	2.447	0.519	0.140	–0.021	0.156	0.127	0.073	0.034
DNSS – TVL2	0.694	0.186	0.038	–0.040	0.130	0.075	0.022	0.000
DNSS – GARCH – TVL1	–2.670	–0.077	–0.031	–0.059	–0.807	–0.048	–0.016	–0.019
DNSS – GARCH – TVL2	–2.670	–0.077	–0.03	–0.059	–0.807	–0.048	–0.015	–0.017
Foreacst Combination								
FC – EW	8.905	1.351	0.681	0.341	–0.03	0.056	0.086	0.076
FC – RANK	8.955	1.338	0.68	0.331	–0.034	0.054	0.073	0.068
FC – MSE	8.879	1.359	0.69	0.340	–0.03	0.063	0.09	0.075
FC – RMSE	8.868	1.359	0.69	0.341	–0.033	0.063	0.091	0.077
FC – OLS	8.874	1.135	0.527	0.157	0.018	0.024	0.060	0.058
	6-month ahead				12-month ahead			
Individual Models								
DNS – Macro	0.021	–0.022	–0.018	0.032	–0.003	–0.028	–0.005	0.091
DNS – GARCH	0.033	–0.025	0.003	0.022	0.001	–0.029	–0.017	0.025
DNS – GARCH – Macro	0.009	0.005	0.018	0.044	–0.003	0.005	0.088	0.182
DNS – TVL	–0.345	–0.029	–0.006	–0.004	–0.147	0.004	0.014	0.009
DNS – TVL – Macro	–0.345	–0.031	–0.012	–0.006	–0.146	0.002	0.009	0.023
DNS – GARCH – TVL	–0.344	–0.029	–0.01	–0.015	–0.146	0.004	0.010	–0.007
DNS – GARCH – TVL – Macro	–0.345	–0.031	–0.012	–0.006	–0.146	0.002	0.009	0.023
DNSS	0.024	0.013	0.014	0.026	0.001	0.005	0.015	0.037
DNSS – GARCH	–0.345	–0.029	–0.007	–0.007	–0.146	0.005	0.014	0.005
DNSS – TVL1	0.024	0.058	0.039	0.042	0.002	0.024	0.029	0.037
DNSS – TVL2	0.032	0.036	0.020	0.017	0.000	0.010	0.012	0.016
DNSS – GARCH – TVL1	–0.345	–0.029	–0.009	–0.011	–0.146	0.005	0.013	0.004
DNSS – GARCH – TVL2	–0.344	–0.029	–0.008	–0.010	–0.146	0.004	0.012	0.002
Foreacst Combination								
FC – EW	–0.088	0.136	0.363	0.267	–0.017	0.219	0.413	0.556
FC – RANK	–0.074	0.113	0.388	0.215	–0.018	0.206	0.346	0.477
FC – MSE	–0.086	0.15	0.372	0.243	–0.017	0.228	0.422	0.525
FC – RMSE	–0.087	0.151	0.374	0.252	–0.017	0.231	0.428	0.530
FC – OLS	–0.085	0.142	0.334	0.119	–0.016	0.206	0.386	0.400

Source: Elaborated by the author.

Legend: This table presents the average utility gain (Δ) in the portfolio management fee (in annualized percent return) that an investor with mean-variance preferences and risk aversion coefficient of δ = 5 would be willing to pay to have access to the DNS forecasts.

2.3 CONCLUDING REMARKS

This section summarises the findings and contributions made. The study analyzed the fit and forecasting of DNSS model extensions with time-varying parameters in linear and nonlinear environments using the Kalman filter and Extended Kalman filter, respectively. The extensions incorporate the time-varying volatility (h_t) and the time-varying loading parameter (λ_t). Specifically, DNS-GARCH and DNSS-GARCH are extensions only with time-varying volatility, DNS-TVL, DNSS-TVL1, and DNSS-TVL2 are extensions only with time-varying loading parameter; and DNS-GARCH-TVL, DNSS-GARCH-TVL1, and DNSS-GARCH-TVL2 are extensions with h_t and $\lambda_{p,t}$ varying over time. The DNS and DNSS models are considered base models, and the other models are considered competitors.

The models with time-varying volatility had better in-sample fit against the DNS model according to RMSE, especially in short maturities. Competing models have satisfactory results, despite adding more parameters (19 and more 32, respectively). The extensions with time-varying loading had not better fit in all maturities. These results might be due to these series approximations used in the extended Kalman filter, which could be poor representations of nonlinear equations and probability distributions involved in the process, as discussed in Wan and Van Der Merwe (2000) and Doucet, De Freitas, and Gordon (2001). Considering this, DNS and DNSS models have new extensions with better performance in fit and forecast the yield curve. Therefore, the analysis results found clear support for the inclusion of time-varying volatility and time-varying loading parameters.

The fit with the DNSS-GARCH and DNS-GARCH models outperform the DNS benchmark model in terms of RMSE, reached a similar conclusion Koopman, Mallee, and Van der Wel (2010) for time-varying volatility. The DNS-TVL, DNS-TVL-Macro, DNS-GARCH-TVL, DNS-GARCH-TVL-Macro, and FC-OLS models had statistically significant out-of-sample results according to RMSPE, Giacomini and White test and Model Confidence set. One can note that FC-OLS also outperforms the benchmark model across most maturities and horizons forecasts. In general, almost all Svensson extensions models outperform the benchmark model in the maturity of 3 months for 1-month ahead forecasts. We found similar conclusions to economic performance; however, modeling volatility added value to the extensions.

The limitations of the present studies naturally include both loading parameters varying overtime in the DNSS model. Further research, such as the inclusion of forward-looking macroeconomic factors, as per Vieira *et al.* (2017) and Fernandes & Vieira (2019), can be implemented in this context of time-varying parameters. Other models for capturing volatility dynamics, such as the GJR-GARCH in Glosten, Jagannathan, and Runkle (1993) can be added to the DNS and DNSS models to investigate whether there are improvements in fit and forecast. The state equation configuration, as in

De Pooter (2007), can be changed with autoregressive diagonal parameter matrix or random walk with diagonal equal to 1, decreasing the number of parameters in the estimation.

The fit forecasting may have more robust results with the Iterated Kalman filter, in which successive iterations can improve the linearization in the Taylor approximation. Moreover, the implementation of another Kalman filter for nonlinear models, such as the Unscented Kalman filter (UKF) and Extended Kalman filter with Singular Value Decomposition (EKF) (see Julier and Uhlmann (1997), Bierman (1977), and YM Zhang, GZ Dai, HC Zhang, et al. (1996)), can be used since one of the problems of EKF is the first-order approximation and the calculation of the gradients, in addition to Bayesian inference using the Monte Carlo Markov Chain (MCMC) method to capture stochastic volatility and time-varying loading parameters.

3 FORECASTING THE YIELD CURVES USING MACROECONOMICS EXPECTATIONS AND TIME-VARYING VOLATILITY

The macro-term structure literature in the last two decades has established a stronger link between macroeconomic information and yield curve dynamics. The related literature includes several papers.[1] That advances shed light on two prominent dynamic and latent factor models: affine no-arbitrage term structure and the Nelson and Siegel (1987) models (henceforth DNS). Ang and Piazzesi (2003) incorporated macroeconomic variables under no-arbitrage restrictions and showed that models with macro factors inflation and real activity from principal components outperform models with only unobservable factors. Despite not considering out-of-sample forecasting, Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) (henceforth DRA) find evidence that real activity, inflation, and the monetary policy instrument have a statistically significant effect on yields. Several academic papers also explore surveys and broad macroeconomic information to link them with interest rates.[2] Altavilla, Giacomini, and Ragusa (2017), for instance, use relative entropy to tilt the segments of the yield curve forecasts from the term structure models to match survey expectations and find that anchoring at three-month horizon improves the accuracy of entire yield curve forecasts.

Moench (2008) jointly modeled the dynamics of macroeconomic variables and government bond yields in a dynamic factor model. He finds that information embedded in the macro factors helps provide out-of-sample yield forecasts that outperform the benchmark at intermediate and long horizons and for short and medium-term maturities. De Pooter, Ravazzolo, and Dick JC Van Dijk (2010) compare the forecast performance of several individual term structure models, they suggest that adding macroeconomic information improves interest rate forecasts, especially in and around recession periods. Fernandes and Vieira (2019) reveals that employing financial and macro information to build factors based on high-frequency forward-looking series in a factor-augmented DNS model can improve the predictive performance.

The gains obtained by including macroeconomic information depend on how macro information is incorporated in the model, as argued by Exterkate et al. (2013). Also, they suggest it is useful only for forecasting yields of medium-term maturities (between 1 and 5 years) and that factor-augmented methods perform well in relatively volatile periods, including the crisis period in 2008-2009 when simpler models do not

[1] Fleming and Remolona (2001), Piazzesi (2001, 2005), Qiang Dai and Philippon (2005), Francis X Diebold, Piazzesi, and Glenn D Rudebusch (2005), Gallmeyer, Hollifield, and Zin (2005), Hördahl, Tristani, and Vestin (2006), Ang, Piazzesi, and Wei (2006), Duffee (2006), Dewachter and Lyrio (2006), Dewachter, Lyrio, and Maes (2006), Glenn D Rudebusch and Tao Wu (2007, 2008), Bikbov and Chernov (2010), Bekaert, Cho, and Moreno (2010), Duffee (2011), Gürkaynak and Wright (2012) and Joslin, Priebsch, and Kenneth J Singleton (2014).

[2] Ludvigson and Serena Ng (2009), Piazzesi, Martin Schneider, et al. (2009), Stark et al. (2010), Kim and Orphanides (2012), Chernov and Mueller (2012), Orphanides and Wei (2012), Ehling et al. (2018), Chun (2011), Favero, Niu, and Sala (2012), Chun (2012), Dick Van Dijk et al. (2014) and Altavilla, Giacomini, and Costantini (2014).

suffice. Koopman and Wel (2013) also extending the class of dynamic factor yield curve models perform an out-of-sample forecasting study, and their results suggest that macroeconomic variables can lead to more accurate yield curve forecasts.[3]

The evidence exploring the links between macroeconomic variables and yield curve modeling for Brazil is not a novelty, notwithstanding the scarcity. Almeida and Faria (2014), for instance, evaluates the term structure forecasting as per Moench (2008) using common factors from macroeconomic series from January 2000 to May 2012. Their results suggest better predictive performance compared to the usual benchmarks but presented deterioration of the results with increased maturity. Also, by eliminating the no-arbitrage restrictions, they produced superior forecasting results. Vieira, Fernandes, and Chague (2017) show that the inclusion of forward-looking data set principal components improves the predictive ability of the factor-augmented VAR methodology with the Nelson-Siegel in out-of-sample analysis to the Brazilian term structure of interest rates. Andrade Alves, Abraham, and Marcio Poletti Laurini (2023) investigates whether Brazilian Central Bank communication helps to forecast the yield curve. They include sentiment variables as additional factors in the dynamic Nelson-Siegel term structure model and found that these sentiment variables contain predictive information for yield curve forecasting.

Although the research cited above illustrated different approaches and shreds of evidence, they have some common features and hypotheses. One of them is constant volatility for all maturities throughout the sampling period. The second one, which may justify the first, is that the absence of time-varying volatility would be acceptable in estimation for advanced economies compared with emerging economies. The last is a need to evaluate forecasting results applied to an investment strategy. So, our goal is to explore these gaps to investigate whether modeling time-varying volatility in macro-term structure models to an emerging market improves forecasting performance and whether these forecasts are worth it for an investor who cares about mean and variance.

Our analysis builds on contributions from Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) who use macroeconomic variables, and Koopman, Mallee, and Van der Wel (2010) that use a factor volatility structure for the latent variables with a specification based on GARCH models[4], both based on the Nelson and Siegel (1987) model of the term structure. In an application to Brazilian data, this paper provides evidence that adding backward- and forward-looking macroeconomic information to the

[3] Poncela (2013) comments that what helps to increase the forecasting accuracy of the US term structure of interest rates is to restrict the transition matrix φ of the VAR(1) model proposed for the common factors. The forecasting results for the smooth dynamic factor model (SDFM) with and without macro variables are very similar, meaning that simpler models for the common factors (in the form of uncoupled factors) are preferred for forecasting.

[4] Some ways to overcome this issue were proposed by Francesco Bianchi, Mumtaz, and Surico (2009), Márcio Laurini and Hotta (2010), João F Caldeira, Márcio P Laurini, and Portugal (2010) and Hautsch and Yang (2012).

dynamic factor models with time-varying volatility can produce more accurate forecasts for forecasts of government bond yields. Following the recent literature emphasizing the interaction between the yield curve and other economic variables, we present forecast results from different model specifications, considering both with and without macroeconomic variables and time-varying volatility.

Although our analysis focuses on statistical measures of predictive accuracy, it is essential to evaluate the extent to which the apparent gains in predictive accuracy can be used in real-time to improve investors' economic utility, that is, translate into better investment performance. Given that statistical significance does not necessarily imply economic significance, we follow what was done in Thornton and Valente (2012), Sarno, Paul Schneider, and Wagner (2016), João F. Caldeira, Guilherme V. Moura, and André A. P. Santos (2016), and Gargano, Pettenuzzo, and Timmermann (2019), among others, and assess the economic value of the predictive power of interest rates by investigating the utility gains accrued to investors who exploit the predictability of yield curve relative to the benchmark model. Our results confirm and extend results found in previous literature that add macroeconomic information.

3.1 DYNAMIC FACTOR MODELS FOR THE YIELD CURVE

Dynamic factor models play a major role in econometrics since allowing the explanation of a large set of time series in terms of a small number of unobserved common factors, see Jungbacker, Koopman, and Wel (2014). Many specifications for the yield curve can be viewed as dynamic factor models with restrictions imposed on factor loadings Joslin, Le, and Kenneth J. Singleton (2013). This section discusses the fourteen individual yield curve models beginning with the three-factor DNS model.

3.1.1 Dynamic Nelson-Siegel model

Francis X Diebold and Li (2006) proposed to study and estimate the term structure of interest rates using the model proposed by Nelson and Siegel (1987), assuming that the parameters vary over time. The following equation would describe the dynamics of the term structure

$$y_{i,t}(\tau_i) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}\right), \tag{14}$$

where y_{it} denotes the yield at time t of a security with maturity τ_i, for $t = 1, \ldots, T$ and $i = 1, \ldots, N$, and λ is a decay parameter that can capture a variety of shapes of the yield curve through time, such as upward and downward sloping, and inversely humped. The β_{1t}, β_{2t}, and β_{3t} are time-varying parameters, or the state variables, that can be interpreted as the level, slope, and curvature latent components of the yield curve.

The DNS is our two starting point to model and forecast the yield curve. The dynamic movements or evolution of the yield curve factors, β_{1t},β_{2t}, and β_{3t}, are assumed to follow a vector autoregressive process of first order, which allows for casting the yield curve latent factor model in state-space.

3.1.2 DNS in State-Space Representation

Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) note that DNS framework can be represented as a state space model by treating $\boldsymbol{\beta_t} = \beta_{j,t}$, for $j = 1, \ldots, 3$, as a latent vector. For these purpose, the general specification of the dynamic factor model is given by:

$$y_t = \Lambda(\lambda_t)\boldsymbol{\beta_t} + \boldsymbol{\varepsilon_t}, \qquad \boldsymbol{\varepsilon_t} \sim \mathcal{N}(0, \Sigma_\varepsilon), \tag{15}$$

where Λ is a $N \times K$ matrix of factor loadings, $\boldsymbol{\beta_t}$ is a K-dimensional stochastic process, and $\boldsymbol{\varepsilon_t}$ is the $N \times 1$ vector of measurement errors, whose covariance matrix given by Σ_ε. For any given, strictly positive λ_1, the $N \times K$ factor loading matrix $\Lambda(\lambda_t)$ is given by:

$$\Lambda_{ij}(\lambda_k) = \begin{cases} 1, & j = 1 \\ \psi_{i2} = \dfrac{1 - z_{1i}}{\lambda_1 \tau_i}, & j = 2 \\ \psi_{i3} = \dfrac{1 - z_{1i}}{\lambda_1 \tau_i} - z_{1i}, & j = 3, \end{cases}$$

where $\psi_{1,i} = \exp(-\lambda_1 \tau_i)$.

The state-space framework is achieved by assuming that the dynamic movements or evolution of the yield curve factors $\boldsymbol{\beta_t}$ are modeled by the following first-order vector-autoregressive process:

$$\boldsymbol{\beta_{t+1}} = \boldsymbol{\mu} + \Phi\left(\boldsymbol{\beta_t} - \boldsymbol{\mu}\right) + \eta_t, \qquad \eta_t \sim \mathcal{N}\left(0, \Sigma_\eta\right), \tag{16}$$

where μ is a $K \times 1$ vector of constants, Φ a $K \times K$ coefficient matrix, and Σ_η is the covariance matrix of the disturbance vector η_t, which is independent of the vector of residuals $\boldsymbol{\varepsilon_t}, \forall t$.

The variance-covariance matrix of the innovations to the measurement system Σ_ε is assumed to be diagonal. This assumption implies that deviations of the observed yields of various maturities from those implied by the fitted yield curve are uncorrelated. While the matrix of variance-covariance of the innovations to the transition system Σ_η is unrestricted so that shocks to the three yield-curve factors are correlated.

3.1.3 The Nelson-Siegel Model with Macro Variables

The first extension of the DNS is the inclusion of macro-finance indicators in the model. In addition to the yield data, we have p factors available representing macroeconomic information at the monthly frequency, covering the same period as the yield data. We include these factors in the state vector such that it becomes ($\beta_{1,t}$, $\beta_{2,t}$, $\beta_{3,t}$, $f_{1,t}$, ..., $f_{p,t})' = \boldsymbol{\beta}_t^{FA}$. With the extension of the state factor, the size of the coefficient matrix Φ^{FA} in the state equation increases from 3×3 to $(3 + p) \times (3 + p)$. The resulting state-space form is then given by

$$\mathbf{y}_t = \underbrace{\begin{bmatrix} \boldsymbol{\Lambda}(\lambda) & \mathbf{0}_{N\times p} \end{bmatrix}}_{\boldsymbol{\Lambda}(\lambda)^{FA}} \underbrace{\begin{bmatrix} \boldsymbol{\beta}_t \\ f_t \end{bmatrix}}_{\boldsymbol{\beta}_t^{FA}} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\varepsilon), \tag{17}$$

$$\underbrace{\begin{bmatrix} \boldsymbol{\beta}_{t+1} \\ f_{t+1} \end{bmatrix}}_{\boldsymbol{\beta}_{t+1}^{FA}} = (I_{3+p} - \Phi^{FA}) \begin{bmatrix} \boldsymbol{\mu} \\ \mathbf{0}_{p\times 1} \end{bmatrix} + \Phi^{FA} \underbrace{\begin{bmatrix} \boldsymbol{\beta}_t \\ f_t \end{bmatrix}}_{\boldsymbol{\beta}_t^{FA}} + \boldsymbol{\eta}_t^{FA}, \quad \boldsymbol{\eta}_t^{FA} \sim \mathcal{N}(\mathbf{0}_{3+p}, \boldsymbol{\Sigma}_\eta^{FA}) \tag{18}$$

for t = 1, ..., T , where the dimensions of $\boldsymbol{\Phi}$, $\boldsymbol{\eta}_{t+1}$, and $\boldsymbol{\Sigma}_\eta$ are increased as appropriate. The coefficient matrix structure implies that the macro-factors affect the individual yields through the Nelson-Siegel factors and feedback from the yields to the macro-factors. Therefore, we estimated the DNS-Macro in that framework. The following section explains the algorithm used in the estimation procedure.

3.1.4 Estimation Procedure

Estimating the loading parameters λ in the measurement matrix in Eq. (15) is the key to calculating the state-space model. Keeping λ's fixed over the whole sample period, the equations (15) and (16) characterize a linear and Gaussian state-space model; thus, the Kalman filter can be used to obtain the likelihood function via the prediction error decomposition. The estimation procedures are discussed below.

3.1.4.1 Estimation of Linear State Space Models Based on the Kalman Filter

Assuming that the decay parameters are constant, the measurement equation becomes linear. The DNS model is treated as linear Gaussian state-space model in this case. Given the state-space formulation of the dynamic factor model presented in (15) and (16), the Kalman filter can be used to obtain the likelihood function via the prediction error decomposition. An optimization algorithm is used to maximize the likelihood function estimated by the Kalman filter, an iterative process of calculating and updating the measurement and transition equations until an optimal point is obtained. In short, the filter computes the optimal yield forecasts and the corresponding forecasting errors, after which the Gaussian likelihood function is evaluated using the prediction-error decomposition of the likelihood function for the forecasts and the states. It updates the measurement and transition equations sequentially until an optimal yield forecast is achieved.

Consider the general state-space representation in (15) and (16). This state-space model is estimated using a Kalman filter, a recursive formula running forwards through time to estimate latent factors from past observations. The Kalman filter evaluates the conditional means and variances of the latent factors $\boldsymbol{\beta}_{t+1}$ conditional on the information available up to and including time t, denoted as $\hat{b}_{t+1|t}$ and $P_{t+1|t}$ respectively. Using the transition equation in (16), the optimal predicted estimates is then given by

$$\hat{b}_{t+1|t} = \boldsymbol{\mu} + \Phi \left(b_{t|t} - \boldsymbol{\mu} \right), \tag{19}$$

$$P_{t+1|t} = \Phi P_t \Phi' + \Sigma_\eta, \tag{20}$$

where $P_{t+1|t}$ is mean square error (MSE), or covariance, matrix. Hence, the optimal filtered estimates $\hat{b}_{t+1}$ and P_{t+1} is given by

$$\hat{b}_{t+1} = \hat{b}_{t+1|t} + P_{t+1|t} \Lambda' F_{t+1|t}^{-1} v_{t+1}, \tag{21}$$

$$P_{t+1} = P_{t+1|t} - P_{t+1|t} \Lambda' F_{t+1|t}^{-1} P_{t+1|t}, \tag{22}$$

where $v_{t+1} = y_{t+1} - \Lambda \hat{b}_{t+1|t}$ is the prediction error, $F_{t+1|t} = \Lambda P_{t+1|t} \Lambda' + \Sigma_\varepsilon$ is the measurement prediction variance, and $P_{t+1|t} \Lambda' F_{t+1|t}^{-1}$ is called the Kalman gain.

The Kalman filter iterative process is initialized by using the unconditional mean and variance of $\boldsymbol{\beta}_t$. For this purpose, we carry out the 2-step procedure as described in Francis X Diebold and Li (2006). Specifically, the unconditional mean and covariance matrix of the state vector is started as follows

$$b_{1|0} = \mathbb{E} \left[\boldsymbol{\beta}_t \right] = \boldsymbol{\mu} \quad \text{and} \quad P_{1|0} = \mathbb{E} \left[\boldsymbol{\beta}_t \boldsymbol{\beta}_t' \right] = \Sigma_\beta,$$

where the unconditional covariance matrix of the state vector is the solution of $\Sigma_\beta - \Phi \Sigma_\beta \Phi = \Sigma_\eta$, which we can solve using the properties of the vectorization operator *vec*, see Bent Jesper Christensen and Wel (2019).

Let the vector θ collects all unknown coefficients in the in the VAR parameter matrix Φ, variance matrices Σ_ε and Σ_η, and Λ and $\boldsymbol{\mu}$. To estimate the parameters vector θ, the likelihood function is constructed from the update step by assuming that the forecasting errors v_t are Gaussian. The Gaussian log-likelihood function is computed as

$$\ell(\theta) = -\frac{NT}{2} \log 2\pi - \frac{1}{2} \sum_{t=1}^{T} \log |F_t| - \frac{1}{2} \sum_{t=1}^{T} v_t' F_t^{-1} v_t. \tag{23}$$

As a result, $\ell(\theta)$ can be evaluated by Kalman filter for a given value of θ. By maximizing this log-likelihood function with respect to the parameters (collective represented as a vector θ) using a quasi-Newton optimization method results in maximum likelihood estimates of the parameters. The algorithm BFGS is used to maximize the log-likelihood function specified in (23) to obtain the estimates of the parameters θ.

Under this framework, we estimated the linear models. In the next sections, we extend the DNS model by considering conditional heteroskedasticity in the yield processes and treating the loading parameters as a stochastically time-varying latent factor.

3.1.4.2 Time-Varying Volatility

In the DNS model, we assume volatility is constant over time, which may be a restrictive assumption since yield curves are related to trading in the financial markets, then changes in volatility may occur. In general, heteroscedasticity is a constant problem in economics, especially finance. The Kalman filter can not handle this problem; the filter works under the hypothesis that the variance matrices are constant or at least known. Assuming the GARCH structure, the matrix Σ_ε is time-varying.

To allow for conditional heteroscedasticity in the yield processes, we modify the DNS model by following Koopman, Mallee, and Van der Wel (2010), who propose capturing yield curve volatility allowing for a common variance component jointly affecting all individual yields. The common variance component is modeled as a generalized autoregressive conditional heteroscedasticity (GARCH) process. Andrew C Harvey, Ruiz, and Sentana (1992) already provides an extensive framework for incorporating this GARCH(1,1) model into unobserved component time series models and how to deal with corresponding implications for estimation procedures. This factor can be interpreted as the volatility of an underlying bond market portfolio according to Engle and Victor K Ng (1993). The error in the measurement equation (15) is decomposed as

$$\varepsilon_t = \Gamma_\varepsilon \varepsilon_t^* + \varepsilon_t^\dagger, \quad t = 1, \ldots, T, \tag{24}$$

where Z_ε and $\varepsilon_t^\dagger$ are $N \times 1$ vectors of loadings and noise component respectively, and ε_t^* is a scalar representing the common disturbance term. The error components are mutually independent of each other and are distributed as follows

$$\varepsilon_t^* \sim \text{NID}(0, h_t), \quad \text{and} \quad \varepsilon_t^\dagger \sim \text{NID}(\mathbf{0}, \Sigma_\varepsilon^\dagger), \quad t = 1, \ldots T, \tag{25}$$

where $\Sigma_\varepsilon^\dagger$ is a diagonal matrix and h_t is the variance specified as a GARCH process, according to Bollerslev (1986). In this case, we have

$$h_{t+1} = \gamma_0 + \gamma_1 \varepsilon_t^{*2} + \gamma_2 h_t, \quad t = 1, \ldots, T, \tag{26}$$

and the estimated parameters have the constraints $\gamma_0 > 0$, $0 < \gamma_1 < 0$, $0 < \gamma_2 < 0$, $h_1 = \gamma_0(1 - \gamma_1 - \gamma_2)^{-1}$, and $(\gamma_1 + \gamma_2) < 1$. The weights vector Γ_ε can be normalized to avoid identification problems, such that $\Gamma_\varepsilon' \Gamma_\varepsilon = 1$, but we follow Koopman, Mallee, and Van der Wel (2010) and fixed γ_0 at 1×10^{-4}. The resulting time-varying variance matrix for ε_t is given by

$$\Sigma_\varepsilon(h_t) = h_t \Gamma_\varepsilon \Gamma_\varepsilon' + \Sigma_\varepsilon^\dagger, \tag{27}$$

where $\Sigma_\varepsilon(h_t)$ depends on a single factor described by the GARCH process in (47). The (unconditional) time-varying variance matrix of y_t is $\Lambda(\lambda)\Sigma_\beta\Lambda(\lambda)' + \Sigma_\varepsilon(h_t)$, where Σ_β is the solution of $\Sigma_\beta - \Phi\Sigma_\beta\Phi' = \Sigma_\eta$. The GARCH factor ε_t^* is incorporated in the measure equation (15), which is treated as a latent factor. Hence, we include ε_t^* in the state vector alongside the DNS factors.

The resulting observation and state equations of the DNS models with time-varying volatility can be rewritten into the state-space formulation as

$$y_t = \begin{bmatrix}\Lambda(\lambda) & \Gamma_\varepsilon\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}_t \\ \varepsilon_t^*\end{bmatrix} + \varepsilon_t^\dagger, \quad \varepsilon_t^\dagger \sim \mathcal{N}(0,\Sigma_\varepsilon^\dagger), \tag{28}$$

$$\begin{bmatrix}\boldsymbol{\beta}_{t+1} \\ \varepsilon_{t+1}^*\end{bmatrix} = \begin{bmatrix}(I_j - \Phi_j)\boldsymbol{\mu} \\ 0\end{bmatrix} \begin{bmatrix}\Phi_j & 0_{j\times 1} \\ 0_{1\times j} & 0\end{bmatrix}\begin{bmatrix}\boldsymbol{\beta}_t \\ \varepsilon_t^*\end{bmatrix} + \begin{bmatrix}\boldsymbol{\eta}_t \\ \varepsilon_{t+1}^*\end{bmatrix}, \qquad \begin{bmatrix}\boldsymbol{\eta}_t \\ \varepsilon_{t+1}^*\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}\mathbf{0} \\ \mathbf{0}\end{bmatrix}, \begin{bmatrix}\Sigma_\eta & 0_{j\times 1} \\ 0_{1\times j} & h_{t+1}\end{bmatrix}\right), \tag{29}$$

for $t = 1, \ldots, T$. The addition of GARCH disturbances and extra parameters requires applying certain adjustments to the estimation procedure.

Since h_{t+1} in (47) is a function of its past values and unobserved values of ε_t^*, it is not possible to calculate the values required for h_{t+1} at time t. Specifically, Andrew C Harvey, Ruiz, and Sentana (1992) explains that, although the models are not conditionally Gaussian because knowledge of past observations does not imply knowledge of past GARCH errors, we may treat the models as though they are conditionally Gaussian. Because of that, in the presence of GARCH errors, the Kalman filter can be regarded as a quasi-optimal filter instead of optimal. Andrew C Harvey, Ruiz, and Sentana (1992) propose to take the expectation of the latent term in the volatility specification such that we obtain an estimate for h_{t+1}, given by

$$\hat{h}_{t+1|t} = \gamma_0 + \gamma_1\mathbb{E}\left[\varepsilon_t^{*2}|\Im_t\right] + \gamma_2\hat{h}_{t|t-1}, \quad t = 1, \ldots, T, \tag{30}$$

where $\Im_t$ denotes all information available up to and including time t. To calculate the expectation term we note that

$$\varepsilon_t^* = \mathbb{E}\left[\varepsilon_{t-1}|\Im_t\right] + \left(\varepsilon_t^* - \mathbb{E}\left[\varepsilon_t^*|\Im_t\right]\right).$$

By squaring and taking conditional expectations we can shown that

$$\begin{aligned}\mathbb{E}\left[\varepsilon_t^{*2}|\Im_t\right] &= \mathbb{E}\left[\varepsilon_t^*|\Im_t\right]^2 + \mathbb{E}\left[\left(\varepsilon_t^* - \mathbb{E}\left[\varepsilon_t^*|\Im_t\right]\right)^2\right], \\ &= \hat{\varepsilon}_{t|t}^{*\,2} + P_{t|t}^\varepsilon,\end{aligned} \tag{31}$$

where $\hat{\varepsilon}_{t|t}^*$ is the last element of the filtered state $b_{t|t}$ and $P_{t|t}^{\varepsilon^*}$ is the last diagonal element of the $P_{t|t}$, the filtered variance of $\hat{\varepsilon}_{t|t}^*$. Then, we substitute the expression $\mathbb{E}\left[\varepsilon_t^{*2}|\Im_t\right]$ into (30) to obtain a prediction for the volatility component h_{t+1}. Lastly, we insert the predicted value h_{t+1} in the position (j,j) of the variance matrix Σ_η, corresponding to the location of ε_t^* in the state vector. With this framework we estimated the DNS-GARCH and extensions

3.2 OUT-OF-SAMPLE ANALYSIS

In order to forecasts h-months ahead, the steps below follow after the filtering step and estimation of the optimal set of parameters $\boldsymbol{\theta}$ throughout the sample, see Durbin and Koopman (2012) for more details. Therefore, we have the following

$$\begin{aligned} \mathbf{y}_{t+1} &= \boldsymbol{\Lambda}(\lambda)\mathbb{E}(\boldsymbol{\beta}_{t+1}|Y_t), \\ \bar{\mathbf{y}}_{t+1} &= \boldsymbol{\Lambda}(\lambda)\bar{\mathbf{b}}_{t+1}, \end{aligned} \tag{32}$$

where $\bar{\mathbf{b}}_{t+1}$ is the state vector and $\bar{\mathbf{B}}_{t+1}$ the variance matrix of the states calculated by the Kalman filter in ((19)) and ((20)). For other forecasts, the filter can be rewritten to h = 2, . . . , H, as follows

$$\bar{\mathbf{b}}_{t+h} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\bar{\mathbf{b}}_{t+1} - \boldsymbol{\mu}), \tag{33}$$

$$\bar{\mathbf{B}}_{t+h} = \boldsymbol{\Phi}\bar{\mathbf{B}}_{t+1}\hat{\boldsymbol{\Phi}}' + \boldsymbol{\Sigma}_\eta, \tag{34}$$

$$\bar{\mathbf{y}}_{t+h} = \boldsymbol{\Lambda}(\lambda)\bar{\mathbf{b}}_{t+h}, \tag{35}$$

where the states vector and the variance matrix of the previous estimation states are used to calculate the predictions in the step h+1. With the time-varying loading parameter, the difference for predictions lies in the loading matrix that multiplies the state vector, that is, $\mathbf{Z}_t(\mathbf{a}_{t|t-1})$ instead of $\boldsymbol{\Lambda}(\lambda)$ in ((32)) and ((35)).

The out-of-sample predictions are assessed by the relative sizes of the root mean square error (RMSPE) of all considered models relative to those from the DNS baseline model. The RMSPE is calculated as follows:

$$R(h,\tau) = \sqrt{\frac{1}{n}\sum_t [\hat{y}_{t+h|t}(\tau) - y_{t+h}(\tau)]^2}, \tag{36}$$

where n is the number of forecasts previously defined in 312. The drawback of using RMSPE is that this is a single statistic summarizing individual forecasting errors over an entire sample. Although often used, they do not give any insight as to where in the sample a particular model makes its largest and smallest forecast errors. Therefore, we also graphically analyze the cumulative squared forecast errors (CSFE) proposed by Welch and Goyal (2008). These cumulative prediction errors series clearly depicts when a model outperforms or underperforms a given benchmark and could motivate the use of adaptive forecast combination schemes. The CSFE is given by:

$$\text{CSFE}_m(h,\tau) = \sum_t \left[\left(\hat{y}_{t+h|t,\mathrm{bench}}(\tau) - y_{t+h}(\tau)\right)^2 - \left(\hat{y}_{t+h|t,m}(\tau) - y_{t+h}(\tau)\right)^2\right]. \tag{37}$$

In the case a model outperforms the benchmark, the $\text{CSFE}_m(h,\tau)$ will be an increasing series. If the benchmark produces more accurate forecasts, then $\text{CSFE}_m(h,\tau)$ will tend to be decreasing.

We use the Giacomini and White (2006) test to assess whether the forecasts of two competing models are statistically different. The Giacomini-White (GW) test is a test of conditional forecasting ability and is constructed under the assumption that forecasts are generated using a moving data window. This is a test of equal forecasting accuracy and as such can handle forecasts based on both nested and non-nested models, regardless from the estimation procedures used in the derivation of the forecasts.

Lastly, we implement the Model Confidence Set (MCS), approach developed by Peter R. Hansen, Lunde, and James M. Nason (2011b), which consist on a sequence of tests which permits to construct a set of 'superior' models, where the null hypothesis of Equal Predictive Ability (EPA) is not rejected at a certain confidence level. The EPA statistic tests is calculated for an arbitrary loss function, in our case we test squared errors of DNS model against competing models.

3.2.1 The Economic Value of the Yield Curve Predictability

Although our analysis is focused on statistical measures of predictive accuracy, it is important to evaluate the extent to which the apparent gains in predictive accuracy can be used in real time to improve investors' economic utility, that is, translate into better investment performance. Given that statistical significance does not necessarily imply economic significance, see Thornton and Valente (2012), Sarno, Paul Schneider, and Wagner (2016), João F. Caldeira, Guilherme V. Moura, and André A. P. Santos (2016) and Gargano, Pettenuzzo, and Timmermann (2019), we assess the economic value of the predictive power of interest rates by investigating the utility gains accrued to investors who exploit the predictability of yield curve relative to a no-predictability alternative associated with the random-walk model.

In this section, we explore the empirical evidence linking statistical forecasting evaluation with economic utility. To this purpose, we consider a mean-variance investor with quadratic utility and relative risk aversion γ who allocates her portfolio on a risky bond with τ periods to maturity versus a one-month T-bill that pays the risky free rate, see Rapach and Zhou (2013). At the end of t, the investor allocates the following share of her portfolio to bond with maturity τ_i during $t+1$:

$$w_{i,t} = \left(\frac{1}{\gamma}\right)\left(\frac{\hat{r}_{t+h}^{(\tau_i)}}{\hat{\sigma}_{t+h}^{2,(\tau_i)}}\right) \tag{38}$$

where $\hat{r}_{t+h}^{(\tau_i)} = \tau_i y_t^{\tau_i} - (\tau_i - h)\hat{y}_{t+h}^{\tau_i - h}$ is a return forecast for the bond with maturity τ_i in time t and $\hat{\sigma}_i^2$[5] is a forecast of the variance of bond returns. Over the forecast evaluation

[5] We follow the strategy of Rapach and Zhou (2013) and estimate the variance of bond returns using

period, the investor realizes the average utility,

$$\hat{v}_i = \hat{\mu}_i - 0.5\gamma\hat{\sigma}_i^2, \tag{39}$$

where $\hat{\mu}_i \left(\hat{\sigma}_i^2\right)$ is the sample mean (variance) of the portfolio formed on the basis of $\hat{r}_{t+h}^{(\tau_i)}$ and $\hat{\sigma}_i^2$ over the forecast evaluation period. The resulting sequences of allocation weights are next used to calculate realized utilities. For each model *m*, the realized utility are converted into equivalent returns CER, i.e., the difference between utility (39) with model *m* and the DNS represents the utility gain accruing to using the competitors models forecast of the bond yields in place of the DNS benchmark forecast in the asset allocation decision. This utility gain (certainty equivalent return) can be interpreted as the portfolio management fee that an investor would be willing to pay to have access to the information in the model forecast relative to the information in the benchmark DNS model.

3.3 DATA AND RESULTS

3.3.1 Data

This paper's data set consists of monthly closing prices observed for yields of future DI contracts. Based on the observed rates for the available maturities, the data were converted to fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 36, 42, 48, and 60 months through interpolations using cubic splines. The database contains the maturities with the highest liquidity for January 2003 through December 2019 (T = 204 observations) and represents the most liquid DI contracts negotiated during the analyzed period. We assess the model's performance by splitting the sample into two parts: the first includes 132 observations used to estimate all models' parameters. The second part is used to analyze the performance out-of-sample of bond portfolios obtained from the model, with 72 observations.

Table 10 displays the descriptive statistics for the Brazilian interest rate curve. For each 14-time series, we report the average, standard deviation, minimum, maximum, and the last three columns containing sample autocorrelations at displacements of 1, 6, and 12 months. Descriptive statistics presented in Table 10 seem to confirm key stylized facts about yield curves: the sample average curve is upward-sloping and concave, volatility decreases with maturity, and autocorrelations are very high and increase with maturity. Also, there is a high persistence in the yields: the first-order autocorrelation for all maturities is above 0.87 for each maturity.

Figure 12 presents a three-dimensional plot of the data set and illustrates how yield levels and spreads vary substantially throughout the sample. Although the yield series change heavily over time for each of the maturities, a strong common pattern

the sample variance computed from a one-year (252-obs) rolling window of historical returns.

in the 14 series over time is apparent. The sample contains 204 monthly observations with maturities of τ = 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 36, 39, 48, and 60 months.

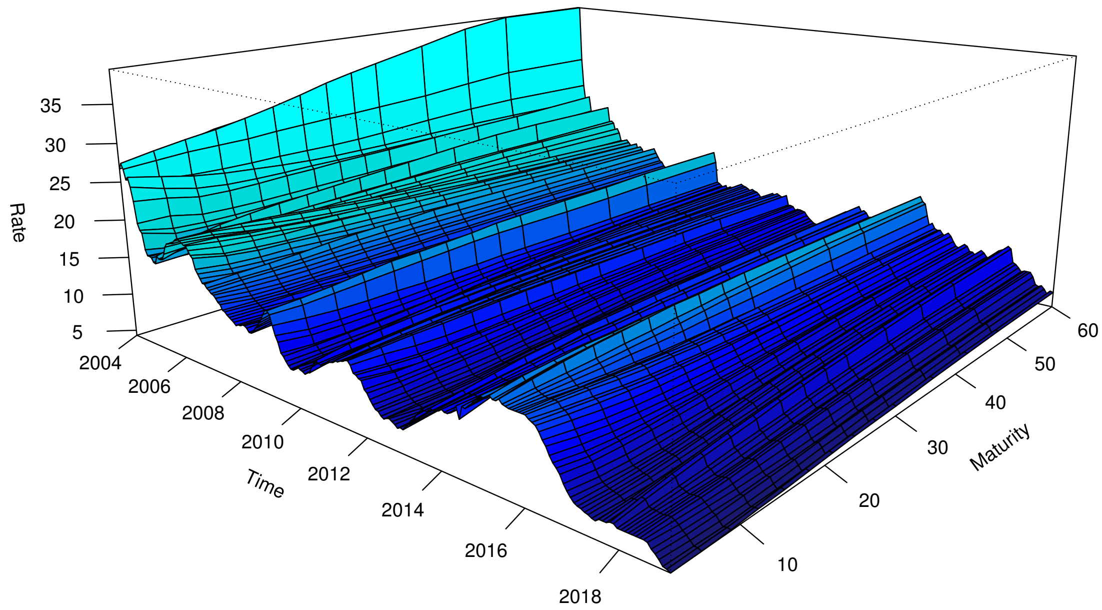
Table 10 – Descriptive statistics for the term structure of interest rates

Maturity	Mean	Std.dev.	Min	Max	Skewness	Kurtosis	$\hat{\rho}_1$	$\hat{\rho}_6$	$\hat{\rho}_{12}$
M3	12.19	4.52	4.30	27.50	0.90	4.11	0.96	0.70	0.48
M6	12.20	4.47	4.30	28.30	0.90	4.24	0.96	0.69	0.48
M9	12.24	4.42	4.40	29.00	0.92	4.56	0.95	0.68	0.48
M12	12.30	4.39	4.50	29.60	0.98	4.93	0.94	0.67	0.47
M15	12.38	4.34	4.50	30.40	1.06	5.39	0.94	0.65	0.46
M18	12.48	4.31	4.60	31.30	1.16	5.95	0.93	0.64	0.45
M21	12.57	4.28	4.70	32.30	1.29	6.64	0.93	0.62	0.44
M24	12.65	4.27	4.90	33.40	1.43	7.41	0.92	0.61	0.43
M27	12.74	4.25	5.00	34.30	1.57	8.15	0.92	0.60	0.42
M30	12.80	4.26	5.10	35.10	1.69	8.86	0.91	0.59	0.41
M36	12.92	4.28	5.30	36.70	1.95	10.35	0.90	0.57	0.40
M42	13.02	4.34	5.60	38.40	2.20	11.93	0.90	0.55	0.38
M48	13.10	4.38	5.70	39.40	2.37	12.96	0.89	0.54	0.38
M60 (Level)	13.18	4.35	6.00	39.40	2.44	13.12	0.89	0.55	0.38
Slope	0.99	2.10	-3.90	11.90	0.75	6.77	0.84	0.35	0.01
Curvature	-0.24	1.17	-3.00	3.10	0.18	2.86	0.87	0.40	0.08

Source: Elaborated by the author.

Legend: The table reports summary statistics for Brazil's yield curve from 2003 to 2019. We examine monthly data constructed using the spline method. For each maturity, we show mean, standard deviation, skewness, kurtosis, minimum, maximum, and three auto-correlations coefficients, $\hat{\rho}_1$, $\hat{\rho}_6$, $\hat{\rho}_{12}$. Also, the table reports proxy estimates for the level, slope, and curvature of the yield curve. The proxies are defined as follows: for level, the highest maturity bond (60 months); for slope, the difference between the bond of 60 months and the bond of 3 months; and for curvature, two times the bond of 18 months minus the sum of a bond of 3 months and bond of 60 months.

Figure 12 – 3D Brazilian yield curves.



Source: Elaborated by the author.

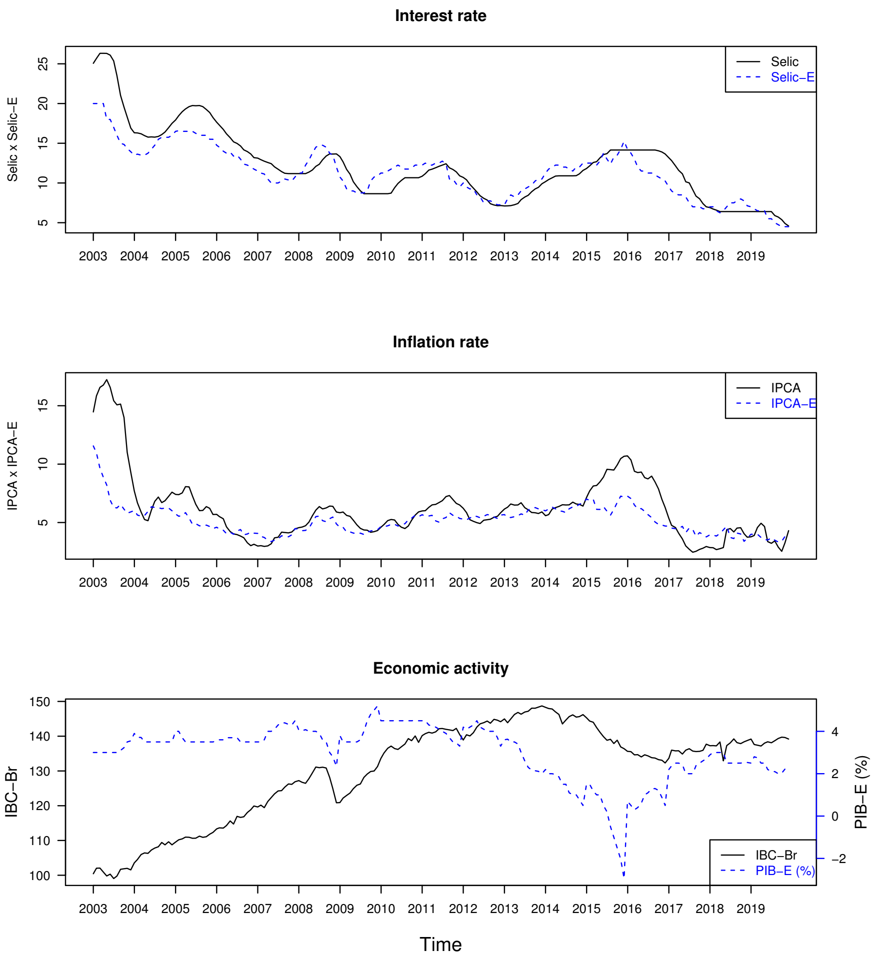
Legend: The figure plots the evolution of the term structure of interest rates (based on DI-future contracts) for the time horizon of 2003:01-2019:12. The sample consisted of the monthly yields for the maturities of 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 36, 42, 48 and 60 months.

We use three macroeconomic factors: the Selic interest rate is the monetary policy interest rate, i.e., the key tool used by the Central Bank of Brazil (BCB) in implementing the monetary policy. The Selic rate, or 'over Selic', is the Brazilian federal funds rate. Selic rate is the weighted average interest rate of the overnight interbank operations, collateralized by federal government securities, carried out at the Special System for Settlement and Custody (Selic).

Under the inflation-targeting regime, the BCB's Monetary Policy Committee (Copom) regularly sets the target for the Selic rate. Within the relevant horizon for the monetary policy, Copom aims to keep the Extended National Consumer Price Index (IPCA inflation rate) around the target and anchor inflation expectations. Accordingly, the BCB performs daily open market operations to keep the effective Selic rate at the target set by Copom.

Brazil has several price indexes that differ significantly in scope, depending on their particular purposes. A price index can be designed to reflect the cost of living for a specific group of households, but each household will have its price index based on its consumer basket. In this sense, there can be different inflation perceptions between what the citizen notices as inflation and what the variation of several price indexes shows. The inflation index adopted is the IPCA, the reference for the Brazilian inflation-targeting system. The BCB ensures that the IPCA's annual inflation is centered at the inflation target set by the National Monetary Council (CMN). The IPCA is also the price index of the National Treasury's Notes Series B (NTN-B) (see Figure 13).

Figure 13 – Interest and inflation rates and economic activity Index.



Source: Elaborated by the author.

Legend: This figure shows the SELIC interest rate, the Extended National Consumer Price Index (IPCA), and the Index of Economic Activity of the Central Bank (IBC-Br) in solid lines and his expectations (GDP for IBC-Br) of one year in dashed lines, respectively. The sample contains 194 monthly observations from January 2003 through December 2019 for realized series and from 2002 through 2018 for expectations.

The IBC-Br is an indicator of the monthly periodicity, which incorporates the pathway of the variables considered as proxies to the development of the economic sectors such as Agriculture and livestock, Industry, and Services. The well-known adherence of the trajectory of the IBC-Br to the GDP behavior confirms the importance of monitoring the indicator to understand better and anticipate the activity analysis.

We use market expectations factors from the BCB's Market Expectations System, which monitors market expectations regarding the main macroeconomic variables, providing important inputs for the monetary policy decision-making process.

The BCB carries out the "Focus Survey", compiling forecasts of about 140 banks, asset managers, and other institutions (real sector companies, brokers, consultancies, etc.). The Survey daily monitors the market expectations for several inflation indices, the GDP and industrial production growth, the exchange rate, the Selic rate, fiscal indicators, and external sector variables. Based on this Survey, the BCB compiles daily – and releases weekly – the Focus Market Readout, which summarizes the statistics calculated over the information collected. We use the IPCA inflation accumulated median percent change, Over-Selic Target median percent p.y., and total GDP median percent change over the next 12 months.

3.4 RESULTS

We use a rolling estimation window of 72 monthly observations (i.e., six years) for computing our results. We produce forecasts for 1-month, 3-month, 6-month, and 12-month-ahead. We calculate the root mean square forecast error (RMSFE) to compare the performance of out-of-sample forecasts. Moreover, the Giacomini and White (2006) test (GW-test) assesses whether each model outperforms the DNS. Table 14 reports statistical measures of the out-of-sample forecasting performance at various horizons. The first row of entries in each panel of the tables reports the value of RMSFE (expressed in basis points) for the DNS model, while all other rows report statistics relative to the DNS.

The statistical results for the out-of-sample forecasts in terms of RMSPE are displayed in Table 14. This table is divided into four panels, each corresponding to a different forecast horizon $(1, 3, 6,$ and 12 steps ahead). The first row in each panel contains the RMSPE of the RW baseline forecasts, whereas the remaining rows report RMSPE of a given model relative to those of the benchmark. Therefore any number below one indicates outperformance relative to the benchmark, whereas any number larger than one indicates underperformance. Asterisks to the right of entries suggest that, at the 10% level of significance, the null hypothesis of the GW test is not rejected. Bold type indicates that the model belongs to $\widehat{M}^*_{0.75}$, the set of superior models containing the best models with probability no less than 75%.

Table 11 – Relative Root Mean Squared Forecast Errors

Model	Maturity													
	M3	M6	M9	M12	M15	M18	M21	M24	M27	M30	M36	M42	M48	M60
Panel A: 1-month ahead forecasts														
RW	41.26	39.69	42.97	47.79	51.04	54.47	56.47	57.74	58.80	60.69	61.63	63.07	64.53	66.56
DNS	0.933	0.992	0.999	0.995	0.982	0.972	0.971	0.963	0.960	0.959	0.967	0.966	0.972	0.971
DNS – Macro	0.940	1.010	0.989	0.976	0.970	0.971	0.977	0.976	0.978	0.978	0.992	0.991	0.996	0.995
DNS – MacroE1	0.881	0.892*	0.978	1.009	1.010	1.009	1.018	1.016	1.020	1.019	1.033	1.034	1.049	1.063
DNS – GARCH	0.964	0.978	1.040	1.056	1.055	1.044	1.050	1.048	1.048	1.042	1.045	1.044	1.047	1.045
DNS – GARCH – Macro	0.940	1.010	0.989	0.976	0.970	0.971	0.976	0.976	0.978	0.978	0.992	0.991	0.996	0.995
DNS – GARCH – MacroE1	0.881	0.892*	0.979	1.009	1.010	1.009	1.018	1.016	1.020	1.019	1.034	1.034	1.049	1.063
Panel B: 3-months ahead forecasts														
RW	90.21	90.41	94.52	99.69	103.88	108.63	111.44	112.61	114.64	117.00	120.34	121.97	124.01	125.32
DNS	0.824	0.892	0.934	0.959	0.953	0.941	0.932	0.925	0.914	0.905	0.887*	0.880*	0.878*	0.878*
DNS – Macro	0.795	0.877	0.913	0.938	0.937	0.935	0.936	0.936	0.932	0.928	0.920*	0.920*	0.922*	0.929*
DNS – MacroE1	0.668	0.800	0.877	0.93	0.942	0.945	0.953	0.955	0.951	0.947	0.934	0.935	0.939	0.952
DNS – GARCH	0.818	0.929	0.977	0.997	0.999	0.993	0.992	0.996	0.992	0.985	0.972	0.974	0.974	0.981
DNS – GARCH – Macro	0.795	0.877	0.913	0.938	0.937	0.935	0.936	0.936	0.932	0.928	0.920*	0.920*	0.922*	0.929*
DNS – GARCH – MacroE1	0.668	0.800*	0.877	0.930	0.942	0.945	0.953	0.956	0.951	0.947	0.934	0.935	0.939	0.952
Panel C: 6-months ahead forecasts														
RW	168.06	166.53	167,34	169.92	172.07	174.62	176.65	177.09	178.23	180.37	184.08	186.97	187.46	188.79
DNS	0.760	0.835	0.892	0.922	0.926	0.919	0.907	0.896	0.884	0.868	0.834	0.815	0.806	0.790
DNS – Macro	0.722	0.819	0.886	0.925	0.936	0.939	0.935	0.930	0.924	0.914	0.889	0.877	0.873	0.864
DNS – MacroE1	0.669	0.782	0.858	0.906	0.922	0.928	0.929	0.926	0.923	0.914	0.885	0.873	0.867	0.86
DNS – GARCH	0.849	0.916	0.949	0.966	0.97	0.965	0.958	0.956	0.948	0.936	0.911	0.903	0.899	0.895
DNS – GARCH – Macro	0.722	0.819	0.886	0.925	0.936	0.939	0.935	0.930	0.924	0.914	0.889	0.877	0.873	0.864
DNS – GARCH – MacroE1	0.669	0.782	0.858	0.906	0.923	0.928	0.929	0.926	0.923	0.914	0.885	0.873	0.867	0.860
Panel D: 12-months ahead forecasts														
RW	296.67	294.33	290.84	286.67	281.94	277.09	272.77	269.43	267.11	265.25	264.99	263.29	262.12	260.45
DNS	0.823	0.875	0.914	0.940	0.951	0.954	0.951	0.941	0.929	0.918	0.884	0.861	0.844	0.819
DNS – Macro	0.830	0.924	0.991	1.036	1.059	1.072	1.073	1.068	1.057	1,0479	1.014	0.990	0.973	0.944
DNS – MacroE1	0.869	0.944	0.996	1.031	1.044	1.047	1.045	1.034	1.019	1.006	0.966	0.937	0.917	0.885
DNS – GARCH	0.894	0.923	0.951	0.969	0.976	0.982	0.977	0.970	0.959	0.951	0.924	0.909	0.899	0.886
DNS – GARCH – Macro	0.830	0.924	0.991	1.036	1.059	1.072	1.074	1.068	1.057	1.048	1.014	0.990	0.973	0.944
DNS – GARCH – MacroE1	0.869	0.944	0.996	1.031	1.044	1.048	1.045	1.034	1.019	1.006	0.966	0.938	0.917	0.885

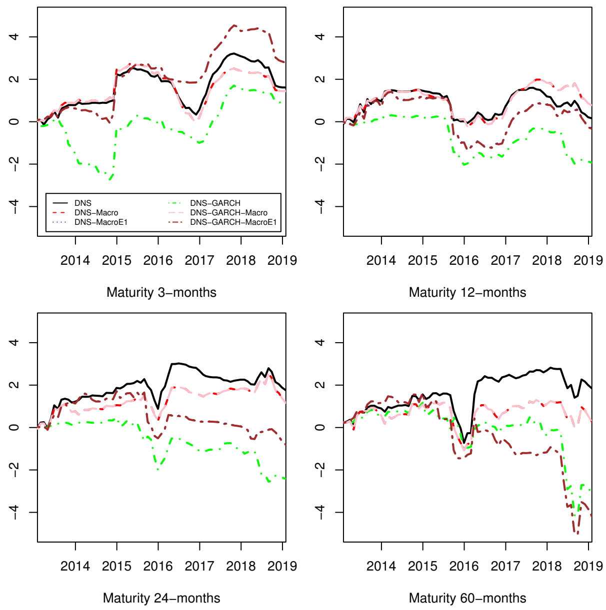
Source: Elaborated by the author.

Legend: The Table reports the relative root mean squared forecast errors (RMSFE) relative to the Random-Walk (RW) model for the 1-month, 3-months, 6-months, and 12-months forecast horizons. The evaluation sample is 2014:1 to 2019:12 (73 out-of-sample forecasts). The first line in each panel of the Table reports the value of RMSFE (expressed in basis points) for the RW model, while all other lines report statistics relative to the RW. The following model abbreviations are used in the Table: DNS-Macro model for realized macroeconomic factors, DNS-MacroE1 model for market expectations of macroeconomic factors, and the last three models have time-varying volatility (-GARCH). Numbers smaller than one indicate that models outperform the RW, whereas larger numbers indicate underperformance. The * on the right of the cell entries indicates the level at which the Giacomini and White (2006) test rejects the null of equal forecasting accuracy at least 10% level. Shaded values indicate that the model belongs to Model Confidence Set (MCS) Peter R Hansen, Lunde, and James M Nason (2011a).

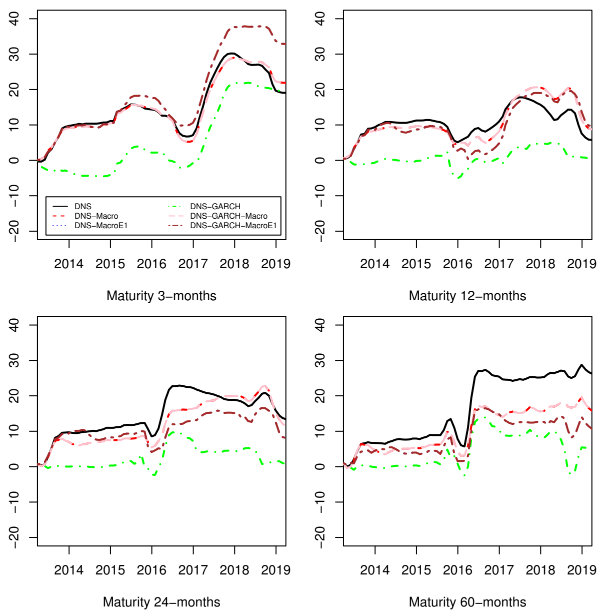
To analyze the accuracy of the forecasts in different time intervals, we calculate the difference in cumulative square forecast errors between each of the prediction models and the RW along the out-of-sample evaluation period, see Figure 14 and Figure 15.

Figure 14 – Cumulative squared forecast errors (1- and 3-month ahead)

(a) One-month ahead



(b) Six-month ahead

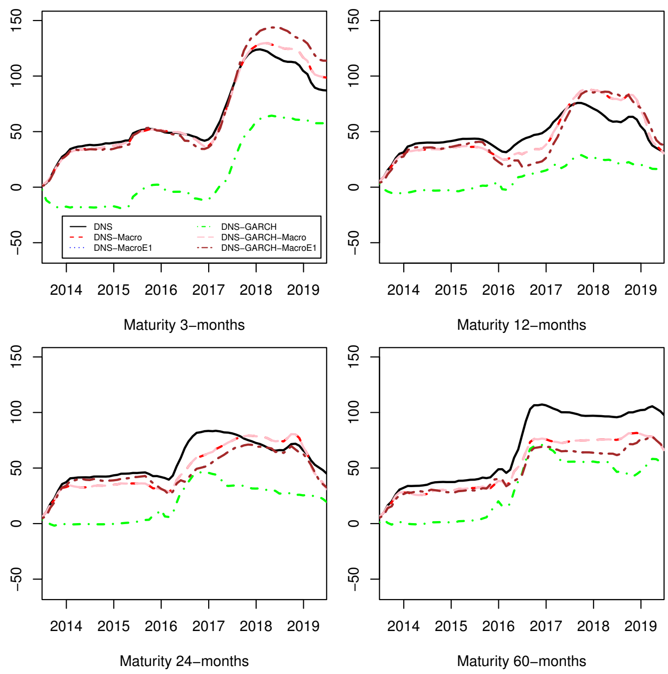


Source: Elaborated by the author.

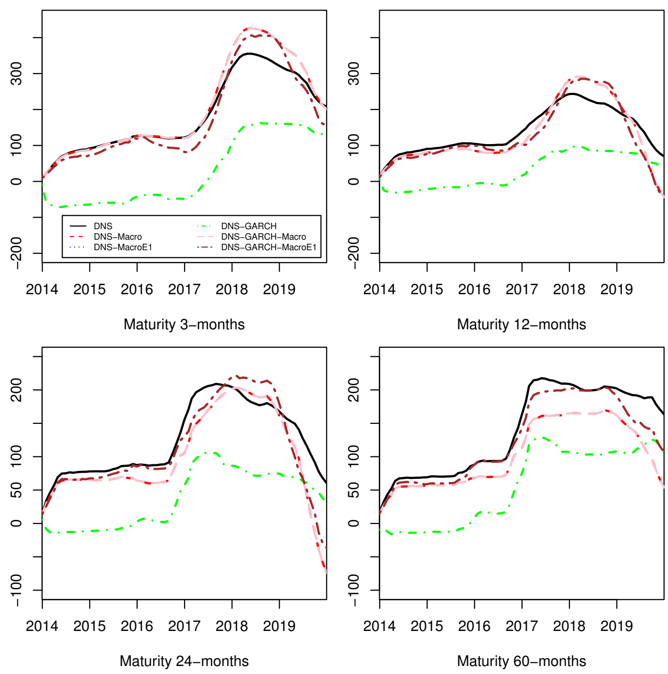
Legend: Figures show the cumulative squared forecast errors (CSFE) of Nelson-Siegel Extensions relative to RW. Figure shows CSFEs for a 1- and 6-month forecast horizon. The evaluation sample is from January 2014 through December 2019 (73 out-of-sample forecasts).

Figure 15 – Cumulative squared forecast errors (6- and 12-month ahead)

(a) Twelve-month ahead



(b) Six-month ahead



Source: Elaborated by the author.

Legend: Figures show the cumulative squared forecast errors (CSFE) of Nelson-Siegel Extensions relative to RW. Figure shows CSFEs for a 1- and 6-month forecast horizon. The evaluation sample is from January 2014 through December 2019 (73 out-of-sample forecasts).

3.4.1 Economic Evaluation Results

Table 12 reports certainty equivalent (average utility gains in annualized percent return) for a mean- variance investor with $\gamma = \{0.1, 0.5, 1, 5\}$ who allocates among 1 to 5 years bonds and risk-free rate using forecasts based on competitors models in place of DNS forecasts.

The performance of economic evaluation is evaluated in terms of average utility gain (δ) excess return relative to the risk-free rate, and we consider the risk free rate to be the interbank rate CDI.

Table 12 – Out-of-sample economic evaluation of the yield curve forecasting

Model	$\gamma = 0.1$					$\gamma = 0.5$					$\gamma = 1$					$\gamma = 5$				
	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$	$\tau = 1$	$\tau = 2$	$\tau = 3$	$\tau = 4$	$\tau = 5$
horizon = 1-*month ahead*																				
DNS	197.70	2.403	–0.279	–0.901	–0.643	50.16	1.843	0.516	0.138	0.072	31.72	1.773	0.615	0.268	0.161	16.97	1.717	0.695	0.372	0.233
DNS – Macro	226.09	7.118	1.431	–0.296	–0.504	54.65	2.587	0.786	0.234	0.094	33.22	2.021	0.705	0.300	0.169	16.07	1.568	0.641	0.353	0.229
DNS – MacroE1	213.77	8.047	3.377	1.657	1.286	52.70	2.734	1.093	0.542	0.377	32.57	2.07	0.808	0.403	0.263	16.46	1.538	0.579	0.291	0.172
DNS – GARCH	449.44	34.41	12.92	6.247	3.711	89.91	6.896	2.600	1.267	0.760	44.97	3.457	1.31	0.644	0.391	9.019	0.706	0.278	0.146	0.095
DNS – GARCH – Macro	449.54	34.48	12.98	6.272	3.701	89.93	6.907	2.609	1.271	0.758	44.98	3.461	1.313	0.646	0.39	9.016	0.704	0.276	0.145	0.096
DNS – GARCH – MacroE1	449.41	34.40	12.92	6.258	3.726	89.91	6.895	2.601	1.268	0.762	44.97	3.457	1.310	0.645	0.391	9.020	0.706	0.278	0.146	0.095
horizon = 3-*month ahead*																				
DNS	38.21	7.202	4.000	2.762	2.271	13.71	2.135	1.054	0.673	0.512	10.65	1.501	0.685	0.412	0.292	8.204	0.995	0.391	0.203	0.116
DNS – Macro	54.28	10.606	5.521	3.529	2.668	16.25	2.672	1.294	0.794	0.574	11.50	1.680	0.765	0.452	0.313	7.697	0.887	0.343	0.179	0.103
DNS – MacroE1	47.23	10.08	5.635	3.899	3.122	15.14	2.590	1.312	0.852	0.646	11.13	1.653	0.771	0.472	0.337	7.919	0.904	0.339	0.167	0.089
DNS – GARCH	181.10	23.13	9.533	5.22	3.377	36.28	4.649	1.927	1.061	0.686	18.17	2.34	0.977	0.541	0.350	3.692	0.492	0.216	0.125	0.081
DNS – GARCH – Macro	181.30	23.25	9.65	5.309	3.423	36.31	4.668	1.946	1.075	0.694	18.18	2.346	0.983	0.546	0.353	3.685	0.488	0.212	0.122	0.080
DNS – GARCH – MacroE1	181.04	23.17	9.609	5.326	3.497	36.27	4.656	1.939	1.078	0.705	18.17	2.342	0.981	0.547	0.356	3.694	0.490	0.214	0.122	0.077
horizon = 6-*month ahead*																				
DNS	12.64	14.251	8.698	6.667	5.362	5.676	2.809	1.637	1.218	0.953	4.806	1.379	0.754	0.537	0.402	4.110	0.235	0.048	–0.008	–0.039
DNS – Macro	20.07	15.762	9.355	6.940	5.422	6.850	3.048	1.741	1.261	0.963	5.197	1.459	0.789	0.551	0.405	3.875	0.187	0.028	–0.017	–0.041
DNS – MacroE1	17.61	15.845	9.206	6.741	5.257	6.461	3.061	1.717	1.23	0.937	5.067	1.463	0.781	0.541	0.397	3.952	0.185	0.032	–0.010	–0.036
DNS – GARCH	85.76	13.442	6.843	4.953	4.096	17.22	2.682	1.344	0.947	0.753	8.654	1.337	0.657	0.447	0.335	1.800	0.261	0.107	0.046	0.001
DNS – GARCH – Macro	85.96	13.50	6.872	4.954	4.031	17.25	2.691	1.349	0.947	0.743	8.665	1.34	0.658	0.447	0.332	1.794	0.259	0.106	0.046	0.003
DNS – GARCH – MacroE1	85.69	13.576	7.027	5.190	4.380	17.21	2.703	1.373	0.985	0.798	8.650	1.344	0.666	0.459	0.350	1.803	0.256	0.101	0.039	–0.008
horizon = 12-*month ahead*																				
DNS	120.24	55.49	21.874	13.88	10.21	23.99	9.131	3.294	2.034	1.480	11.969	3.337	0.972	0.553	0.388	2.344	–1.298	–0.886	–0.632	–0.485
DNS – Macro	120.84	53.57	19.94	12.146	8.716	24.09	8.829	2.989	1.76	1.244	12.00	3.237	0.870	0.462	0.310	2.326	–1.238	–0.825	–0.577	–0.437
DNS – MacroE1	120.72	54.56	20.53	12.469	8.892	24.07	8.985	3.081	1.811	1.272	11.99	3.288	0.901	0.479	0.319	2.329	–1.269	–0.844	–0.587	–0.443
DNS – GARCH	118.88	14.42	3.330	4.670	6.527	23.78	2.648	0.366	0.579	0.898	11.89	1.176	–0.004	0.068	0.195	2.387	–0.001	–0.301	–0.341	–0.368
DNS – GARCH – Macro	119.27	14.72	3.402	4.635	6.398	23.85	2.695	0.377	0.574	0.878	11.92	1.192	–0.001	0.066	0.188	2.375	–0.011	–0.303	–0.340	–0.364
DNS – GARCH – MacroE1	118.99	15.36	4.483	6.039	8.130	23.80	2.796	0.548	0.796	1.151	11.90	1.225	0.056	0.140	0.279	2.384	–0.031	–0.337	–0.384	–0.419

Source: Elaborated by the author.

Legend: This table reports the average utility gain (δ) of the portfolio management fee (in annualized percent return) that an investor with mean-variance preferences and risk aversion coefficient of 0.1 to 5 would be willing to pay to have access to the forecasting method relative to the DNS benchmark forecast. The following model abbreviations are used in the table: DNS-Macro model for realized macroeconomic factors, DNS-MacroE1 model for market expectations of macroeconomic factors, and the last three models have time-varying volatility (-GARCH). The sample started on January 2003, and the evaluation period is from January 2013 to December 2019.

3.5 CONCLUDING REMARKS

This research aims to explore forward-looking macroeconomic data in the factor-augmented DNS model to evaluate forecastability performance. Also, we include time-varying volatility to evaluate the out-of-sample performance. We assess the forecastability of the DNS with time-varying volatility and forward-looking macroeconomic factors. Our first goal was to run the DNS with observed macroeconomic factors such as interest rate, economic activity, and inflation rate. Also, evaluate the DNS-GARCH model in the same manner. Our second goal was to run the initial framework, including forward-looking macroeconomic factors, then examine out-of-sample results against the DNS benchmark model.

The results in statistical analysis suggest the DNS with macroeconomics expectations outperforms the benchmark model in short maturities and the first three horizons of forecasts(1 month, 3 months, and 6 months). The DNS with macroeconomic expectations and time-varying volatility also had similar results. The economic evaluations s, on the other hand, indicate that time-varying volatility improved the results. In other words, the volatility aggregates value to an investor, even in high-risk aversion.

4 THE TREND IS YOUR FRIEND: DIRECTIONAL INTEREST RATE FORECASTS WITH THE DNS MODEL

Interest rate forecasts are a fundamental ingredient to guide monetary policy decision-making and support investors in fixed-income portfolio allocation and pricing derivatives. In this context, the dynamic version of the Nelson and Siegel (1987) three-factor model (hereafter DNS) proposed in Francis X Diebold and Li (2006) and revisited in Koopman, Mallee, and Van der Wel (2010) has become a major workhorse and a widely used tool in the financial community as it can deliver good in-sample fit and accurate out-of-sample forecasts, see BIS (2005) and ECB (2018). However, the literature has mainly focused on using the DNS model's variants to obtain point forecasts, see De Pooter (2007), Francis X Diebold, Li, and Yue (2008), Yu and Salyards (2009), De Pooter, Ravazzolo, and Dick JC Van Dijk (2010), Jens HE Christensen, Francis X Diebold, and Glenn D Rudebusch (2011), Márcio Laurini and Hotta (2010) and João F Caldeira, Guilherme V Moura, and André AP Santos (2016), therefore neglecting the ability of this specification to generate direction-of-change forecasts.

In this paper, we use the directional forecasting framework developed by Peter F Christoffersen and Francis X Diebold (2006) and Peter Christoffersen et al. (2006) to show how to employ the conditional moments extracted from a heteroskedastic DNS model to obtain directional interest rate forecasts. We find that the directional forecasts obtained for a large panel of interest rates outperform those obtained with a baseline non-parametric specification and a logistic model when information about the skewness and kurtosis of the term structure is taken into account.

During the last two decades, many works have been devoted to developing alternative modeling and forecasting methodologies for the term structure. One of the most prominent methods consists of modeling the dynamic evolution of the yield curve using the three-factor model proposed by Nelson and Siegel (1987), according to which the factors represent the level, slope, and curvature of the curve and depend on a single exponential decay rate parameter. Assuming that the exponential decay rate is known and constant over time, Francis X Diebold and Li (2006) propose a simple two-step estimation procedure. Koopman, Mallee, and Van der Wel (2010) extend it in two directions. First, they suggest allowing the discount parameter to evolve over time; see Márcio Laurini and Hotta (2010) and Hevia et al. (2015) for another proposal in which the discount parameter is allowed to change over time. Second, they specify a common volatility factor that a GARCH model represents. A common denominator to these studies is that allowing for time-varying parameters and time-varying volatilities leads to improved performance regarding in-sample fit and out-of-sample forecasts.

The accuracy of yield curve forecasts, not only of magnitude but specifically of the direction of change in financial assets, emerges as a research source in the prediction field. Mark Greer (2003), for instance, after conducting tests on the directional accuracy of long-term interest rates forecasts published on *The Wall Street Journal*, suggests these predictions, in general, could be performed with the same precision as flipping a coin, even if there was a combination of the best predictions, see Mark R Greer (2005).

On the other hand, authors such as Peter F Christoffersen and Francis X Diebold (2006) found that stock returns' volatility forecasts generated a significant forecastability of positive return probability. In other words, volatility predictability produces predictability in the direction of stock returns. They also suggest it would be more likely to find forecastability in intermediate return horizons such as monthly frequency. The authors argued the results are significant to academic studies and by market practitioners, who usually use *timing* strategies linked to volatility movements, see Rattray and Balasubramanian (2003). Peter F Christoffersen and Francis X Diebold (2006) and Peter Christoffersen et al. (2006) were the first to thoroughly investigate the conditional volatility dynamics and the positive return probability in financial assets forecasts.

We aim to explore the literature gap on forecasting the yield curves using conditional means and conditional volatility forecasts as inputs to predict the return direction in the fixed income. The results suggest that incorporating information about the skewness and kurtosis of yield curve returns leads to directional forecasts that outperform the benchmark model, mainly for long maturities and short-term forecasts. The ingredients for this research are the results presented in the first and second sections. The following section introduces the method.

4.1 THE DYNAMIC NELSON-SIEGEL MODEL

The factors model for the yield curve can represent the forms usually associated with the yield curve, that is, monotonic, curved, and S. Francis X Diebold and Li (2006) modified the Nelson and Siegel (1987) model by incorporating time-varying factors in the following way[1]

$$y_{i,t}(\tau_i) = \beta_{1,t} + \beta_{2,t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i}\right) + \beta_{3,t}\left(\frac{1-e^{-\lambda\tau_i}}{\lambda\tau_i} - e^{-\lambda\tau_i}\right), \tag{40}$$

where y_t denotes the yields at time t and τ_i the maturity of the bond to $\{t\}_{i=1}^{T}$ and $\{i\}_{i=1}^{N}$, respectively. The parameter λ determines the exponential decay rate, i.e., small λ values results in slow decay and can better fit the curve for longer maturities; on the other hand, large λ values produce rapid decay and can better fit the curve for shorter maturities.

In the Francis X Diebold and Li (2006) model, λ is kept fixed while parameters $\beta_{1,t}, \beta_{2,t}, \beta_{3,t}$ are estimated by ordinary least squares for each period. Cross-section estimates can be obtained whenever there are sufficient interest rates for different

1 The Equation (40) corresponds to equation (2) in the paper of Nelson and Siegel (1987). According to Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006), the following notations are adopted: τ for maturity instead of *m*, and the loading parameter λ equal to $\frac{1}{\tau}$.

maturities in time. In Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006), on the other hand, λ is estimated. In the following Section, I presented the models through the state-space representation.

4.1.1 The Dynamics of the Latent Factors

Francis X Diebold, Glenn D Rudebusch, and Aruoba (2006) advanced by proposing that the NS model framework can be represented as a state-space model by treating vector $\boldsymbol{\beta_t} = (\beta_{1,t},\beta_{2,t},\beta_{3,t})'$ as a latent vector. The model equation can be written as follows:

$$\begin{bmatrix} y_t(\tau_1) \\ \vdots \\ y_t(\tau_N) \end{bmatrix} = \begin{bmatrix} 1 & x_{1,2} & x_{1,3} \\ \vdots & \vdots & \vdots \\ 1 & x_{N,2} & x_{N,3} \end{bmatrix} \begin{bmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{bmatrix} + \begin{bmatrix} \varepsilon_t(\tau_1) \\ \vdots \\ \varepsilon_t(\tau_N) \end{bmatrix}, \tag{41}$$

where

$$x_{i,2} = \frac{1-z_i}{\lambda\tau_i}, \quad x_{i,3} = \frac{1-z_i}{\lambda\tau_i} - z_i,$$

$$z_i = \exp(-\lambda\tau_i).$$

The observation equation in Equation (41) relates the observed interest rates of the $i = 1,\ldots,N$ maturities with the latent factors $\boldsymbol{\beta}_t$.

The vector autoregressive of order 1 of the factors that govern the dynamics of the state equation is defined as follows:

$$\begin{bmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} + \begin{bmatrix} \varphi_{1,1} & \varphi_{1,2} & \varphi_{1,3} \\ \varphi_{2,1} & \varphi_{2,2} & \varphi_{2,3} \\ \varphi_{3,1} & \varphi_{3,2} & \varphi_{3,3} \end{bmatrix} \begin{bmatrix} \beta_{1,t}-\mu_1 \\ \beta_{2,t}-\mu_2 \\ \beta_{3,t}-\mu_3 \end{bmatrix} + \begin{bmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{bmatrix}, \tag{42}$$

in matrix notation, the Equation (41) and Equation (42) can be rewritten as follows

$$\mathbf{y}_t = \boldsymbol{\Lambda}(\lambda)\boldsymbol{\beta}_t + \varepsilon_t, \tag{43}$$

$$\boldsymbol{\beta}_{t+1} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\boldsymbol{\beta}_t - \boldsymbol{\mu}) + \boldsymbol{\eta}_t, \tag{44}$$

where y_t is a vector $N \times 1$, $\boldsymbol{\Lambda}(\lambda)$ is a loading matrix $N \times 3$, $\boldsymbol{\Phi}$ is a VAR(1) parameters matrix 3×1, $\boldsymbol{\beta}_t$ and $\boldsymbol{\mu}$ are vectors 3×1. We assumed that $\boldsymbol{\eta}_t$ and ε_t are orthogonal to each other.

The variance matrix of the observation errors $\boldsymbol{\Sigma}_\varepsilon$ is diagonal. This assumption implies that interest rate deviations for different maturities are not correlated, which facilitates model estimation by reducing the number of parameters. On the other hand, the assumption that the state errors variance matrix $\boldsymbol{\Sigma}_\eta$ is unrestricted allows the shocks in the three factors to be correlated.

We have the representation of the DNS model in the state-space form. In this study, we used the Kalman filter. This algorithm is a recursive procedure to calculate the optimal estimator of the state vector t, based on the available information at time t, and make forecasts for the state vector at $t+1$ based on t. In the following sections, we introduced the time-varying volatility extension.

4.1.2 Time-Varying Volatility

In the DNS model, we assume that volatility is constant, which may be a flexible assumption since yield curves are related to trading in the financial markets. Thus, volatility in these markets may change over time; in general, heteroscedasticity is a constant problem in economics, especially in finance. On the other hand, the Kalman filter can not solve this problem; that is, the filter works under the hypothesis that the variance and covariance matrix is constant or at least known. Assuming the GARCH structure, the array is unknown; in other words, it is time-varying.

Therefore, this section presents how to modify the Kalman filter to incorporate the GARCH approach to perform parameter and volatility estimations in a single step. Therefore, the DNS model class has a common volatility component modeled by a univariate GARCH process according to Andrew C Harvey, Ruiz, and Sentana (1992) and Koopman, Mallee, and Van der Wel (2010). The error vector, in the Equation (43), is decomposed as follows:

$$\varepsilon_t = \boldsymbol{\Gamma}_\varepsilon \varepsilon_t^\star + \varepsilon_t^+, \tag{45}$$

where $\boldsymbol{\Gamma}_\varepsilon$ and ε_t^+ are defined as a vector of weights and an error vector of dimensions N$\times$1, respectively, and $\varepsilon_t^\star$ a scalar error factor. The error components are independent of each other as follows

$$\varepsilon_t^\star \sim \mathcal{N}(0,h_t), \quad \varepsilon_t^+ \sim \mathcal{N}(\mathbf{0},\boldsymbol{\Sigma}_\varepsilon^+), \quad t = 1,...T, \tag{46}$$

where $\boldsymbol{\Sigma}_\varepsilon^+$ is a diagonal matrix and h_t is the variance specified as a GARCH process, according to Bollerslev (1986). In this case, I have the following

$$h_{t+1} = \gamma_0 + \gamma_1 \varepsilon_t^{\star 2} + \gamma_2 h_t, \quad t = 1,..., T, \tag{47}$$

and the estimated parameters have the constraints $\gamma_0 > 0$, $0 < \gamma_1 < 0$, $0 < \gamma_2 < 0$, $h_1 = \gamma_0(1-\gamma_1-\gamma_2)^{-1}$ and $(\gamma_1+\gamma_2) < 1$. The vector of weights Γ_ε can be normalized to avoid identification problems, such that $\Gamma_\varepsilon' \Gamma_\varepsilon = 1$, however, this restriction can be replaced by γ_0 fixed at 1×10^{-4}. This last restriction, therefore, I used in estimation. The variance matrix of ε_t in Equation (47) is time-varying as follows

$$\boldsymbol{\Sigma}_\varepsilon(h_t) = h_t \Gamma_\varepsilon \Gamma_\varepsilon' + \boldsymbol{\Sigma}_\varepsilon^+, \tag{48}$$

where it depends on a single factor described by the GARCH process in Equation (47). The unknown parameters in the GARCH specification, $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \Gamma_\varepsilon')$, are grouped in the parameter vector $\boldsymbol{\theta}$.

The state in Equation (44) has one more unobservable component, i.e., $\varepsilon_t^\star$ is now calculated as a latent state. The state-space representation of the observation Equation (43) and the state Equation (44) have some modifications as follows

$$\mathbf{y}_t = \underbrace{\begin{bmatrix}\boldsymbol{\Lambda}(\lambda) & \boldsymbol{\Gamma}_\varepsilon\end{bmatrix}}_{\boldsymbol{\Lambda}^\star(\boldsymbol{\lambda})} \underbrace{\begin{bmatrix}\boldsymbol{\beta}_t \\ \varepsilon_t^\star\end{bmatrix}}_{\boldsymbol{\beta}_t^\star} + \varepsilon_t^+, \quad \varepsilon_t^+ \sim \mathcal{N}(0, \boldsymbol{\Sigma}_\varepsilon^+), \tag{49}$$

$$\underbrace{\begin{bmatrix}\boldsymbol{\beta}_{t+1} \\ \varepsilon_{t+1}^\star\end{bmatrix}}_{\boldsymbol{\beta}_{t+1}^\star} = \underbrace{\begin{bmatrix}(I_j - \Phi_j)\boldsymbol{\mu} \\ 0\end{bmatrix}}_{\boldsymbol{\mu}^\star} \underbrace{\begin{bmatrix}\boldsymbol{\Phi}_j & \mathbf{0}_{j\times 1} \\ \mathbf{0}_{1\times j} & 0\end{bmatrix}}_{\boldsymbol{\Phi}^\star} \underbrace{\begin{bmatrix}\boldsymbol{\beta}_t \\ \varepsilon_t^\star\end{bmatrix}}_{\boldsymbol{\beta}_t^\star} + \underbrace{\begin{bmatrix}\boldsymbol{\eta}_t \\ \varepsilon_{t+1}^\star\end{bmatrix}}_{\boldsymbol{\eta}_t^\star}, \tag{50}$$

$$\underbrace{\begin{bmatrix}\boldsymbol{\eta}_t \\ \varepsilon_{t+1}^\star\end{bmatrix}}_{\boldsymbol{\eta}_t^\star} \sim \mathcal{N}\left(\begin{bmatrix}\mathbf{0} \\ \mathbf{0}\end{bmatrix}, \underbrace{\begin{bmatrix}\boldsymbol{\Sigma}_\eta & \mathbf{0}_{j\times 1} \\ \mathbf{0}_{1\times j} & h_{t+1}\end{bmatrix}}_{\boldsymbol{\Sigma}_\eta^\star}\right), \tag{51}$$

to $t = 1, ...,$ T e $j = 1, 2, 3$ refer to the DNS-GARCH model. Since h_{t+1} in Equation (47) is a function to its past values and unobserved values of $\varepsilon_t^\star$, it is not possible to calculate the values required for h_{t+1} in time t. Andrew C Harvey, Ruiz, and Sentana (1992) propose to replace the square of the error term in Equation (47) by their expected value. Therefore, h_{t+1} can be replaced by its estimate based on observations $y_1, ..., y_t$ as follows

$$\hat{h}_{t+1|t} = \gamma_0 + \gamma_1 \mathbb{E}[\varepsilon_t^{\star 2} | Y_t] + \gamma_2 \hat{h}_{t|t-1}, \quad t = 1, ..., T, \tag{52}$$

in which the expected value can be calculated by the recursions of the Kalman filter using the increased state vector with $\varepsilon_t^\star$ filtered in the last element of vector $\mathbf{b}_{t|t}$, in the Equation (56). The expected value follows

$$\mathbb{E}[\varepsilon_t^{\star 2} | Y_t] = \hat{\varepsilon}_{t|t}^{\star 2} + B_{t|t}^{\varepsilon}, \tag{53}$$

where $\hat{\varepsilon}_{t|t}$ is the filtered estimate of ε_t, and $B_{t|t}^{\varepsilon}$ is the variance of ε_t, which are computed for all states during recursions of the Kalman filter, given the observations until period t. Because of the substitution of $\hat{h}_{t|t-1}$ in h_{t+1} in the Equation (51), which in the filter is inserted into the (j, j) element of the $\boldsymbol{\Sigma}_\eta^\star$ matrix in the Equation (57), the filter and likelihood estimates are sub-optimal, see Andrew C Harvey, Ruiz, and Sentana (1992) for more details.

Therefore the (unconditional) time-varying variance matrix of y_t is $\boldsymbol{\Lambda}^\star(\lambda)\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^\star\boldsymbol{\Lambda}^\star(\lambda)' + \boldsymbol{\Sigma}_\varepsilon^\star(h_t)$, where $\boldsymbol{\Sigma}_\beta^\star$ is the solution of $\boldsymbol{\Sigma}_\beta^\star - \boldsymbol{\Phi}^\star\boldsymbol{\Sigma}_\beta^\star\boldsymbol{\Phi}^{\star\prime} = \boldsymbol{\Sigma}_\eta^\star$. To estimate the DNS-GARCH-Macro model, the Equation (49), Equation (50), and Equation (51) are increased as appropriate. In the following Section, we presented the procedure for the estimation.

4.2 ESTIMATION BASED ON THE KALMAN FILTER

The model equations of the Equation (43) and Equation (44) is linear and Gaussian. The estimation is therefore based on the Kalman filter. This recursive algorithm uses the data information at time t to construct state estimates at time $t + 1$. The proposed estimation method combines the filter with the Maximum Likelihood (ML) estimation.

The procedure for calculating latent values and unknown parameters is recursive; we started the process by making an initial assumption about the unknown parameters $\boldsymbol{\theta}_1$ to execute the algorithm. The prediction error vector is calculated, $\mathbf{v}_t$, and the prediction error matrix, $\mathbf{F}_t$, in the Equation (54) and Equation (55), respectively, to analyze the log-likelihood in Equation (60). Updating the state vector $\mathbf{b}_{t|t}$ and the variance matrix $\mathbf{B}_{t|t}$, in the Equation (56) and Equation (57), is done in the filtering step in t given the set of information up to t. Therefore, consider the model defined in Equation (43) and Equation (44), and define $\mathbf{b}_{t|s}$ as minimum mean squared error linear estimators of $\boldsymbol{\beta}_t$ given y_t, ..., y_s to $s = t-1$, t in the following recursion:

$$\mathbf{v}_t = \mathbf{y}_t - \boldsymbol{\Lambda}(\boldsymbol{\lambda})\mathbf{b}_{t|t-1}, \tag{54}$$

$$\mathbf{F}_t = \boldsymbol{\Lambda}(\boldsymbol{\lambda})\mathbf{B}_{t|t-1}\boldsymbol{\Lambda}(\boldsymbol{\lambda})' + \boldsymbol{\Sigma}_{\varepsilon}, \tag{55}$$

$$\mathbf{b}_{t|t} = \mathbf{b}_{t|t-1} + \mathbf{B}_{t|t-1}\boldsymbol{\Lambda}(\boldsymbol{\lambda})'\mathbf{F}_t^{-1}\mathbf{v}_t, \tag{56}$$

$$\mathbf{B}_{t|t} = \mathbf{B}_{t|t-1} - \mathbf{B}_{t|t-1}\boldsymbol{\Lambda}(\boldsymbol{\lambda})'\mathbf{F}_t^{-1}\boldsymbol{\Lambda}(\boldsymbol{\lambda})\mathbf{B}_{t|t-1}, \tag{57}$$

$$\mathbf{b}_{t+1|t} = \boldsymbol{\mu} + \boldsymbol{\Phi}(\mathbf{b}_{t|t} - \boldsymbol{\mu}), \tag{58}$$

$$\mathbf{B}_{t+1|t} = \boldsymbol{\Phi}\mathbf{B}_{t|t}\boldsymbol{\Phi}' + \boldsymbol{\Sigma}_{\eta}, \tag{59}$$

where the parameters in the coefficient matrix of the VAR, $\boldsymbol{\Phi}$, the matrices of variances $\boldsymbol{\Sigma}_{\varepsilon}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\eta}}$, the mean vector $\boldsymbol{\mu}$ and the parameter λ are treated as unknown coefficients and grouped into the parameter vector $\boldsymbol{\theta}$, as previously mentioned. The forecast of $\mathbf{b}_{t+1|t}$ and $\mathbf{B}_{t+1|t}$, that is, a step forward is calculated in the filter prediction step in Equation (58) and Equation (59). The results of the prediction error vector, $\mathbf{v}_t$, and the prediction error matrix, $\mathbf{F}_t$, are again used as inputs into the log-likelihood function so that the estimate can be conducted to obtain new estimates of the unknown parameters θ_2. These steps are then iterated until the parameter values of the $\boldsymbol{\theta}_{\boldsymbol{MV}}$ are found, for which the log-likelihood function is maximized.

The estimation of $\boldsymbol{\theta}$ is based on the numerical maximization of the log-likelihood function via the prediction error decomposition, see Andrew C Harvey (1989). Therefore, log-likelihood follows by form:

$$\log L(Y_n) = -\frac{NT}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}(\log|\mathbf{F}_t| + \mathbf{v}_t'\mathbf{F}_t^{-1}\mathbf{v}_t), \tag{60}$$

where $\mathbf{v}_t$ and $\mathbf{F}_t$ are calculated recursively by the Kalman filter Equation (54) to Equation (59) for a given set of $\boldsymbol{\theta}$, such that log $L(Y_n)$ is computed using the filter result.

The calculations required for implementation were made through the R language maintained by the R Core Team (2018), and the minimization of the log-likelihood function was obtained by the `nlminb` optimization function.

The initial parameters were calculated in the estimation in two steps according to Francis X Diebold and Li (2006), namely, $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}_\varepsilon$ diagonal matrices, $\boldsymbol{\Sigma}_\eta$ upper triangular matrix, and $\boldsymbol{\Phi}$ VAR parameters matrix. According to Koopman, Mallee, and Van der Wel (2010) and Bent Jesper Christensen, Wel, et al. (2010), $\mathbf{b}_{1|0}$ and $\boldsymbol{\Sigma}_\beta$ of the model can be calculated as follows, according to the distribution of $\beta_{j,1}$, given by

$$\beta_1 \sim \mathcal{N}(\boldsymbol{\mu},\boldsymbol{\Sigma}_\beta), \tag{61}$$

in which the unconditional covariance matrix of the state vector, $\boldsymbol{\Sigma}_\beta$, can be started as follows

$$\begin{aligned}
\boldsymbol{\Sigma}_\beta - \boldsymbol{\Phi}\boldsymbol{\Sigma}_\beta\boldsymbol{\Phi}' &= \boldsymbol{\Sigma}_\eta,\\
\text{vec}(\boldsymbol{\Sigma}_\beta) - \text{vec}(\boldsymbol{\Phi}\boldsymbol{\Sigma}_\beta\boldsymbol{\Phi}') &= \text{vec}(\boldsymbol{\Sigma}_\eta),\\
\mathbf{I}_{j^2}\cdot\text{vec}(\boldsymbol{\Sigma}_\beta) - (\boldsymbol{\Phi}\otimes\boldsymbol{\Phi})\cdot\text{vec}(\boldsymbol{\Sigma}_\beta) &= \text{vec}(\boldsymbol{\Sigma}_\eta),\\
[\mathbf{I}_{j^2} - (\boldsymbol{\Phi}\otimes\boldsymbol{\Phi})]\cdot\text{vec}(\boldsymbol{\Sigma}_\beta) &= \text{vec}(\boldsymbol{\Sigma}_\eta),\\
\text{vec}(\boldsymbol{\Sigma}_\beta) &= [\mathbf{I}_{j^2} - (\boldsymbol{\Phi}\otimes\boldsymbol{\Phi})]^{-1}\cdot\text{vec}(\boldsymbol{\Sigma}_\eta),
\end{aligned} \tag{62}$$

then, under the assumption of a stationary process, the initial value of the state vector is equal to the unconditional mean, $\boldsymbol{\beta}_1 = \boldsymbol{\mu}$, and the initial value of the unconditional covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ is equal to Equation (62).

The DNS-GARCH model generates forecasts of yields and their volatilities. Thus, to map interest rates returns, we calculate $\hat{r}^{(\tau_i)}_{t+h} = \tau_i y^{\tau_i}_t - (\tau_i - h)\hat{y}^{\tau_i - h}_{t+h}$ where is a return forecast for the bond with maturity τ_i in time t and $\hat{\sigma}^2_i$ is a forecast of the variance of bond returns[2] of bond returns to models with constant volatility and $\Sigma_{r_{t|t-1}} = \tau'\tau \otimes [\Lambda(\lambda)\Sigma_\beta\Lambda(\lambda)' + \Sigma_\varepsilon(h_t)]$ to time-varying volatility.

4.3 DIRECTION-OF-CHANGE MODELS

Let R_t be a series of returns and Ω_t be the information set available at time t. $\Pr[R_t > 0]$ is the probability of a positive return at time t. The conditional mean and variance are denoted, respectively, as $\mu_{t+1|t} = \text{E}\left[R_{t+1} \mid \Omega_t\right]$ and $\sigma^2_{t+1|t} = \text{Var}\left[R_{t+1} \mid \Omega_t\right]$. The return series is said to display conditional mean predictability if $\mu_{t+1|t}$ varies with Ω_t; conditional variance predictability is defined similarly. If $\Pr[R_t > 0]$ exhibits conditional dependence, i.e., $\Pr\left[R_{t+1} > 0 \mid \Omega_t\right]$ varies with Ω_t, then we say the return series is sign predictable (or the price series is direction-of-change predictable).

2 We follow the strategy of Rapach and Zhou (2013) and estimate the variance of bond returns using the sample variance computed from a one-year (252-obs) rolling window of historical returns.

Suppose $\mu_{t+1|t} = \mu$ for all t and $\sigma^2_{t+1|t}$ varies with t in a predictable manner. Denoting $\mathcal{D}(\mu,\sigma^2)$ as a generic distribution dependent only on its mean μ and variance σ^2, assume

$$R_{t+1} \mid \Omega_t \sim D\left(\mu, \sigma^2_{t+1|t}\right).$$

Then the conditional probability of a positive return is

$$\begin{aligned}
\Pr(R_{t+1} > 0 \mid \Omega_t) &= 1 - \Pr(R_{t+1} \leq 0 \mid \Omega_t), \\
&= 1 - \Pr\left(\frac{R_{t+1} - \mu}{\sigma_{t+1|t}} \leq \frac{-\mu}{\sigma_{t+1|t}}\right), \\
&= 1 - F\left(\frac{-\mu}{\sigma_{t+1|t}}\right), \qquad (63)
\end{aligned}$$

where F is the distribution function of the “standardized" return $(R_{t+1|t} - \mu)/\sigma_{t+1|t})$. If the conditional volatility is predictable, then the sign of the return is predictable even if the conditional mean is unpredictable, provided $\mu \neq 0$. Note Also, if the distribution is asymmetric, the sign can be predictable even if the mean is zero: time-varying skewness can be driving sign prediction in this case.

Interaction between volatility and higher-ordered conditional moments can similarly affect the potency of conditional volatility as a predictor of return signs. We follow Peter F Christoffersen and Francis X Diebold (2006) and use

$$\Pr(R_{t+1} > 0 \mid \Omega_t) = 1 - F\left(\frac{-\mu_{t+1|t}}{\sigma_{t+1|t}}\right) \qquad (64)$$

to explore the sign predictability of one-, -two, three-, and six-month returns in yield curves[3]. We also use an extended version of Equation (64) that explicitly considers the interaction between volatility and higher-ordered conditional moments.

4.3.1 Baseline model

As Peter Christoffersen et al. (2006), we evaluate the forecasting performance of two sets of forecasts and compare them against forecasts from a baseline model. Our baseline forecasts are generated using the empirical cumulative distribution function (cdf) of the R_t using data from the beginning of our sample period right up to the time the forecast is made, i.e., at period k, we compute

$$\widehat{\Pr}\left(R_{k+1|k} > 0\right) = \frac{1}{k}\sum_{t=1}^{k} I(R_t > 0), \qquad (65)$$

where $I(\cdot)$ is the indicator function.

[3] where $\hat{\mu}^{(\tau_i)}_{t+h} = \tau_i y^{\tau_i}_t - (\tau_i - h)\,\hat{y}^{\tau_i - h}_{t+h}$ is a return forecast for the bond with maturity τ_i in time t and $\Sigma_{\hat{\mu}_{t+h}} = \tau'\tau \otimes \Sigma_{y_{t+h}}$ is their conditional covariance matrix.

4.3.2 Logistic model

A key idea of binary response models is that the return is transformed into a binary sign return indicator I_t that is used as the dependent variable. Let $R_{t+1:t+h}$ be the h-day return, and define the "positive return" indicator as $I_{t+h} = 1$ if $R_{t+1:t+h} > 0$ and $I_{t+h} = 0$ otherwise. We want to forecast I_{t+h}, and Peter Christoffersen et al. (2006) suggests using a model of the form,

$$I_{t+h} = F\left(\frac{\mu}{\sigma_t}\right) + e_{t+h}, \tag{66}$$

where $F(\cdot)$ is a monotone function with a left limit of zero and a right limit of one, μ is the h-day expected return, and σ_t is a forecast of h-day return volatility. In the logistic regression model, the relationship between x and the probability of the event of interest is described by:

$$F(x) = \frac{\exp(x)}{1 + \exp(x)}. \tag{67}$$

4.3.3 Non-parametric model

Our first forecasting model makes direct use of Equation (65). Using all available data at time k, we first regress R_t on a constant, $\log(\hat{\sigma}_t)$, and $[\log(\hat{\sigma}_t)]^2$, and compute

$$\hat{\mu}_t = \hat{\beta}_0 + \hat{\beta}_1 \log(\hat{\sigma}_t) + \hat{\beta}_2 [\log(\hat{\sigma}_t)]^2, \quad t = 1, \ldots, k \tag{68}$$

where $\hat{\sigma}_t$ is the square root of (actual, not forecasted) realized volatility. The period $k+1$ forecast is then generated by

$$\begin{aligned} \widehat{\Pr}\left(R_{k+1|k} > 0\right) &= 1 - \hat{F}\left(-\frac{\hat{\mu}_{k+1|k}}{\hat{\sigma}_{k+1|k}}\right), \\ &= 1 - \frac{1}{k}\sum_{t=1}^{k} I\left(\frac{R_t - \hat{\mu}_t}{\hat{\sigma}_t} \leq \frac{\hat{\mu}_{k+1|k}}{\hat{\sigma}_{k+1|k}}\right), \end{aligned} \tag{69}$$

i.e., $\hat{F}$ is the empirical cdf of $(R_t - \hat{\mu}_t)/\hat{\sigma}_t$. The forecasts of conditional mean $\hat{\mu}_{k+1|k}$ and conditional variance $\hat{\sigma}_{k+1|k}$ are from all DNS models and extensions.

4.3.4 Extended model

The second model is an extension of Equation (64) and explicitly considers the interaction between volatility, skewness, and kurtosis. We use the Gram-Charlier expansion:

$$1-F\left(\frac{-\mu_{t+1|t}}{\sigma_{t+1|t}}\right)\approx 1-\Phi\left(\frac{-\mu_{t+1|t}}{\sigma_{t+1|t}}\right)+\Phi\left(\frac{-\mu_{t+1|t}}{\sigma_{t+1|t}}\right)\left[\frac{\gamma_{3,t+1|t}}{3!}\left(\frac{\mu^2_{t+1|t}}{\sigma^2_{t+1|t}}-1\right)\right.$$
$$\left.+\frac{\gamma_{4,t+1|t}}{4!}\left(\frac{\mu^3_{t+1|t}}{\sigma^3_{t+1|t}}+\frac{3\mu_{t+1|t}}{\sigma_{t+1|t}}\right)\right],$$

where $\Phi(.)$ is the distribution function of a standard normal, and γ_3 and γ_4 are, respectively, the skewness and excess kurtosis, with the usual notation for conditioning on Ω_t. This equation can be rewritten as

$$1-F(-\mu_{t+1|t}x_{t+1})\approx 1-\Phi(-\mu_{t+1|t})(\beta_{0t}+\beta_{1t}x_{t+1}+\beta_{2t}x^2_{t+1}+\beta_{3t}x^3_{t+1}),$$

with β_{0t}=1+$\gamma_{3,t-1|t}/6$, β_{1t}=-$\gamma_{4,t+1|t}\mu_{t+1|t}/8$, β_{2t}=-$\gamma_{3,t+1|t}\mu^2_{t+1|t}/6$ e β_{3t}=$\gamma_{4,t+1|t}\mu^3_{t+1|t}/24$, where for notational convenience, we denote $x_{t+1}=1/\sigma_{t+1|t}$.

Whether $\mu_{t+1|t}$ is small, as in the case of short investment horizons, then β_{2t} and β_{3t} cac be safely be ignored, resulting in

$$1-F(-\mu_{t+1|t}x_{t+1})\approx 1-\Phi(-\mu_{t+1|t})(\beta_{0t}+\beta_{1t}x_{t+1}).$$

Thus, conditional skewness affects sign predictability through β_{0t}, and conditional kurtosis affects sign predictability through β_{1t}. When there are no conditional dynamics in skewness and kurtosis, the above equation is reduced to

$$1-F(-\mu_{t+1|t}x_{t+1})\approx 1-\Phi(-\mu_{t+1|t}x_{t+1})(\beta+\beta x_{t+1}), \tag{70}$$

for some time-invariant quantities β_0 and β_1.

We use Equation (70) as our second model for sign prediction, i.e., we generate forecasts of the probability of positive returns as

$$\widehat{Pr}(R_{t+1|t}>0=x_{t+1})\approx 1-\Phi(-\widehat{\mu}_{t+1|t}\hat{x}_{t+1})(\widehat{\beta}_0+\widehat{\beta}_1\widehat{x}_{t+1}), \tag{71}$$

where $\hat{x}_{t+1|t}=1/\hat{\sigma}_{t+1|t}$, and where $\hat{\mu}_{t+1|t}$ and $\hat{\sigma}_{t=t+1|t}$ are as defined earlier. We refer to these as forecasts from the “extended" model. The parameters β_0 and β_1 are estimated by regressing $1-I(R_t>0)$ on $\Phi\left(-\hat{\mu}_t\hat{x}_t\right)$ and $\Phi\left(-\hat{\mu}_t\hat{x}_t\right)\hat{x}_t$ for $t=1,\ldots,$k. Although we have not explicitly placed any constraints on this model to require $\Phi\left(-\hat{\mu}_t\hat{x}_t\right)\left(\hat{\beta}_0+\hat{\beta}_1\hat{x}_t\right)$ to lie between 0 and 1, we ensure it by applying the logistic function in extended model results to forecasts lie between 0 and 1.

4.4 FORECAST EVALUATION

We perform out-of-sample comparison of the forecast performance of Equation (69), Equation (71), and the logistic function[4] for the sign of return. Both are compared against baseline forecasts [Equation (64)]. This is done for one-, -three, six-month, and twelve-month returns. We assess the performance of the forecasting models using Brier scores:

$$\text{Brier(Sq)} = \frac{1}{T-k}\sum_{t=k}^{T} 2\left(\widehat{\Pr}\left(R_{t+1|t} > 0\right) - z_{t+1}\right)^2,$$
$$\text{Brier(Abs)} = \frac{1}{T-k}\sum_{t=k}^{T} \left|\widehat{\Pr}\left(R_{t+1|t} > 0\right) - z_{t+1}\right|,$$

where $z_{t+1} = I(R_{t+1} > 0)$. The latter is the traditional Brier score for evaluating the performance of probability forecasts and is analogous to the usual RMSFE. A score of 0 for Brier(Sq) occurs when perfect forecasts are made: where at each period, correct probability forecasts of 0 or 1 are made. The worst score is 1 and occurs if at each period probability forecasts of 0 or 1 are made but turn out to be wrong each time.

Note that if we follow the usual convention where a correct probability forecast of $I(R_{t+1} > 0)$ is 1 that is greater than 0.5, then accurate forecasts will have an individual Brier(Sq) score between 0 and 0.5. In contrast, incorrect predictions have individual scores between 0.5 and 1. We only consider a modified version of the Brier score, which we call Brier(Abs). In the following sections, we introduced the data set and empirical findings.

4.5 DATA AND RESULTS

The data set consists of monthly closing prices observed for yields of future DI contracts. Based on the observed rates for the available maturities, the data were converted to fixed maturities of 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 36, 48, and 60 months through interpolations using cubic splines. The database contains the maturities with the highest liquidity for January 2004 through December 2021 (T = 216 observations) and represents the most liquid DI contracts negotiated during the analyzed period. We assess the model's performance by splitting the sample into two parts: the first includes 108 observations used to estimate the parameters. The second part is used to analyze the performance out-of-sample of bond portfolios obtained from the model, with 108 observations.

[4] We also add the logistic function as Peter F Christoffersen and Francis X Diebold (2006) did: $F(x) = \frac{\exp(x)}{1+\exp(x)}$ where $x = \frac{\hat{\mu}_{t+1|t}}{\hat{\sigma}_{t+1|t}}$.

The Table 13 reports summary statistics for Brazil's yield curve. We examine monthly data constructed using the spline method. For each maturity, we show mean, standard deviation, skewness, raw kurtosis, minimum, maximum, and three auto-correlations coefficients, $\hat{\rho}_1$, $\hat{\rho}_6$, $\hat{\rho}_{12}$. Also, the table reports proxy estimates for the level, slope, and curvature of the yield curve. The proxies are defined as follows: for level, the highest maturity bond (60 months); for slope, the difference between the bond of 60 months and the bond of 3 months; and for curvature, two times the bond of 18 months minus the sum of a bond of 3 months and bond of 60 months.

Also, Figure 16 presents a plot of the data set and illustrates how yield levels evolve throughout the sample. Although the yield series change heavily over time for each maturity, a robust common pattern in the 14 series is apparent.

Table 13 – Descriptive statistics for the term structure of interest rates

Maturity	Mean	Std.dev.	Min	Max	Skewness	Kurtosis	$\hat{\rho}_1$	$\hat{\rho}_6$	$\hat{\rho}_{12}$
M3	12.19	4.52	4.30	27.50	0.90	4.11	0.96	0.70	0.48
M6	12.20	4.47	4.30	28.30	0.90	4.24	0.96	0.69	0.48
M9	12.24	4.42	4.40	29.00	0.92	4.56	0.95	0.68	0.48
M12	12.30	4.39	4.50	29.60	0.98	4.93	0.94	0.67	0.47
M15	12.38	4.34	4.50	30.40	1.06	5.39	0.94	0.65	0.46
M18	12.48	4.31	4.60	31.30	1.16	5.95	0.93	0.64	0.45
M21	12.57	4.28	4.70	32.30	1.29	6.64	0.93	0.62	0.44
M24	12.65	4.27	4.90	33.40	1.43	7.41	0.92	0.61	0.43
M27	12.74	4.25	5.00	34.30	1.57	8.15	0.92	0.60	0.42
M30	12.80	4.26	5.10	35.10	1.69	8.86	0.91	0.59	0.41
M36	12.92	4.28	5.30	36.70	1.95	10.35	0.90	0.57	0.40
M42	13.02	4.34	5.60	38.40	2.20	11.93	0.90	0.55	0.38
M48	13.10	4.38	5.70	39.40	2.37	12.96	0.89	0.54	0.38
M60 (Level)	13.18	4.35	6.00	39.40	2.44	13.12	0.89	0.55	0.38
Slope	0.99	2.10	-3.90	11.90	0.75	6.77	0.84	0.35	0.01
Curvature	-0.24	1.17	-3.00	3.10	0.18	2.86	0.87	0.40	0.08

Source: Elaborated by the author.

Legend: The table reports summary statistics for Brazil yield curve over the period 2004-2021. We examine monthly data, constructed using the spline method. For each maturity we show mean, standard deviation, skewness, raw kurtosis, minimum, maximum, and three auto-correlations coefficients, $\hat{\rho}_1$, $\hat{\rho}_6$, $\hat{\rho}_{12}$. Also the table reports proxy estimates for level, slope, and curvature of the yield curve. The proxies are defined as follows: for level, the highest maturity bond (60 months); for slope, the difference between the bond of 60 months and the bond of 3 months; and for curvature, two times the bond of 18 months minus the sum of bond of 3 months and bond of 60 months.

Figure 16 – Brazilian yield curves

Source: Elaborated by the author.
Legend: Monthly Brazil yield curves from January 2004 through December 2022.

We also present the absolute Brier score results and the cumulative Brier score. Table 14 shows the out-of-sample comparison of the forecast performance of Equation (69), Equation (71), and the logistic function for the sign of return. We assess the performance of the forecasting models using Brier scores. The worst score is 1. Correct forecasts have individual scores between 0 and 0.5, whereas incorrect forecasts carry scores between 0.5 and 1. The Logistic, Non-parametric, and Extended are relative to Baseline, so values below 1 outperform the benchmark model. The results suggest that models with information of skewness and kurtosis of returns outperform the benchmark model, mainly in long maturities and short horizons of forecasts.

Figure 17, Figure 18, Figure 19,Figure 20, and Figure 21 present the cumulative absolute brier score relative to the benchmark model to uncover the path of all models throughout the out-of-sample period. Hence, the competitor model outperforms the benchmark for values above zero. In all month's horizons forecasts, the results suggest the extended model outperforms the benchmark model when the Selic rate rises. However, it is not a rule. The results of the extended model are consistent for short horizons and become better for long horizons.

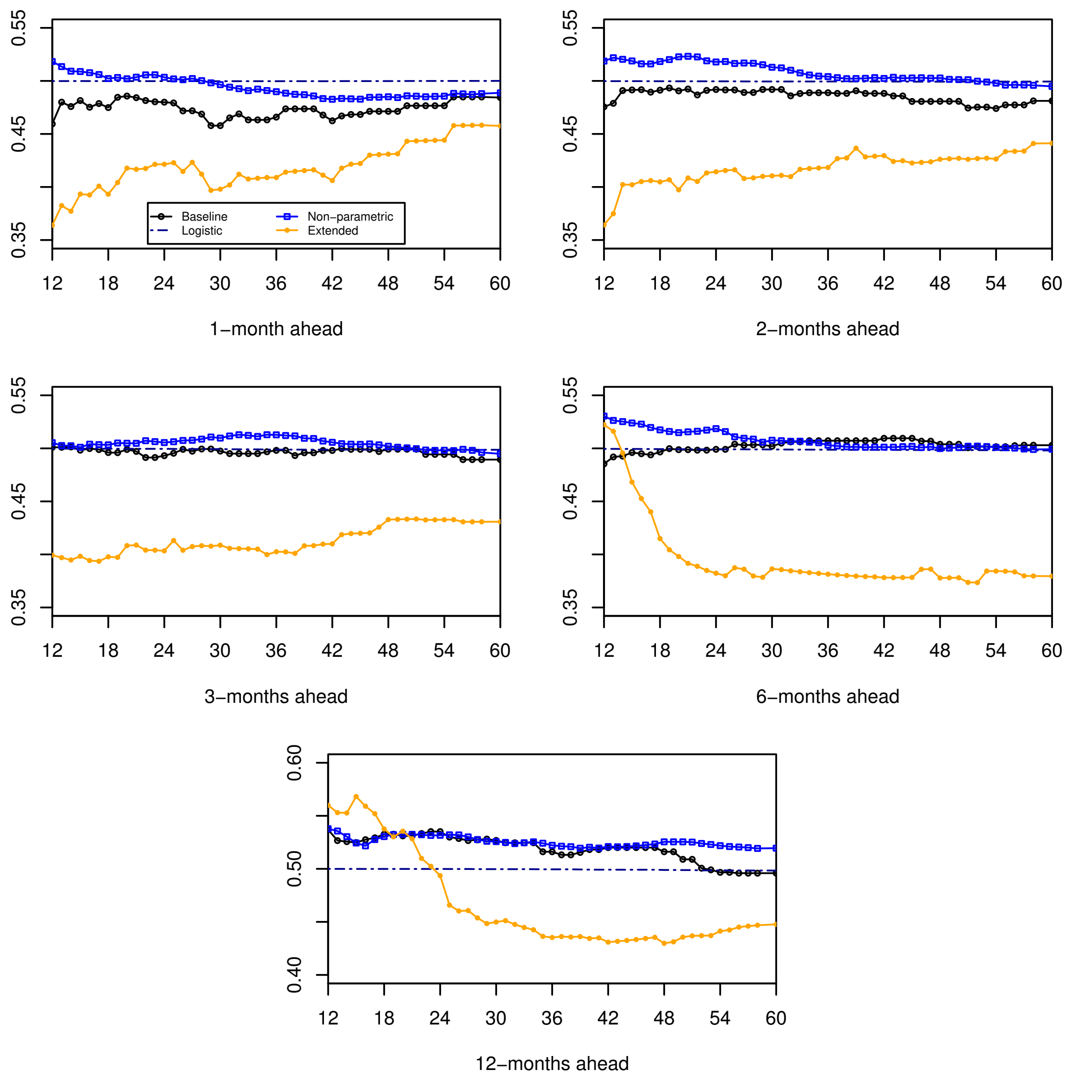
Table 14 – Absolute Brier score results

Model	Maturity							
	M12	M18	M24	M30	M36	M42	M48	M60
Panel A: 1-month ahead forecasts								
Baseline	0.460	0.475	0.480	0.458	0.466	0.462	0.471	0.484
Logistic	0.499	0.499	0.499	0.499	0.499	0.499	0.500	0.500
Non – parametric	0.518	0.502	0.504	0.496	0.489	0.482	0.484	0.489
Extended	0.364	0.393	0.421	0.397	0.408	0.406	0.431	0.457
Panel B: 2-months ahead forecasts								
Baseline	0.476	0.491	0.492	0.492	0.489	0.488	0.481	0.481
Logistic	0.498	0.498	0.499	0.499	0.500	0.500	0.500	0.500
Non – parametric	0.519	0.518	0.518	0.512	0.503	0.500	0.500	0.492
Extended	0.364	0.405	0.414	0.410	0.418	0.429	0.425	0.441
Panel C: 3-months ahead forecasts								
Baseline	0.501	0.496	0.493	0.498	0.498	0.498	0.499	0.489
Logistic	0.498	0.498	0.498	0.498	0.498	0.498	0.499	0.499
Non – parametric	0.505	0.503	0.505	0.509	0.513	0.504	0.500	0.493
Extended	0.399	0.397	0.403	0.408	0.401	0.409	0.432	0.429
Panel D: 6-months ahead forecasts								
Baseline	0.485	0.497	0.499	0.502	0.507	0.510	0.504	0.503
Logistic	0.498	0.498	0.498	0.498	0.498	0.500	0.500	0.500
Non – parametric	0.530	0.516	0.515	0.504	0.502	0.499	0.497	0.496
Extended	0.522	0.414	0.381	0.385	0.379	0.375	0.375	0.376
Panel D: 12-months ahead forecasts								
Baseline	0.537	0.533	0.535	0.527	0.516	0.520	0.516	0.496
Logistic	0.497	0.497	0.497	0.497	0.498	0.498	0.498	0.499
Non – parametric	0.538	0.530	0.532	0.525	0.522	0.519	0.520	0.513
Extended	0.560	0.537	0.493	0.448	0.433	0.428	0.428	0.449

Source: Elaborated by the author.

Legend: We perform an out-of-sample comparison of the forecast performance of equations Equation (69), Equation (71), and the logistic function for the sign of return. Both are compared against baseline forecasts [Equation (64)]. We assess the performance of the forecasting models using Brier scores. The worst score is 1. Correct forecasts have individual scores between 0 and 0.5, whereas incorrect forecasts carry scores between 0.5 and 1. The Logistic, Non-parametric, and Extended are relative to Baseline, so values below 1 outperform the benchmark model.

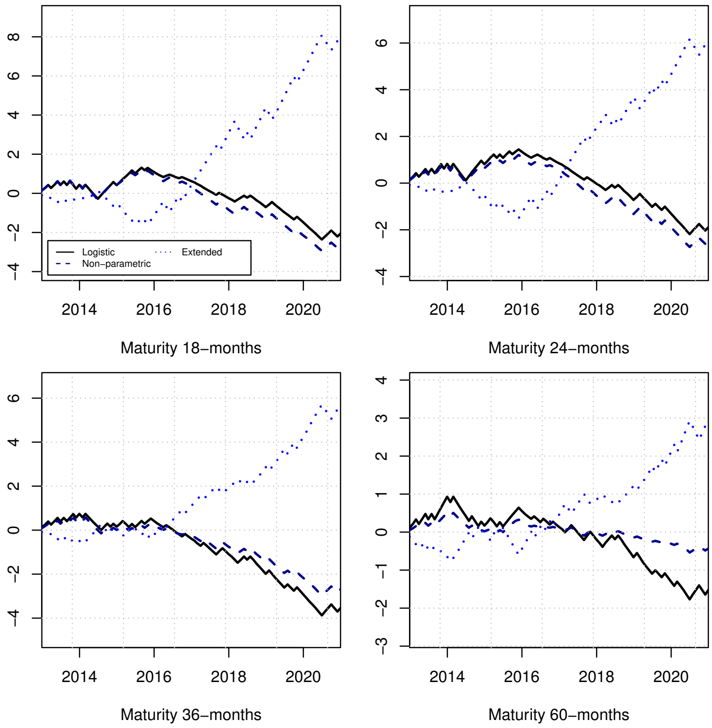
Figure 17 – Absolute Brier score results



Source: Elaborated by the author.

Legend: This figure reports the cumulative brier score relative to the baseline model at one, two, three, six, and twelve months ahead of maturities between 12 and 60 months.

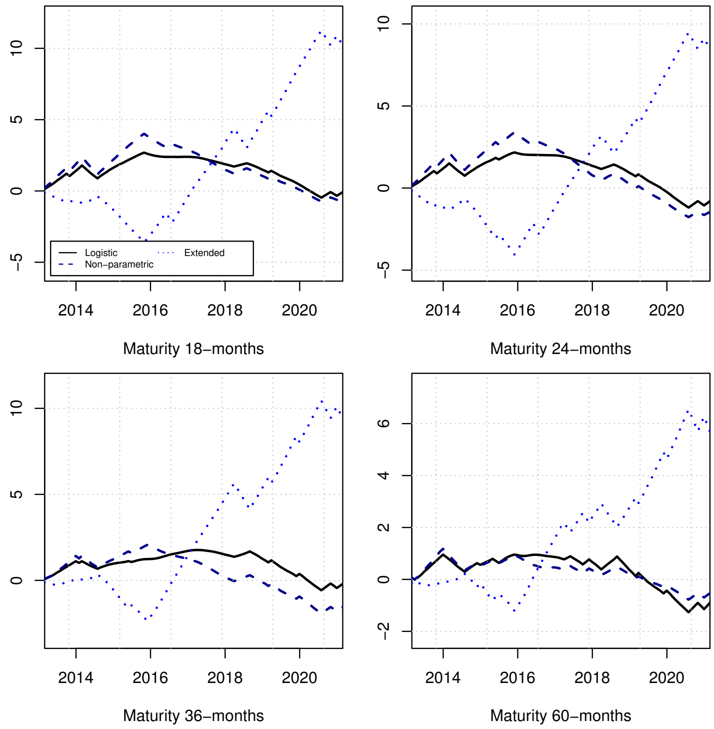
Figure 18 – Cumulative Absolute Brier Scores: 1-month ahead



Source: Elaborated by the author.

Legend: We use the cumulative absolute brier score relative to the benchmark model to uncover the path of all models throughout the out-of-sample period. Hence, the competitor model outperforms the benchmark for values above zero. The top left chart presents results of maturity of 18 months, the top right present results of maturity of 24 months, the bottom left presents results of maturity of 36 months, bottom right presents results of maturity of 60 months.

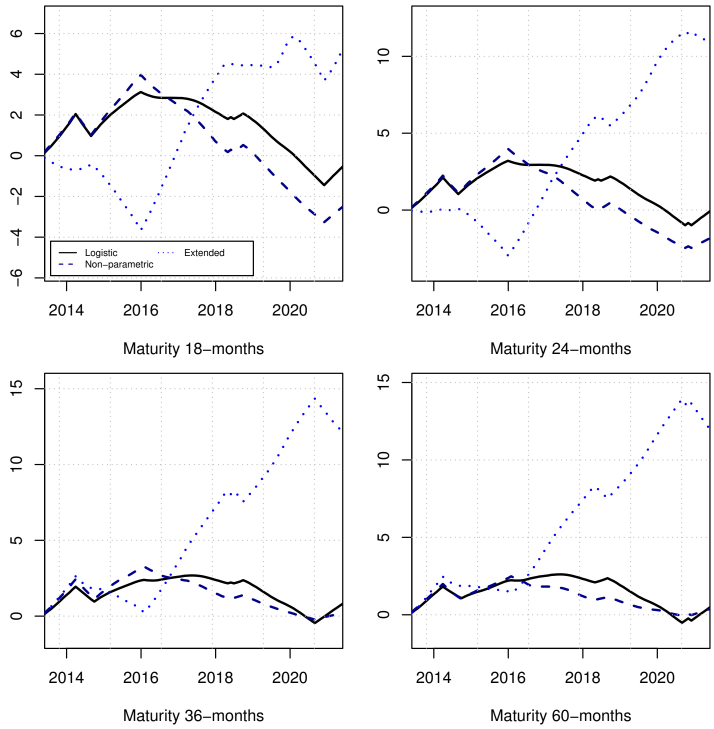
Figure 19 – Cumulative Absolute Brier Scores: 3-months ahead



Source: Elaborated by the author.

Legend: We use the cumulative absolute brier score relative to the benchmark model to uncover the path of all models throughout the out-of-sample period. Hence, the competitor model outperforms the benchmark for values above zero. The top left chart presents results of maturity of 18 months, the top right present results of maturity of 24 months, the bottom left presents results of maturity of 36 months, bottom right presents results of maturity of 60 months.

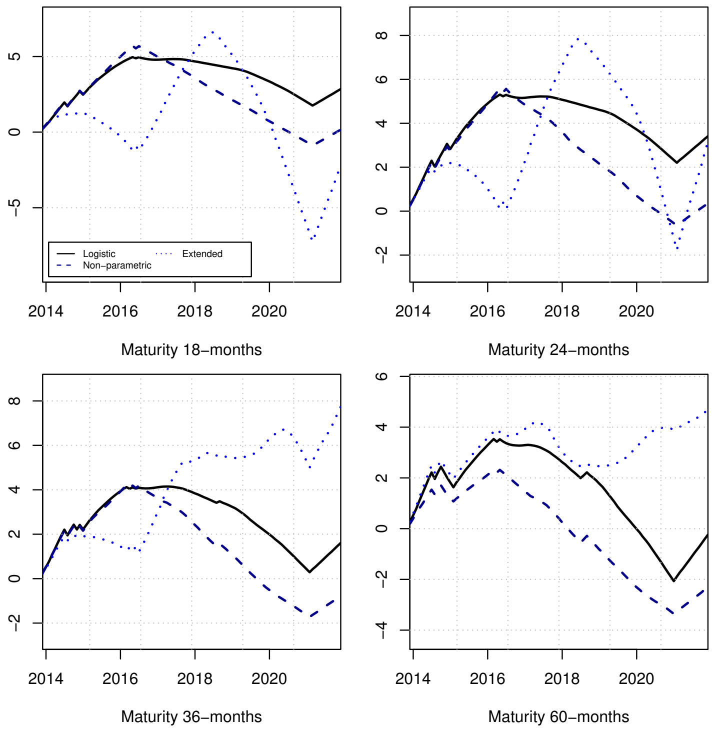
Figure 20 – Cumulative Absolute Brier Scores: 6-months ahead



Source: Elaborated by the author.

Legend: We use the cumulative absolute brier score relative to the benchmark model to uncover the path of all models throughout the out-of-sample period. Hence, the competitor model outperforms the benchmark for values above zero. The top left chart presents results of maturity of 18 months, the top right present results of maturity of 24 months, the bottom left presents results of maturity of 36 months, bottom right presents results of maturity of 60 months.

Figure 21 – Cumulative Absolute Brier Scores: 12-months ahead



Source: Elaborated by the author.

Legend: We use the cumulative absolute brier score relative to the benchmark model to uncover the path of all models throughout the out-of-sample period. Hence, the competitor model outperforms the benchmark for values above zero. The top left chart presents results of maturity of 18 months, the top right present results of maturity of 24 months, the bottom left presents results of maturity of 36 months, bottom right presents results of maturity of 60 months.

4.6 CONCLUDING REMARKS

The accuracy of yield curve forecasts, not only of magnitude but specifically of the direction of change in financial assets, emerges as a research source in the prediction field. We aim to explore the literature gap on forecasting the yield curves using conditional means and conditional volatility forecasts as inputs to predict the return direction in the fixed income.

This paper assesses the direction-of-change forecasts based on conditional variance from the Dynamic Nelson-Siegel model. Although the literature focuses on forecasting the level of yield curves, which is a difficult task, we propose forecasts for the direction-of-change of the yield curve returns. The results suggest that models with information of skewness and kurtosis of returns outperform the benchmark model, mainly in long maturities and short forecast horizons. Also, In all month's horizons forecasts, the results suggest all models have stable performance against the benchmark model when the Selic rate rises; however, it is not a rule.

5 DOES THE DECOMPOSITION OF THE BREAKEVEN INFLATION HELP US TO FORECASTING INFLATION? EVIDENCE OF EXPECTED INFLATION AND INFLATION RISK PREMIUM IN THE BRAZILIAN ECONOMY

The estimation of the inflation risk premium has proved to be a challenging problem, see Breach, D'Amico, and Orphanides (2020). Models with different specifications or analyzing different periods have found different results. For instance, with data before the 2008 financial crisis, estimations involving structural models obtained results with a high magnitude of inflation risk premium, see Ang, Bekaert, and Wei (2008), Bekaert and Wang (2010) and Chernov and Mueller (2012). On the other hand, studies using more recent data suggest premiums for the risk of inflation of smaller magnitude and sometimes even negative, see Grishchenko and Huang (2013), Michael Abrahams et al. (2016) and Breach, D'Amico, and Orphanides (2020). Since the relation of yield curves and the macroeconomics, see Litterman and Scheinkman (1991), Cochrane and Piazzesi (2005), Ang, Piazzesi, and Wei (2006), Cochrane and Piazzesi (2009), Cieslak and Povala (2011), Crump, Eusepi, and Moench (2018) and Bernanke (1990), it is essential to understand of the movements the term structure to improve forecasting, derivatives pricing, hedging, and fiscal and monetary policy.

In this paper, we use an arbitrage-free affine Gaussian model for the term structure (ATSM) to jointly model nominal and real interest rates, decompose the breakeven inflation, analyze the term premium dynamics, and forecast inflation. For model estimation, we use the recent approach for asset pricing based on linear regressions proposed by Adrian, Crump, and Moench (2015) and Michael Abrahams et al. (2016). They present a method that allows computational gains in estimating factor models for the term structure while allowing the term premium to vary over time and serial dependence on the factors. Several other studies that estimated ATSM-class models for other economies used maximum likelihood, see Joyce, Lildholdt, and Sorensen (2010), Kaminska (2013) and d'Amico, Kim, and Wei (2018), which involves high-dimensional nonlinear optimization over a maximum likelihood function that can have many maxima locations, see Hamilton and Jing Cynthia Wu (2012). The approach proposed by Michael Abrahams et al. (2016) considerably reduces these difficulties in estimating models of this class. To our knowledge, no study uses this procedure to address this question in the Brazilian economy.

The literature on the Brazilian economy suggests that the risk premium varies over time; see Lima and Issler (2003), Benjamin Miranda Tabak and Andrade (2003), Marçal and Pereira (2007) and Benjamin Tabak (2009) for early references. Vicente and Graminho (2015) and João F Caldeira (2020), for instance, suggest that the inflation risk premium is time-varying. However, Vicente and Graminho (2015) does not find evidence of a liquidity premium in Brazil. Also, They suggest that the inflation risk premium is small for short horizons and is time-varying for long horizons, and inflation expectations are the main component of breakeven inflation. Surveys of expected inflation naturally emerge as a predictor of future inflation. See, among other Ang, Bekaert, and Wei (2007) and Chun (2012). The FOCUS survey, conducted by the Central Bank of Brazil, emerges as the main competitor to forecasting inflation in our research, see Carvalho, Minella, et al. (2009). Also, breakeven inflation is naturally a competitor; see Vicente and Guillen (2013) and João F Caldeira and Furlani (2013).

Our innovation is in the method used for the estimation, which lies in the approach of Michael Abrahams et al. (2016). Our main findings can be summarized as follows. First, we disentangle the influence of term premiums on nominal and real rates. Then, we show the decomposition of the BEIR in expected inflation and inflation risk premium. Lastly, we use the derivation of the expected inflation to predict the IPCA. Following the literature, the results suggest that the premiums are time-varying and increase along maturities. The inflation risk premium is also time-varying, with negative values in specific periods. The expected inflation and the Focus survey outperform the RW forecasts; however, the Focus approach is a major workhorse. The second section introduces the Michael Abrahams et al. (2016) AFNS model estimation following this introductory section. In the third section, we present in-sample results, term premium results, and out-of-sample inflation forecasts. In the fourth section, we conclude.

5.1 AFFINE GAUSSIAN MODELS FOR TERM STRUCTURE

Affine term structure models (ATSMs), since Duffee (2002), are the most commonly used class of models in the literature for decomposing interest rates on government bonds. More recently, the approach developed by Adrian, Crump, and Moench (2015) has been widely used to decompose interest rates into their components: expectation and forward premium. This section presents an ATMS model specification following the exposition of M. Abrahams et al. (2015) and Michael Abrahams et al. (2016).

The price, at time t, of a zero-coupon bond with maturity n is denoted by $P_t^{(n)}$. As is common in Gaussian models for the term structure, it is assumed that the vector of state variables is governed by an autoregressive process of the type VAR(1):

$$X_{t+1} - \mu_X = \Phi\left(X_t - \mu_X\right) + \nu_{t+1}, \qquad \nu_{t+1} \sim \mathcal{N}(0,\Sigma) \tag{72}$$

where the shocks ν_{t+1} are conditionally Gaussian, homoscedastic and independent over time. A single pricing mechanism is introduced to enforce the absence of arbitrage which governs all traded assets:

$$P_t^{(n)} = \mathbb{E}\left\{M_{t+1} P_{t+1}^{(n-1)}\right\}. \tag{73}$$

The stochastic discount factor M_t (pricing kernel) is a function of the short-term interest rate and the risk perceived by the market:

$$M_{t+1} = \exp\left(-r_t - \frac{1}{2}\lambda_t'\lambda_t - \lambda_t'\Sigma^{-1/2}\nu_{t+1}\right), \tag{74}$$

where $r_t = \ln P_t^{(n)}$ denotes the risk-free interest rate that is continuously compounded. In Gaussian ATSMs the log price, $P_t^{(n)}$, of a risk-free discount bond with remaining time to maturity n follows $\log P_t^{(n)} = A_n + B_n^{prime} X_t$ which implies that:

$$r_t = \delta_0 + \delta_1' X_t. \tag{75}$$

The risk market price vector, λ_t, is an essentially affine function of the factors, as in Duffee (2002):

$$\lambda_t = \Sigma^{-1/2}(\lambda_0 + \lambda_1 X_t), \tag{76}$$

where λ_0 and λ_1 have dimensions $K \times 1$ and $K \times K$, respectively. Further defines:

$$\tilde{\mu} = (I_K - \Phi)\mu_X - \lambda_0, \tag{77}$$

$$\tilde{\Phi} = \Phi - \lambda_1. \tag{78}$$

These parameters govern the dynamics of the pricing factors under the risk-neutral and feature prominently in the recursive pricing relationships derived below.

Given the above assumptions, it can be shown that interest rates on zero-coupon bonds are affine functions of the factors, see Ang and Piazzesi (2003):

$$y_t^{(n)} = -\frac{1}{n}\left(A_n + B_n' X_t\right), \tag{79}$$

where the coefficients A_n and B_n follow the recursive equations:

$$A_n = A_{n-1} + B_{n-1}'\tilde{\mu} + \frac{1}{2}B_{n-1}'\Sigma B_{n-1} - \delta_0, \quad A_0 = 0 \tag{80}$$

$$B_n' = B_{n-1}'\tilde{\Phi} - \delta_1', \quad B_0 = 0_{K\times 1}. \tag{81}$$

Recently, there has been a growing interest in the literature in recovering expectations about future inflation rates from the nominal and real term structure of interest rates, see Michael Abrahams et al. (2016) and Breach, D'Amico, and Orphanides (2020). Let Q_t be a time price index t and let $P_{t,R}^{(n)}$ be the price in t of an inflation-indexed bond with face value 1, which pays the amount $\frac{Q_{t+n}}{Q_t}$ at maturity, $t+n$. The price of such a title satisfies the following:

$$P_{t,R}^{(n)} = \mathbb{E}_t\left\{\exp\left(-r_t - \ldots - r_{t+n-1}\right)\frac{Q_{t+n}}{Q_t}\right\}. \tag{82}$$

Denote the log-inflation for one period by $\pi_t = \ln\left(\frac{Q_t}{Q_{t-1}}\right)$, therefore:

$$\frac{Q_{t+n}}{Q_t} = \exp\left(\sum_{i=1}^{n}\pi_{t+i}\right). \tag{83}$$

As in the case of nominal bonds, the prices of inflation-indexed bonds are exponentially affine in terms of pricing factors:

$$\log P_{t,R}^{(n)} = A_{n,R} + B_{n,R}' X_t. \tag{84}$$

Thus, one-period inflation is also a linear function of the state variables:

$$\pi_t = \pi_0 + \pi_1' X_t,$$

where π_0 is a scalar and π_1 is a vector of dimension ($K \times 1$). According to Michael Abrahams et al. (2016), it is possible to derive recursions for the prices of inflation-linked bonds by rewriting the equation (82) in terms of the price of another inflation-linked bond traded one period ahead:

$$P_{t,n}^{R} = \mathbb{E}_t \left\{ \exp\left(-r_t + \pi_{t+1}\right) P_{t+1,R}^{(n-1)} \right\}. \tag{85}$$

Solving this equation and combining the coefficients, we arrive at the coefficients of Equation (84), which are determined by the following system of equations in differences:

$$A_{n,R} = A_{n-1,R} + B_{n-1,R}^{\pi\prime} \tilde{\mu} + \frac{1}{2} B_{n-1,R}^{\pi\prime} \Sigma B_{n-1,R}^{\pi} - \delta_{0,R}, \quad A_{0,R} = 0 \tag{86}$$

$$B_{n,R}' = B_{n-1,R}^{\pi\prime} \tilde{\Phi} - \delta_1', \quad B_{0,R} = 0_{K\times 1}. \tag{87}$$

where $\delta_{0,R} = \delta_0 - \pi_0$ and $B_{n,R}^{\pi} = (B_{n,R} + \pi_1)\ \forall n$. Making the parameters referring to the risk market price, λ_0 and λ_1, equal to zero in the systems of equations (80)-(81) and **??**-(87), we obtain the risk-adjusted pricing parameters (makes the mapping of the risk-neutral measure, $\mathbb{Q}$, to the physical measure, $\mathbb{P}$).

5.2 ESTIMATION

5.2.1 Nominal bonds returns

Recall that log excess one-period holding returns are defined as

$$rx_{t+1}^{(n-1)} = \log P_{t+1}^{(n-1)} - \log P_t^{(n)} - r_t. \tag{88}$$

Plugging Equation (84) we obtain

$$rx_{t+1}^{(n-1)} = \left(A_{n-1} - A_n - \delta_0\right) - \left(B_n' + \delta_1'\right) X_t + B_{n-1}' X_{t+1} \tag{89}$$

Thus, imposing the recursive equations yields Equation (80) and Equation (81)

$$rx_{t+1}^{(n-1)} = \alpha_{n-1} - B_{n-1}' \tilde{\Phi} X_t + B_{n-1}' X_{t+1}, \tag{90}$$

where

$$\alpha_{n-1} = -\left(B_{n-1}' \tilde{\mu} + \frac{1}{2} B_{n-1}' \Sigma B_{n-1} \right). \tag{91}$$

5.2.2 Inflation-indexed bonds returns

Log excess one period holding returns on inflation indexed securities are then given by

$$rx_{t+1,R}^{(n-1)} = \log P_{t+1,R}^{(n-1)} - \log P_{t,R}^{(n)} - r_t. \tag{92}$$

Thus, imposing the recursive equations yields Equation (86) and Equation (87)

$$rx_{t+1,R}^{(n-1)} = \alpha_{n-1,R} - \left(B_{n-1,R} + \pi_1\right)' \tilde{\Phi} X_t + B'_{n-1,R} X_{t+1}, \tag{93}$$

where

$$\alpha_{n-1,R} = -\left(\pi_0 + \left(B_{n-1,R} + \pi_1\right)' \tilde{\mu} + \frac{1}{2}\left(B_{n-1,R} + \pi_1\right)' \Sigma \left(B_{n-1,R} + \pi_1\right)\right). \tag{94}$$

5.2.3 Initial Conditions

To obtain initial conditions note that adding inflation to both sides of equation Equation (92) and combining with equations Equation (86), and Equation (87), we obtain

$$rx_{t+1,R}^{(n-1)} + \pi_{t+1} = \alpha_{n-1,R}^{\pi} - \left(B_{n-1,R} + \pi_1\right)' \tilde{\Phi} X_t + \left(B_{n-1,R} + \pi_1\right)' X_{t+1}, \tag{95}$$

where

$$\alpha_{n-1,R}^{\pi} = -\left(\left(B_{n-1,R} + \pi_1\right)' \tilde{\mu} + \frac{1}{2}\left(B_{n-1,R} + \pi_1\right)' \Sigma \left(B_{n-1,R} + \pi_1\right)\right). \tag{96}$$

Stacking log excess holding period returns on nominal bonds from equation Equation (90) and on inflation-indexed bonds from equation Equation (95) into the vector R^{π}, we thus obtain

$$R_{t+1}^{\pi} = \alpha - B\tilde{\Phi} X_t + B X_{t+1}, \tag{97}$$

where

$$\alpha = -\left(B\tilde{\mu} + \frac{1}{2}\gamma\right), \tag{98}$$

$$B = \left(B_1, \ldots, B_{N_N}, B_{1,R} + \pi_1, \ldots, B_{N_R,R} + \pi_1\right)', \tag{99}$$

$$\gamma = \left(B_1' \Sigma B_1, \ldots, B'_{N_N} \Sigma B_{N_N}, \left(B_{1,R} + \pi_1\right)' \Sigma \left(B_{1,R} + \pi_1\right), \ldots, \left(B_{N_R,R} + \pi_1\right)' \Sigma \left(B_{N_R,R} + \pi_1\right)\right)'. \tag{100}$$

For initial conditions we use an approach similar to Adrian, Crump, and Moench (2015). To provide initial estimates of our parameters we stack the observed return data as

$$\mathbf{R}^{\pi} = \alpha^{\pi} \iota'_T - B\tilde{\Phi} X_{-} + B X + E \tag{101}$$

where R^{π} is $N \times T, X_{-}$and X are $K \times T$ matrices of the stacked X_{t-1} 's and X_t 's, respectively, and ι_T is a $T \times 1$ vector of ones. Using the estimated residuals, $\hat{E}_{\text{ols}}$, from this regression we obtain $\hat{\Sigma}_e = T^{-1} \cdot \hat{E}_{\text{ols}} \hat{E}'_{\text{ols}}$. Our initial value for $\tilde{\Phi}$ is

$$\hat{\tilde{\Phi}}_{\text{gls}} = -\left(\hat{B}'_{\text{ols}} \hat{\Sigma}_e^{-1} \hat{B}_{\text{ols}}\right)^{-1} \hat{B}'_{\text{ols}} \hat{\Sigma}_e^{-1} \widehat{B\tilde{\Phi}_{\text{ols}}} \tag{102}$$

We then run an additional SUR on ι_T and $\left(-\hat{\tilde{\Phi}}_{\text{gls}} X_{-} + X\right)$ to obtain initial values for α and B which we label $\hat{\alpha}_{\text{gls}}$ and $\hat{B}_{\text{gls}}$. Finally, we produce an initial value for $\tilde{\mu}$ as

$$\hat{\tilde{\mu}}_{\text{gls}} = -\left(\hat{B}'_{\text{gls}} \hat{\Sigma}_e^{-1} \hat{B}_{\text{gls}}\right)^{-1} \hat{B}'_{\text{gls}} \hat{\Sigma}_e^{-1} \left(\hat{\alpha}_{\text{gls}} + \frac{1}{2}\hat{\gamma}_{\text{gls}}\right), \tag{103}$$

where $\hat{\gamma}_{\text{gls}}$ is formed using $\hat{B}_{\text{gls}}$ and $\hat{\Sigma}$ (see equation Equation (100)). We also need initial values for the parameters (δ_0, δ'_1) governing the nominal short rate. Since the nominal short rate is directly observed, this is simply achieved by performing an OLS regression of the short rate onto a constant and the vector of pricing factors as in Adrian, Crump, and Moench (2015).

The parameters $\tilde{\mu}$ and $\tilde{\Phi}$ are related to the market price of risk parameters λ_0 and λ_1 via the relationships $\tilde{\mu} = (I_K - \Phi)\mu_X - \lambda_0$ and $\tilde{\Phi} = \Phi - \lambda_1$. Since the pricing factors X are observed and follow the joint vector autoregression given by Equation (72), the OLS estimator of μ_X is simply given by the sample mean of the factors X and the OLS estimator of Φ is obtained by regressing the demeaned observations of X on their one period lags equation by equation. We stack the estimated innovations into the matrix $\hat{V}$ and construct an estimator of the state variable variance-covariance matrix $\hat{\Sigma} = T^{-1} \cdot \hat{V}\hat{V}'$. Given estimates $\hat{\mu}_X$ and $\hat{\Phi}$, we then obtain estimates of the market price of risk parameters via

$$\begin{aligned} \hat{\lambda}_0 &= \left(I_K - \hat{\Phi}\right)\hat{\mu}_X - \hat{\tilde{\mu}}_{\text{gls}}, \\ \hat{\lambda}_1 &= \hat{\Phi} - \hat{\tilde{\Phi}}_{\text{gls}}. \end{aligned} \tag{104}$$

In our empirical application we skip the estimation of parameters via numerical maximization of the likelihood as per Michael Abrahams et al. (2016) and use the values of OLS initial conditions estimation. Also we use the sum of squared real return fitting errors as the criterion function to estimate π_0 and π_1 as per M. Abrahams et al. (2015). We provide explicit expressions for real yields as linear-quadratic functions of π_0 and π_1 (given estimates for $\hat{\tilde{\Phi}}_{\text{gls}}, \hat{\tilde{\mu}}_{\text{gls}}, \hat{\delta}_{0,\text{ols}}, \hat{\delta}_{1,\text{ols}}$) which may be used for numerical optimization. We then solve for the estimated π_0 and π_1 with the initial conditions via,

$$(\hat{\pi}_0, \hat{\pi}'_1)' = \arg\min_{\pi_0, \pi_1} \sum_{i=1}^{N_R} \sum_{t=1}^{T} \left(rx_{t+1,R}^{(n-1)} - g\left(\pi_0, \pi_1; \hat{\tilde{\Phi}}_{\text{gls}}, \hat{\tilde{\mu}}_{\text{gls}}, \hat{\delta}_{0,\text{ols}}, \hat{\delta}_{1,\text{ols}}, n_i, t\right)\right)^2 \tag{105}$$

where $g(\cdot)$ can be found by using the recursive equations Equation (86) and Equation (87). Next section presents the dataset and results.

5.3 DATA AND RESULTS

We use end-of-month values from 2006:01 to 2022:04 for a total of T = 196 monthly observations. In the estimation, a cross-section of N_N = 11 one-month excess holding period returns for nominal rates with maturities n = 6,12,24,...,120 months and N_R = 9 excess returns on NTNB′s with maturities n = 24,...,120 months is used. The SELIC rate is used as the nominal risk-free rate. The price index Q_t used to calculate NTNB′s payouts is IPCA index, which is available from the IBGE site. See Table Table 15 for statistics of interest rates.

Table 15 – Descriptive Statistics

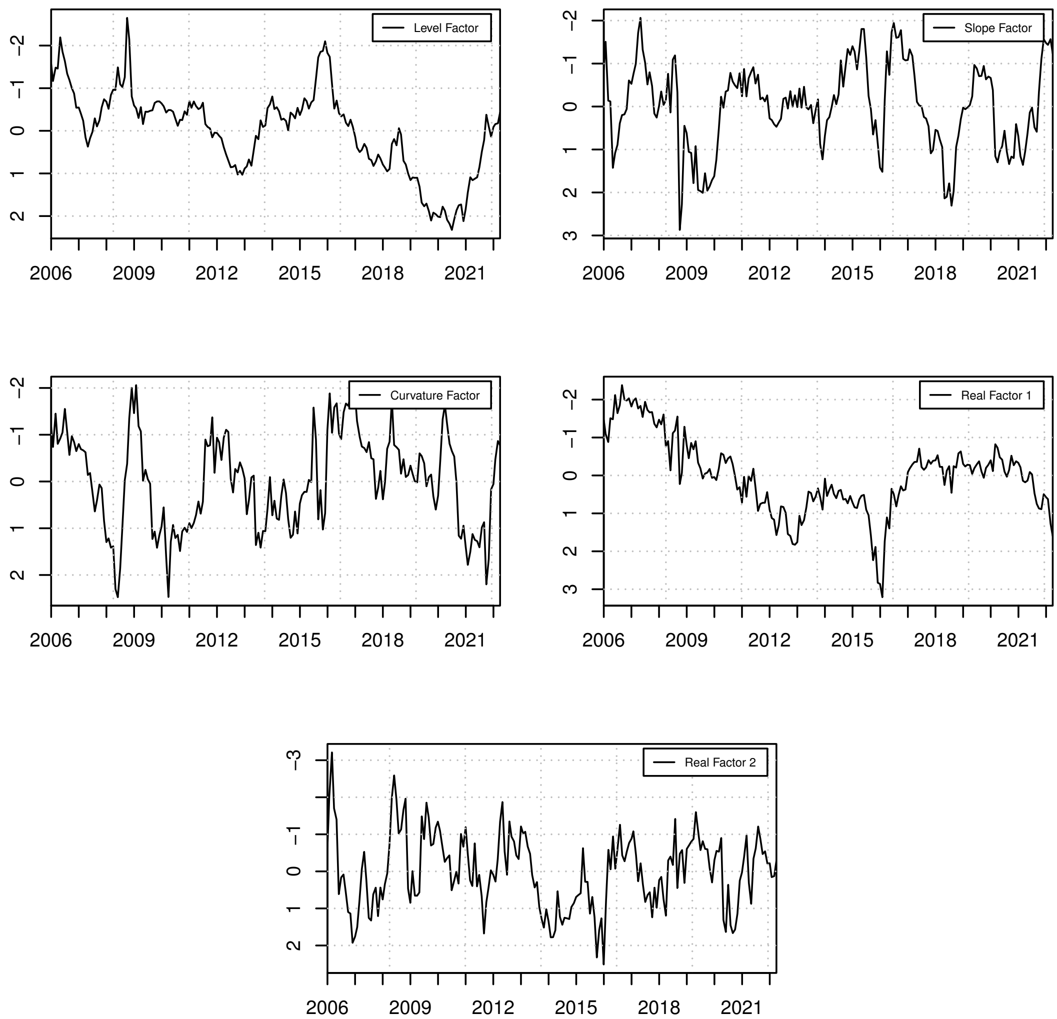
	n = 12	n = 24	n = 36	n = 60	n = 120
Nominal Interest Rates					
Avg.	0,103	0,107	0,110	0,114	0,117
Std. Dev.	0,034	0,030	0,027	0,024	0,022
Skewness	–0,514	–0,485	–0,404	–0,237	–0,052
kurtosis	2,484	2,671	2,828	3,053	3,205
$\rho(1)$	0,977	0,970	0,961	0,948	0,935
$\rho(6)$	0,775	0,767	0,751	0,713	0,649
Real Interest Rates					
Avg.	0,048	0,053	0,055	0,057	0,058
Std. Dev.	0,027	0,024	0,022	0,018	0,014
Skewness	–0,097	0,019	0,064	0,108	0,076
kurtosis	2,783	2,791	2,899	2,950	2,839
$\rho(1)$	0,962	0,964	0,964	0,966	0,962
$\rho(6)$	0,732	0,773	0,778	0,774	0,758

Source: Elaborated by the author.

Legend: The table reports summary statistics for Brazil yield curve over the period 2006-2022 to real and nominal yield curves. For each maturity we show mean, standard deviation, skewness, raw kurtosis, and two auto-correlations coefficients, $\hat{\rho}_1$ and $\hat{\rho}_6$.

Following Michael Abrahams et al. (2016) and various other authors, see Adrian, Crump, and Moench (2015), Joslin, Kenneth J Singleton, and Zhu (2011) and Wright (2011), we calculate principal components from yields and used as pricing factors in the model. Specifically, two sets of principal components are used. First, K_N = 3 principal components are extracted from nominal yields of maturities n = 6,12,24,...,120 months. Then additional factors are obtained as the first K_N = 2 principal components from the residuals of regressions of NTNB′s yields of maturities n = 24,...,120 months on the K_N nominal principal components. This orthogonalization step reduces the unconditional collinearity among the pricing factors. In sum, $K = K_N + K_R = 5$ model factors. See Figure 22.

Figure 22 – Pricing factors: observed time series



Source: Elaborated by the author.

Legend: This figure plots the time series of the factors of our model. These are the first three principal components extracted from the cross-section of end-of-month observations of nominal yields of maturities $n = 6,12,24,\ldots,120$ months. The fourth and fifth factors are the first two principal components extracted from the cross-section of orthogonalized real yields of maturities $n = 24,\ldots,120$, the residuals from regressing real yields on the first three principal components of the nominal yield curve.

We show the in-sample results in Figure 23 and Figure 24 and Figure 25 . In general, the model fits better to long maturities. The BEIR decomposition suggests that the real and nominal term premiums increase along maturities. Also, the same happens with inflation risk premiums, which account for the most movements of the BEIR in long maturities. Thus the expected inflation is quite flat for long maturities and is highly correlated to BEIR in short maturities. See Figure 26, Figure 27, and Figure 28.

In Inflation forecasting, we use the model-implied inflation expectations as a predictor, representing breakeven inflation rates adjusted for risk premia. For instance, we use the six-month maturity to predict inflation 6-months ahead, and so on. The same is done to unadjusted NTNB breakevens, which is a predictor of future inflation as well. The third is a simple random walk forecast, which takes the average realized inflation over the prior *n* months as a prediction of average inflation over the next *n* months. Forecasts are performed over horizons from 6 to 36 months, and forecasting errors are computed using overlapping observations. The panel reports out-of-sample results, using an eleven-year "learning period" over the period 2006:01–2016:06 and forecasting over the period 20016:07–2022:04. So, 6-months ahead has 70 forecasts, 12-months ahead has 64 forecasts, 24 months ahead has 52 forecasts, and 24-months ahead has 40 forecasts. See Figure 29 and Figure 30. The results suggest it is difficult to outperform the Focus survey; however, the model-implied forecast follows closely. The next section presents concluding remarks.

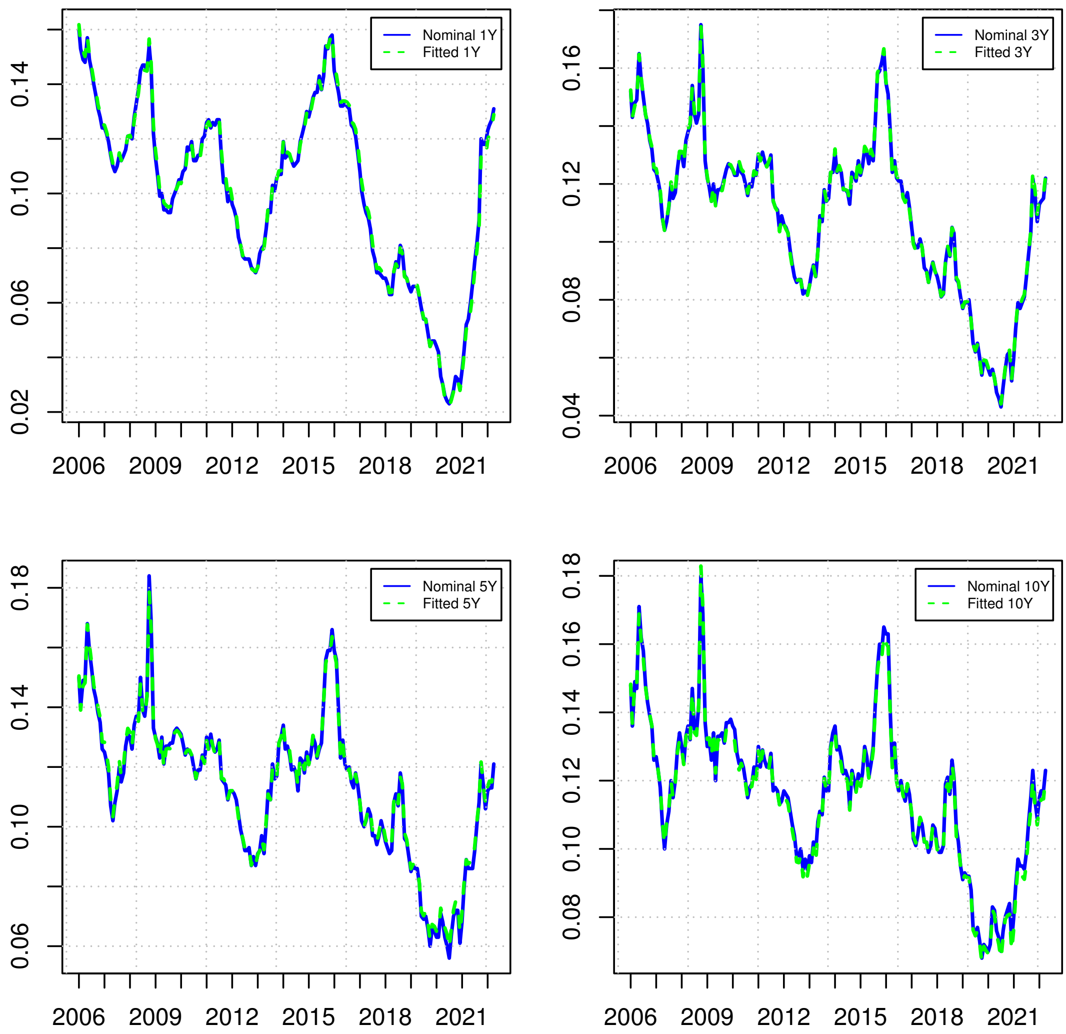
Figure 23 – In-Sample results

Measure	**Maturity**			
	n = 12	n = 36	n = 60	n = 120
RMSE : Nominal	0.157	0.073	0.135	0.148
RMSE : Real	0.472	0.096	0.132	0.070
MAE : Nominal	0.123	0.058	0.104	0.117
MAE : Real	0.375	0.075	0.099	0.057

Source: Elaborated by the author.

Legend: This table compares the root mean squared error and mean absolute error of nominal and real yield curves at one-year, three-year, five-year, and ten-year maturities. The first panel reports in-sample results for the entire sample from 2006:01 to 2022:04.

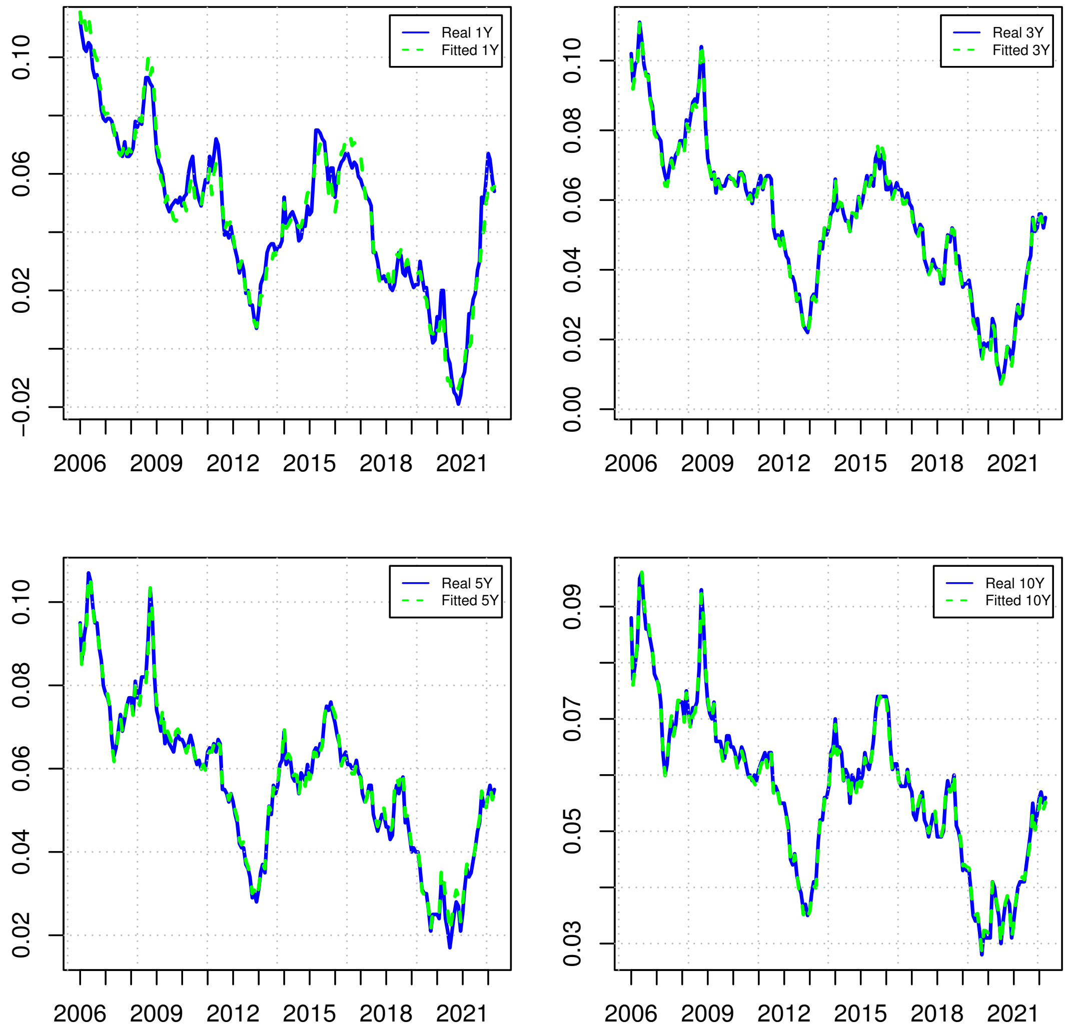
Figure 24 – Observed and nominal model-implied time series



Source: Elaborated by the author.

Legend: This figure provides time series plots of observed and model-implied nominal yields at one-year, three-year, five-year, and ten-year maturities. The observed yields are plotted by solid blue lines, whereas dashed green lines correspond to model-implied yields.

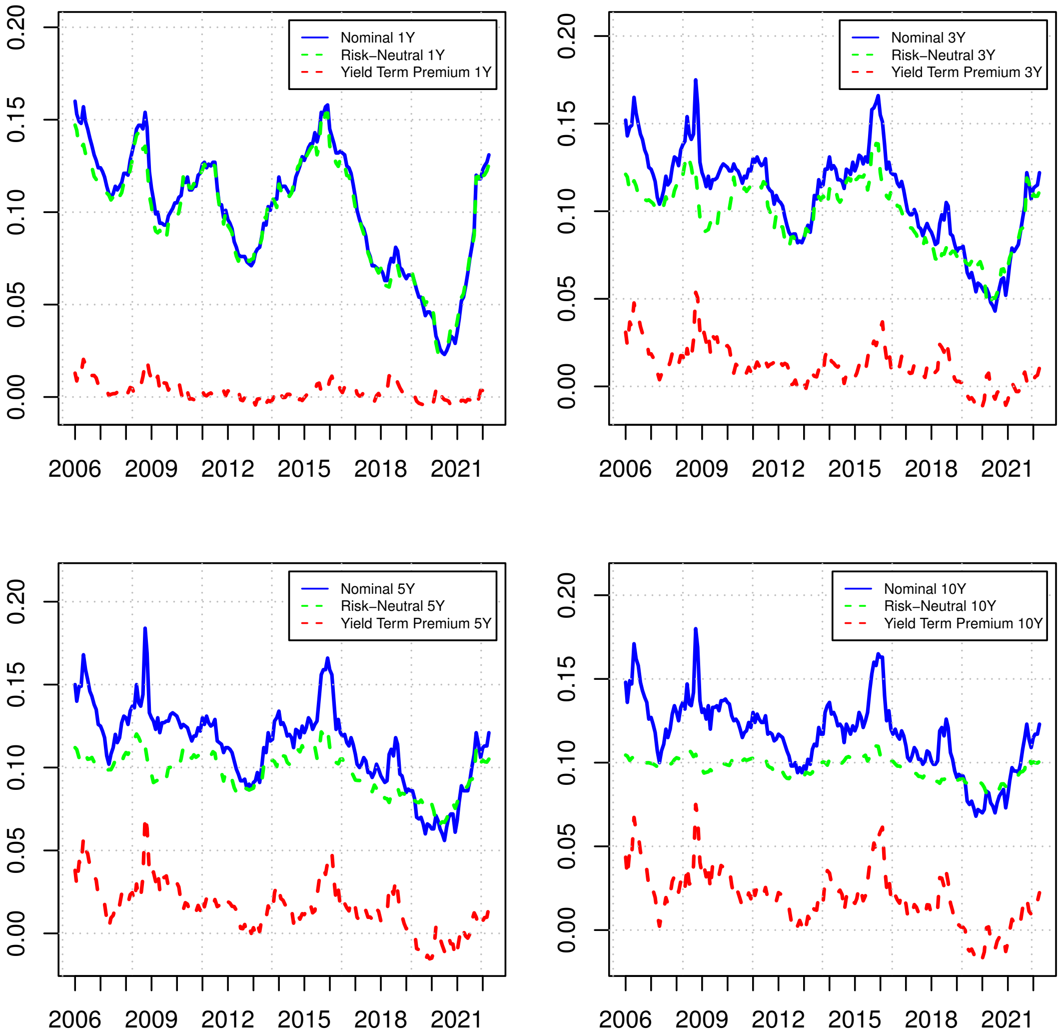
Figure 25 – Observed and real model-implied time series



Source: Elaborated by the author.

Legend: This figure provides time series plots of observed and model-implied real yields at one-year, three-year, five-year, and ten-year maturities. The observed yields are plotted by solid blue lines, whereas dashed green lines correspond to model-implied yields.

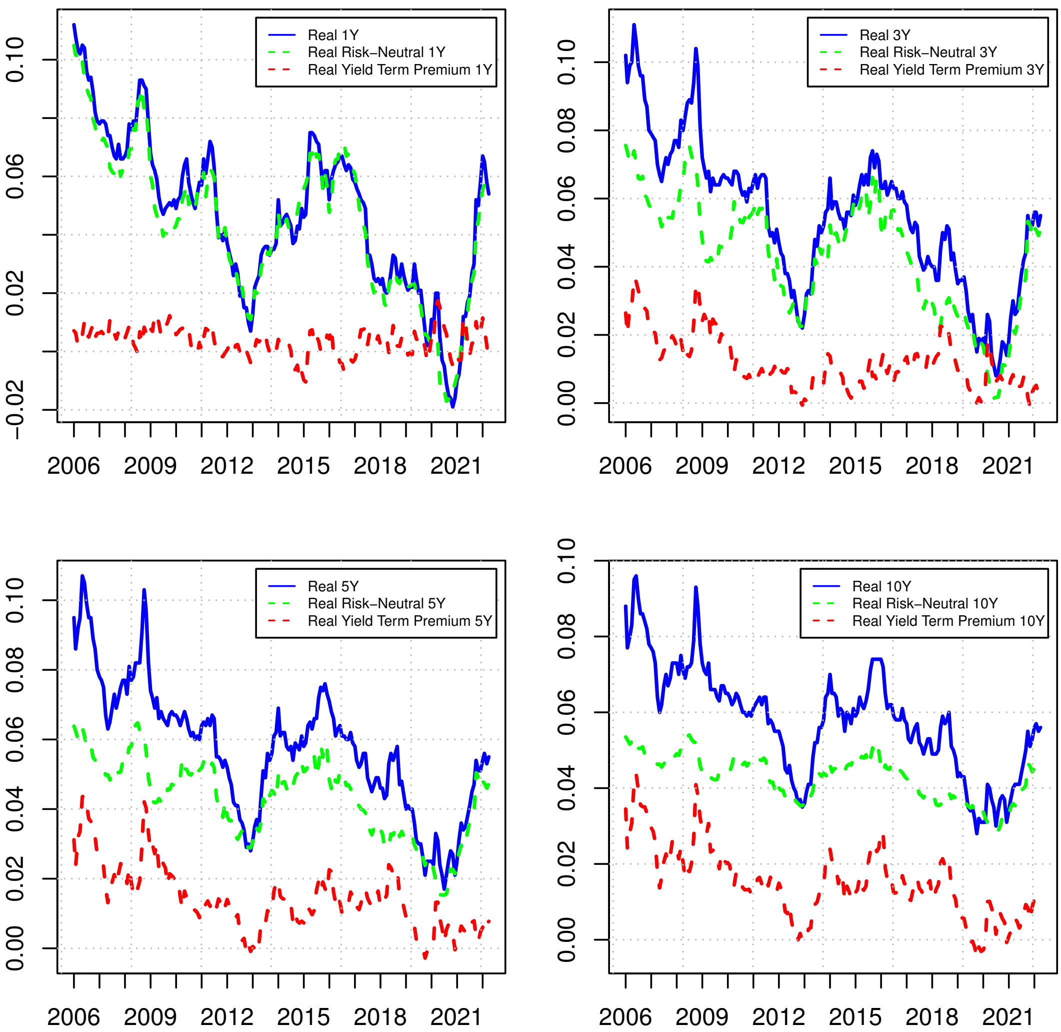
Figure 26 – Nominal term premium



Source: Elaborated by the author.

Legend: This figure provides time series plots of the decomposition of the observed nominal yield curves in risk-neutral yield and yield term premium at one-year, three-year, five-year, and ten-year maturities.

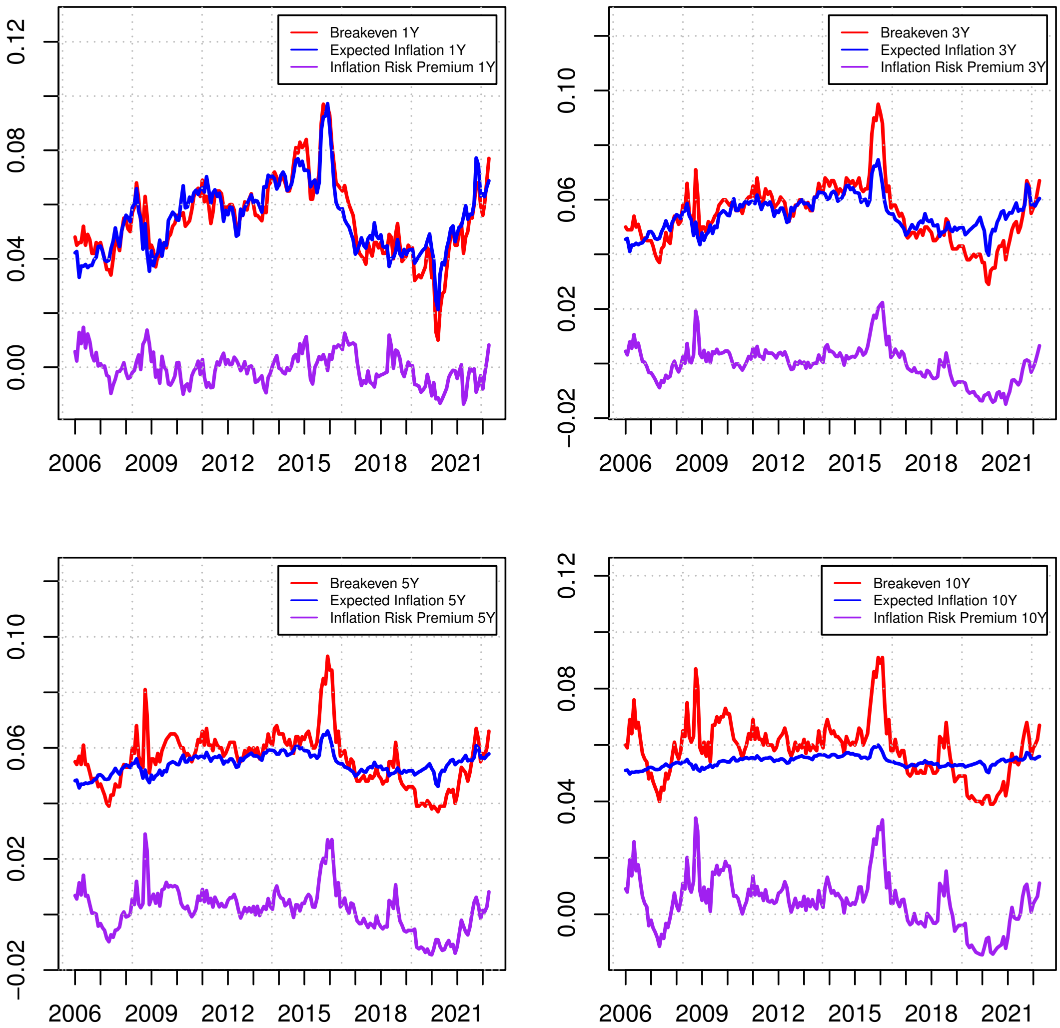
Figure 27 – Real term premium



Source: Elaborated by the author.

Legend: This figure provides time series plots of the decomposition of the observed nominal yield curves in risk-neutral yield and yield term premium at one-year, three-year, five-year, and ten-year maturities.

Figure 28 – BEIR Decomposition



Source: Elaborated by the author.

Legend: This figure shows the decomposition of breakeven inflation rates into the model-implied expected inflation and the inflation risk premium. The panels show this decomposition at one-year, three-year, five-year, and ten-year maturities.

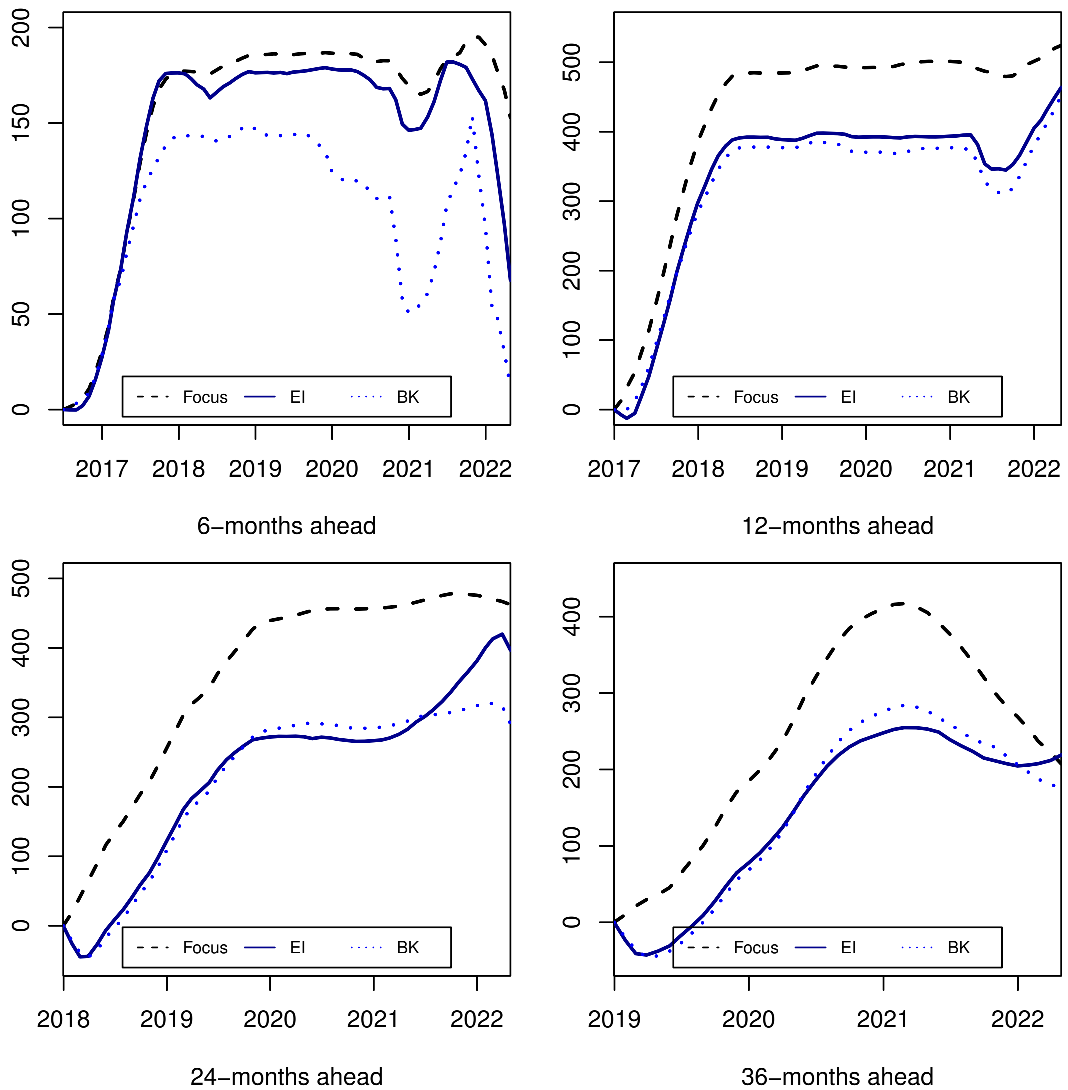
Figure 29 – Inflation Forecasting

Model	**Horizon**			
	n = 6	n = 12	n = 24	n = 36
RandonWalk	2.743	4.320	4.487	4.295
Focus	**0.842**	**0.749**	**0.747**	**0.848**
Modelforecast	0.933	0.782	**0.788**	0.839
Breakevens	0.989	0.788	0.850	0.873

Source: Elaborated by the author.

Legend: This table compares three models' root mean squared error for predicting future inflation (IPCA). The first uses the model-implied inflation expectations derived in Section 2. These represent breakeven inflation rates adjusted for risk premia. The second method takes unadjusted breakevens as a predictor of future inflation. The third is a simple random walk forecast, i.e., it takes the average realized inflation over the prior *n* months as a prediction of average inflation over the next *n* months. Forecasts are performed over horizons from 6 to 36 months, and forecasting errors are computed using overlapping observations. The panel reports out-of-sample results, utilizing an eleven-year "learning period" over the period 2006:01–2016:06 and forecasting over the period 20016:07–2022:04. So, 6-months ahead has 70 forecasts, 12-months ahead has 64 forecasts, 24-months ahead has 52 forecasts, and 24-months ahead has 40 forecasts. Bold values are statistically significant by at least 5%, according to (GIACOMINI; WHITE, 2006) test.

Figure 30 – Cumulative Squared Prediction Error



Source: Elaborated by the author.

Legend: This figure shows the cumulative squared prediction error of Random Walk, Focus, Model-Implied Expected Inflation, and BEIR forecasts at one-year, three-year, five-year, and ten-year maturities.

5.4 CONCLUDING REMARKS

We estimate an arbitrage-free Gaussian model for the term structure of the yield curve that allows joint modeling of nominal and real interest rates. The model enables the decomposition of BEIR into expectations for inflation and risk premium. In-sample results suggest that the term premiums are time-varying and increase along maturities, which include negative values. The risk-adjusted inflation expectations outweigh unadjusted BEIRs and a Random Walk in the out-of-sample inflation forecast. The Focus survey is a benchmark challenging to outperform. However, the model-implied predictions have better results in long horizons.

BIBLIOGRAPHY

ABRAHAMS, M.; ADRIAN, T.; CRUMP, R.K.; MOENCH, E. **Decomposing Real and Nominal Yield Curves. Staff Report 570**. [S.l.], 2015.

ABRAHAMS, Michael; ADRIAN, Tobias; CRUMP, Richard K; MOENCH, Emanuel; YU, Rui. Decomposing real and nominal yield curves. **Journal of Monetary Economics**, Elsevier, v. 84, p. 182–200, 2016.

ADRIAN, Tobias; CRUMP, Richard K; MOENCH, Emanuel. Regression-based estimation of dynamic asset pricing models. **Journal of Financial Economics**, Elsevier, v. 118, n. 2, p. 211–244, 2015.

AIOLFI, Marco; TIMMERMANN, Allan. Persistence in forecasting performance and conditional combination strategies. **Journal of Econometrics**, Elsevier, v. 135, n. 1-2, p. 31–53, 2006.

ALMEIDA, Caio Ibsen Rodrigues de; FARIA, Adriano. Forecasting the Brazilian term structure using macroeconomic factors. **Brazilian Review of Econometrics**, Sociedade Brasileira de Econometria, v. 34, n. 1, p. 45–77, 2014.

ALTAVILLA, Carlo; GIACOMINI, Raffaella; COSTANTINI, Riccardo. Bond returns and market expectations. **Journal of Financial Econometrics**, Oxford University Press, v. 12, n. 4, p. 708–729, 2014.

ALTAVILLA, Carlo; GIACOMINI, Raffaella; RAGUSA, Giuseppe. Anchoring the yield curve using survey expectations. **Journal of Applied Econometrics**, Wiley Online Library, v. 32, n. 6, p. 1055–1068, 2017.

ANDRADE ALVES, Cassio Roberto de; ABRAHAM, Kuruvilla Joseph; LAURINI, Marcio Poletti. Can Brazilian Central Bank Communication help to predict the yield curve? **Journal of Forecasting**, Wiley Online Library, 2023.

ANG, Andrew; BEKAERT, Geert; WEI, Min. Do macro variables, asset markets, or surveys forecast inflation better? **Journal of monetary Economics**, Elsevier, v. 54, n. 4, p. 1163–1212, 2007.

ANG, Andrew; BEKAERT, Geert; WEI, Min. The term structure of real rates and expected inflation. **The Journal of Finance**, Wiley Online Library, v. 63, n. 2, p. 797–849, 2008.

ANG, Andrew; PIAZZESI, Monika. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. **Journal of Monetary economics**, Elsevier, v. 50, n. 4, p. 745–787, 2003.

ANG, Andrew; PIAZZESI, Monika; WEI, Min. What does the yield curve tell us about GDP growth? **Journal of econometrics**, Elsevier, v. 131, n. 1-2, p. 359–403, 2006.

BAUER, Michael D. Restrictions on risk prices in dynamic term structure models. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 36, n. 2, p. 196–211, 2018.

BEKAERT, Geert; CHO, Seonghoon; MORENO, Antonio. New Keynesian macroeconomics and the term structure. **Journal of Money, Credit and Banking**, Wiley Online Library, v. 42, n. 1, p. 33–62, 2010.

BEKAERT, Geert; WANG, Xiaozheng. Inflation risk and the inflation risk premium. **Economic Policy**, Oxford University Press, v. 25, n. 64, p. 755–806, 2010.

BERNANKE, Ben S. **The federal funds rate and the channels of monetary transnission**. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1990.

BERNANKE, Ben S. The new tools of monetary policy. **American Economic Review**, v. 110, n. 4, p. 943–83, 2020.

BIANCHI, Daniele; BÜCHNER, Matthias; TAMONI, Andrea; NIEUWERBURGH, Stijn Van. Bond Risk Premiums with Machine Learning. **Review of Financial Studies**, v. 34, n. 2, p. 1046–1089, 2021.

BIANCHI, Francesco; MUMTAZ, Haroon; SURICO, Paolo. The great moderation of the term structure of UK interest rates. **Journal of Monetary Economics**, Elsevier, v. 56, n. 6, p. 856–871, 2009.

BIERMAN, Gerald J. **Factorization methods for discrete sequential estimation**. [S.l.]: New York: Academic Press, 1977.

BIKBOV, Ruslan; CHERNOV, Mikhail. No-arbitrage macroeconomic determinants of the yield curve. **Journal of Econometrics**, Elsevier, v. 159, n. 1, p. 166–182, 2010.

BIS. **Zero-coupon yield curves: Technical documentation**. [S.l.: s.n.], 2005.

BIS. **Zero-coupon yield curves. Technical Documentation**. [S.l.]: Bank for International Settlements, 1999.

BOLLERSLEV, Tim. Generalized autoregressive conditional heteroskedasticity. **Journal of econometrics**, Elsevier, v. 31, n. 3, p. 307–327, 1986.

BREACH, Tomas; D'AMICO, Stefania; ORPHANIDES, Athanasios. The term structure and inflation uncertainty. **Journal of Financial Economics**, Elsevier, v. 138, n. 2, p. 388–414, 2020.

BRIGO, Damiano; MERCURIO, Fabio, et al. **Interest rate models: theory and practice**. [S.l.]: Springer, 2001. v. 2.

CALDEIRA, João F. Investigating the expectation hypothesis and the risk premium dynamics: new evidence for Brazil. **Empirical Economics**, Springer, v. 59, n. 1, p. 395–412, 2020.

CALDEIRA, João F; FURLANI, Luiz GC. Inflação implícita e o prêmio pelo risco: uma alternativa aos modelos VAR na previsão para o IPCA. **Estudos Econômicos (São Paulo)**, SciELO Brasil, v. 43, p. 627–645, 2013.

CALDEIRA, João F; LAURINI, Márcio P; PORTUGAL, Marcelo S. Bayesian inference applied to dynamic Nelson-Siegel model with stochastic volatility. **Brazilian Review of Econometrics**, v. 30, n. 1, p. 123–161, 2010.

CALDEIRA, João F; MOURA, Guilherme V; SANTOS, André AP. Bond portfolio optimization using dynamic factor models. **Journal of Empirical Finance**, Elsevier, v. 37, p. 128–158, 2016.

CALDEIRA, João F.; MOURA, Guilherme V.; SANTOS, André A. P. Predicting the Yield Curve Using Forecast Combinations. **Computional Statistics Data Analysis**, Elsevier Science Publishers B. V., v. 100, n. 3, p. 79–98, Aug. 2016. ISSN 0167-9473.

CARRIERO, Andrea; CLARK, Todd E; MARCELLINO, Massimiliano. No-arbitrage priors, drifting volatilities, and the term structure of interest rates. **Journal of Applied Econometrics**, Wiley Online Library, v. 36, n. 5, p. 495–516, 2021.

CARVALHO, Fabia A de; MINELLA, André, et al. Market Forecasts in Brazil: performance and determinants. **Central Bank of Brazil Working Paper Series**, n. 185, 2009.

CHERNOV, Mikhail; MUELLER, Philippe. The term structure of inflation expectations. **Journal of financial economics**, Elsevier, v. 106, n. 2, p. 367–394, 2012.

CHRISTENSEN, Bent Jesper; WEL, Michel van der. An asset pricing approach to testing general term structure models. **Journal of Financial Economics**, v. 134, n. 1, p. 165–191, 2019.

CHRISTENSEN, Bent Jesper; WEL, Michel van der, et al. An asset pricing approach to testing general term structure models including Heath-Jarrow-Morton specifications and affine subclasses. **Department of Economics and Business Economics, Aarhus University**, 2010.

CHRISTENSEN, Jens HE; DIEBOLD, Francis X; RUDEBUSCH, Glenn D. The affine arbitrage-free class of Nelson–Siegel term structure models. **Journal of Econometrics**, Elsevier, v. 164, n. 1, p. 4–20, 2011.

CHRISTOFFERSEN, Peter; DIEBOLD, Francis X; MARIANO, Roberto S; TAY, Anthony S; TSE, Yiu Kuen. Direction-of-change forecasts based on conditional variance, skewness and kurtosis dynamics: international evidence, 2006.

CHRISTOFFERSEN, Peter F; DIEBOLD, Francis X. Financial asset returns, direction-of-change forecasting, and volatility dynamics. **Management Science**, INFORMS, v. 52, n. 8, p. 1273–1287, 2006.

CHUN, Albert Lee. Expectations, bond yields, and monetary policy. **The Review of Financial Studies**, Society for Financial Studies, v. 24, n. 1, p. 208–247, 2011.

CHUN, Albert Lee. Forecasting interest rates and inflation: Blue chip clairvoyants or econometrics?, 2012.

CIESLAK, Anna; POVALA, Pavol. Understanding bond risk premia. **Unpublished working paper. Kellogg School of Management, Evanston, IL**, Citeseer, p. 677–691, 2011.

CLARK, Todd E; MCCRACKEN, Michael W. Improving forecast accuracy by combining recursive and rolling forecasts. **International Economic Review**, Wiley Online Library, v. 50, n. 2, p. 363–395, 2009.

COCHRANE, John H; PIAZZESI, Monika. Bond risk premia. **American economic review**, v. 95, n. 1, p. 138–160, 2005.

COCHRANE, John H; PIAZZESI, Monika. Decomposing the yield curve. In: AFA 2010 Atlanta Meetings Paper. [S.l.: s.n.], 2009.

COX, John C; INGERSOLL JR, Jonathan E; ROSS, Stephen A. An intertemporal general equilibrium model of asset prices. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 363–384, 1985.

CRUMP, Richard K; EUSEPI, Stefano; MOENCH, Emanuel. The term structure of expectations and bond yields. FRB of NY Staff Report, 2018.

D'AMICO, Stefania; KIM, Don H; WEI, Min. Tips from TIPS: the informational content of Treasury Inflation-Protected Security prices. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 53, n. 1, p. 395–436, 2018.

DAI, Qiang; PHILIPPON, Thomas. **Fiscal policy and the term structure of interest rates**. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 2005.

DE POOTER, Michiel. Examining the Nelson-Siegel class of term structure models: In-sample fit versus out-of-sample forecasting performance, 2007.

DE POOTER, Michiel; RAVAZZOLO, Francesco; VAN DIJK, Dick JC. Term structure forecasting using macro factors and forecast combination, 2010.

DEWACHTER, Hans; LYRIO, Marco. Macro factors and the term structure of interest rates. **Journal of Money, Credit and Banking**, JSTOR, p. 119–140, 2006.

DEWACHTER, Hans; LYRIO, Marco; MAES, Konstantijn. A joint model for the term structure of interest rates and the macroeconomy. **Journal of Applied Econometrics**, Wiley Online Library, v. 21, n. 4, p. 439–462, 2006.

DIEBOLD, Francis X; LI, Canlin. Forecasting the term structure of government bond yields. **Journal of econometrics**, Elsevier, v. 130, n. 2, p. 337–364, 2006.

DIEBOLD, Francis X; LI, Canlin; YUE, Vivian Z. Global yield curve dynamics and interactions: a dynamic Nelson–Siegel approach. **Journal of Econometrics**, Elsevier, v. 146, n. 2, p. 351–363, 2008.

DIEBOLD, Francis X; PIAZZESI, Monika; RUDEBUSCH, Glenn D. Modeling bond yields in finance and macroeconomics. **American Economic Review**, v. 95, n. 2, p. 415–420, 2005.

DIEBOLD, Francis X; RUDEBUSCH, Glenn D, et al. Yield curve modeling and forecasting: the dynamic Nelson-Siegel approach. **Economics Books**, Princeton University Press, 2012.

DIEBOLD, Francis X; RUDEBUSCH, Glenn D; ARUOBA, S Boragan. The macroeconomy and the yield curve: a dynamic latent factor approach. **Journal of econometrics**, Elsevier, v. 131, n. 1-2, p. 309–338, 2006.

DIEBOLD, Francis X.; RUDEBUSCH, Glenn D. **Yield Curve Modeling and Forecasting: The Dynamic Nelson-Siegel Approach**. [S.l.]: Princeton University Press, 2013. (Economics Books, 9895).

DOUCET, Arnaud; DE FREITAS, Nando; GORDON, Neil. An introduction to sequential Monte Carlo methods. In: SEQUENTIAL Monte Carlo methods in practice. [S.l.]: Springer, 2001. P. 3–14.

DUFFEE, Gregory R. Information in (and not in) the term structure. **The Review of Financial Studies**, Oxford University Press, v. 24, n. 9, p. 2895–2934, 2011.

DUFFEE, Gregory R. Term premia and interest rate forecasts in affine models. **The Journal of Finance**, Wiley Online Library, v. 57, n. 1, p. 405–443, 2002.

DUFFEE, Gregory R. Term structure estimation without using latent factors. **Journal of Financial Economics**, Elsevier, v. 79, n. 3, p. 507–536, 2006.

DUFFEE, Gregory R; STANTON, Richard H. Estimation of dynamic term structure models. **The Quarterly Journal of Finance**, World Scientific, v. 2, n. 02, p. 1250008, 2012.

DUFFIE, Darrell; KAN, Rui. A yield-factor model of interest rates. **Mathematical finance**, Wiley Online Library, v. 6, n. 4, p. 379–406, 1996.

DURBIN, James; KOOPMAN, Siem Jan. **Time series analysis by state space methods**. [S.l.]: Oxford University Press, 2012. v. 38.

ECB. **Yield curve modelling and a conceptual framework for estimating yield curves: evidence from the European Central Bank's yield curves**. [S.l.], 2018.

EHLING, Paul; GALLMEYER, Michael; HEYERDAHL-LARSEN, Christian; ILLEDITSCH, Philipp. Disagreement about inflation and the yield curve. **Journal of Financial Economics**, Elsevier, v. 127, n. 3, p. 459–484, 2018.

ENGLE, Robert; NG, Victor K. Time-Varying Volatility and the Dynamic Behavior of the Term Structure. **Journal of Money, Credit and Banking**, v. 25, n. 3, p. 336–49, 1993.

EXTERKATE, Peter; DIJK, Dick Van; HEIJ, Christiaan; GROENEN, Patrick JF. Forecasting the yield curve in a data-rich environment using the factor-augmented Nelson–Siegel model. **Journal of Forecasting**, Wiley Online Library, v. 32, n. 3, p. 193–214, 2013.

FAMA, Eugene F. Term-structure forecasts of interest rates, inflation and real returns. **Journal of Monetary Economics**, North-Holland, v. 25, n. 1, p. 59–76, 1990.

FAVERO, Carlo A; NIU, Linlin; SALA, Luca. Term structure forecasting: no-arbitrage restrictions versus large information set. **Journal of Forecasting**, Wiley Online Library, v. 31, n. 2, p. 124–156, 2012.

FERNANDES, Marcelo; VIEIRA, Fausto. A dynamic Nelson–Siegel model with forward-looking macroeconomic factors for the yield curve in the US. **Journal of Economic Dynamics and Control**, Elsevier, v. 106, p. 103720, 2019.

FILIPOVIC, Damir. **Term-Structure Models. A Graduate Course.** [S.l.]: Springer, 2009.

FISHER, Irving. **Appreciation and Interest: A Study of the Influence of Monetary Appreciation and Depreciation on the Rate of Interest with Applications to the Bimetallic Controversy and the Theory of Interest**. [S.l.]: American economic association, 1896. v. 11.

FISHER, Irving. **The rate of interest**. [S.l.]: (New York: Macmillan, 1907.

FLEMING, Michael J; REMOLONA, Eli M. The term structure of announcement effects. BIS Working Paper, 2001.

GALLMEYER, Michael F; HOLLIFIELD, Burton; ZIN, Stanley E. Taylor rules, McCallum rules and the term structure of interest rates. **Journal of Monetary Economics**, Elsevier, v. 52, n. 5, p. 921–950, 2005.

GARGANO, Antonio; PETTENUZZO, Davide; TIMMERMANN, Allan. Bond Return Predictability: Economic Value and Links to the Macroeconomy. **Management Science**, v. 65, n. 2, p. 508–540, 2019.

GIACOMINI, Raffaella; WHITE, Halbert. Tests of conditional predictive ability. **Econometrica**, Wiley Online Library, v. 74, n. 6, p. 1545–1578, 2006.

GLOSTEN, Lawrence R; JAGANNATHAN, Ravi; RUNKLE, David E. On the relation between the expected value and the volatility of the nominal excess return on stocks. **The journal of finance**, Wiley Online Library, v. 48, n. 5, p. 1779–1801, 1993.

GOLIŃSKI, Adam; SPENCER, Peter. Estimating the term structure with linear regressions: Getting to the roots of the problem. **Journal of Financial Econometrics**, Oxford University Press, v. 19, n. 5, p. 960–984, 2021.

GRANGER, Clive WJ. Invited review combining forecasts—twenty years later. **Journal of forecasting**, Wiley Online Library, v. 8, n. 3, p. 167–173, 1989.

GRANGER, Clive WJ; JEON, Yongil. Thick modeling. **Economic Modelling**, Elsevier, v. 21, n. 2, p. 323–343, 2004.

GREER, Mark. Directional accuracy tests of long-term interest rate forecasts. **International Journal of Forecasting**, Elsevier, v. 19, n. 2, p. 291–298, 2003.

GREER, Mark R. Combination forecasting for directional accuracy: An application to survey interest rate forecasts. **Journal of Applied Statistics**, Taylor & Francis, v. 32, n. 6, p. 607–615, 2005.

GRISHCHENKO, Olesya V; HUANG, Jing-Zhi. The inflation risk premium: Evidence from the TIPS market. **The Journal of Fixed Income**, Institutional Investor Journals Umbrella, v. 22, n. 4, p. 5–30, 2013.

GÜRKAYNAK, Refet S; WRIGHT, Jonathan H. Macroeconomics and the term structure. **Journal of Economic Literature**, American Economic Association, v. 50, n. 2, p. 331–367, 2012.

HAMILTON, James D; WU, Jing Cynthia. The effectiveness of alternative monetary policy tools in a zero lower bound environment. **Journal of Money, Credit and Banking**, Wiley Online Library, v. 44, p. 3–46, 2012.

HANSEN, Peter R; LUNDE, Asger; NASON, James M. The model confidence set. **Econometrica**, Wiley Online Library, v. 79, n. 2, p. 453–497, 2011.

HANSEN, Peter R.; LUNDE, Asger; NASON, James M. The Model Confidence Set. **Econometrica**, v. 79, n. 2, p. 453–497, Mar. 2011.

HARVEY, Andrew C. **Forecasting, Structural Time Series Models and the Kalman filter**. [S.l.]: Cambridge University Press, Cambridge, 1989.

HARVEY, Andrew C; RUIZ, Esther; SENTANA, Enrique. Unobserved component time series models with ARCH disturbances. Elsevier, 1992.

HARVEY, Campbell R. Forecasts of economic growth from the bond and stock markets. **Financial Analysts Journal**, Taylor & Francis, v. 45, n. 5, p. 38–45, 1989.

HAUTSCH, Nikolaus; YANG, Fuyu. Bayesian inference in a stochastic volatility Nelson–Siegel model. **Computational Statistics & Data Analysis**, Elsevier, v. 56, n. 11, p. 3774–3792, 2012.

HEVIA, Constantino; GONZALEZ-ROZADA, Martin; SOLA, Martin; SPAGNOLO, Fabio. Estimating and forecasting the yield curve using a Markov switching dynamic Nelson and Siegel model. **Journal of Applied Econometrics**, Wiley Online Library, v. 30, n. 6, p. 987–1009, 2015.

HICKS, J. Value and Capital. Oxford University Press Oxford, UK, 1939.

HODGES, Stewart D; SCHAEFER, Stephen M. A model for bond portfolio improvement. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 12, n. 2, p. 243–260, 1977.

HÖRDAHL, Peter; TRISTANI, Oreste; VESTIN, David. A joint econometric model of macroeconomic and term-structure dynamics. **Journal of Econometrics**, Elsevier, v. 131, n. 1-2, p. 405–444, 2006.

JAZWINSKI, Andrew H. **Stochastic Processes and Filtering Theory**. [S.l.]: Academic Press, 1970. v. 64. (Mathematics in Science and Engineering).

JOSLIN, Scott; LE, Anh; SINGLETON, Kenneth J. Why Gaussian macro-finance term structure models are (nearly) unconstrained factor-VARs. **Journal of Financial Economics**, v. 109, n. 3, p. 604–622, 2013.

JOSLIN, Scott; PRIEBSCH, Marcel; SINGLETON, Kenneth J. Risk premiums in dynamic term structure models with unspanned macro risks. **The Journal of Finance**, Wiley Online Library, v. 69, n. 3, p. 1197–1233, 2014.

JOSLIN, Scott; SINGLETON, Kenneth J; ZHU, Haoxiang. A new perspective on Gaussian dynamic term structure models. **The Review of Financial Studies**, Oxford University Press, v. 24, n. 3, p. 926–970, 2011.

JOYCE, Michael AS; LILDHOLDT, Peter; SORENSEN, Steffen. Extracting inflation expectations and inflation risk premia from the term structure: a joint model of the UK nominal and real yield curves. **Journal of Banking & Finance**, Elsevier, v. 34, n. 2, p. 281–294, 2010.

JULIER, Simon J; UHLMANN, Jeffrey K. New extension of the Kalman filter to nonlinear systems. In: INTERNATIONAL SOCIETY FOR OPTICS and PHOTONICS. SIGNAL processing, sensor fusion, and target recognition VI. [S.l.: s.n.], 1997. v. 3068, p. 182–194.

JUNGBACKER, Borus; KOOPMAN, Siem Jan; VAN DER WEL, Michel. Smooth dynamic factor analysis with application to the US term structure of interest rates. **Journal of Applied Econometrics**, Wiley Online Library, v. 29, n. 1, p. 65–90, 2014.

JUNGBACKER, Borus; KOOPMAN, Siem Jan; WEL, Michel. Smooth Dynamic Factor Analysis With Application To The Us Term Structure Of Interest Rates. **Journal of Applied Econometrics**, v. 29, n. 1, p. 65–90, Jan. 2014.

KAMINSKA, Iryna. A no-arbitrage structural vector autoregressive model of the UK yield curve. **Oxford Bulletin of Economics and Statistics**, Wiley Online Library, v. 75, n. 5, p. 680–704, 2013.

KEYNES, J. M. **The General Theory of Employment, Interest and Money**. [S.l.]: Macmillan, 1936. 14th edition, 1973.

KIM, Don H; ORPHANIDES, Athanasios. Term structure estimation with survey data on interest rate forecasts. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 47, n. 1, p. 241–272, 2012.

KOOPMAN, Siem Jan; MALLEE, Max IP; VAN DER WEL, Michel. Analyzing the term structure of interest rates using the dynamic Nelson–Siegel model with time-varying parameters. **Journal of Business & Economic Statistics**, Taylor & Francis, v. 28, n. 3, p. 329–343, 2010.

KOOPMAN, Siem Jan; WEL, Michel van der. Forecasting the US term structure of interest rates using a macroeconomic smooth dynamic factor model. **International Journal of Forecasting**, Elsevier, v. 29, n. 4, p. 676–694, 2013.

KRIPPNER, Leo et al. **A theoretical foundation for the Nelson and Siegel class of yield curve models, and an empirical application to US yield curve dynamics**. [S.l.], 2010.

KUTTNER, Kenneth N. Outside the box: Unconventional monetary policy in the great recession and beyond. **Journal of Economic Perspectives**, American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2418, v. 32, n. 4, p. 121–146, 2018.

LAURINI, Márcio; HOTTA, Luiz Koodi. Bayesian extensions to diebold-li term structure model. **International Review of Financial Analysis**, Elsevier, v. 19, n. 5, p. 342–350, 2010.

LAURINI, Márcio P; CALDEIRA, João F. A macro-finance term structure model with multivariate stochastic volatility. **International Review of Economics & Finance**, Elsevier, v. 44, p. 68–90, 2016.

LAURINI, MÁrcio Poletti; HOTTA, Luiz Koodi. Forecasting the term structure of interest rates using integrated nested Laplace approximations. **Journal of Forecasting**, Wiley Online Library, v. 33, n. 3, p. 214–230, 2014.

LIMA, Alexandre Maia Correia; ISSLER, João Victor. A hipótese das expectativas na estrutura a termo de juros no Brasil: uma aplicação de modelos de valor presente. **Revista brasileira de economia**, SciELO Brasil, v. 57, p. 873–898, 2003.

LITTERMAN, Robert; SCHEINKMAN, Jose. Common factors affecting bond returns. **Journal of fixed income**, v. 1, n. 1, p. 54–61, 1991.

LIU, Yan; WU, Jing Cynthia. Reconstructing the Yield Curve. **Journal of Financial Economics**, 2021. ISSN 0304-405X.

LONGSTAFF, Francis A; SCHWARTZ, Eduardo S. Interest rate volatility and the term structure: A two-factor general equilibrium model. **The Journal of Finance**, Wiley Online Library, v. 47, n. 4, p. 1259–1282, 1992.

LUDVIGSON, Sydney C; NG, Serena. Macro factors in bond risk premia. **The Review of Financial Studies**, Oxford University Press, v. 22, n. 12, p. 5027–5067, 2009.

LUTZ, Friedrich A. The structure of interest rates. **The Quarterly Journal of Economics**, MIT Press, v. 55, n. 1, p. 36–63, 1940.

MARÇAL, Emerson Fernandes; PEREIRA, Pedro Luiz Valls. A estrutura a termo das taxas de juros no brasil: Testando a hipótese de expectativas. Instituto de Pesquisa Econômica Aplicada (Ipea), 2007.

MISHKIN, Frederic S. A multi-country study of the information in the shorter maturity term structure about future inflation. **Journal of International Money and Finance**, Elsevier, v. 10, n. 1, p. 2–22, 1991.

MISHKIN, Frederic S. The information in the longer maturity term structure about future inflation. **The Quarterly Journal of Economics**, MIT Press, v. 105, n. 3, p. 815–828, 1990.

MODIGLIANI, Franco; SUTCH, Richard. Innovations in interest rate policy. **The American Economic Review**, JSTOR, v. 56, n. 1/2, p. 178–197, 1966.

MOENCH, Emanuel. Forecasting the yield curve in a data-rich environment: A no-arbitrage factor-augmented VAR approach. **Journal of Econometrics**, Elsevier, v. 146, n. 1, p. 26–43, 2008.

MORLEY, James. Macro-Finance Linkages. **Journal of Economic Surveys**, Wiley Online Library, v. 30, n. 4, p. 698–711, 2016.

NELSON, Charles R; SIEGEL, Andrew F. Parsimonious modeling of yield curves. **q Journal of Business**, v. 60, n. 4, p. 473–489, 1987.

NEWBOLD, Paul; HARVEY, David I. Forecast combination and encompassing. **A companion to economic forecasting**, Wiley Online Library, v. 1, p. 620, 2002.

ORPHANIDES, Athanasios; WEI, Min. Evolving macroeconomic perceptions and the term structure of interest rates. **Journal of Economic Dynamics and Control**, Elsevier, v. 36, n. 2, p. 239–254, 2012.

PESARAN, M Hashem; TIMMERMANN, Allan. Selection of estimation window in the presence of breaks. **Journal of Econometrics**, Elsevier, v. 137, n. 1, p. 134–161, 2007.

PIAZZESI, Monika. **An econometric model of the yield curve with macroeconomic jump effects**. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 2001.

PIAZZESI, Monika. Bond yields and the Federal Reserve. **Journal of Political Economy**, The University of Chicago Press, v. 113, n. 2, p. 311–344, 2005.

PIAZZESI, Monika; SCHNEIDER, Martin, et al. **Trend and cycle in bond premia**. [S.l.]: Citeseer, 2009. v. 424.

PONCELA, Pilar. Comments on “Forecasting the US term structure of interest rates using a macroeconomic smooth dynamic factor model” by Koopman and van der Wel. **International Journal of Forecasting**, Elsevier, v. 29, n. 4, p. 695–697, 2013.

R CORE TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2018. Available from: `https://www.R-project.org/`.

RAPACH, David; ZHOU, Guofu. Forecasting Stock Returns. In: ELLIOTT, G.; GRANGER, C.; TIMMERMANN, A. (Eds.). **Handbook of Economic Forecasting**. [S.l.]: Elsevier, 2013. v. 2. (Handbook of Economic Forecasting). chap. 0, p. 328–383.

RATTRAY, Sandy; BALASUBRAMANIAN, Venkatesh. The new VIX as a market signal: It still works. **Equity Derivatives Strategy: Options & Volatility**, 2003.

RONN, Ehud I. A new linear programming approach to bond portfolio management. **Journal of Financial and Quantitative Analysis**, Cambridge University Press, v. 22, n. 4, p. 439–466, 1987.

RUDEBUSCH, Glenn D; WU, Tao. A macro-finance model of the term structure, monetary policy and the economy. **The Economic Journal**, Oxford University Press Oxford, UK, v. 118, n. 530, p. 906–926, 2008.

RUDEBUSCH, Glenn D; WU, Tao. Accounting for a shift in term structure behavior with no-arbitrage and macro-finance models. **Journal of Money, Credit and Banking**, Wiley Online Library, v. 39, n. 2-3, p. 395–422, 2007.

SARNO, Lucio; SCHNEIDER, Paul; WAGNER, Christian. The economic value of predicting bond risk premia. **Journal of Empirical Finance**, v. 37, n. 100, p. 247–267, 2016.

SHIN, Minchul; ZHONG, Molin. Does realized volatility help bond yield density prediction? **International Journal of Forecasting**, Elsevier, v. 33, n. 2, p. 373–389, 2017.

STARK, Tom et al. Realistic evaluation of real-time forecasts in the Survey of Professional Forecasters. **Federal Reserve Bank of Philadelphia Research Rap, Special Report**, v. 1, 2010.

SVENSSON, Lars EO. **Estimating and interpreting forward interest rates: Sweden 1992-1994**. [S.l.], 1994.

TABAK, Benjamin. Testing the expectations hypothesis in the Brazilian term structure of interest rates: a cointegration analysis. **Applied Economics**, Taylor & Francis, v. 41, n. 21, p. 2681–2689, 2009.

TABAK, Benjamin Miranda; ANDRADE, Sandro Canesso de. Testing the expectations hypothesis in the Brazilian term structure of interest rates. **Brazilian Review of Finance**, v. 1, n. 1, pp–19, 2003.

THORNTON, Daniel L.; VALENTE, Giorgio. Out-of-Sample Predictions of Bond Excess Returns and Forward Rates: An Asset Allocation Perspective. **Review of Financial Studies**, v. 25, n. 10, p. 3141–3168, 2012.

TIMMERMANN, Allan. Forecast combinations. **Handbook of economic forecasting**, Elsevier, v. 1, p. 135–196, 2006.

VAN DIJK, Dick; KOOPMAN, Siem Jan; VAN DER WEL, Michel; WRIGHT, Jonathan H. Forecasting interest rates with shifting endpoints. **Journal of Applied Econometrics**, Wiley Online Library, v. 29, n. 5, p. 693–712, 2014.

VASICEK, Oldrich. An equilibrium characterization of the term structure. **Journal of financial economics**, Elsevier, v. 5, n. 2, p. 177–188, 1977.

VICENTE, José Valentim Machado; GRAMINHO, Flávia Mourão. Decompondo a inflação implícita. **Revista Brasileira de Economia**, SciELO Brasil, v. 69, p. 263–284, 2015.

VICENTE, José Valentim Machado; GUILLEN, Osmani Teixeira de Carvalho. Do inflation-linked bonds contain information about future inflation? **Revista Brasileira de Economia**, SciELO Brasil, v. 67, p. 251–260, 2013.

VIEIRA, Fausto; FERNANDES, Marcelo; CHAGUE, Fernando. Forecasting the Brazilian yield curve using forward-looking variables. **International Journal of Forecasting**, Elsevier, v. 33, n. 1, p. 121–131, 2017.

WAN, Eric A; VAN DER MERWE, Rudolph. The unscented Kalman filter for nonlinear estimation. In: IEEE. ADAPTIVE Systems for Signal Processing, Communications, and Control Symposium 2000. AS-SPCC. The IEEE 2000. [S.l.: s.n.], 2000. P. 153–158.

WELCH, Ivo; GOYAL, Amit. A Comprehensive Look at The Empirical Performance of Equity Premium Prediction. **Review of Financial Studies**, v. 21, n. 4, p. 1455–1508, July 2008.

WRIGHT, Jonathan H. Term premia and inflation uncertainty: Empirical evidence from an international panel dataset. **American Economic Review**, American Economic Association, v. 101, n. 4, p. 1514–1534, 2011.

YU, Wei-Choun; SALYARDS, Donald M. Parsimonious modeling and forecasting of corporate yield curve. **Journal of Forecasting**, Wiley Online Library, v. 28, n. 1, p. 73–88, 2009.

ZHANG, YM; DAI, GZ; ZHANG, HC, et al. A SVD-based extended kalman filter and application to flight state and parameter estimation of aircraft. **Control Theory and Applications**, v. 13, n. 1, p. 107–144, 1996.