ISSN 1518-3548

BANCO CENTRAL DO BRASIL

Working Paper Series 288

Forecasting Bond Yields with Segmented Term Structure Models

Caio Almeida, Axel Simonsen and José Vicente

July, 2012

ISSN 1518-3548
CNPJ 00.038.166/0001-05

Working Paper Series	Brasília	n. 288	Jul.	2012	p. 1-44

Working Paper Series

Edited by Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Benjamin Miranda Tabak – E-mail: benjamin.tabak@bcb.gov.br
Editorial Assistant: Jane Sofia Moita – E-mail: jane.sofia@bcb.gov.br
Head of Research Department: Adriana Soares Sales – E-mail: adriana.sales@bcb.gov.br

The Banco Central do Brasil Working Papers are all evaluated in double blind referee process.

Authorized by Carlos Hamilton Vasconcelos Araújo, Deputy Governor for Economic Policy.

General Control of Publications

Banco Central do Brasil

Secre/Comun/Cogiv

SBS – Quadra 3 – Bloco B – Edifício-Sede – 1º andar

Caixa Postal 8.670

70074-900 Brasília – DF – Brazil

Phones: +55 (61) 3414-3710 and 3414-3565

Fax: +55 (61) 3414-1898

E-mail: editor@bcb.gov.br

Consumer Complaints and Public Enquiries Center

Banco Central do Brasil

Secre/Comun/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Fax: +55 (61) 3414-2553

Internet: <http://www.bcb.gov.br/?english>

Forecasting Bond Yields with Segmented Term Structure Models *

Caio Almeida[†] Axel Simonsen[‡] José Vicente[§]

Abstract

Recent empirical analysis of interest rate markets documents that bond demand and supply directly affect yield curve movements and bond risk premium. Motivated by those findings we propose a parametric interest rate model that allows for segmentation and local shocks in the term structure. We split the yield curve in segments presenting their own local movements that are globally interconnected by smoothing conditions. Two classes of segmented exponential models are derived and compared to successful term structure models based on a sequence of out-of-sample forecasting exercises. Adopting U.S. interest rates data available from 1985 to 2008, the segmented models present overall better forecasting performance suggesting that local shocks might indeed be important determinants of yield curve dynamics.

Keywords: Yield curve forecasting, segmented models, preferred-habitat theory.
JEL Classification: C53, C58, G12.

*We would like to thank Frank Diebold, Siem Koopman, Glenn Rudebusch, Michel Van Der Wel (discussant) and the participants of the Workshop on Yield Curve Modelling and Forecasting at Erasmus University Rotterdam (2010).

[†]E-mail: calmeida@fgv.br, Graduate School of Economics, Getulio Vargas Foundation, Rio de Janeiro, Brazil.

[‡]E-mail: axel.simonsen@fgvmail.br, Graduate School of Economics, Getulio Vargas Foundation, Rio de Janeiro, Brazil.

[§]E-mail: jose.valentim@bcb.gov.br, Central Bank of Brazil.

1 Introduction

There is now considerable empirical evidence that Treasury supply and bond clientele demand have large effects on a variety of yield spreads, term structure movements, and even on bond risk premium.[1] Those findings are clearly in accordance with the preferred-habitat theory of the term structure (Culbertson (1957), and Modigliani and Sutch (1966)), which advocates that interest rates for each maturity may be influenced by local shocks.[2]

If demand and supply shocks appear to be important components of interest rate dynamics, is it possible to explore this fact to better forecast future interest rates? Can we combine empirical evidence and economic theory to provide better econometric performance? In this paper, we try to address those two questions.

We are certainly not the first to adopt economic theory to improve statistical models.[3] Among others, Campbell and Thompson (2008) impose theoretically motivated restrictions on predictive regressions improving their out-of-sample performance on equity markets. Almeida, Graveline and Joslin (2011) include options on the estimation of Dynamic Affine Term Structure Models improving the econometric identification of bond risk premia. Also, an important strand of the empirical interest rate literature has tested how no-arbitrage restrictions affect model forecasting ability.[4] Motivated by the preferred-habitat theory, the first contribution of this paper is to build segmented models with local term structure shocks and analyze their econo-

[1]See Krishnamurthy (2002), Longstaff F. (2004), Greenwood and Vayanos (2009), and Krishnamurthy and Vising-Jorgensen (2010).

[2]The preferred-habitat theory has been recently reconsidered by Vayanos and Vila (2009) via a modern equilibrium term structure model; and it has been successfully adopted by Greenwood and Vayanos (2010) to empirically examine how government debt maturities affect bond yields and excess returns. For an alternative formal examination of the preferred-habitat theory see Cox, Ingersoll and Ross (1981).

[3]See Elliott and Timmermann (2008) for a discussion on the use of economic theory in forecasting.

[4]While Ang and Piazessi (2003), Almeida and Vicente (2008), Favero, Niu and Sala (2010), Carriero (2011) and Cristensen, Diebold, and Rudebusch (2011) find that the imposition of no-arbitrage improves Root Mean Squared Error forecasts (RMSE), Duffee (2011a), Joslin, Singleton and Zhou (2010), and Joslin, Le and Singleton (2011) find that no-arbitrage is not important, at least for Gaussian models. Carriero and Giacomini (2011) find no-arbitrage to be important when using an economic measure of accuracy. Almeida and Vicente (2008) interestingly find that no-arbitrage helps to improve forecasting ability especially for parametric polynomial models with **stochastic volatility**.

metric performance. Based on a sequence of forecasting experiments, we compare the proposed models to established benchmarks and find that they consistently outperform in out-of-sample forecasting results.

The Segmented-Loading term structure models (SL) split the original set of maturities into segments separated by knots, and within each segment, the term structure dynamics can be generated by a different set of basis functions. On each segment, the local dynamic factors are the time-varying coefficients on the linear combination of basis functions. They are local factors by construction since a shock to any of those factors affects only maturities within that segment. However, by imposing that the yield curve is a smooth function of maturities, all the local dynamic factors are interconnected in a global way. With this trick, we design an econometric model that corresponds to a reduced-form version of the equilibrium model proposed by Vayanos and Vila (2009). In their model each bond maturity has its own clientele and substitution across maturities is performed by risk-averse arbitrageurs.

Initially, within each segment there are as many dynamic factors as the number of basis functions. However, after imposing the smoothing restrictions, the number of dynamic factors decreases and becomes equal to the number of knots segmenting the term structure. The existence of different segments allows local factors to capture specific information about local dynamics. Once the factors are restricted, they transfer local information to the whole yield curve. In contrast, in traditional term structure models, all the movements are directly related to the whole set of maturities in the spirit of principal components as in Litterman and Scheinkman (1991).

There is a lot of flexibility in terms of model specification within the class of SL models: Choosing the number of knots splitting segments, the number and parametric form of basis functions within segments, if each segment will or not have the same parametric functions, and the choice of latent factor dynamics. In this paper we concentrate in how the choice of parametric functions affects forecasting.[5] We choose basis functions to be either polynomial or exponential functions. The Exponential Segmented Models are designed with Nelson and Siegel (1987) exponential-type functions, directly generalizing the model proposed by Diebold and Li (DL, 2006). The polynomial models adopt the Bowsher and Meeks (BM, 2008) piecewise natural cubic spline functions.

[5]An alternative way to test the importance of segmentation could be obtained by varying the number and position of knots.

Another important contribution of this paper is methodological: We design a set of forecasting experiments to compare models in a robust way. In general, researchers evaluate and compare term structure models by choosing only one specific in-sample and one out-of-sample period (both ad-hoc). In contrast, we work with several experiments were the starting points of the in-sample (out-of-sample) windows roll from January of 1985 (1994) to January of 1993 (2002). This allows us to obtain a large number of out-of-sample experiments: A total of 96 experiments that are truly based on out-of-sample data. We report information on the average of all experiments (Table 1) and also on each individual experiment (Table 2 and pictures 6, 7, 8 and 5). Each individual forecasting experiment adopts a Diebold and Mariano (1995) test with quadratic loss and a 5% significance level. It should be clear that as we vary the in-sample (out-of-sample) windows on a monthly basis, there will exist some correlation between the results of out-of-sample experiments that adopt similar windows. Nevertheless, with this procedure we obtain a dynamic view on each competing model and produce results that are robust to data mining criticisms.[6]

We compare out-of-sample forecasts of two versions of the exponential SL model, and one of the polynomial SL model (BM) against two very successful competitors according to the term structure literature: The Random Walk (RW), and the DL model.[7] Comparisons are based on a dataset of U.S zero coupon Treasury yields with monthly frequency, ranging from January of 1985 to December of 2008. We first verify that the three SL models have consistently smaller RMSEs than the DL for all maturities and forecasting horizons. Considering the average of all out-of-sample experiments, the SL models have RMSEs similar to the RW for shorter forecasting horizons (1-, 6-month) but achieve significantly smaller RMSEs for the 12-month forecasting horizon. On the other hand, considering each individual experiment based on a Diebold and Mariano test, the SL models present smaller RMSEs than the RW for most horizons and maturities, with a small number of ties.

Several papers deal with the issue of forecasting future interest rates.[8]

[6]Hansen and Timmermann (2011) propose a methodology robust to mining over the sample split point in out-of-sample forecast evaluation tests. Rossi and Inoue (2011) propose a test robust to the window size choice.

[7]Carriero, Kapetanios, and Marcellino (2010) show that a large number of term structure models are outperformed by the RW (see section 3.2 for more details).

[8]For seminal papers, see Fama and Bliss (1987) and Campbell and Shiller (1991). See also Duffee (2002), Ang and Piazessi (2003), DL (2006), Diebold, Rudebusch and Aruoba

Duffee (2002) was one of the first to consider model forecasting ability based in out-of-sample data. He analyzed the class of arbitrage-free affine term structure models, showing that those models are not especially useful to forecast future interest rates: They are outperformed by the RW, which basically repeats the last observation as a forecast. Diebold and Li (2006) adopted a parsimonious three-factor exponential term structure model that produced smaller RMSEs than the RW, specially for long-horizon forecasts.[9] Bowsher and Meeks (2008) proposed a latent cubic spline approach whose goal was to capture high-dimensional term structures. Their model was able to produce short-horizon (1-month) forecasts better than the RW, directly complementing the results obtained by DL. Jungbaker, Koopman and V.D. Wel (2010) provide formal statistical procedures to decide for the optimal number of knots in a smooth version of a general dynamic factor model. Based on a panel of U.S. data, they find that the Smooth Dynamic Factor Model (SDFM) with a large number of spline knots is statistically supported by in-sample data. In addition, the SDFM produces slightly better out-of-sample forecasts than other versions of the dynamic factor model, including a state-space version of the DL model.

The rest of the paper is organized as follows. Section 2 describes the theoretical model. Section 3 presents a description of the data adopted, model estimation, and forecasting results. Section 4 concludes. The Appendix contains technical details about the model derivation.

2 Segmented Loading Model

Our model builds on the work of Bowsher and Meeks (2008), who define the yield curve as a piecewise polynomial function plus an error term. Letting $[T_m, T_M]$ be the interval for maturities of the yields, consider the following

(2006), Favero, Niu and Sala (2007), Almeida and Vicente (2008), BM (2008), Mönch (2008), Exterkate et al. (2010), Joslin, Singleton and Zhou (2010), Jungbaker, Koopman and V.D. Wel (2010), Almeida, Graveline and Joslin (2011), Carriero (2011), Carriero and Giacomini (2011), Cristensen, Diebold, and Rudebusch (2009, 2011), Duffee (2011a), Joslin, Le and Singleton (2011), among others. For the sake of clarity, we choose to discuss only the ones closest to our work.

[9]For a list of papers that were inspired by the DL model see Diebold and Rudebusch (2011).

partition fixed over time:

$$\phi = \{T_m = x_0 < x_1 < ... < x_k = T_M\}, \tag{1}$$

where the x_i's are denominated knots of the model. The yield with maturity $x \in [T_m, T_M]$ at time t will be expressed by:

$$y_t(x) = \sum_{i=1}^{k} f_t^i(x)\chi(A_i), \tag{2}$$

where $\chi(A_i)$ is the indicator function of set A_i which is defined by:

$$A_i = \left\{ \ \{x \,|x_{i-1} < x \leq x_i\}, \qquad i = 1, ..., k \ \right\} \tag{3}$$

and

$$f_t^i(x) = a_t^i + b_t^i g_i(x) + c_t^i h_i(x) + d_t^i z_i(x), \qquad i = 1, ..., k. \tag{4}$$

The interpolation functions g_i, h_i and z_i represent the local loadings of the yield curve within each segment, and are also fixed over time. To ensure smoothness, we require those functions to be of class C^2 in the closure of A_i and to satisfy the splines constraints, i.e., f_t^{i-1} and f_t^i share common values of derivatives of order zero (the function), one and two at each knot.

Let $\beta_t = [\beta_t^1 \ \cdots \beta_t^k] = [a_t^1 \ b_t^1 \ c_t^1 \ d_t^1 \ \cdots \ a_t^k \ b_t^k \ c_t^k \ d_t^k]'$ be the vector of parameters and $Y_t(\tau)$ and τ the $m \times 1$ vectors of the observable yields and respective maturities. The model can be represented in the following matrix specification:

$$Y_t(\tau) = W(\tau)\beta_t + \varepsilon_t(\tau) \quad \text{s.a.} \quad R(\tau)\beta_t = 0. \tag{5}$$

where W and R are $m \times 4k$ and $3k - 1 \times 4k$, matrices, and ε is a residual term. Each line of W defines the spline function for each specific maturity (details appear in the Appendix). Condition $R(\tau)\beta_t = 0$ represents the spline restrictions guaranteeing that even though coefficients and functional loadings may change with segments, the term structure is smooth at the domain of maturities $[T_m, T_M]$.

We distinguish two forms of segmentation conveniently denominated weak and strong. Under the former, only factor local dynamics within each segment (β's) contribute to local movements since factor loadings are the same across segments (functions g_i, h_i, and z_i are independent of i). It may be viewed as a reduced-form version of the equilibrium model proposed in

Vayanos and Vila (2009). Segmentation arises because the same factor loading (or term structure movement) may present different weights across segments A_i's. On the other hand, under the strong segmentation form, not only factors dynamics vary across segments but also the functional form of the loadings, i.e., both the coefficients a^i, b^i, c^i and d^i and the functions g_i, h_i, and z_i are not independent of i.

Since there are only equality restrictions, we are able to rewrite Equation (5) in an unconstrained form, reducing the dimensionality of the model from $4k$ to $k+1$. After some algebraic manipulations (see the Appendix), observable yields must satisfy the following equation:

$$Y_t(\tau) = Z(\tau)\theta_t + \varepsilon_t(\tau), \tag{6}$$

where the dimensions of Z and θ_t are $m \times k+1$ and $k+1 \times 1$, respectively. Therefore, like in BM (2008), our model can be specified in a restricted form (Equation (5)) or in an unrestricted form (Equation (6)).

Equation (6) is valid for any pair of maturities and yields. In particular, if the yields corresponding to the knots ϕ are obtained without error, then $Y_t(\phi) = Z(\phi)\theta_t$. In general, the yields $Y_t(\phi)$ are latent. However, since the matrix $Z(\phi)$ is square and invertible, there is a relationship between $Y_t(\tau)$ and $Y_t(\phi)$:

$$Y_t(\tau) = Z(\tau)\left(Z(\phi)\right)^{-1} Y_t(\phi) + \varepsilon_t(\tau) \Longrightarrow Y_t(\tau) = \Pi(\tau,\phi)Y_t(\phi) + \varepsilon_t(\tau). \tag{7}$$

Thus, the latent yields can be estimated in the cross-section by OLS:

$$\widehat{Y}_t(\phi) = Z(\phi)(Z(\tau)'Z(\tau))^{-1}Z(\tau)'Y_t(\tau). \tag{8}$$

If we set $g_i(x) = x$, $h_i(x) = x^2$ and $z_i(x) = x^3$ for all i in Equation (7) we have the cross-section specification of the BM model. In this work, we propose two other versions of Equation (7) using exponential loadings instead of polynomials. The first (NS4) is based on the Nelson and Siegel (1987) model, adapted to a dynamic context by Diebold and Li (2006), with the following loadings:

$$g_i(x) = \frac{(1-\mathrm{e}^{-\lambda_1 x})}{\lambda_1 x}, \qquad h_i(x) = \frac{(1-\mathrm{e}^{-\lambda_1 x})}{\lambda_1 x} - \mathrm{e}^{-\lambda_1 x}, \qquad z_i(x) = \frac{(1-\mathrm{e}^{-\lambda_2 x})}{\lambda_2 x}, \qquad \forall i,$$

where λ_1 and λ_2 are constants that control the decaying speed on each exponential. The BM and NS4 models are examples of segmentation of the yield curve in the weak form according to the definition above.

The second version (NS4E) is also based on the Nelson and Siegel (1987) / DL model. However, we allow different functional forms for g_i, h_i, and z_i within each segment. Despite the fact that we introduce a discontinuity in the loadings at the knots, the smoothing restrictions guarantee that the yield curve will remain smooth and continuous. Under this model, each segment has its own dynamics and functional loadings, while smoothing constraints connect local to global dynamics across maturities, reinforcing again the analogy with the preferred-habitat theory. We propose the following form for the interpolation functions:

$$\begin{aligned} g_i(x) &= \frac{(1-\mathrm{e}^{-\lambda_1\Lambda_i(x)})}{\lambda_1\Lambda_i(x)}, \\ h_i(x) &= \frac{(1-\mathrm{e}^{-\lambda_1\Lambda_i(x)})}{\lambda_1\Lambda_i(x)} - \mathrm{e}^{-\lambda_1\Lambda_i(x)}, \\ z_i(x) &= \frac{(1-\mathrm{e}^{-\lambda_2 x})}{\lambda_2 x}, \end{aligned}$$

where Λ_i's are functions that introduce discontinuities in the loadings at the knots, $i = 1, \ldots, k$. Note that the NS4 model is a special case of NS4E where $\Lambda_i(x) = x$ for all i.

Although there are many possibilities for Λ, in the empirical section, we propose a linear functional form given by:

$$\Lambda_i(x) = x - x_{i-1}(1-p), \quad x \in A_i \text{ and } p \in \mathbb{R}. \tag{9}$$

It is important to observe that while the parameters λ_1 and λ_2 control the decay rates on the exponential loadings, the parameter p has a very different role. It controls the degree of loading segmentation.

The SL models can be classified in two different ways: with respect to the functional form of the loadings, there are polynomial versions (BM model) and exponential versions (NS4 and NS4E); with respect to the loading form, models may present strong (NSE4) or weak (BM and NS4) segmentation.

We follow the cointegration-based yield curve literature (see Hall, Anderson and Granger (1992), and Bowsher and Meeks (2008)) specifying an Error Correction Model (ECM) for the dynamics of the latent yields. The representation of the model in a state-space form is:

$$Y_t(\tau) = \Pi\,(\tau, \phi)\, Y_t(\phi) + \varepsilon_t(\tau), \tag{10}$$

$$\Delta Y_{t+1}(\phi) = \alpha(\rho' Y_t(\phi) - \mu_s) + \Psi \Delta Y_t(\phi) + \nu_t, \tag{11}$$

for $t = 1, 2, \ldots$ where α, ρ and Ψ are $k+1 \times k+1$ matrices, and μ_s is a $k+1$ vector. ρ is a cointegration matrix that here is fixed such that $\rho' Y_t(\phi) - \mu_s$ is a stationary mean-zero vector of cointegrating relations: More specifically, $\rho' Y_t(\phi)$ are the k spreads between the knot yields. The error terms ε and ν are such that $E(\varepsilon_t(\tau)\varepsilon_t(\tau)') = \Sigma_\varepsilon$, $E(\nu_t(\tau)\nu_t(\tau)') = \Sigma_\nu$, $E(\varepsilon_t(\tau)\nu_t) = 0$ as in Bowsher and Meeks (2008).

The system (10) and (11) can be estimated with the use of a Kalman filter (Bowsher and Meeks (2008)). Alternatively, Diebold and Li (2006) suggest a simpler two-step estimation procedure that produces very strong out-of-sample forecasting results. In addition, according to Diebold and Rudebusch (2011) "little is lost in practice by using two-step estimation because there is typically enough cross-sectional variation...". By privileging simplicity, we adopt the two-step procedure proposed by DL, by first estimating the latent yields running an OLS (see Equations (7) and (8)), and then estimating the ECM in Equation (11) to obtain the parameters $\{\alpha, \mu_s, \Psi, \Sigma_\nu\}$.

3 Empirical Results

3.1 Data

We adopt a time series provided by the Federal Reserve Bank of St. Louis of monthly U.S. Treasury yields starting in January 1985 ending in December 2008. Figure 1 presents the time evolution of those yields, with maturities of 3-, 6-, 12-, 24- 36-, 60-, 84- and 120-months. Observe that on a lower frequency (annual) the yields decline over time, but on a monthly basis there are many periods where the yields oscillate around the same region.[10] The yields achieve a maximum of 10% at the beginning of the sample and a minimum of 2% at the end, with term structures almost always having an upward sloping shape.

3.2 Competing Models

The goal of our exercise is to identify if SL term structure models generate better out-of-sample forecasting results than traditional models adopted by

[10] Oscillation favors the RW model while trends favor autoregressive models.

interest rate researchers. Although we could choose from a large pool of models, we preferred to concentrate on a small number of well-established benchmarks.

Carriero, Kapetanios, and Marcellino (2010) showed in a recent study that a large number of term structure models fail to outperform the RW, for most maturities, for either 1-,3-, 6-, or 12-month forecasting horizons. Moreover, the superiority of the RW is especially strong for more recent data (2001-2008). The outperformed models include one-dimensional and multivariate autoregressive models for yields (ARs and VARs), the Stock and Watson (2002) factor model, forward rate models (like Fama and Bliss (1987); Cochrane and Piazzesi (2005)), the DL model, and to a less extent, Dynamic Affine Term Structure Models (Duffie and Kan (1996); Duffee (2002); Ang and Piazzesi (2003)).

Inspired by their negative results, we concentrate on comparisons of SL term structure models with the RW and the DL model. We decided to keep the DL model because at the sample analyzed by Diebold and Li (2006) it strongly outperformed the RW in out-of-sample forecasts, especially for long forecasting horizons. The decision to exclude affine arbitrage-free models came from the findings that no-arbitrage is not important for forecasting purposes at least when the models are affine and Gaussian (see Duffee (2011a); Joslin, Singleton and Zhou (2010)).

In summary, we provide a fundamental comparison of segmented models to a no-change hard-to-beat benchmark (RW) and to one of the most successful dynamic factor models (DL, see also Diebold and Rudebusch (2011)).

3.3 Model Estimation

An important issue is to identify the appropriate number of latent factors that will drive term structure dynamics. Traditionally, Litterman and Scheinkman (1991) have showed that three factors describe more than 95% of the variability of a whole set of interest rates.[11] Building on their result, most term structure models adopt three or less factors to capture the dynamics of the yield curve[12]. On the other hand, Cochrane and Piazzesi (2005) identify that

[11] In fact, this result is robust for different datasets and periods of time.

[12] See, for instance, Dai and Singleton (2000), Duffee (2002) and references therein for arbitrage-free models. For parametric exponential term structure models, see Nelson and Siegel (1987), and Diebold and Li (2006) and for polynomials models see Almeida, Duarte and Fernandes (1998).

more than three factors are necessary to capture future short-term rate dynamics and bond risk premia.[13] Ludvigson and Ng (2009) and Joslin, Priebsch and Singleton (2011) identify macroeconomic risks that are not spanned by the traditional three movements of the term structure and propose models with a larger number of latent factors.[14] Duffee (2011b) implements a five-factor Gaussian affine filtered model to monthly U.S. Treasury yields, finding a latent factor with almost no effect in cross-sectional fitting that explains around 30% of the total variance in expected bond returns.

All those papers together suggest that it is not unreasonable to search in a range of 3 to 5 factors to determine the optimal number of latent factors in a term structure model.[15] Since the number of factors in our model directly relates to the number of term structure knots, by fixing the number of knots we fix the number of latent factors and vice versa. In this paper, we follow Duffee (2011b) and Joslin, Priebsch and Singleton (2011) and adopt five factors for forecasting purposes. This will directly translate into five knots, and 16 loadings in the unrestricted form of the model.

Rather then trying to find the best forecasting model by extensive econometric search, we are more interested in providing a basic test for the importance of segmentation for forecasting purposes. Therefore, we arbitrarily fix the knots positions at 1, 6, 21, 84 and 120 months, to represent four different segments of the term structure: two short-maturity segments, one medium-maturity segment and one long-maturity segment. An alternative would be to search for the optimal position of knots, considering one of a number of possible techniques. Bowsher and Meeks (2008) identify, by using an in-sample selection procedure, the position and number of knots in their spline model by minimizing the residual sum of squares for cross-sectional fitting of yields.[16] Jungbaker, Koopman and Van Der Wel (2010) suggest an

[13] In particular, Cochrane and Piazzesi (2005) find that a linear combination of forward rates that has an important component unrelated to the traditional level, slope and curvature movements, has a strong forecasting power on expected excess returns on bonds.

[14] Based in the methodology proposed by Joslin, Singleton and Zhou (2010), Joslin, Priebsch and Singleton (2011) propose five-factor Gaussian affine models, where three factors are latent and two are macroeconomic variables that have no effect on the cross section of yields but directly affect their dynamics.

[15] Among other articles that work with more than three factors, we can cite the four factor arbitrage-free models of Collin-Dufresne, Goldstein, and Jones (2009), and Joslin (2007), and some dynamic versions of the Svensson (1994) model (see Koopman, Mallee and Van der Wel (2008), De Pooter (2007) and Almeida et al. (2009)).

[16] We have also tested with the original knots proposed by Bowsher and Meeks (2008).

optimal procedure also implemented based on in-sample data. Alternatively, we could split the in-sample data between training and testing samples and test the forecasting ability of the SL models based on different combinations of knots. We leave those extensions for future research.

For the exponential models, in this paper, rather than estimating each decay parameter lambda we fix both at values that will concentrate the curvature factors in important areas of maturities. By following DL, we fix $\lambda_1 = 0.0609$, making the first curvature factor loading to achieve its maximum at 30 months, directly relating to a medium-term factor. Following Almeida at el (2009), we decided to fix λ_2 at a higher value to make the second curvature term to achieve a maximum at the short end of the yield curve. In fact, Almeida, et al (2009), adopt a Svensson dynamic model and estimate the two decay parameters (lambdas) that minimize in-sample RMSEs. They find that a second curvature factor controlling short-term movements is very important in out-of-sample forecasting exercises. Therefore we fix $\lambda_2 = 0.24$ making the second curvature to achieve its maximum around 7 months, becoming responsible for driving short-term movements before spline restrictions. We will see that this short-maturity peek for the second curvature factor will generate important short-term movements after the spline smoothing restrictions are imposed.

For the NS4E model, we have to either estimate or fix the parameter p from Equation (9) that allows for differences in the decaying structure of the exponential functions. Note that p=0 corresponds to maximum loading segmentation and p=0.95 to minimum segmentation, while if p=1 the model becomes the NS4 model, with no loading segmentation. For simplicity, we have arbitrarily decided for a moderate level of segmentation by picking up p= 0.5. However, before proceeding with a blind view, we were interested in verifying if different levels of loading segmentation imply much difference in forecasting results. Therefore, we have experienced with a grid of values for p, ranging from 0 to 0.95 with increments of 0.05. For values of p close to 0.95, the NS4E model produces slightly worse forecasts than for p=0.5. However, for values of p close to 0 the results are mixed. For instance, while the NS4E(p=0) produces better long-horizon forecasts than NS4E(p=0.5) with a decrease of 2 to 3% on average RMSEs, it produces worse results for short-horizon forecasts with an increase of around 1% on average RMSEs.

The results obtained with their set of five knots are qualitatively the same, and are available upon request.

In summary, apparently a strong degree of loading segmentation behaves slightly better but we keep the moderate segmentation in our analysis.[17]

Figure 2 shows the unrestricted loadings for the BM, NS4 and NS4E models, before the transformation that imposes the smoothing restrictions to the term structure at each knot. Vertical black lines indicate the positions of knots. Note that each of the three models contains a set of four loading functions within each segment. For the BM and NS4 models, we keep the loading functions equal across the segments. Then, the sixteen loadings that appear considering all segments are equivalent to four unrestricted loadings for the whole set of maturities, exactly as in traditional term structure models. The difference lies on the yield curve response to shocks in the loadings: On the BM and NS4 models the response is dependent on the segment, while in traditional models the response depends on the whole set of maturities. On the other hand, NS4E model presents unrestricted discontinuous loadings (for the slope and first curvature) implying that the responses to shocks are not only dependent on the particular segments of the term structure, but also may affect the segments with different intensities.

Figure 3 shows the restricted loadings. Now instead of having sixteen ($4k$) loadings we will have only five ($k+1$). First note that all models present three of the five factors with loadings very similar to the traditional level, slope and curvature of Litterman and Scheinkman (1991). The level appears in red and is the function that is constant for all maturities (equal to one). Slope is represented by a cyan line in the BM and NS4 model and by a magenta line in the NS4E model. It varies a little bit across models as expected. For the exponential models, the slope is non-linear being almost constant on the NS4E model for short-term maturities. Curvature (magenta line for the BM and NS4 models and cyan line for the NS4E) appears more pronounced in the exponential models than in the BM.

The most interesting differences come from the two extra factors that decay to zero after some range of maturities. Note that under the exponential models both factors (blue and green) have a faster decay, specially the green factor. This fast decay is related to the role that the second curvature factor plays at the short-end of the yield curve, before imposing spline smoothing restrictions. Indeed, while the green factor achieves zero for a maturity of 60 months under the BM model, it achieves zero for a maturity around 45 months, under the exponential models. These loadings indicate the idea

[17] Results under different values of p are available upon request.

of segmentation that we want to test. By creating segments and imposing smoothing restrictions we generate loadings that are important for only certain sub-sets of maturities. Those special loadings may help to pick up some additional dynamics that is not captured by the traditional non-segmented factor models.

Figure 4 shows examples of how the models capture some cross sectional data of yields. As expected, due to a larger number of factors, the SL models work better than the Diebold and Li (2006) model, what may suggest some overfitting. However, on Section 3.4 we show that the SL models strongly outperform the DL model in the out-of-sample yield curve forecasting exercises.

3.4 Forecasting Exercise

When evaluating the econometric performance of interest rate models, researchers work in general with a unique out-of-sample window, sometimes considering a small number of subsamples for robustness purposes.[18] Alternatively, in this paper, we compare models based on a set of multiple in-sample and out-of-sample windows, providing results less subject to data mining issues.

Although we propose the use of multiple out-of-sample windows, we still have to choose between two usual approachs adopted in comparisons of out-of-sample forecasts: Rolling-window versus recursive method (see MacCracken (2007)). The rolling-window method performs model re-estimation with a fixed-size window, while the recursive method re-estimates models with cumulative data. Giacomini and White (2006) show that the rolling-window method produces substantial forecasting accuracy gains relative to the recursive method, for important economic time series and a number of forecasting models. Therefore, based on their results we adopt the rolling-window method for each of the 96 out-of-sample experiments in our dataset.

In a recent paper, Hansen and Timmermann (2011) observe how it is important to control for the split point of a dataset into estimation and evaluation periods in out-of-sample evaluation tests, and propose a test statistic robust to mining over the start of the out-of-sample period. On a related issue, Rossi and Inoue (2011) propose a robust approach to data snoop-

[18] Papers that adopt subsample robustness tests include De Pooter (2007), Jungbaker, Koopman and Wel (2010), Koopman and Wel (2011).

ing across the length of the estimation window, in rolling-window forecasting evaluations.[19] In this paper, we do not examine variations of window-size nor mining over the sample split. Instead, we vary the datasets creating different experiments but keeping both window-size and sample split constants, with in-sample window-size equal to $108 - h + 1$ months (h, forecasting horizon) and out-of-sample window-size equal to 84 months.

An experiment consists of a sequence of 84 out-of-sample monthly forecasts constructed with the rolling-window method. For each experiment, and each forecasting horizon, we define the Root Mean Square Error (RMSE) of the m-maturity yield by:

$$RMSE(T_E, h, m) = \frac{\sum_{i=1}^{84} \left(\hat{Y}_{T_E-84+i}(m) - Y_{T_E-84+i}(m) \right)^2}{84}, \qquad (12)$$

where T_E is the last prediction date and $\hat{Y}_{T_E-84+i}(m) - Y_{T_E-84+i}(m)$ is the prediction error, since $\hat{Y}_{T_E-84+i}(m) = E_{T_E-84+i-h}\left(Y_{T_E-84+i}(m)\right)$ is the model expected value for the m-maturity yield. To calculate this expected value, the model is estimated based on data between $T_E - 192 + i$ and $T_E - 84 + i - h$ (in-sample period). For example, suppose T_E = January 2001 and $h = 6$. In this experiment, the first prediction is done for February 1994, with an in-sample period from February 1985 to August 1993. Similarly, the eighty-fourth (last) prediction is done for January 2001 with an in-sample period from January 1992 to July 2000. The set of all 96 different experiments is obtained by changing T_E monthly between December of 2000 and December of 2008.

Our methodology based on the analysis of multiple datasets is related to Giacomini and Rossi (2008). They develop statistical tests to examine the stability over time of out-of-sample relative forecasting performance of a pair of models in the presence of an unstable environment. While they provide an empirical application to exchange rate markets, our objective is to identify possible time-varition in model forecasting ability with an application to interest rate markets.

Figure 5 presents the dynamics of the RMSE for the RW along the 96 experiments based on different out-of-sample datasets.[20] For each experiment,

[19] See also Pesaran and Timmermann (2005) who analyze small sample properties of autoregressive models under the presence of breaks and offer practical recommendations regarding the length of estimation window for such models.

[20] Although, some pairs of neighbor datasets (among the 96) share a large subset of

the element in the x-axis is T_E, and the value in the y-axis corresponds to the average of the 84 out-of-sample squared prediction errors (see Equation (12)). For the sake of clarity, although our models are estimated based on the whole set of maturities, we only provide results on the 3-,12-, 36-, 60-, and 120-month maturities.[21]

Note that the RMSEs for the 1-month forecasting horizon are approximately constant through time, ranging between 20 and 30 basis points (bps) depending on the maturity. In contrast, the RMSEs for longer forecasting horizons (6-, and 12-month) are clearly time-varying, increasing for shorter maturities like 3-, and 12-month and strongly decreasing for longer maturities like 60- and 120-month. For instance, for a 12-month forecasting horizon the RMSE for the 120-month maturity starts around 100 bps ending 8 years later below 75 bps. This decrease along time in RMSE values for longer maturities makes it much more difficult for models to beat the RW in individual experiments.

We are particularly interested in making three kinds of comparisons: Segmented models versus the RW, segmented models versus traditional models (represented by DL here), and exponential versus polynomial segmented models. To that end, Figures 6, 7 and 8 present the time series of RMSEs (relative to the RW) for 1-, 6-, and 12-month forecasting horizons. In addition, Table 1 aggregates information on different experiments by providing on each entry (per maturity and horizon) the average of all 96 different relative RMSEs.

Observing the pictures and Table 1 we are able to extract important information relative to the performance of each term structure model under analysis. For short-maturities like 3- and 12-month, the exponential SL models have an impressive performance when compared to the RW model, for all forecasting horizons. In special, for the 1-month forecasting-horizon, in 94 out of 96 different experiments, the relative RMSE of SL exponential models is smaller than one (see Figure 6). For longer forecasting horizons, in more than 80% of the experiments, the RMSE is smaller than one, achieving values like 0.7, i.e., 30% smaller than the one obtained by the RW (see Figures 7 and 8). Those are difficult results to observe, in special, their time persistence. In contrast, considering the same maturities (3- and 12-month), the DL relative

observations, the farther a pair of datasets is in time the less information they share. In particular, the first dataset shares no observations at all with the last 12 datasets.

[21]Results for the 6-,24-, and 84-month maturities are available upon request.

RMSE is above one for most experiments. Interestingly, the strongest performance of the SL models in forecasting short-maturity yields appears to come from the two local factors driving specifically short-maturities (see Figure 3) that are not existent in traditional term structure models. Those findings may suggest the existence of local-maturity shocks driving term structure models as suggested in Vayanos and Vila (2009).

For longer maturities (36-, 60-, 120-month), all models behave similarly for the 1-month forecasting horizon. Figure 6 shows that all the time series of relative RMSEs oscillate around one, with the RW outperforming in the first half of experiments and the opposite happening in the second half. For the 6-month forecasting horizon, the RW outperforms the DL model and the BM model, but has a tie with the exponential SL models (see Figure 7 and Table 1), except for the most recent 10 experiments where the RMSE of the RW is around 10% smaller than all SL models.[22] On the other hand, when considering the 12-month forecasting horizon, the three SL models strongly outperform the RW. For experiments where T_E is in 2006 and 2007, the RMSEs are sometimes more than 20% smaller than the corresponding RMSE obtained by the RW.

By observing Figures 7 and 8, we note that, for 6- and 12-month horizons, the SL models have smaller RMSEs than DL in more than 90% of the experiments, especially in those including more recent data. The DL model behaves better than the RW benchmark in less than 10% of the experiments, while SL models consistently outperform the RW.

We note that the exponential SL models have lower RMSEs than the polynomial model (BM). For all forecasting horizons and maturities, the mean relative RMSE of the BM model is equal or higher than the mean relative RMSE of both exponential versions (see Table 1). In particular RMSEs of exponential models are more than 5% smaller for a 12-month forecasting horizon and get up to 10% smaller for short maturities. Given that very recent papers are exploring the structure of factor models with polynomial splines (see Bowsher and Meeks (2008), Jungbaker, Koopman and Wel (2010), and Koopman and Wel (2011)), our novel results suggest that exponential spline models should definitely be further explored in forecasting problems.

Table 2 presents results of individual experiments. Each individual fore-

[22]Precisely in those datasets the RW achieves very small RMSEs when forecasting future long maturity yields as shown in Figure 5.

casting experiment adopts a Diebold and Mariano (1995) test with quadratic loss and a 5% significance level. Candidate models (DL, BM, NS4, or NS4E) are tested against the RW. The first number in each entry of the table provides the percentage of experiments were the candidate model achieves statistically significant smaller RMSE than the RW. The second number on the same entry represents the percentage of experiments were the RW achieves statistically significant smaller RMSE than the candidate model. 100 minus the sum of those two numbers gives the percentage of experiments were there is a statistical tie between RW and candidate model.

The results in this table corroborate the findings based on the average experiment but also reveal some interesting additional information. They show, for instance, that the SL models are in fact much stronger than the RW in a robust way. In particular, BM presents RMSE statistically smaller than the RW according to the Diebold and Mariano test in around 20% of the experiments (19 in 96) for short-term maturities, and the SL exponential models outperform the RW in more than 40% of the experiments (38 in 96). The RW outperform the SL models in only a very small number of tests considering maturities and horizons, and whenever it is better, it presents statistically smaller RMSEs in less than 10% of the experiments. For the 1-month forecasting horizon the exponential models have an impressive performance, sometimes outperforming the RW model in more than 70% of the experiments. We also observe a result that is consistent with the findings by Diebold and Li (2006). While the RW achieves much better performance than the DL model for the 1-month forecasting horizon, for longer forecasting horizons, the DL model improves. Indeed, for longer forecasting horizons, DL is better than the RW for 3-,12-, and 36-month maturities and worse for longer maturities (60-, 120-month). However, in general, the RW achieves better results than the DL model.

Finally, results are mixed when comparing the two different exponential SL models. By looking at Table 1 we observe that for the 1-month forecasting horizon, the strong segmented version (NS4E) is superior, while for the 12-month horizon the weak segmented model (NS4) is better. For the 6-month horizon NS4E outperforms NS4 for short maturities, but for long maturities the opposite happens. Therefore, based on Table 1 we are not able to identify the best exponential model. However, Table 2 reveals results slightly more in favor of the NS4 model since for longer maturities (36-, 60- and 120-month) it

presents higher percentage of experiments were NS4 outperforms the RW.[23]

4 Conclusion

Motivated by recent empirical studies that evidence the preferred-habitat theory of the term structure, we propose a family of segmented term structure models with local movements to forecast future yields. Based in a series of 96 different out-of-sample experiments with U.S. Treasury yields, we identify that the segmented models produce significantly smaller RMSEs than important benchmarks like the Random Walk and the Diebold and Li (2006) model. We also show that the exponential segmented models provide better forecasting results than polynomial segmented models. Future research suggests analyzing a larger number of loading functional forms when testing how segmentation affects forecasting.

[23]Note however that we only test a very special form of strong segmentation while many different functions could be adopted instead of a linear function on maturity ($\Lambda(x)$) to create the discontinuity of the loadings

References

[1] Almeida C., A. Duarte, and C. Fernandes (1998). Decomposing and Simulating the Movements of Term Structures in Emerging Eurobonds Markets. *Journal of Fixed Income*, **1**, 21-31.

[2] Almeida C., R. Gomes, A. Leite, A. Simonsen and J. Vicente (2009). Does Curvature Enhance Forecasting? *International Journal of Theoretical and Applied Finance*, **12**, 8, 1171-1196.

[3] Almeida C., J. Graveline and S. Joslin (2011). Do Interest Rate Options Contain Information About Excess Returns? *Journal of Econometrics*, **164**, 35-44.

[4] Almeida C. and J. Vicente (2008). The Role of No-arbitrage on Forecasting: Lessons from a Parametric Term Structure Model. *Journal of Banking and Finance*, **32**, 12, 2695-1705.

[5] Ang, A., Piazzesi, M., 2003. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary Economics*, **50**, 745-787.

[6] Bowsher C. and R. Meeks (2008). The Dynamics of Economic Functions: Modelling and Forecasting the Yield Curve. *Journal of the American Statistical Association*, **103**, 1419-1437.

[7] Campbell J. and R. Shiller (1991). Yield Spreads and Interest Rate Movements: A Bird's Eye View. *Review of Economic Studies*, **58**, 3, 495-514.

[8] Campbell J. and S. Thompson (2008). Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average? *Review of Financial Studies*, **21**, 4, 1509-1531.

[9] Carriero A. (2011). Forecasting the Yield Curve Using Priors from No-arbitrage Affine Term Structure Models. *International Economic Review*, forthcoming.

[10] Carriero A., G. Kapetanios, and M. Marcellino (2010). Forecasting Government Bond Yields with Large Bayesian VARs. Working Paper, Queen Mary University of London.

[11] Carriero A., and R. Giacomini (2011). How Useful are No-arbitrage Restrictions for Forecasting the Yield Curve?. *Journal of Econometrics*, forthcoming.

[12] Christensen, J.H., Diebold, F., Rudebusch, G.D., (2009). An Arbitrage-Free Generalized Nelson-Siegel Term Structure Model. *The Econometrics Journal*, 12, 33-64.

[13] Christensen, J.H., Diebold, F., Rudebusch, G.D., (2011). The Affine Arbitrage-Free Class of Nelson-Siegel Term Structure Models. *Journal of Econometrics*, forthcoming.

[14] Cochrane J. and M. Piazzesi (2005). Bond Risk Premia. *American Economic Review*, 138-160.

[15] Collin-Dufresne P. , R. Goldstein and C. Jones (2009). Can Three Factor Affine Models Explain the Volatility of the Short Term Rate?, *Journal of Financial Economics*.

[16] Cox J.C., J. E. Ingersoll, and S. A. Ross (1981). A Re-Examination of Traditional Hypotheses About the Term Structure of Interest Rates. *Journal of Finance*, **36**, 4, 769-799.

[17] Culbertson J. (1957). The Term Structure of Interest Rates. *Quarterly Journal of Economics*, **71**, 485-517.

[18] Dai Q. and K. Singleton (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance*, **LV**, 5, 1943-1977.

[19] De Pooter M. (2007). Examining the Nelson-Siegel Class of Term Structure Models. Working Paper, Tinbergen Institute Discussion Paper.

[20] Diebold F.X. and C. Li (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics*, 130, 337-364.

[21] Diebold F.X. and R.S. Mariano (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, **13**, 253-263.

[22] Diebold F.X., G.D. Rudebusch, and B. Aruoba (2006). The Macroeconomy and the Yield Curve: A Dynamic Latent Factor Approach. *Journal of Econometrics*, **131**, 309-338.

[23] Diebold F. and G. Rudebusch (2011). *The Dynamic Nelson-Siegel Approach to Yield Curve Modeling and Forecasting*. Princeton University Press.

[24] Duffee G. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance*, **57**, 405-443.

[25] Duffee G. (2011a). Forecasting with the Term Structure: The Role of No-arbitrage. Working Paper, John Hopkins University.

[26] Duffee G. (2011b). Information in (and not in) the Term Structure. *Review of Financial Studies*, forthcoming.

[27] Duffie D. and Kan R. (1996). A Yield Factor Model of Interest Rates. *Mathematical Finance*, **6**, 4, 379-406.

[28] Elliott G. and A. Timmermann (2008). Economic Forecasting. *Journal of Economic Literature*, **46**, 3-56.

[29] Exterkate P., D. van Dijk, C. Heij, and P.J.F. Groenen (2010). Forecasting the Yield Curve in a Data-Rich Environment using the Factor-Augmented Nelson-Siegel Model. Working Paper, Econometric Institute, Erasmus University Rotterdam.

[30] Fama E. and R. Bliss (1987). The Information in Long-Maturity Forward Rates. *American Economic Review*, **77**, 4, 680-692.

[31] Favero, A.C., Niu, L., Sala, L., (2010). Term Structure Forecasting: No-arbitrage Restrictions Versus Large Information Set. *Journal of Forecasting*, forthcoming.

[32] Giacomini R. and B. Rossi (2008). Forecast Comparisons in Unstable Environments. *Journal of Applied Econometrics*, forthcoming.

[33] Giacomini R. and H. White (2006). Tests of Conditional Predictive Ability. *Econometrica*, **76**, 4, 1545-1578.

[34] Greenwood R. and D. Vayanos (2009). Price Pressure in the Government Bond Market. Working Paper, Harvard Business School.

[35] Greenwood R. and D. Vayanos (2010). Bond Supply and Excess Bond Returns. Working Paper, Harvard Business School.

[36] Hall A.D., H.M. Anderson, and C. Granger (1992). A Cointegration Analysis of Treasury Bill Yields. *The Review of Economics and Statistics*, 74, 116-126.

[37] Hansen P.R. and A. Timmermann (2011). Choice of Sample Split in Out-of-Sample Forecast Evaluation. working paper, Stanford University and UCSD.

[38] Joslin S. (2007). Pricing and Hedging Volatility in Fixed Income Markets, Working Paper, Stanford University.

[39] Joslin S., K. Singleton, and H. Zhou (2010). A New Perspective on Gaussian Dynamic Term Structure Models. *Review of Financial Studies*, forthcoming.

[40] Joslin S., A. Le, and K. Singleton (2011). Why Gaussian Macro-Finance Term Structure Models Are (Nearly) Unconstrained Factor-VARs. Working Paper, Univeristy of Southern California, University of North Carolina, and Stanford University.

[41] Joslin S., M. Priebsch, and K. Singleton (2011). Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks. Working Paper, MIT Sloan, and Stanford University.

[42] Jungbaker B., S.J. Koopman, M. V.d. Wel (2010). Dynamic Factor Models with Smooth Loadings for Analyzing the Term Structure of Interest Rates. Working Paper, CREATES, Aarhus University.

[43] Koopman S. J., M.I.P. Mallee and M. van der Wel (2010). Analyzing the Term Structure of Interest Rates using the Dynamic Nelson-Siegel Model with Time-Varying Parameters. *Journal of Business and Economic Statistics*, **28**, 329 - 343.

[44] Koopman S. J. and M. van der Wel (2011). Forecasting the U.S. Term Structure of Interest Rates using a Macroeconomic Smooth Dynamic Factor Model. Working Paper, Tinbergen Institute Discussion Paper.

[45] Krishnamurthy A. (2002). The Bond/Old Bond Spread. *Journal of Financial Economics*, **66**, 463-506.

[46] Krishnamurthy A. and A. Vissing-Jorgensen (2010). The Demand for Treasury Debt. Working Paper, Northwestern University.

[47] Litterman R. and Scheinkman J. A. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income*, **1**, 54-61.

[48] Longstaff F. (2004). The Flight to Liquidity Premium in U.S. Treasury Bond Prices. *Journal of Business*, **77**, 511-526.

[49] Ludvigson S. and S. Ng (2009). Macro Factors in Bond Risk Premia. *Review of Financial Studies*, **22**, 12, 5027-5067.

[50] McCraken M.W. (2007). Asymptotics for Out-of-Sample Tests of Granger Causality. *Journal of Econometrics*, **140**, 719-752.

[51] Modigliani F. and R. Sutch (1966). Innovations in Interest-Rate Policy. *American Economic Review*, **56**, 178-197.

[52] Mönch, E. (2008). Forecasting the Yield Curve in a Data-Rich Environment: A No-arbitrage Factor-augmented VAR Approach. *Journal of Econometrics*, **146**, 26-43.

[53] Nelson C. and A. Siegel (1987). Parsimonious Modeling of Yield Curves. *Journal of Business*, 60, 4, 473-489.

[54] Pesaran M.H. and A. Timmermann (2005). Small Sample Properties of Forecasts from Autoregressive Models under Structural Breaks. *Journal of Econometrics*, 129, 183-217.

[55] Rossi B. and A. Inoue (2011). Out-of-Sample Forecast Tests Robust to the Window Size Choice. Working paper, Duke University and North Carolina State University.

[56] Stock J.H. and M.W. Watson (2002). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American Statistical Association*, **97**, 1167-1179.

[57] Svensson L. (1994). Monetary Policy with Flexible Exchange Rates and Forward Interest Rates as Indicators. *Institute for International Economic Studies*, Stockholm University.

[58] Vayanos D. and J.-L. Vila (2009). A Preferred-Habitat Model of the Term Structure of Interest Rates. Working Paper, London School of Economics.

Appendix

In this Appendix, we describe the construction of matrices W and R of Equation (5), and Z of Equation (6). In order to simplify the notation, we drop the time index t.

The splines restrictions can be written as

$$f^i(x_i) = Y(x_i), \qquad i = 0, 1, ..., k \tag{13}$$

$$\tag{14}$$

and

$$f^i(x_i) = f^{i+1}(x_i), \qquad i = 1, ..., k-1 \tag{15}$$

$$f^{i\prime}(x_i) = f^{i+1\prime}(x_i), \qquad i = 1, ..., k-1 \tag{16}$$

$$f^{i\prime\prime}(x_i) = f^{i+1\prime\prime}(x_i), \qquad i = 1, ..., k-1. \tag{17}$$

The first set of restrictions (Equation (13)) states that $f(x)$ must pass exactly at the latent yields when the maturities are the knots. The second set (Equations (15), (16) and (17)) guarantees continuity and smoothness at the knots.

Note that the system (13), (15), (16) and (17) has $4k-2$ equations and $4k$ variables $\beta = [a^1\ b^1\ c^1\ d^1\ \cdots\ a^k\ b^k\ c^k\ d^k]'$. So, it is necessary to impose two additional restrictions to have only one solution. One of the ways suggested by the literature is to impose that the second derivatives of $f(x)$ are set equal to zero at the endpoints of the interval $[0, T]$,

$$f^{1\prime\prime}(x_0) = 0 \tag{18}$$

$$f^{k\prime\prime}(x_k) = 0, \tag{19}$$

Let us define the following vector functions

$$X_1^i(x) = \begin{bmatrix} 1 & g_i(x) & h_i(x) & z_i(x) \end{bmatrix} \tag{20}$$

$$X_2^i(x) = \begin{bmatrix} 0 & g_i'(x) & h_i'(x) & z_i'(x) \end{bmatrix} \tag{21}$$

$$X_3^i(x) = \begin{bmatrix} 0 & g_i''(x) & h_i''(x) & z_i''(x) \end{bmatrix}. \tag{22}$$

Then the R matrix is given by:

$$R = \begin{bmatrix} X_1^1(x_1) & -X_1^2(x_1) & 0_{4\times1} & \dots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} \\ 0_{4\times1} & X_1^2(x_2) & -X_1^3(x_2) & \dots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \dots & 0_{4\times1} & X_1^{k-1}(x_{k-1}) & -X_1^k(x_{k-1}) \\ X_2^1(x_1) & -X_2^2(x_1) & 0_{4\times1} & \dots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \dots & 0_{4\times1} & X_2^{k-1}(x_{k-1}) & -X_2^k(x_{k-1}) \\ X_3^1(x_1) & -X_3^2(x_1) & 0_{4\times1} & \dots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \dots & 0_{4\times1} & X_3^{k-1}(x_{k-1}) & -X_3^k(x_{k-1}) \\ X_3^1(x_0) & 0_{4\times1} & 0_{4\times1} & \dots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} \\ 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \dots & 0_{4\times1} & 0_{4\times1} & X_3^k(x_k) \end{bmatrix}. \tag{23}$$

The latent yields are measured without error, that is,

$$Y(\phi) = W(\phi)\beta, \quad \text{s.t. } R\beta = 0, \tag{24}$$

where the $W(\phi)$ matrix has the following form,

$$\begin{bmatrix} X_1^1(x_0) & 0_{4\times1} & \cdots & 0_{4\times1} & 0_{4\times1} \\ X_1^1(x_1) & 0_{4\times1} & \cdots & 0_{4\times1} & 0_{4\times1} \\ 0_{4\times1} & X_1^2(x_2) & \cdots & 0_{4\times1} & 0_{4\times1} \\ \vdots & \vdots & \ddots & X_1^{k-1}(x_{k-1}) & 0_{4\times1} \\ 0_{4\times1} & 0_{4\times1} & \cdots & 0_{4\times1} & X_1^k(x_k) \end{bmatrix}.$$

Note that the form of matrix W allows one to have different factor loadings defined by $X_1^i(\cdot)$ within each segment.

In contrast, the observed yields are measured with noises:

$$Y(\tau) = W(\tau)\beta + \varepsilon \quad s.t. \quad R\beta = 0. \tag{25}$$

For each $(y_j, t_j) \in Y(\tau) \times \tau$, $\exists i$ such that $t_j \in A_i$. The j-th row of $W(\tau)$ is given by[24]

$$W(\tau)(j,:) = \begin{bmatrix} 0_{4\times1} & \cdots & 0_{4\times1} & X_1^i(\tau_j) & 0_{4\times1} & \cdots & 0_{4\times1} \end{bmatrix}.$$

[24] Assuming $m - 1 > i_{(t_j)} > 2$

For instance, let $Y(\tau) = \{y_1, y_2, y_3\}$ be a set of yields with respective maturities $\tau = \{t_1, t_2, t_3\}$ where

$$\begin{aligned} 0 &< t_1 < x_1 \\ x_{i-1} &< t_2 < x_i \\ x_{k-1} &< t_3 < x_k \end{aligned}$$

then

$$W(\tau) = \begin{bmatrix} X_1^1(t_1) & 0_{4\times1} & \cdots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \cdots & 0_{4\times1} \\ 0_{4\times1} & 0_{4\times1} & \cdots & 0_{4\times1} & X_1^i(t_2) & 0_{4\times1} & \cdots & 0_{4\times1} \\ 0_{4\times1} & 0_{4\times1} & \cdots & 0_{4\times1} & 0_{4\times1} & 0_{4\times1} & \cdots & X_1^k(t_3) \end{bmatrix}.$$

The dimension of the restriction matrix R is $3k-1\times4k$. By construction, all rows are linearly independents, hence the rank of R is equal to $3k-1$. Thus, it is possible to partition the R matrix into two submatrix: a squared full-rank matrix R_1 with dimension $3k-1$, and the complement matrix R_2 with dimension $3(k-1)\times k+1$. The system of restrictions can be written as:

$$R_1\hat{\theta} + R_2\theta = 0,$$

where $R = [R_1 \ R_2]$ and $\beta = [\hat{\theta} \ \theta]$. Using the fact that R_1 is invertible, we solve for $\hat{\theta}$:

$$\hat{\theta} = -R_1^{-1}R_2\theta. \tag{26}$$

Reorganizing the measurement equation we have:

$$Y(\tau) = W_1(\tau)\theta + W_2(\tau)\hat{\theta} + \varepsilon(\tau).$$

Then the measurement equation with the restrictions included becomes:

$$Y(\tau) = Z(\tau)\theta + \varepsilon(\tau),$$

where $Z(\tau)$ is the restricted loading matrix given by:

$$Z(\tau) = (W_1(\tau) - W_2(\tau)R_1^{-1}R_2). \tag{27}$$

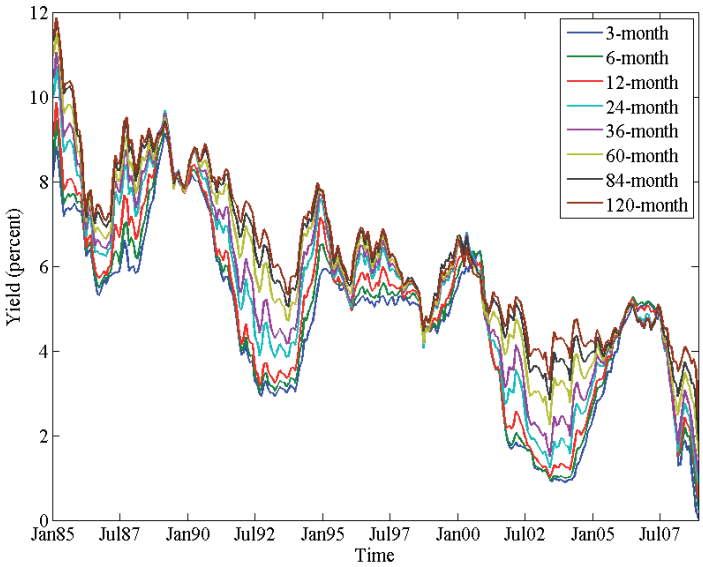

Figure 1: U.S. Treasury Yields.

This figure contains time series of monthly U.S. Treasury yields from January of 1985 to December of 2008, with maturities of 3, 6, 12, 24, 36, 60, 84 and 120 months.

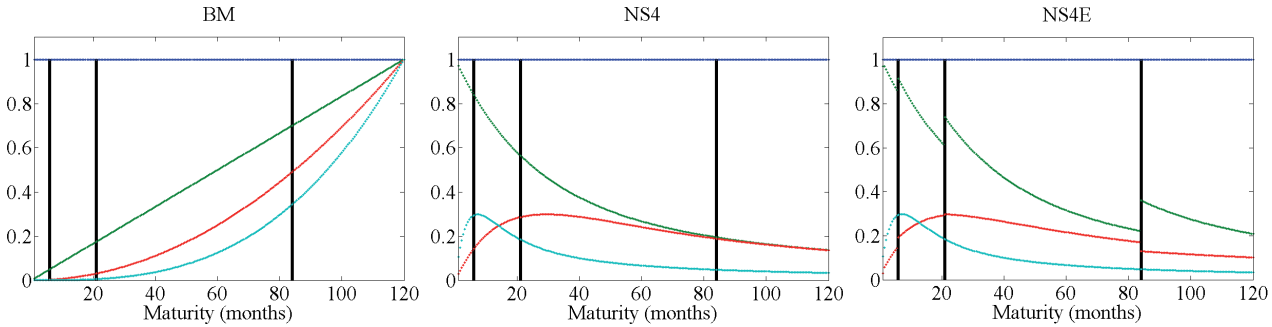

Figure 2: Unrestricted Loadings.

This picture presents the original loadings before smoothing transformation, for the Segmented Loading models (BM, NS4 and NS4E). The BM model is implemented via a natural cubic splines, the NS4 model with fixed exponential loadings, and the NS4E model uses exponential loadings with different functional forms across segments. BM, NS4 and NS4E models have the same set of knots $(1, 6, 21, 84, 120$ months). Within each pair of knots (separated by black vertical lines) there is a total of four basis functions representing loadings.

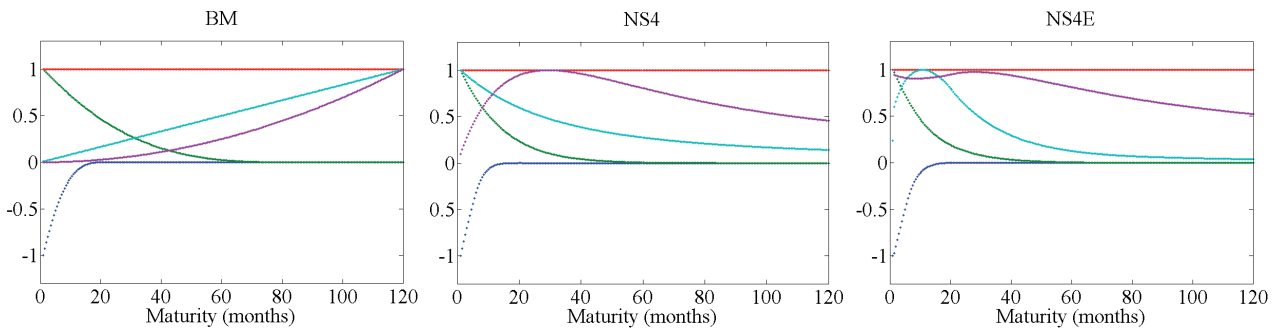

Figure 3: Restricted Loadings.

This picture presents loadings after smoothing transformation, for the Segmented Loading models (BM, NS4 and NS4E). The BM model is implemented via a natural cubic splines, the NS4 model with fixed exponential loadings, and the NS4E model uses exponential loadings with different functional forms across segments. BM, NS4 and NS4E models have the same set of knots $(1, 6, 21, 84, 120$ months$)$.

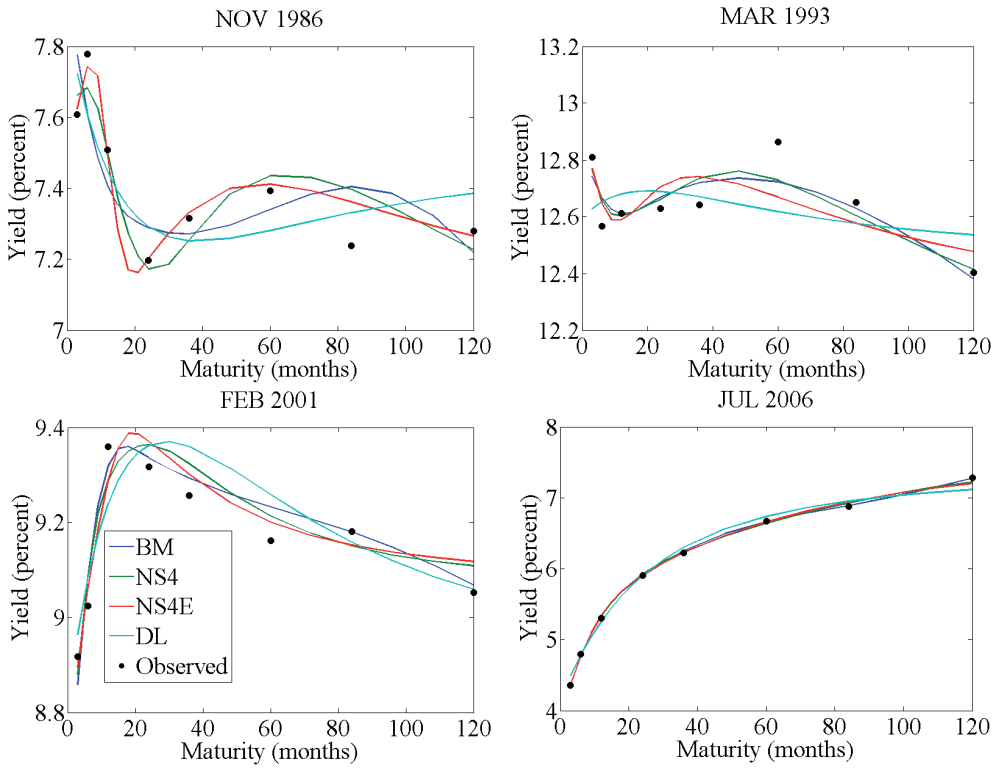

Figure 4: Cross Sectional Fits of the U.S. Term Structure. This picture presents cross-sectional fits for the Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E). The BM, NS4 and NS4E models have the same set of knots $(1, 6, 21, 84, 120$ months).

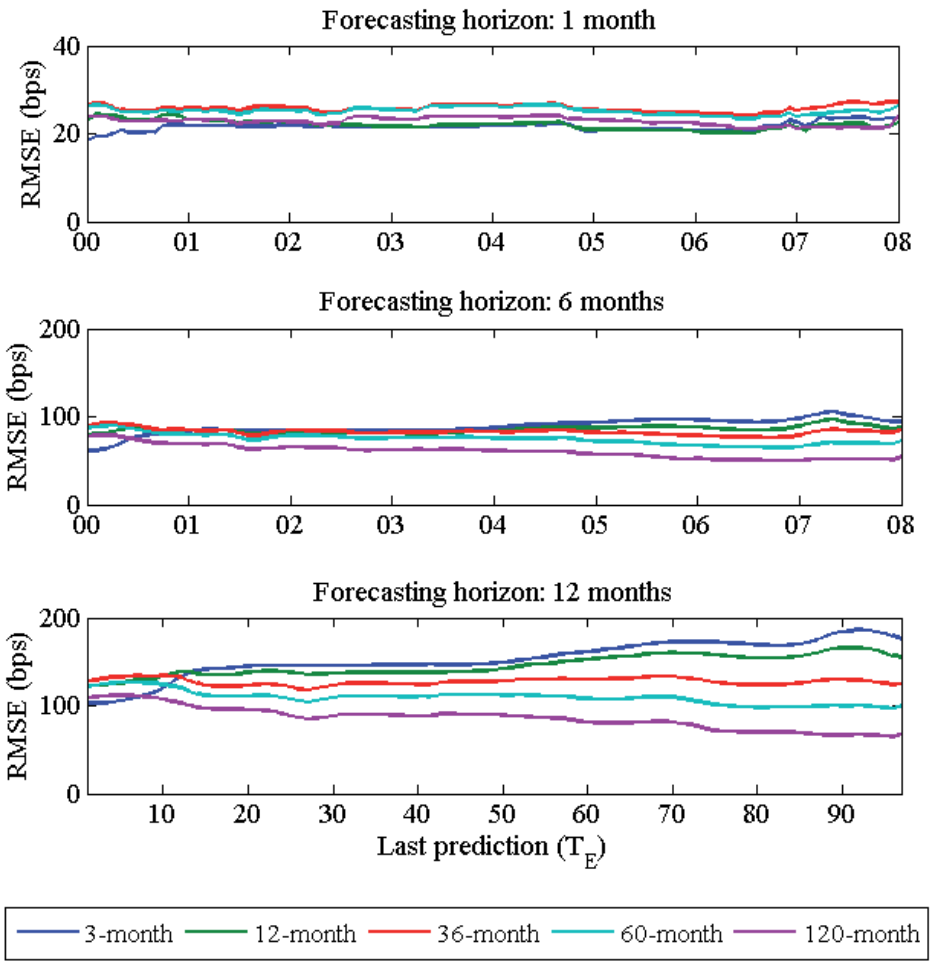

Figure 5: RMSE of random walk.

This picture presents the RMSE of random walk for all experiments. Each experiment is defined by the month at which the last prediction is done (T_E). T_E varies on a monthly basis between December 2000 and December 2008.

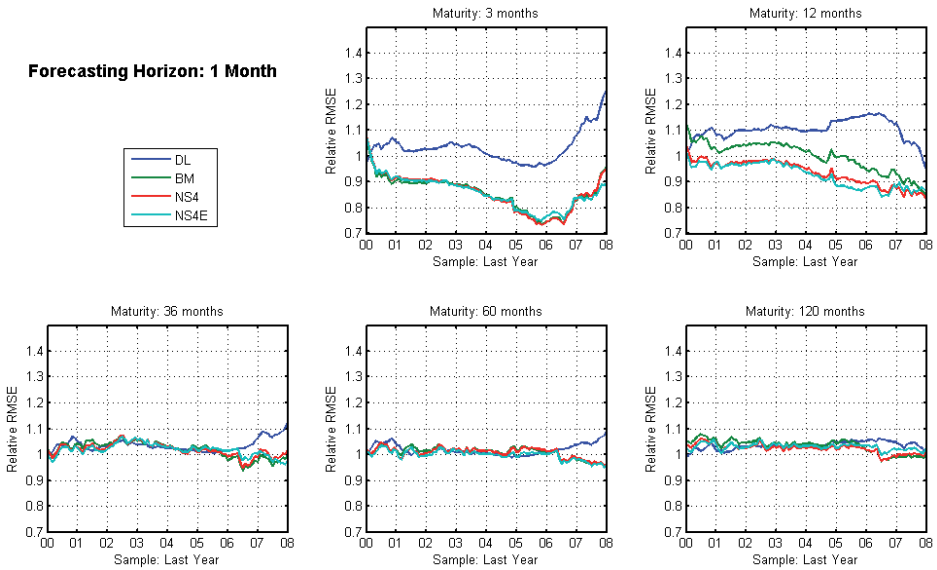

Figure 6: Time Series of RMSE for 1-month forecasting horizon.
This picture presents the mean of the RMSE relative to the RW for the 1-month forecasting horizon for the models studied in this paper. This mean is computed for each experiment which is defined by the date for which the last prediction is done (represented in the graphs above by the x-axis). An experiment consists of a sequence of forecasts constructed such that when a new observation is available, the in-sample window is shifted across time one month ahead, the model is reestimated and a new forecast is computed. For all experiments, we run a total of 84 monthly forecasts. The models studied in this work are the Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E).

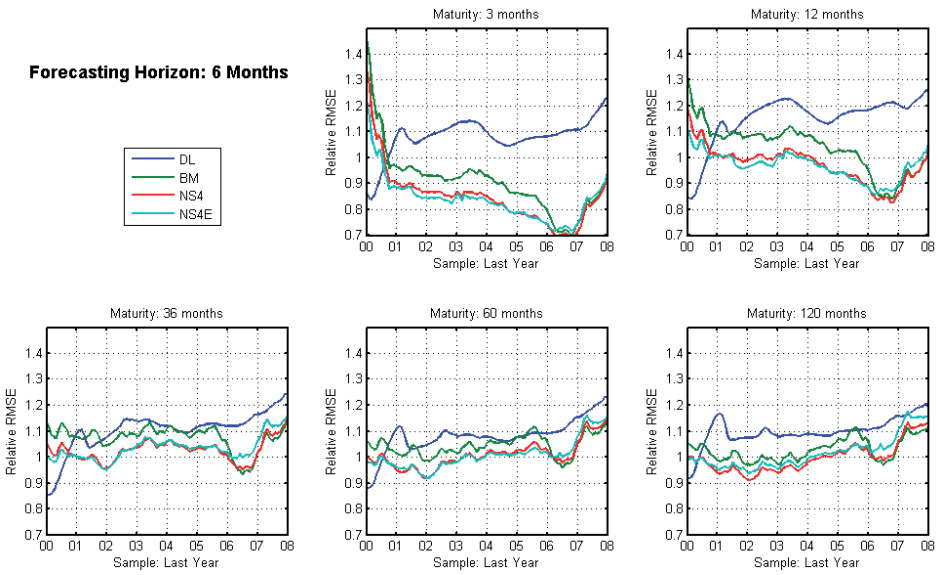

Figure 7: Time Series of RMSE for 6-month forecasting horizon.
This picture presents the mean of the RMSE relative to the RW for the 1-month forecasting horizon for the models studied in this paper. This mean is computed for each experiment which is defined by the date for which the last prediction is done (represented in the graphs above by the x-axis). An experiment consists of a sequence of forecasts constructed such that when a new observation is available, the in-sample window is shifted through time one month ahead, the model is reestimated and a new forecast is computed. For all experiments, we run a total of 84 monthly forecasts. The models studied in this work are the Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E).

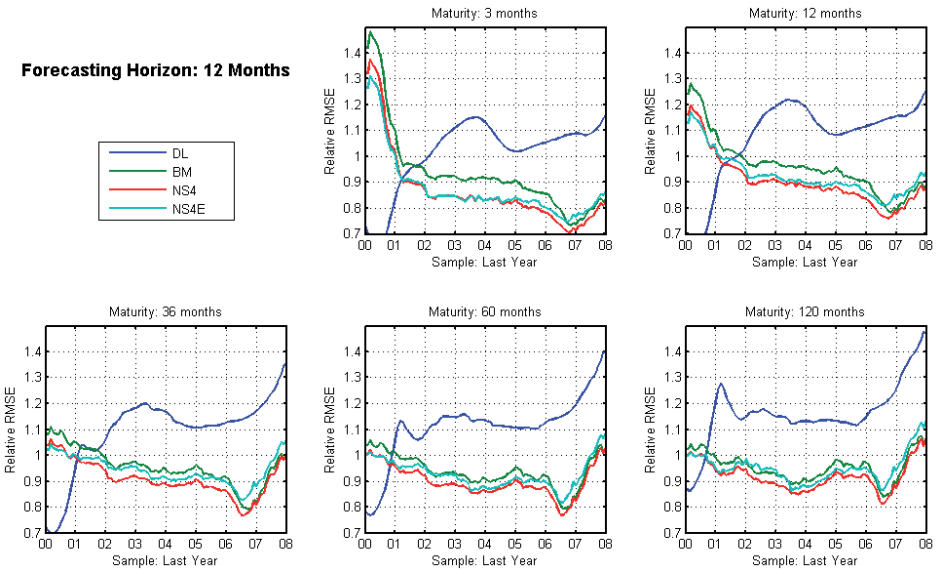

Figure 8: Time Series of RMSE for 12-month forecasting horizon.
This picture presents the mean of the RMSE relative to the RW for the 12-month forecasting horizon for the models studied in this paper. This mean is computed for each experiment which is defined by the date for which the last prediction is done (represented in the graphs above by the x-axis). An experiment consists of a sequence of forecasts constructed such that when a new observation is available, the in-sample window is shifted through time one month ahead, the model is reestimated and a new forecast is computed. For all experiments, we run a total of 84 monthly forecasts. The models studied in this work are the Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E).

Horizon (in months)	Model				
	RW	DL	BM	NS4	NS4E
		3-month yield			
1	21.58	1.04	**0.86**	**0.86**	**0.85***
6	88.74	1.09	**0.90**	**0.84**	**0.83***
12	152.90	1.03	**0.92**	**0.85***	**0.86**
		12-month yield			
1	21.99	1.10	**0.99**	**0.94**	**0.93***
6	85.37	1.15	1.04	0.97	**0.96***
12	145.57	1.08	**0.96**	**0.90***	**0.92**
		36-month yield			
1	25.77	1.03	1.02	1.02	1.01*
6	82.69	1.10	1.07	1.02*	1.03
12	127.40	1.09	**0.96**	**0.91***	**0.94**
		60-month yield			
1	25.14	1.02	1.01	1.01	**0.99***
6	74.42	1.08	1.04	1.01*	1.01
12	109.27	1.10	**0.94**	**0.90***	**0.93**
		120-month yield			
1	22.80	1.03	1.03	1.02*	1.03
6	60.41	1.09	1.03	**0.99***	1.01
12	85.77	1.15	**0.96**	**0.92***	**0.95**

Table 1: RMSE Mean for Out-of-Sample Forecast Errors.
This table presents the mean over T_E of the relative RMSE with respect to the RW for Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E). At the second column we present the absolute RMSE of the RW model. Bold face values represent a RMSE lower than RW while an * represents the lowest RMSE for each fixed horizon and maturity. The mean is obtained over 96 different out-of-sample experiments.

Horizon (in months)	Model			
	DL	BM	NS4	NS4E
		3-month yield		
1	19/50	24/10	43/9	47/8
6	8/0	17/4	20/5	21/5
12	0/0	16/0	18/4	19/4
		12-month yield		
1	18/73	22/7	45/5	47/5
6	15/0	18/0	21/0	20/0
12	4/0	17/0	19/0	18/0
		36-month yield		
1	18/73	20/0	53/0	21/0
6	14/1	11/0	13/0	0/0
12	11/0	10/0	10/0	0/0
		60-month yield		
1	14/73	16/0	82/0	73/0
6	4/11	7/0	6/0	0/0
12	0/1	5/0	5/0	0/0
		120-month yield		
1	8/77	41/0	70/0	51/5
6	0/36	0/0	1/0	0/0
12	0/24	0/0	0/0	0/0

Table 2: Diebold and Mariano Test.

This table presents the Diebold and Mariano (1995) test with a signficance of 5% using quadratic loss for the Diebold and Li (2006) model (DL), the cubic-spline model of Bowsher and Meeks (2008) (BM), and two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4E). The first number in each cell represents the percentage of times that the model produces statistically significant better forecasts than the RW considering all 96 out-of-sample experiments. The second number is the percentage of times that the RW produces statistically significant better forecasts than the model considering all 96 out-of-sample experiments.

Banco Central do Brasil

Trabalhos para Discussão

Os Trabalhos para Discussão do Banco Central do Brasil estão disponíveis para download no website http://www.bcb.gov.br/?TRABDISCLISTA

Working Paper Series

The Working Paper Series of the Central Bank of Brazil are available for download at http://www.bcb.gov.br/?WORKINGPAPERS

245	**Pesquisa Trimestral de Condições de Crédito no Brasil** *Clodoaldo Aparecido Annibal e Sérgio Mikio Koyama*	Jun/2011
246	**Impacto do Sistema Cooperativo de Crédito na Eficiência do Sistema Financeiro Nacional** *Michel Alexandre da Silva*	Ago/2011
247	**Forecasting the Yield Curve for the Euro Region** *Benjamim M. Tabak, Daniel O. Cajueiro and Alexandre B. Sollaci*	Aug/2011
248	**Financial regulation and transparency of information: first steps on new land** *Helder Ferreira de Mendonça, Délio José Cordeiro Galvão and Renato Falci Villela Loures*	Aug/2011
249	**Directed clustering coefficient as a measure of systemic risk in complex banking networks** *B. M. Tabak, M. Takami, J. M. C. Rocha and D. O. Cajueiro*	Aug/2011
250	**Recolhimentos Compulsórios e o Crédito Bancário Brasileiro** *Paulo Evandro Dawid e Tony Takeda*	Ago/2011
251	**Um Exame sobre como os Bancos Ajustam seu Índice de Basileia no Brasil** *Leonardo S. Alencar*	Ago/2011
252	**Comparação da Eficiência de Custo para BRICs e América Latina** *Lycia M. G. Araujo, Guilherme M. R. Gomes, Solange M. Guerra e Benjamin M. Tabak*	Set/2011
253	**Bank Efficiency and Default in Brazil: causality tests** *Benjamin M. Tabak, Giovana L. Craveiro and Daniel O. Cajueiro*	Oct/2011
254	**Macroprudential Regulation and the Monetary Transmission Mechanism** *Pierre-Richard Agénor and Luiz A. Pereira da Silva*	Nov/2011
255	**An Empirical Analysis of the External Finance Premium of Public Non-Financial Corporations in Brazil** *Fernando N. de Oliveira and Alberto Ronchi Neto*	Nov/2011
256	**The Self-insurance Role of International Reserves and the 2008-2010 Crisis** *Antonio Francisco A. Silva Jr*	Nov/2011

257	**Cooperativas de Crédito: taxas de juros praticadas e fatores de viabilidade** *Clodoaldo Aparecido Annibal e Sérgio Mikio Koyama*	Nov/2011
258	**Bancos Oficiais e Crédito Direcionado – O que diferencia o mercado de crédito brasileiro?** *Eduardo Luis Lundberg*	Nov/2011
259	**The impact of monetary policy on the exchange rate: puzzling evidence from three emerging economies** *Emanuel Kohlscheen*	Nov/2011
260	**Credit Default and Business Cycles: an empirical investigation of Brazilian retail loans** *Arnildo da Silva Correa, Jaqueline Terra Moura Marins, Myrian Beatriz Eiras das Neves and Antonio Carlos Magalhães da Silva*	Nov/2011
261	**The relationship between banking market competition and risk-taking: do size and capitalization matter?** *Benjamin M. Tabak, Dimas M. Fazio and Daniel O. Cajueiro*	Nov/2011
262	**The Accuracy of Perturbation Methods to Solve Small Open Economy Models** *Angelo M. Fasolo*	Nov/2011
263	**The Adverse Selection Cost Component of the Spread of Brazilian Stocks** *Gustavo Silva Araújo, Claudio Henrique da Silveira Barbedo and José Valentim Machado Vicente*	Dec/2011
264	**Uma Breve Análise de Medidas Alternativas à Mediana na Pesquisa de Expectativas de Inflação do Banco Central do Brasil** *Fabia A. de Carvalho*	Jan/2012
265	**O Impacto da Comunicação do Banco Central do Brasil sobre o Mercado Financeiro** *Marcio Janot e Daniel El-Jaick de Souza Mota*	Jan/2012
266	**Are Core Inflation Directional Forecasts Informative?** *Tito Nícias Teixeira da Silva Filho*	Jan/2012
267	**Sudden Floods, Macroprudention Regulation and Stability in an Open Economy** *P.-R. Agénor, K. Alper and L. Pereira da Silva*	Feb/2012
268	**Optimal Capital Flow Taxes in Latin America** *João Barata Ribeiro Blanco Barroso*	Mar/2012
269	**Estimating Relative Risk Aversion, Risk-Neutral and Real-World Densities using Brazilian Real Currency Options** *José Renato Haas Ornelas, José Santiago Fajardo Barbachan and Aquiles Rocha de Farias*	Mar/2012
270	**Pricing-to-market by Brazilian Exporters: a panel cointegration approach** *João Barata Ribeiro Blanco Barroso*	Mar/2012
271	**Optimal Policy When the Inflation Target is not Optimal** *Sergio A. Lago Alves*	Mar/2012

272	**Determinantes da Estrutura de Capital das Empresas Brasileiras: uma abordagem em regressão quantílica** *Guilherme Resende Oliveira, Benjamin Miranda Tabak, José Guilherme de Lara Resende e Daniel Oliveira Cajueiro*	Mar/2012
273	**Order Flow and the Real: Indirect Evidence of the Effectiveness of Sterilized Interventions** *Emanuel Kohlscheen*	Apr/2012
274	**Monetary Policy, Asset Prices and Adaptive Learning** *Vicente da Gama Machado*	Apr/2012
275	**A geographically weighted approach in measuring efficiency in panel data: the case of US saving banks** *Benjamin M. Tabak, Rogério B. Miranda and Dimas M. Fazio*	Apr/2012
276	**A Sticky-Dispersed Information Phillips Curve: a model with partial and delayed information** *Marta Areosa, Waldyr Areosa and Vinicius Carrasco*	Apr/2012
277	**Trend Inflation and the Unemployment Volatility Puzzle** *Sergio A. Lago Alves*	May/2012
278	**Liquidez do Sistema e Administração das Operações de Mercado Aberto** *Antonio Francisco de A. da Silva Jr.*	Maio/2012
279	**Going Deeper Into the Link Between the Labour Market and Inflation** *Tito Nícias Teixeira da Silva Filho*	May/2012
280	**Educação Financeira para um Brasil Sustentável** **Evidências da necessidade de atuação do Banco Central do Brasil em educação financeira para o cumprimento de sua missão** *Fabio de Almeida Lopes Araújo e Marcos Aguerri Pimenta de Souza*	Jun/2012
281	**A Note on Particle Filters Applied to DSGE Models** *Angelo Marsiglia Fasolo*	Jun/2012
282	**The Signaling Effect of Exchange Rates: pass-through under dispersed information** *Waldyr Areosa and Marta Areosa*	Jun/2012
283	**The Impact of Market Power at Bank Level in Risk-taking: the Brazilian case** *Benjamin Miranda Tabak, Guilherme Maia Rodrigues Gomes and Maurício da Silva Medeiros Júnior*	Jun/2012
284	**On the Welfare Costs of Business-Cycle Fluctuations and Economic-Growth Variation in the 20th Century** *Osmani Teixeira de Carvalho Guillén, João Victor Issler and Afonso Arinos de Mello Franco-Neto*	Jul/2012
285	**Asset Prices and Monetary Policy – A Sticky-Dispersed Information Model** *Marta Areosa and Waldyr Areosa*	Jul/2012
286	**Information (in) Chains: information transmission through production chains** *Waldyr Areosa and Marta Areosa*	Jul/2012

287 Some Financial Stability Indicators for Brazil Jul/2012
Adriana Soares Sales, Waldyr D. Areosa and Marta B. M. Areosa