Journal of Financial Econometrics, 2018, Vol. 16, No. 1, 1–33
doi: 10.1093/jjfinec/nbx002
Advance Access Publication Date: 14 February 2017
Article

Forecasting Bond Yields with Segmented Term Structure Models*

Caio Almeida[1], Kym Ardison[1], Daniela Kubudi[1], Axel Simonsen[2] and José Vicente[3,4]

[1]FGV-EPGE, [2]Vinci Partners School, [3]Ibmec Business School and [4]IBMEC Business School and Banco Central do Brasil

Address correspondence to Caio Almeida, FGV-EPGE, Praia de Botafogo, 190, Rio de Janeiro, Brazil, or e-mail: caio.almeida@fgv.br.

Received May 29, 2015; revised January 3, 2017; editorial decision January 4, 2017; accepted January 6, 2017

Abstract

Inspired by the preferred habitat theory, we propose parametric interest rate models that split the term structure into segments. The proposed models are compared with successful term structure benchmarks based on out-of-sample forecasting exercises using U.S. Treasury data. We show that segmentation can improve long-horizon term structure forecasts when compared with nonsegmentation. Additionally, introducing cointegration in latent factor dynamics of segmented models makes them particularly strong to forecast short-maturity yields. Better forecasting is justified by the segmented models' ability to accommodate idiosyncratic shocks in the cross-section of yields.

Key words: Error Correction Model, exponential splines, local shocks, model selection, preferred habitat theory

JEL classification: C53, C58, E43, G17

The preferred habitat theory of the term structure (Modigliani and Sutch, 1966) advocates that local shocks may influence interest rates for each maturity. Empirical evidence related to

* We would like to thank two anonymous referees, the editor (Federico Bandi), Francis Diebold, Raffaella Giacomini, Allan Timmermann, Michel van der Wel (discussant) and participants at the 2010 Workshop on Yield Curve Modeling and Forecasting at Erasmus University, the 2011 Brazilian Econometric Society Meeting, the 2012 SoFiE Annual Conference at Oxford, and the 2012 Brazilian Meeting of Finance for useful comments and suggestions. Caio Almeida and José Vicente gratefully acknowledge financial support from CNPq. Kym Ardison thanks FAPERJ and ANBIMA for financial support. The views expressed in this paper are those of the authors and do not necessarily reflect those of Banco Central do Brasil.

this theory reveals that U.S. Treasury bonds' supply and demand shocks have nonnegligible effects on yield spreads, term structure movements, and bond risk premium.[1]

In an attempt to formalize the preferred habitat theory, Vayanos and Vila (2009) propose an equilibrium model in which demand directly influences and determines all yields in the term structure, in a dynamic way. Their structural design offers many theoretical conjectures that inspire testable hypotheses to ascertain how the maturity structure of debt and supply factors affect monetary policy (Hamilton and Wu, 2012), the yield curve (Li and Wei, 2013) and bond returns (Greenwood and Vayanos, 2014).

Inspired by the preferred habitat theory, and more specifically by its recent formalization by Vayanos and Vila (2009), we propose a class of models that separate the yield curve into segments which present their own local shocks, but which are simultaneously interconnected, composing the whole yield curve. We apply the economic idea of yield curve segmentation to a practical reduced-form statistical model that highlights the existence of local factors. These factors are responsible for capturing idiosyncratic movements within the term structure of interest rates. The main rationale behind our reduced-form model is that if yields carry information about demand shocks, then allowing for the existence of local movements within the term structure should help in extracting specific information about those shocks. Such refined information about demand contributes to a better econometric identification of the yield curve dynamics.

The segmented proposed models split the original set of maturities into a group of segments, whose dynamics are driven by local factors. Each shock to a local factor affects maturities specific to its own segment. The model is completed by imposing spline smoothing conditions to the yield curve, which interconnect all local factors as part of a unique global system.

In the empirical section, we compare out-of-sample forecasts of segmented models against successful competitors in the term structure literature. We aim to assess the empirical forecasting performance of segmented term structure models, and to verify the role of factors' dynamics [Autoregressive (AR) versus Error Correction Models (ECM)] on models' forecasting performance.[2]

Our segmented models generate compelling empirical results. First, these models have superior performance at the short end of the yield curve under all different forecasting horizons analyzed, and under the longest forecasting horizon for all maturities analyzed. Second, adopting ECM instead of AR dynamics in latent factors' dynamics improves the forecasting performance of all models analyzed. Although the primary focus of our empirical analysis is on how yield curve segmentation affects forecasting, documenting the superior forecasting ability of ECM over AR dynamics in interest rate factor models is itself an interesting result, for the fixed-income literature. More details on ECM's superiority appear in the empirical section.

Let us associate, for a moment, higher forecasting ability of future yields (or bond returns) to a better identification of yield curve dynamics. In this context, we may conjecture that the superior performance of segmented models could be related to intermediation

1 See Krishnamurthy (2002), Greenwood and Vayanos (2010, 2014), and Krishnamurthy and Vising-Jorgensen (2012).

2 The majority of previously proposed term structure models are based on Vector AR (VAR) models of order one (Duffee, 2002; Diebold and Rudebusch, 2013). Important exceptions include those of Hall, Anderson, and Granger (1992) and Bowsher and Meeks (2008) who adopt ECM factor dynamics.

frictions in the U.S. Treasury bond market. Such frictions may manifest themselves through different channels, two of them emphasized here: funding liquidity, as suggested by Fontaine and Garcia (2012), and slow-moving capital, as analyzed by Duffie (2010).

Fontaine and Garcia (2012) identify a liquidity factor, negatively correlated to bond risk premia, extracted from bonds with similar cash flows but different ages. This factor suggests that U.S. Treasury bonds are useful to hedge against funding liquidity shocks. Since we have a reduced-form model that is estimated based only on U.S. Treasury yields, it is hard to pin down structural results. Notwithstanding, we try to relate the improvement obtained by the segmented models in forecasting short-term maturities to a novel idea: a model that is more flexible in the cross-section of yields and better captures idiosyncratic shocks might be, in reality, capturing the dynamics of liquidity issues and frictions, which appear in the repo markets. Some weak empirical evidence supporting this idea is revealed by one of our five-factor segmented models. We calculate the correlation of these factors with Fontaine and Garcia's liquidity factor and observe that the two local factors whose loadings are associated with short-maturity movements present much higher correlation with the liquidity factor (twice as much) compared to the three additional factors related to traditional term structure movements (level, slope, and curvature).

Duffie (2010) shows that patterns of price responses to supply and demand shocks in the U.S. Treasury market involve a sharp initial reaction, followed by a subsequent slower reversal. The speed of this reversal depends on the amount of intermediary capital needed to be raised. These patterns may be related to the improvement obtained for longer horizon forecasts (1 year) achieved by our segmented models. Our maturity segmentation combined with ECM factor dynamics could be picking up slower moving responses to supply and demand shocks happening in the U.S. Treasury market. Adrian, Moench, and Shin (2010) show that intermediary balance sheet aggregates contain strong predictive power for excess returns on Treasury instruments. Based on their findings, we suggest that the existence of risk premia reversals can impact our local factors' dynamics and end up affecting longer horizon forecasting of the proposed segmented term structure models.

It is important to keep in mind that our intention in this work is neither to structurally estimate the Vayanos and Vila (2009) model nor to implement its testable empirical implications, as done by Kaminska, Vayanos, and Zinna (2011) and Zinna (2016). Indeed, in order to either estimate the model or directly test its empirical implications, additional data on supply, mortgage, real rates, and/or liquidity should be added to the yield curve. Instead, we propose a statistical model using the central concept of *local shocks* and how they can transmit information about demand factors. We use yields data to estimate our model. Throughout the article, descriptions of the behavior of the proposed segmented models in capturing the cross-section of yields and their dynamics are included. Our model's flexibility in capturing local movements is the key to better identification of the yield curve, when considering both its cross-sectional and dynamic components.

1 Economic Motivation

To develop our term structure model, we build on the central insights of the preferred habitat theory formalized by Vayanos and Vila (2009). One of their fundamental hypotheses is that each investor may demand bonds of specific maturities. Therefore, the equilibrium interest rate for a given maturity results from the interaction of supply and demand forces

related to that particular maturity (Gürkaynak and Wright, 2012). Greenwood and Vayanos (2010) advocate in favor of this argument, based on two recent empirical episodes: the U.K. 2004 pension reform and the U.S. Treasury buyback programs of 2000 and 2001. Despite differing in policy, in both cases a long-lasting shift in the long-term interest rates was documented. Greenwood and Vayanos (2010) show that interest rate movements incorporate information extrapolated from those contained in the short-term interest rate, inflation, and macroeconomic variables. Empirical evidence favors the clientele-based demand/supply approach as an important driver of term structure movements.

In light of the recent U.S. Federal Reserve repurchase programs and of the evidence provided by Greenwood and Vayanos (2010), several authors have analyzed the impacts of Large Scale Asset Purchase (LSAP) programs on the U.S. term structure. Among them, D'Amico and King (2013) and Li and Wei (2013) build on the theory of Vayanos and Vila (2009) to provide empirical evidence in favor of market segmentation. In contrast, our goal is to construct a new reduced-form statistical term structure model, to use the idea of segmentation to extract previously ignored information about demand and supply shocks from idiosyncratic yields movements. The remainder of this section explores the nuances of Vayanos and Vila's model and relates its stylized facts to our proposed term structure model, introduced in Section 2.1.

The essence of their proposition is the interaction between investors and arbitrageurs. Investors represent all possible types of clientele and have preferences for specific bond maturities. For example, pension funds might demand long-term bonds, whereas speculators might prefer short-term bonds (Greenwood and Vayanos, 2010). In contrast, arbitrageurs invest in bonds with any maturity and aim to maximize a mean–variance utility function. Therefore, arbitrageurs play a significant role in guaranteeing some smoothness among yields with different maturities.

Introducing some notation, the demand at time t for a bond with maturity τ is assumed to be a linear function of the bond yield $y_{t,\tau}$, given by

$$D_{t,\tau} = \alpha(\tau)\tau(y_{t,\tau} - \beta_{t,\tau}) \tag{1}$$

where $\alpha(\tau)$ is a positive constant and $\beta_{t,\tau}$, the inelastic component of the investors' demand, is a function that may include multiple risk factors:

$$\beta_{t,\tau} = \overline{\beta} + \sum_{k=1}^{K} \theta_k(\tau)\beta_{k,t}. \tag{2}$$

In Equation (2), each $\{\beta_{k,t}\}_{k=1,\ldots,K}$ represents a distinct factor that captures demand risk, whereas $\{\theta_k(\tau)\}_{k=1,\ldots,K}$ characterizes how each factor impacts the demand for bonds with maturity τ. The model is completed with a stochastic differential equation for the short-term rate r_t and for the demand risk factors $\beta_{k,t}$, all following mean-reverting Gaussian processes:

$$dr_t = \kappa_r(\overline{r} - r_t)dt + \sigma_r dB_{r,t}\,, \text{ and} \tag{3}$$

$$d\beta_{k,t} = -\kappa_{\beta_k}\beta_{k,t}dt + \sigma_{\beta_k}dB_{\beta_k,t}. \tag{4}$$

Given a linear demand function for yields, arbitrageurs that maximize a mean–variance utility, market clearing conditions, and Gaussian affine dynamics for the short-rate and

demand risk factors, the yield of a bond with maturity τ is an affine function of the short-rate and demand risk factors[3]:

$$y_{t,\tau} = a_r(\tau) r_t + \sum_{k=1}^{K} a_{\beta_k}(\tau) \beta_{k,t}. \tag{5}$$

The original loadings of risk factors $\{\theta_k(\tau)\}_{k=1,\ldots,K}$ in the demand function play an important role in determining the new loadings $a_{\beta_k}(\tau)_{k=1,\ldots,K}$ that relate yield $y_{t,\tau}$ to demand factors $\beta_{k,t}$. Determining the loadings $a_{\beta_k}(\tau)$ as a function of the original demand loadings $\{\theta_k(\tau)\}$ is strongly affected by the existence of arbitrageurs. The no-arbitrage condition is what connects (possibly local) demand shocks to the cross-section of yields. For example, in the limiting case with no arbitrageurs, yield $y_{t,\tau}$ would be given directly by the demand function for τ-maturity bonds: $y_{t,\tau} = \overline{\beta} + \sum_{k=1}^{K} \theta_k(\tau)\beta_{k,t}$. This could represent an extreme case of segmentation, if functions $\{\theta_k(\tau)\}_{k=1,\ldots,K}$ were chosen to be single-peaked around specific maturities. In such a case, $\{\beta_{k,t}\}_{k=1,\ldots,K}$ would represent local demand shocks.

Empirically, results by D'Amico and King (2013) indicate that Treasuries with similar maturities are close substitutes for each other, whereas this substitutability decreases the higher the maturity difference. Therefore, responses to local shocks restricted to a group of securities—in their particular case, the LSAP of Treasuries—are concentrated in segments of the yield curve consisting of close maturity yields.

Linking the arbitrageurs in the Vayanos and Vila (2009) model to the empirical findings of D'Amico and King (2013), we note that arbitrageurs should play two crucial roles. First, they are responsible for generating the links between demand factors and bond yields. Second, they smooth the yield curve, guaranteeing that bond yields with close maturities have close values. Without arbitrageurs, in theory, yields could be completely different for different maturities, and arbitrage opportunities would become available.

In this article, we offer a new perspective on market segmentation. We propose a reduced-form segmented model which accommodates clientele effects by allowing the term structure to have local shocks. The arbitrageurs' roles are paralleled with a spline second-order smoothness condition imposed on the model. In our model, the smoothness condition distributes information from local shocks to other parts of the term structure, exactly how Vayanos and Vila's (2009) arbitrageurs transform the original demand loadings $\{\theta_k(\tau)\}$ into yields loadings $a_{\beta_k}(\tau)$, propagating local demand information to the whole curve.

Our reduced-form model builds on three insights from the preferred habitat environment. (i) Preferred habitat investors allocate their resources in bonds of specific maturities. (ii) Each maturity is potentially subjected to difference factors driving yields' dynamics. (iii) The existence of arbitrageurs in the market guarantees the absence of arbitrages, linking demand factors from all different maturities. On the other hand, trading arbitrages is risky, that is, exposed to duration risk, and thus local shocks to specific maturities might not be entirely eliminated.

It is important to highlight that our model does not impose no-arbitrage. Instead, we simply link different segments of the yield curve through a set of splines restrictions. From an empirical perspective, Joslin, Singleton, and Zhou (2011) show that neither cross-

3 Moreover, market prices of risk in this economy are also affine functions of the short-rate and demand risk factors.

section nor dynamic no-arbitrage conditions improve forecasting ability in Gaussian affine models.[4]

From an empirical perspective, the flexibility of our dynamic model allows us to evaluate the benefits of market segmentation in terms of an important specific feature from the fixed income literature: predictability of yields and returns.

2 Segmented Loading Model

2.1 The General Model

Our model builds on the piecewise polynomial methodology proposed by Bowsher and Meeks (2008) and generalized by Jungbaker, Koopman, and van der Wel (2014). Bowsher and Meeks (2008) assume that the yield curve is a perturbed version of a polynomial cubic spline function. In their methodology, latent variables are yields with maturities coinciding with the knots of the spline, while observed yields are assumed to be contaminated with error. They show that visualizing the term structure as a perturbed spline is useful to provide accurate short-term forecasts of future yields, especially when a broad cross-section of correlated yields is available.

We expand Bowsher and Meeks' perturbed spline idea by considering exponential-type functions as an alternative to the cubic polynomials that appear within knots in their original splines formulation. Our exponential splines naturally generalize the model of Diebold and Li (2006) and others similar to it, obtaining segmented versions of four-factor exponential models. In addition, we look at these spline functions from two different perspectives: first, similarly to Bowsher and Meeks, as perturbed splines with latent yields at knots; second, as models with subsets of local factors that drive the dynamics of specific segments of the yield curve. In this sense, we connect the functional time series idea of Bowsher and Meeks to the preferred habitat theory formalized by Vayanos and Vila (2009).

Introducing a general notation, let $\tilde{\tau}$ denote the $m \times 1$ vector containing the maturities of the m observed yields, that is, $\tilde{\tau} = (\tilde{\tau_1}, \ldots, \tilde{\tau_m})$. We denote by $y_t(\tau)$ the time t yield with maturity τ. Very generally, we model the collection of all the observed yields $Y_t(\tilde{\tau})$ as a factor model:

$$Y_t(\tilde{\tau}) = W(\tilde{\tau})\beta_t + \varepsilon_t(\tilde{\tau}) \tag{6}$$

where $W(\tilde{\tau})$ is a $m \times K$ matrix of factor loadings, β_t is a vector of K unobserved factors, ε is a residual term, and K is the total number of factors.

This general framework is common to several term structure models, including those of Diebold and Li (2006), Bowsher and Meeks (2008), and Jungbaker, Koopman, and van der Wel (2014), among others. The main difference between these papers and ours resides in the assumptions about the structure of the matrix $W(\tilde{\tau})$ (and also about the factors' dynamics, which will be further discussed in Section 2.3). While Diebold and Li (2006) assume that factor loadings are exponential functions, Bowsher and Meeks (2008) work with a polynomial basis. In contrast, Jungbaker, Koopman, and van der Wel (2014) introduce a more

4 The fact that our model is Gaussian but differs from classical affine factor models due to spline segmentation raises the possibility of analyzing whether no-arbitrage improves forecasting ability, but this is beyond the scope of this article.

general framework in which they estimate underlying factors and factor loadings via a Kalman Filter approach, without imposing a fixed structure for $W(\tilde{\tau})$.

For the general model in this section, we propose to incorporate in Equation (6) the stylized facts of the preferred habitat theory. In particular, we want our model (i) to allow for *local shocks in the term structure* and (ii) to present a *smooth transition between segments*. To accomplish this, we proceed in two stages: first, we impose restrictions to the matrix $W(\tilde{\tau})$ in a way such that each segment of the yield curve ends up being governed by the same set of "movements"; second, we impose smoothness restrictions across segments of the curve, guaranteeing smoothness in the final yield curve as a whole. The remainder of this section links these two restrictions to the general model in Equation (6).

Considering first the issue of local shocks, we split the yield curve into segments of close maturities. This is similar to the empirical approach of Zinna (2016) who considers a simplified version of Vayanos and Vila's (2009) model focusing on long-term investors. Additionally, this formulation is motivated by the empirical findings of D'Amico and King (2013), identifying the existence of substitutability between Treasuries. In what follows, each segment of the yield curve will be subjected to local shocks, as in the preferred habitat model. Letting $[T_m, T_M]$ be the interval for maturities of the yields, consider the following fixed partition over time:

$$\phi = \{T_m = \tau_0 < \tau_1 < \ldots < \tau_k = T_M\}. \tag{7}$$

The τ_i's are denominated knots of the model, and $\boldsymbol{k}$ is the exogenously chosen **number of segments** in the yield curve. The yield of a bond with maturity $\tau \in [T_m, T_M]$ at time t will be expressed by

$$y_t(\tau) = \sum_{i=1}^{k} f_t^i(\tau) I(A_i), \tag{8}$$

where $I(A_i)$ is the indicator function of set A_i,[5] and A_i is defined by

$$A_i = \{\, \tau \in \Re : \tau_{i-1} \leq \tau \leq \tau_i,\ i = 1, \ldots, k \,\}. \tag{9}$$

Related to the general model in Equation (6), Equation (8) imposes an important restriction to $W(\tilde{\tau})\beta_t$: for two yields with maturities $\tilde{\tau_1}$ and $\tilde{\tau_2}$ such that $\tilde{\tau_1}, \tilde{\tau_2} \in A_i$, for a certain arbitrary i, it must be the case that $W(\tilde{\tau_1})\beta_t = f_t^i(\tilde{\tau_1})$ and $W(\tilde{\tau_2})\beta_t = f_t^i(\tilde{\tau_2})$. This assures that two yields within the same segment are governed by the same function $f_t^i(\cdot)$. Note that, for now, the function $f_t^i(\cdot)$ can vary across segments, indexed by i, but within each segment, $f_t^i(\cdot)$ is completely determined by the maturity parameter τ.

The shapes of local loadings $f_t^i(\tau)$ at each segment of the yield curve should be seen as important ingredients of the model. In principle, by assuming a dynamic for β_t, one could follow Jungbaker, Koopman, and van der Wel (2014) and estimate Equation (6) using the Kalman Filter method. In contrast, we follow the parametric term structure literature and impose an additional parametric structure on $f_t^i(\tau)$. Generalizing the idea of a polynomial cubic spline of Bowsher and Meeks (2008), we allow for four different types of local movements within each segment. To that end, we let functions 1, g_i, h_i, and z_i represent the time-invariant loadings at segment i. Paralleling the preferred habitat model of Section 1, these

5 $I(A_i) = 1$, if $\tau \in A_i$, and 0, otherwise.

functions resemble the loadings $\theta_k(\tau)$'s of the demand risk factors in a specific case: the case in which $\theta_k(\tau)$'s are single-peaked around specific maturities, each representing a segment of the yield curve, and the arbitrageurs exhibit infinite risk aversion.[6]

With the formulation above, at each segment, f^i is given by

$$f_t^i(\tau) = a_t^i + b_t^i g_i(\tau) + c_t^i h_i(\tau) + d_t^i z_i(\tau),\ i = 1, ..., k, \tag{10}$$

where $[a_t^i\ b_t^i\ c_t^i\ d_t^i]'$ represent the time-varying local yield curve factors of segment i. Stacking the factors for each segment together, following the notation of Equation (6), we have that $\beta_t = [a_t^1, b_t^1, \ldots, c_t^k, d_t^k]$.

Up to this point, the proposed model can be viewed as a version of the preferred habitat model of Vayanos and Vila (2009) with arbitrageurs exhibiting infinite risk aversion. That is, each segment of the curve is affected only by local shocks and there is no propagation of shocks across segments. However, we also want to ensure smoothness of the yield curve. In the habitat model, this is achieved by introducing active arbitrageurs, that is, arbitrageurs with finite risk aversion that ensure no-arbitrage conditions in the market. Here, we do not impose no-arbitrage restrictions. Instead, we require functions f^i to be of class C^2 in the closure of A_i, that is, to satisfy the splines constraints. Functions f_t^i and f_t^{i+1} and their derivatives of order one and two must have the same value, at each internal knot $\tau_i, i = 1, 2, ..., k-1$. In order to exemplify how these splines restrictions take place in our environment, consider the first interior knot τ_1 in the partition ϕ. It must be the case that

$$[1,\ g_1(\tau_1),\ h_1(\tau_1),\ z_1(\tau_1)]\beta_t^1 - [1,\ g_2(\tau_1),\ h_2(\tau_1),\ z_2(\tau_1)]\beta_t^2 = 0, \tag{11}$$

$$[0,\ g_1'(\tau_1),\ h_1'(\tau_1),\ z_1'(\tau_1)]\beta_t^1 - [0,\ g_2'(\tau_1),\ h_2'(\tau_1),\ z_2'(\tau_1)]\beta_t^2 = 0, \tag{12}$$

$$[0,\ g_1''(\tau_1),\ h_1''(\tau_1),\ z_1''(\tau_1)]\beta_t^1 - [0,\ g_2''(\tau_1),\ h_2''(\tau_1),\ z_2''(\tau_1)]\beta_t^2 = 0, \text{ and} \tag{13}$$

$$[1,\ g_1(\tau_1),\ h_1(\tau_1),\ z_1(\tau_1)]\beta_t^1 = y_t(\tau_1). \tag{14}$$

Restriction (11) implies that $f^1(\cdot)$ should be equal to $f^2(\cdot)$ at the knot τ_1. Restrictions (12) and (13) guarantee that the first and second derivatives of $f^1(\cdot)$ and $f^2(\cdot)$ are equal at knot τ_1. And finally, restriction (14) imposes the condition that the latent yield at knot τ_1, $y_t(\tau_1)$, should be perfectly captured by the spline function $f^1(\tau_1) = y(\tau_1)$, meaning that noise only exists for yields in the interior of yield curve segments. This is an important condition, since it guarantees that the factors in the model are the latent yields at the knots. This noise restriction is the only one that is also satisfied by the extreme knots τ_0 and τ_k.

More generally, the splines restrictions given above for knot τ_1 should be valid for all internal knots $\tau_i, i = 1, 2, ...k-1$. That is, for every internal knot $\tau_i \in \phi, i = 1, 2, ..., k-1$, we impose that $f^i(\tau_i) = f^{i+1}(\tau_i)$, $f^{i\prime}(\tau_i) = f^{i+1\prime}(\tau_i)$, $f^{i\prime\prime}(\tau_i) = f^{i+1\prime\prime}(\tau_i)$, and $f^i(\tau_i) = y(\tau_i)$. In addition, we impose that $f^1(\tau_0) = y(\tau_0)$ and $f^k(\tau_k) = y(\tau_k)$. Imposing these conditions, we end up with a total of $4k-2$ restrictions. Since vector β_t contains $4k$ unknowns, to avoid a multiplicity of solutions, we follow the spline literature and set the second derivative of $f_i(\cdot)$ at the endpoints of ϕ to be zero. Stacking up all the restrictions for each knot $\tau_i \in \tilde{\phi}, i = 1, 2, ..., k$ and denoting

6 In the case where arbitrageurs are actively eliminating arbitrages across the yield curve, the functions 1, g_i, h_i, and z_i resemble the transformed loadings $a_{\beta_k}(\tau)$'s of the demand risk factors.

by $R(\tilde{\phi})$ the matrix containing the terms multiplying the β_t in the first line of Equation (11), we can write the restriction in matrix form as $R(\tilde{\phi})\beta_t = 0$.[7]

Using the equality $W(\tilde{\tau})\beta_t = f_t^i(\tau)$ and the functional form for $f_t^i(\tau)$, we can write the model, with the spline restrictions, in terms of the general framework on Equation (6). In fact, recall that $\beta_t = [\beta_t^1 \cdots \beta_t^k] = [a_t^1\ b_t^1\ c_t^1\ d_t^1 \cdots a_t^k\ b_t^k\ c_t^k\ d_t^k]'$. Define $W(\tilde{\tau})$ as follows:

$$W(\tilde{\tau})(j,:) = \begin{pmatrix} 0_{1\times 4} & \cdots & 0_{1\times 4} & 1g_i(\tilde{\tau}_j)h_i(\tilde{\tau}_j)z_i(\tilde{\tau}_j) & 0_{1\times 4} & \cdots & 0_{1\times 4} \end{pmatrix}.$$

Note that each line of W defines the spline function for a specific maturity in $\tilde{\tau}$. That is, for each $\tilde{\tau}_j \in \tilde{\tau}$, there exists i such that $\tilde{\tau}_j \in A_i$, implying that the j-th row of $W(\tilde{\tau})$ contains only the local loading functions of segment i. Therefore, $W(\tilde{\tau})$ is of dimension $m \times K$ with $K = 4 \times k$, where k indicates the number of segments in the term structure.

Finally, to ease the computation of the model, we write it in matrix form, following the specification on Equation (6):

$$Y_t(\tilde{\tau}) = W(\tilde{\tau})\beta_t + \varepsilon_t(\tilde{\tau})\,,\ R(\tilde{\phi})\beta_t = 0. \tag{15}$$

Interestingly, while matrix W is built based on the vector $\tilde{\tau}$ of maturities of observed yields, the condition $R(\tilde{\phi})\beta_t = 0$ represents the spline restrictions that should be satisfied at the knots in $\tilde{\phi}$. It guarantees that even though coefficients and functional loadings may change across segments, the term structure is smooth at the domain of maturities $[T_m, T_M]$.

Lastly, we note that our approach is close to that of Jungbaker, Koopman, and van der Wel (2014), who develop a model imposing spline restrictions on the factors loadings matrix. Departing from their model, however, the matrix $W(\tilde{\tau})$ is prespecified, given by the choice of the term structure parametrization, following the lead of Nelson and Siegel (1987), Svensson (1994), and Diebold and Li (2006). We expand on this in Section 2.4. In contrast, Jungbaker, Koopman, and van der Wel (2014) impose spline restrictions directly on $W(\tilde{\tau})$, which is not fixed, but estimated.

2.2 The Unrestricted Model

The previous section presented the general framework of the article. To further simplify the estimation procedure we can rewrite the model in Equation (15) in a more convenient way. Since all restrictions are equalities, Equation (15) can take an unconstrained form, reducing factors' dimensionality from $4k$ to $k+1$. Here, we focus on a simplified description of our reduction procedure. Detailed algebra and additional comments appear in the Online Appendix.

Note first that matrix R is block diagonal. By construction, all rows of R are independent, implying that its rank is equal to $3k+1$. Therefore, it is possible to decompose R into two sub-matrices, one formed of its first $3k+1$ rows and columns, and the other of the complementary columns. Analogously, decomposing both β_t and $W(\tilde{\tau})$ in two subcomponents that match the decomposition on R, we can rewrite the model as follows[8]:

$$\begin{aligned} &Y_t(\tilde{\tau}) = W_1(\tilde{\tau})\theta_t + W_2(\tilde{\tau})\hat{\theta}_t + \varepsilon_t(\tilde{\tau}) \\ &R_1\hat{\theta}_t + R_2\theta_t = 0 \end{aligned} \tag{16}$$

7 In the Online Appendix, we provide in more detail the construction of the restriction matrix.

8 In the Online Appendix, we write explicit formulas for R and W so that the decomposition matches both matrices.

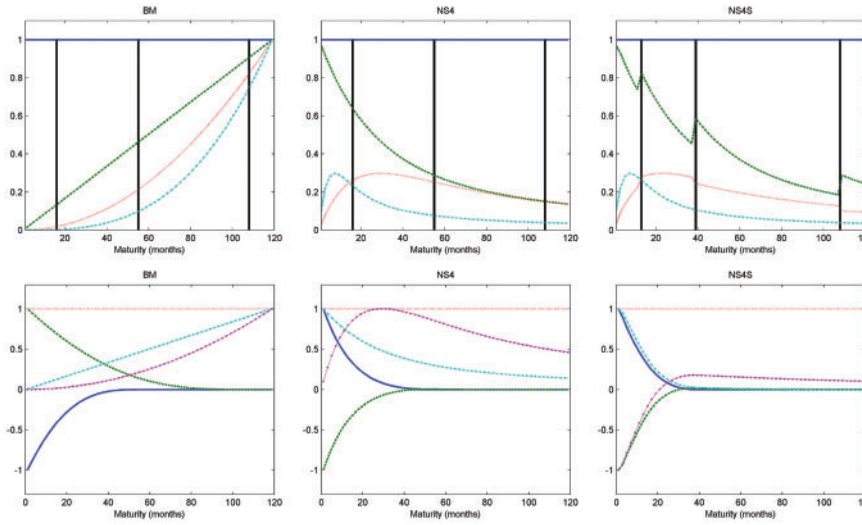

Figure 1. Loadings.

Notes: These graphs present original and transformed loadings, for the segmented models: Bowsher and Meeks (BM), weak segmented (NS4), and strong segmented (NS4S). BM and NS4 models have the same set of knots (1, 16, 55, 108, 120 months), whereas the NS4S model has the following set of knots: 1, 13, 39, 108, 120 months. In between each pair of knots and separated by black vertical lines, there are four functions representing local loadings.

where R_1 is a $3k-1 \times 3k-1$ full-rank square matrix, R_2 is a complementary $3k-1 \times k+1$ matrix, and $\theta_t, \hat{\theta}_t$ are the corresponding sub-vectors of local factors of β_t. Using the fact that R_1 can be inverted, we have that $\hat{\theta}_t = -R_1^{-1}R_2\theta_t$. Thus, setting $Z(\tilde{\tau}) = (W_1(\tilde{\tau}) - W_2(\tilde{\tau})R_1^{-1}R_2)$ we get an *unrestricted form* for the model:

$$Y_t(\tilde{\tau}) = Z(\tilde{\tau})\theta_t + \varepsilon_t(\tilde{\tau}). \tag{17}$$

With this decomposition, we drastically reduce the dimensionality of the problem. In fact, the dimensions of Z and θ_t are $m \times (k+1)$ and $(k+1) \times 1$, respectively.

Figure 1 helps to develop some intuition for the particular case adopted later in the empirical section, where we have four segments of the yield curve. Its top panel shows, for each segmented model, sixteen local factors (four on each segment, separated by vertical lines). Its bottom panel contains transformed loadings after smoothing conditions are imposed. Basically, instead of dealing with sixteen unrestricted loadings, four within each segment that appear in the restricted form [Equation (15)], the unrestricted version of the model [Equation (17)] uses the smoothing conditions to rotate the original factors and reduces their dimension from sixteen to five.

It is important to emphasize that the spline smoothing conditions are specifically responsible for linking the $4k$ untransformed local spline functions 1, g_i, h_i, and z_i within each segment i and represented in matrix form by $W(\tilde{\tau})$ in Equation (15) to the final $k+1$ transformed loadings $Z(\tilde{\tau})$ in Equation (17). This is exactly what happens in the Vayanos

and Vila model where arbitrageurs connect the (local) demand loadings $\{\theta_k(\tau)\}$'s to the transformed yields' loadings $a_{\beta_k}(\tau)$.[9]

Equation (17) should be valid for any pair of maturities and yields. In particular, given that yields with nonobserved maturities corresponding to the knots in ϕ are assumed not to contain an error term (see the set of restrictions on R), it must be true that $Y_t(\phi) = Z(\phi)\theta_t$.[10] Using that matrix $Z(\phi)$ is square and invertible, we obtain a relation between $Y_t(\tilde{\tau})$ and $Y_t(\phi)$:

$$\begin{aligned} Y_t(\tilde{\tau}) &= Z(\tilde{\tau})\Big(Z(\phi)\Big)^{-1} Y_t(\phi) + \varepsilon_t(\tilde{\tau}) \\ \Rightarrow Y_t(\tilde{\tau}) &= \Pi(\tilde{\tau}, \phi) Y_t(\phi) + \varepsilon_t(\tilde{\tau}). \end{aligned} \tag{18}$$

Therefore, the latent yields can be estimated in the cross-section by ordinary least squares (OLS):

$$\hat{Y}_t(\phi) = Z(\phi)(Z(\tilde{\tau})'Z(\tilde{\tau}))^{-1} Z(\tilde{\tau})' Y_t(\tilde{\tau}). \tag{19}$$

2.3 Latent Factor Dynamics

A final step to conclude our model is to specify the dynamics for the latent factors. We need to disentangle the role of imposing cross-sectional segmentation from the role of the choice of dynamics for the latent factors. We propose an analysis of two different important types of latent factors' dynamics previously adopted in the term structure literature: AR versus ECM.

First, following Diebold and Li and the affine literature, we use a classical form of factors' AR dynamics. In this case, we adopt VAR/AR processes for the latent yields, and the model can be represented in the following state-space form:

$$Y_t(\tau) = \Pi(\tau, \phi) Y_t(\phi) + \varepsilon_t(\tau), \tag{20}$$

$$Y_{t+1}(\phi) = c + \gamma Y_t(\phi) + \nu_t, \tag{21}$$

for $t = 1, 2, \ldots$ where c is a $k+1$ vector and γ is a $k+1 \times k+1$ that is either a diagonal or a triangular-diagonal matrix. The error terms ε and ν are such that $E(\varepsilon_t(\tau)\varepsilon_t(\tau)') = \tilde{\Sigma}_\varepsilon$, $E(\nu_t(\tau)\nu_t(\tau)') = \tilde{\Sigma}_\nu$, $E(\varepsilon_t(\tau)\nu_t) = 0$.

It is interesting that while the factors that appear in the work of Diebold and Li and in dynamic affine models are usually rotations of principal components of yields, the latent factors in our model are latent yields that inherit nonstationarity conditions from observed yields.[11] Therefore, to control such potential nonstationarity, in the second approach, we follow the cointegration-based yield curve literature (Hall, Anderson, and Granger, 1992;

9 An important word on notation: note that the restricted form of the model in Equation (15) presents unrestricted loadings, since the restriction is written explicitly in the model formulation, but not applied. On the other hand, the unrestricted form of the model in Equation (17) presents restricted loadings, since the original restriction is applied to the model.

10 Although in general yields $Y_t(\phi)$ are latent, they are not forced to be, for instance, if we choose a knot equal to the maturity of an observed yield. In such cases, we assume that the yield at that knot is observed with error.

11 Principal components and other linear combinations of yields are also sometimes nonstationary, as we see in the empirical section.

Bowsher and Meeks, 2008), and propose an ECM for the dynamics of the latent yields. The model state-space form is given by

$$Y_t(\tau) = \Pi(\tau, \phi) Y_t(\phi) + \varepsilon_t(\tau), \text{ and} \tag{22}$$

$$\Delta Y_{t+1}(\phi) = \alpha(\rho' Y_t(\phi) - \mu_s) + \Psi \Delta Y_t(\phi) + \nu_t, \tag{23}$$

for $t = 1, 2, \ldots$, where α and ρ are $k+1 \times k$ matrices, Ψ is a $k+1 \times k+1$ matrix, and μ_s is a $k+1$ vector. Here ρ is a cointegration matrix that 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 knot yields.[12] The error terms ε and ν are such that $E(\varepsilon_t(\tau)\varepsilon_t(\tau)') = \Sigma_\varepsilon$, $E(\nu_t(\tau)\nu_t(\tau)') = \Sigma_\nu$, and $E(\varepsilon_t(\tau)\nu_t) = 0$. For more details on this specification we refer to Bowsher and Meeks (2008).

Both systems [(20) and (21)] and [(22) and (23)] 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, in general, produces strong out-of-sample forecasting results. In addition, according to Diebold and Rudebusch (2013), "little is lost in practice by using two-step estimation because there is typically enough cross-sectional variation." For the sake of simplicity, in the empirical section, we adopt the two-step procedure proposed by Diebold and Li: first estimating the latent yields running an OLS regression [see Equations (18) and (19)], and then estimating the VAR/AR(1) in Equation (21) and the ECM in Equation (23) to obtain the parameters $\{c, \gamma, \tilde{\Sigma}_\varepsilon, \tilde{\Sigma}_\nu\}$ and $\{\alpha, \mu_s, \Psi, \Sigma_\varepsilon, \Sigma_\nu\}$, respectively.

2.4 Specification of Loadings in Segmentation

The previous sections introduced the general segmented model's three fundamental aspects: the partition ϕ of the yields maturities, the function $f_i(\cdot)$ that determines the term structure movements within each segment, and the dynamics of the underlying factors. Although, in the general model, one does not need to assume any specific functional form for $f_i(\cdot)$, we specialize our empirical application to two specific cases: the polynomial model of Bowsher and Meeks (2006) and the exponential one of Nelson and Siegel (1987), Svensson (1994), and Diebold and Li (2006). An alternative is presented by Jungbaker, Koopman, and van der Wel (2014), who impose spline restriction for knots of the term structure, but allow segments to be independent of each other.

Additionally, for the exponential version, we distinguish two forms of segmentation, conveniently designated weak and strong. Under the former, only local factors (a^i, b^i, c^i, and d^i) change across segments but loadings (g_i, h_i, and z_i) remain the same, and are, therefore, independent of i. On the other hand, under the strong segmentation form, not only factors' dynamics vary across segments, but also the functional form of the loadings, that is, neither coefficients a^i, b^i, c^i, and d^i nor functions g_i, h_i, and z_i are independent of i. We start by developing the strong form of segmentation and then proceed to write the weak model as a special case of the strong segmented model.

12 The i-th column of ρ contains -1 at the i-th line, 1 at the $(i+1)-$th line, and zero for all other entries.

The strong segmentation factor loadings are based on the Nelson and Siegel (1987), Svensson (1994), and Diebold and Li (2006) models. However, we allow different functional forms for g_i, h_i, and z_i within each segment. Despite possible discontinuities in the loading functions at the knots, the smoothing restrictions guarantee that the yield curve remains continuous and smooth. In this model, each segment has its own dynamics and functional loadings, while smoothing constraints connect local to global dynamics across maturities, reinforcing the analogy with the preferred habitat theory. We propose the following form for the local loading functions:

$$g_i(\tau) = \frac{(1 - \exp^{-\lambda_1 \Lambda_i(\tau)})}{\lambda_1 \Lambda_i(\tau)}, \tag{24}$$

$$h_i(\tau) = \frac{(1 - \exp^{-\lambda_1 \Lambda_i(\tau)})}{\lambda_1 \Lambda_i(\tau)} - \exp^{-\lambda_1 \Lambda_i(\tau)}, \text{ and} \tag{25}$$

$$z_i(\tau) = \frac{(1 - \exp^{-\lambda_2 \tau})}{\lambda_2 \tau} - \exp^{-\lambda_2 \tau} \tag{26}$$

where Λ_i's are functions that introduce discontinuities in the loadings at the knots, and $i = 1, \ldots, k$.[13] Although there are many possibilities for Λ, for our first assessment of how this kind of segmentation could affect forecasting, we adopt the following linear functional form in the empirical section:

$$\Lambda_i(\tau) = \tau - \tau_{i-1}(1 - p),\ \tau \in A_i\,, \tau_{i-1} \in \phi \,\text{and}\, p \in [0, 1]. \tag{27}$$

While the parameters λ_1 and λ_2 control the decay rates on the exponential loadings, the parameter p has a different role. It controls the degree of loading segmentation. To go from the strong segmentation model to the weak version we simply need to set $\Lambda_i(\tau) = \tau$ for all i.

The segmented loading models can be classified in two different ways. First, with respect to the parametric family of functions that define local loadings, there are polynomial versions (Bowsher and Meeks' model) and exponential versions (given above). With respect to the degree of segmentation of loading functions, models may present strong or weak segmentation. At this point, it is useful to introduce some notation. We refer to the polynomial version of the model as the Bowsher and Meeks (BM) model. The exponential versions of the model are, respectively, designated as the weak segmented model (NS4, for Nelson and Siegel's four-factor model) and the strong segmented model (NS4S, "S" for strong).

We end this section describing the general methodology we propose with the particular cases studied in the literature. In particular, if we set $g_i(\tau) = \tau$, $h_i(\tau) = \tau^2$ and $z_i(\tau) = \tau^3$ for $i = 1, 2, ..., k$, we have the cross-section specification of the BM model. If we force the coefficients a^i, b^i, c^i, and d^i to be equal across segments and do not impose the smoothness conditions, we narrow to the four-factor Svensson model. If we further restrict $d^i = 0$ for all segments, we get the three-factor Nelson and Siegel model.

13 The choices of how to introduce a discontinuity in the loadings, and in which loadings to introduce such discontinuity, are completely arbitrary. For illustrative purposes we choose to segment only the slope and the first curvature, keeping the level and the second curvature intact.

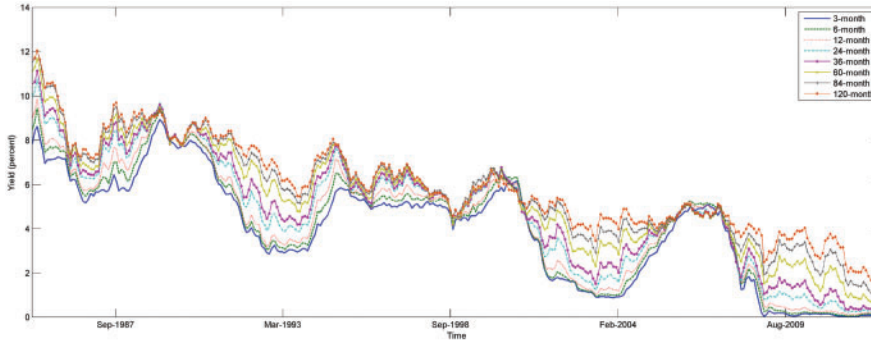

Figure 2. U.S. Treasury bootstrapped zero coupon yields.

Notes: This figure contains a monthly time series of bootstrapped U.S. Treasury yields with maturities of 3, 6, 12, 24, 36, 60, 84, and 120 months. Yields, available from January 1985 to October 2012, were bootstrapped from par yields from the St Louis Federal Reserve data set.

3 Empirical Results

3.1 Data and Competing Models

The analysis provided in this section has two main objectives. First, we assess the empirical performance of segmented term structure models against several benchmarks, capturing term structure information expressed as both cross-sectional and dynamic properties of yields. Second, we verify the role of factors' dynamics restrictions (VAR/AR versus ECM) in models' forecasting performance.

Instead of choosing from a large pool of models to benchmark our results, we concentrate on a small number of well-established ones that have appeared in the term structure literature. These include the Random Walk,[14] Diebold and Li, Dynamic Svensson Model, and a three-factor, essentially affine, Gaussian model.[15]

Several methodologies for the construction of zero coupon yield curves exist, so for robustness purposes, we use three data sets in our forecasting exercise. First, we consider the St Louis Federal Reserve's constant-maturity U.S. Treasury yields on a monthly basis, starting in January 1985 and ending in October 2012 (henceforth, FRED data). This data set consists of par yields, from which we obtain zero-coupon yields by applying a bootstrap procedure with flat-forward interpolation. The two other data sets consist of smoothed and unsmoothed Fama–Bliss yields that are also observed on a monthly basis starting in January 1985, and ending in December 2009. Due to space constraints, we report results obtained with only one of the data sets (FRED), reserving an Online Appendix to report detailed results based on the other two.

Figure 2 plots the time evolution of bootstrapped zero-coupon yields with maturities of 3, 6, 12, 24, 36, 60, 84, and 120 months. At a lower frequency (annually), yields decline

14 This choice was motivated by Carriero, Kapetanios, and Marcellino (2012), who show that it outperforms a large number of term structure models in out-of-sample forecasting exercises.

15 See Diebold and Li (2006) and De Pooter (2007) for results on forecasting performance of exponential models. In addition, Duffee (2002) documents that the affine model performs well in out-of-sample predictions.

over time, while on a monthly basis, there are many periods when yields oscillate around a fixed region. The yields achieve a maximum of 11.6% (for the 84 month maturity yield) at the beginning of the sample and a minimum of 0.04% (for the 3 month maturity yield) at the end of the sample. Throughout the observed period, yield curves frequently have a downward-sloping shape. Nevertheless, at some points, the term structure is increasing, hump-shaped, or inverse.

3.2 Model Estimation

Two important characteristics must be taken into account when detailing the parametric term structure models: choice of the number of latent factors and types of parameters in each model. There is no consensus in the literature on choosing the number of latent factors. Traditionally, Litterman and Scheinkman (1991) show that three factors account for more than 95% of the variability of the U.S. term structure of interest rates. Building on their result, most term structure models adopt only three dynamic factors. In contrast, Cochrane and Piazzesi (2005) identify more than three factors that are necessary to forecast bond risk premia.[16] In addition, Duffee (2011) applies a five-factor Gaussian affine filtered model to monthly U.S. Treasury yields. He finds that a latent factor with almost no effect in cross-sectional fit explains around 30% of the total variance in expected bond returns. Therefore, apparently, the number of optimal latent factors depends on the kind of application desired. It appears reasonable to search in a range from three to five factors when designing a term structure model whose goals include forecasting yields and bond returns.

Drawing on four-factor exponential models such as those of Svensson,[17] all of our segmented models present four local factors within each segment (see Section 2.1). We also know that the number of final dynamic factors, after imposing smoothing restrictions, should equal the number of segments plus one. In an effort to capture yield curve dynamics with factors that may differ from traditional level, slope and curvature factors, we deliberately choose to have five dynamic factors in our segmented models. Our choice automatically imposes a couple of significant restrictions to these models: each will present four segments (five knots) and sixteen unrestricted loadings, with four loadings per segment.[18]

In order to choose the knots' positions, we follow Bowsher and Meeks (2008), fixing an in-sample training period (January 1985 to January 1994). The in-sample stage for knot selection is based on assessing the cross-sectional fit for both segmented models (NS4 and NS4S). Any candidate has extreme knots at 1 and 120 months and avoids combinations in which the distance between neighboring knots is less than 12 months, to avoid clustering.

16 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.

17 For a few dynamic versions of exponential models with more than three factors, see Koopman, Mallee, and van der Wel (2010), De Pooter (2007), and Almeida et al. (2009).

18 In our empirical exercise, we choose the number of dynamic factors, and that determines the number of knots. A different way to proceed is to choose the number of knots based on some fitting statistical procedure like Bowsher and Meeks (2008). Assuming that the yield curve is driven by cubic splines with a parsimonious number of parameters, they search for the best in-sample fit, choosing the positions of knots from all possible splines with five or six knots. For more on this estimation procedure, see Section 2.3 of Bowsher and Meeks (2008).

The best knot vector is the one that minimizes the root mean square error (RMSE) of the panel of yields within the training period, subject to the constraints mentioned. The optimal knot vector for the weak segmented model is $\{1, 16, 55, 108, 120\}$, and for its strong extended version, $\{1, 13, 39, 108, 120\}$, with all maturities measured in months. Those knots define four segments: two short maturities, one medium maturity, and one long maturity.

To accommodate different types of parameters in exponential models (segmented models, Diebold and Li, or Svensson models), we must choose values for λ_1 and/or λ_2, the parameters that govern the decay speeds of exponential loadings. Following Diebold and Li, we fix $\lambda_1 = 0.0609$, making the first curvature loading related to medium-term maturities, with a maximum at 30 months. For the second curvature parameter λ_2, we follow Almeida et al. (2009) and fix $\lambda_2 > \lambda_1$, aiming for a maximum value at the short end of the yield curve.[19] Varying only λ_2, we minimize the in-sample RMSE within the training period, finding a value of $\lambda_2 = 0.24$. With this value, the second curvature achieves a maximum around 7 months, ending up locally driving short-term movements at each segment. In our empirical analysis, this short-maturity peak of the second curvature is responsible for an improvement in forecasts at the short end of the yield curve.

When considering the strong segmented loading model (NS4S), we must choose the value of the parameter p that controls segmentation across loadings. Note that $p = 0$ corresponds to maximum loading segmentation, whereas $p = 1$ corresponds to no segmentation. We set a grid of values for p, ranging from 0 to 0.95, with increments of 0.05, to verify the sensitivity of forecasting results to changes in this parameter. Our analysis indicates that the improvement in forecasting performance is a slowly increasing function of the degree of segmentation p. In our empirical section, we report results for the moderate segmentation case ($p = 0.5$). In the Online Appendix, we discuss in more detail the methodologies adopted to select knots, decaying parameters $\lambda's$, and the degree of segmentation parameter p, and show that the reported results are robust to several variations of these parameters.

3.3 Tightening the Economic Link

Section 1 explores the heuristics behind the preferred habitat theory in developing an economic intuition for segmentation in fixed-income markets. As already mentioned, we incorporate the ideas of this economic intuition into a reduced-form statistical model. Here, we offer a more concrete understanding of how segmentation affects interest rates cross-sectional fitting, and how it identifies latent factors driving yield curve dynamics.[20]

Before getting to the results, we must first understand the elements that define each model's cross-sectional fitting. Overall, both segmented (BM, NS4, and NS4S) and traditional (Nelson and Siegel, i.e., DL, and Svensson, i.e., DSM) models have latent factors that are rotations of observable yields. Putting aside the number of factors driving the yield curve, the key differences across these models consist in how latent factors are extracted from yield curve data. Notably, segmented models allow for local factors whose influence

19 Almeida et al. (2009) adopt a dynamic Svensson 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.

20 Given the peculiarities of the affine model, including how no-arbitrage restricts cross-sectional fitting, we choose to exclude it from the following discussion.

Table 1. Fitting error

Model	Mean absolute basis points error per maturity							
	3	6	12	24	36	60	84	120
DSM	1.46	4.46	4.63	4.88	2.00	4.83	4.19	3.85
	(1.11)	(3.42)	(3.67)	(3.23)	(1.52)	(4.10)	(3.73)	(3.48)
DL	9.52	8.85	6.68	3.69	4.99	6.95	4.20	5.68
	(7.33)	(7.38)	(5.24)	(2.72)	(3.89)	(4.92)	(3.75)	(4.64)
BM	7.76	8.76	4.77	2.92	3.61	2.97	1.04	0.09
	(8.00)	(8.00)	(4.32)	(2.67)	(3.65)	(3.24)	(1.14)	(0.10)
NS4	7.40	8.28	2.64	1.79	4.80	2.11	4.45	1.80
	(6.21)	(7.53)	(1.94)	(1.50)	(4.09)	(1.68)	(3.46)	(1.38)
NS4S	1.39	3.83	4.09	3.28	2.57	2.33	4.12	2.02
	(1.13)	(3.08)	(3.30)	(2.52)	(2.01)	(2.10)	(3.65)	(1.73)

Notes: This table presents the mean absolute error, measured in basis points, for the Svensson model (DSM), the Diebold and Li (2006) model (DL), the Bowsher and Meeks (2008) polynomial segmented model (BM), and the two exponential loading models, one with fixed loadings (NS4) and the other with loadings varying across segments (NS4S). BM and NS4 models have the same set of knots (1, 16, 55, 108, 120 months), whereas the NS4S model has the following set of knots: 1, 13, 39, 108, 120 months. For each model we present the average across all dates of the absolute error, in basis points, segregated by maturity date. We also indicate the respective variance (in parentheses).

spreads out to the yield curve smoothly, through splines restrictions. In contrast, each factor in traditional models affects the whole yield curve via its factor's loadings. This last feature distinguishes the cross-sectional fitting structures of segmented and traditional models: while traditional models cannot capture factors restricted to small subsets of yields, segmented models are designed to do that.

This ability of segmented models to capture local movements allows them to absorb idiosyncratic yield curve movements for specific maturities that happen dynamically over time. After a local absorption happens, splines restrictions propagate the absorbed local movement to the whole yield curve, guaranteeing that in another layer in time, the original idiosyncratic movement will directly affect the yield curve dynamics. This mechanism mimics closely how demand shocks are propagated to the yield curve dynamics of Vayanos and Vila (2009).

Empirically, this can be verified by each model's ability to fit the yield curve. Table 1 summarizes the overall mean absolute error, in basis points, for all models when considering the full sample period.

In general, segmented models have lower fitting errors compared to nonsegmented ones. The only exception is the performance of the Svensson model in the short end of the curve. Notably, the addition of a fourth factor, partially capturing short-maturity yields, improves its fitting, compared to the Nelson and Siegel model. Still, with the exception of the 3 year maturity, the fitting errors of the Svensson model are significantly higher than those of the strong segmented model (NS4S). Indeed, the strong segmented model (NS4S) presents the best fit for the very short end of the curve (i.e., 3 and 6 month maturities). For intermediary maturities, the weak segmented model (NS4) performs well, followed closely by the strong segmented model. For the long end of the curve, the flexibility of the polynomial model is naturally translated into a very precise fit.

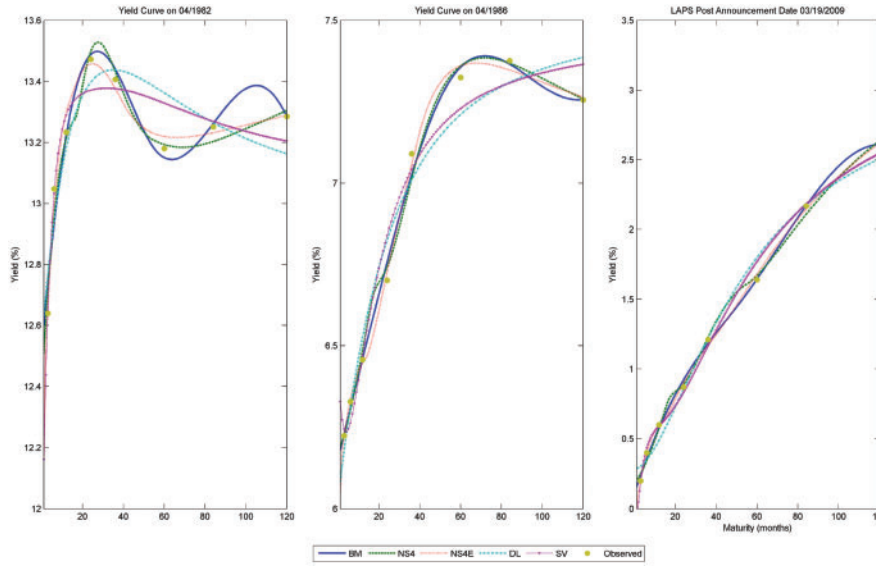

Figure 3. Actual and fitted values for selected dates.

Notes: This picture presents the fitted values for the Svensson model (DSM), Diebold and Li (2006) model (DL), the Bowsher and Meeks (2008) polynomial segmented model (BM), and the two exponential loading models, one with fixed loadings (NS4) and the other with loadings varying across segments (NS4S) for three selected dates. BM and NS4 models have the same set of knots (1, 16, 55, 108, 120 months), whereas the NS4S model has the following set of knots: 1, 13, 39, 108, 120 months.

Results in Table 1 corroborate that allowing for segmentation may improve models' ability to capture the yield curve cross-section. Nevertheless, this summary hides even more interesting information about segmentation. To illustrate this, we select three specific dates: April 1, 1982, April 1, 1986, and March 19, 2009. Notably, the three selected yield curves are clearly influenced by local idiosyncrasies. In particular, the yield curve in April 1982 has a slightly lower yield for the 5 year maturity, compared to the neighboring yields with maturities of 3 and 7 years. In contrast, the term structures in April 1986 and March 2009 are, overall, downward sloping, with the former presenting the 10 year yield slightly below the 7 year.

Figure 3 plots, for the three selected dates, the observed and fitted yields for the Svensson model (DSM), Diebold and Li (2006) model (DL), the Bowsher and Meeks (2008) polynomial segmented model (BM), and the two exponential loading models, one with fixed loadings (NS4), and the other with loadings varying across segments (NS4S). Starting with the left-hand side graph (the 1982 yield curve), we note that while most models have similar performance in the very short end of the curve, their differences are substantial for the 5 and 10 year maturity. In fact, the Svensson and the Nelson and Siegel models underprice the 5 year bond and overprice the 10 year bond. In contrast, segmented models closely price all the available yields, capturing the idiosyncrasies within the long end of the curve. The results for the 1986 yield curve are even more striking. While

segmented models closely match observed yields, results for nonsegmented models deteriorate, compared to the 1982 yield curve.

The right-hand side graph in Figure 3 plots the fitted and observed term structure on March 19, 2009. We select this date, which is not part of our original sample, inspired by D'Amico and King's (2013) evaluation of the LSAP program. On March 18, 2009, the Federal Reserve officially announced the buyback program. Although the program was partially anticipated by market participants, D'Amico and King (2013) argue that the program's scale clearly surprised investors, a reaction reflected in the next day's yield curve. The market response to the program announcement was a heterogeneous shift in the yield curve, with the long end particularly affected. This is the kind of phenomenon that enables us to examine how segmented models accommodate these shocks. Our data reveal that segmented models are capable of closely fitting the observed points on the day subsequent to the announcement. In contrast, the traditional models present a mispricing pattern similar to the one observed in the right-hand side graph: they underprice the 5 year bond and overprice the 10 year bond.

This previous analysis helps us understand how segmented models behave in terms of cross-sectional fitting. Now, to have an idea how this ability to statically capture idiosyncratic yields translates into models' flexibility in capturing local shocks, we propose an additional exercise: to introduce synthetic shocks into the data and see models' abilities to fit the new yield curve. To proceed with this exercise, we select an initial date where all models fit the observed yields extremely well. We perturb the original yields by introducing synthetic local shocks to this data, and reestimate all the models.

Figure 4 provides the results of this illustrative exercise. The top-left panel presents the baseline curve. In the top-right panel, we plot the fitted curves when we impose a 10% positive shock to the 6 month yield. Aligned with the results in Table 1, we note that both the strong segmented (NS4S) and the Svensson (DSM) models capture this shock reasonably well, with the former model slightly outperforming the latter in the short end. Results change drastically when we look at the bottom-left panel, which shows the perturbed yield curves after a 10% positive shock imposed on the 2 year rate. Here, the segmented models partially accommodate the local shock, without completely compromising the fit for the long maturities. In contrast, both the Svensson and the Nelson and Siegel models' fit is significantly affected. More importantly, these latter models are not flexible enough to capture the shape of the perturbed curve.

Zinna (2016) argues that strong demand pressures, for longer maturities, might be related to the inversion of the yield curve. Inspired by his argument, we simulate a 10% negative shock imposed on the 5, 7, and 10 year yields. The bottom-right panel in Figure 4 shows the result of this exercise. Overall, the implications are similar to those previously discussed. Most of the segmented models can clearly accommodate the shock, whereas nonsegmented models' errors can be significant. The results for several other shocks are quite similar and can be summarized as follows: as long as after the local shocks the term structure is not sufficiently smooth, nonsegmented models provide a poor fit.[21]

Given this discussion, what can we say about forecasting? We notice that throughout the analyzed sample, segmented models fit observed data reasonably well, capturing most

21 By a sufficiently smooth yield curve, we mean one that assumes one of the traditional four shapes: downward sloping, increasing, hump-shaped, or inverse.

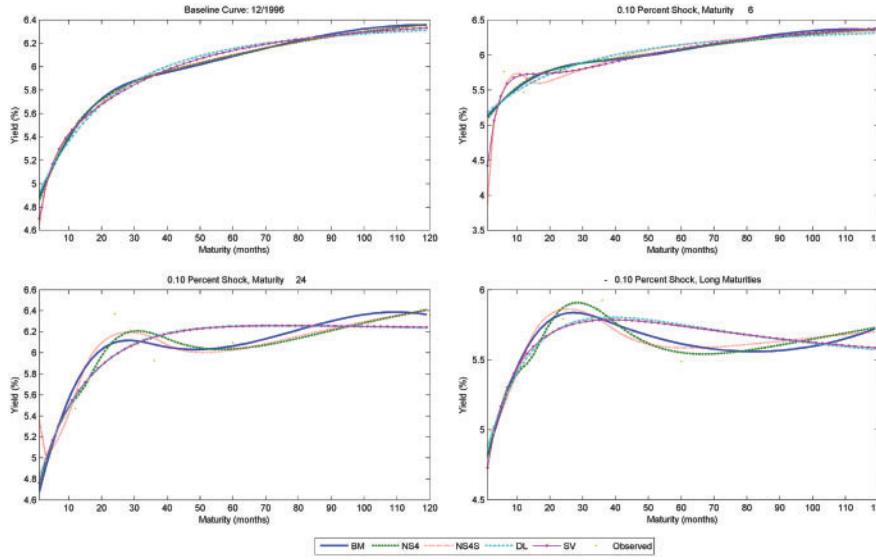

Figure 4. Local shocks yield curve.

Notes: These plots present the fitted values for the Svensson model (DSM), Diebold and Li (2006) model (DL), the Bowsher and Meeks (2008) polynomial segmented model (BM), and the two exponential loading models, one with fixed loadings (NS4) and the other with loadings varying across segments (NS4S) for a selected date and three possible variations due to local shocks in the baseline curve. BM and NS4 models have the same set of knots (1, 16, 55, 108, 120 months), whereas the NS4S model has the following set of knots: 1, 13, 39, 108, 120 months.

of its nuances and accommodating local shocks. By better capturing yields' cross-sectional information, our proposed models can more precisely identify the underlying latent factors driving the yield curve. Intuitively, this more precise identification naturally translates into better model forecasting ability. In fact, as we see later on, the biggest improvements in forecasting provided by segmented models appear in the short end of the yield curve, precisely where the strong segmented model provides enough flexibility to fit data well.

3.4 Factors' Loadings and Impulse Responses

In this section, we present explanations for factors' loadings and factors' dynamics of segmented term structure models. Figure 1 shows the unrestricted (or unsmoothed) loadings in the top panel and restricted (or transformed) loadings in the bottom panel. Vertical lines indicate the positions of knots within each model's unsmoothed loadings. Each model contains a set of four loading functions per segment. For the Bowsher and Meeks and the NS4 segmented models, the total of sixteen unsmoothed loadings that appear when observing the whole set of maturities are visually equivalent to four loadings (level, slope, curvature, and second curvature) observed in traditional term structure models. The difference, however, lies in the yield curve response to shocks in factors attached to each of these loadings: while in the segmented models (Bowsher and Meeks and NS4) the response is potentially local at each segment, in traditional models the response to a shock in any factor directly affects the whole set of maturities. In contrast, the strong segmentation model (NS4S) presents discontinuities at the

unsmoothed loadings of slope and first curvature, implying that the responses to shocks in factors are not only dependent on a particular segment of the term structure, but also may affect segments with intensities that differ from corresponding loadings of traditional four-factor exponential models. We saw in the previous sub-section that the propagation of a shock is very different when comparing segmented to nonsegmented models.

For restricted loadings, all models present as level (dotted red), loadings very similar to a traditional level factor in the spirit of Litterman and Scheinkman (1991). The same is true for the slope factor of the Bowsher and Meeks and the weak segmented NS4 model (dash-dotted cyan). Nevertheless, the slope for the strong segmented model NS4S (triangular magenta line) is nonlinear, approximated by two linear functions, one steeper for short maturities up to 20 months and the other, almost flat, picking up longer maturities. For each model, there is a set of traditional factors, as in Litterman and Scheinkman (1991), and another set of more specific local factors. Local factors are characterized by a quick decay to zero. They are represented by the solid blue and dashed green factors of the three segmented models, and by the dash-dotted cyan factor of the strong segmented model. These local loadings are important only for certain subsets of maturities and, empirically, part of the differences in the forecasting ability of models comes from the existence of local factors and from differences across these factors. They help to capture additional yield dynamics that are not identified by traditional nonsegmented factor models.

We have just provided a static view of the implications of factors' loadings to the cross-section of yields in segmented models. In order to complement this view, we close this section with an impulse response exercise on the restricted factors in the weak segmented model (NS4) with ECM dynamics. In particular, we show how observed yields at different segments of maturities react to shocks applied to the transformed factors of this model.[22] Using the whole sample available, we estimate the dynamics in Equations (22) and (23), apply a one-standard-deviation shock to each restricted dynamic factor, and track the system response for twenty periods. Then, using Equation (22), we translate factors' responses into observable yields' responses.

In Figure 5, let us start by analyzing the first two panels, which present shocks relative to the two local factors of the NS4 model. Those are the factors attached to the solid blue and dashed green loadings appearing in the center panel on the bottom of Figure 1. A shock applied to the dynamic factor attached to the blue loading affects the short-run, short-maturity yields more severely than the other segments. This effect dissipates in the long term and inverts, leading, in the long run, to a smaller response for the short-maturity segment, relative to the other segments. The effect of the shock for longer maturities is smaller in the short run and stronger in the long run, indicating clearly different behaviors for the responses to this shock across segments of the term structure.

For a shock in the dynamic factor attached to the dashed green loading, longer maturity yields are most affected. Although all yields respond to the shock with a similar shape, the intensities and decaying speeds are clearly different between short-maturity and longer maturity yields. In contrast, at the bottom of Figure 5, we present shocks relative to the traditional level, slope and curvature dynamic factors, respectively, attached to the dotted red, dash-dotted cyan, and triangular magenta loadings of the NS4 model on the bottom of Figure 1. Notably, movements of the term structure relative to a shock in any of these factors are very similar, with yields moving much closer to each other than in the top panel.

22 Results for the extended exponential segmented model are similar and available upon request.

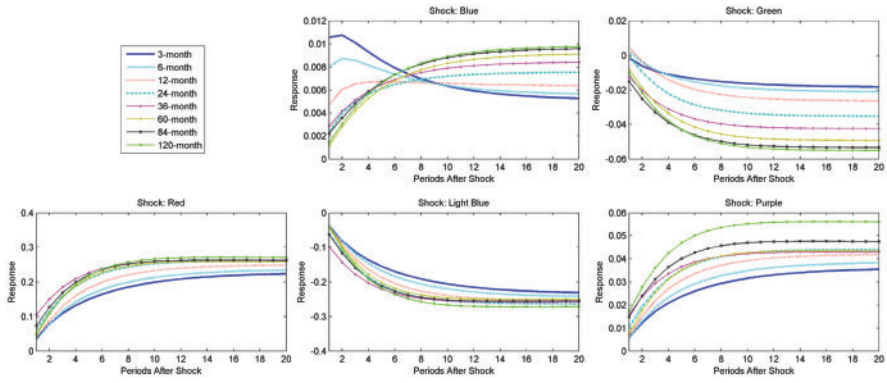

Figure 5. Impulse response function.

Notes: This figure presents five panels, each containing the response of all U.S Treasury yields when a one-standard-deviation shock is applied to one of the five factors of the segmented model NS4 with ECM dynamics. Knots appear in (1, 16, 55, 108, 129) months.

3.5 Forecasting

In this article, we compare models based on multiple forecasting exercises. Given our goal, a choice has to be made between two usual approaches, rolling-window versus recursive methods (McCraken, 2007). Based on Giacomini and White's (2006) findings favoring a rolling-window approach, we adopt this method for the 143 out-of-sample forecasting exercises performed.[23]

Hansen and Timmermann (2012) note the importance of controlling the split point of a data set into estimation and evaluation periods in out-of-sample tests. On a related issue, Rossi and Inoue (2012) propose a robust approach to data snooping across the length of the estimation window, in rolling-window forecasting evaluations. In this article, we do not consider variations of window size, nor examine data mining over the sample split. Instead, we vary the data sets, creating different forecasting exercises, keeping both window size and sample splits constant. We follow Diebold and Li (2006), and fix an in-sample window size equal to $108 - h + 1$ months (h, forecasting horizon) and an out-of-sample window size equal to 84 months. A forecasting exercise consists of a sequence of eighty-four out-of-sample monthly forecasts constructed with the rolling-window method. For each exercise and forecasting horizon h, we define the RMSE of the m-maturity yield by

$$\mathrm{RMSE}(T_E, h, m) = \sqrt{\frac{\sum_{i=1}^{84}\left(\hat{Y}^{h}_{T_E-84+i}(m) - Y_{T_E-84+i}(m)\right)^2}{84}}, \tag{28}$$

where T_E is the last prediction date and $\hat{Y}^{h}_{T_E-84+i}(m) - Y_{T_E-84+i}(m)$ is the prediction error, since $\hat{Y}^{h}_{T_E-84+i}(m) = E_{T_E-84+i-h}\left(Y_{T_E-84+i}(m)\right)$ is the model conditional expectation at time

23 Giacomini and White (2006) adopt a rolling-window forecasting scheme because it is a convenient way to handle instabilities in the data environment. Moreover, they show that the use of rolling windows accounts for estimation uncertainty when adopting the Diebold and Mariano test for forecasting accuracy.

$T_E - 84 + i - h$ for the m-maturity yield at time $T_E - 84 + i$, h months in the future. To calculate this conditional 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 forecasting exercise, the first prediction is done for February 1994, with an in-sample period from February 1985 to August 1993. Similarly, the 84th (last) prediction is done for January 2001 with an in-sample period from January 1992 to July 2000. The set of all 143 different forecasting exercises is obtained by changing T_E monthly between December 2000 and October 2012.[24]

Following Diebold and Li, we estimate all models (except for the affine) with AR dynamics for each latent factor.[25] The three-factor essentially affine Gaussian model (Duffee, 2002) is estimated with a maximum likelihood method described in the Online Appendix.[26]

Table 2 aggregates information on all forecasting exercises, providing for each entry (per maturity and horizon) the average of all 143 RMSEs relative to the Random Walk. The column labeled RW presents the absolute RMSE of the Random Walk in percentage points. For instance, the average RMSE of the Random Walk, when forecasting the 3 month maturity yield 1 month in the future ($h = 1$), is equal to 21.67 basis points. For any other model in the table, an entry with a value lower than one means that the model is (on average) better than the Random Walk, when considering an aggregation of all 143 different forecasting exercises. Surprisingly, the affine Gaussian model outperforms traditional and segmented models for most cases. In addition, the biggest gains in RMSE, when comparing the affine model to the Random Walk, appear for short maturities. Segmented models seem to outperform most traditional models. Comparing polynomial and exponential models, clearly the latter have lower RMSEs for most maturities and forecasting horizons. In addition, corroborating results in Carriero, Kapetanios, and Marcellino (2012), all models with AR dynamics (except for the affine) are outperformed by the Random Walk.

In view of the poor performance of models with AR dynamics, we adopt alternative dynamics in all analyzed models. Based on the work of Hall, Anderson, and Granger (1992) and Bowsher and Meeks (2008), who suggest an ECM approach to latent factors in term structure models, we adopt it in all models.

The use of ECM dynamics in segmented models is justified by noting that most frequently yields, but sometimes spreads and other linear combinations of yields-generating

24 Our methodology based on the analysis of multiple data sets, although less formal from a statistical viewpoint, is related to Giacomini and Rossi (2010). They develop statistical tests to examine the stability over time of out-of-sample relative forecasting performance of a pair of models in an unstable environment.

25 The differences on average RMSEs obtained with VAR dynamics versus univariate independent AR dynamics for latent factors are presented in Table 1 in the Online Appendix. Negative numbers indicate that the VAR methodology outperforms the AR, whereas positive numbers indicate the opposite. We observe that in 70% of the cases the AR methodology outperforms VAR, which encourages us to include results with AR dynamics in the main body of the article.

26 The forecasting performance (average RMSEs) of a four-factor essentially affine Gaussian model is also presented in the Online Appendix, in Table 11. It is outperformed by the three-factor affine model in twenty-one out of twenty-four different combinations of maturities and forecasting horizons.

Table 2. RMSE mean—AR dynamics

Horizons	RW	Affine	DL	DSM	BM AR	NS4 AR	NS4S AR
3 month yields							
1	0.2154	**0.9066***	1.1712	1.0593	1.0489	1.0575	**0.9961****
6	0.8732	1.0076**	1.1232	1.1215	1.0560	1.0073*	1.0300
12	1.5291	**0.9904***	1.0700	1.0613	1.1106	1.0430**	1.0733
6 month yields							
1	0.2153	**0.8817***	1.0501	1.0927	1.0651	1.0249	1.0234**
6	0.9006	**0.9945****	1.0936	1.1365	1.0462	**0.9660***	1.0301
12	1.5702	1.0028**	1.0522	1.0693	1.0966	**0.9933***	1.0793
12 month yields							
1	0.2179	**0.9418***	1.0160	1.0733	1.0106	**0.9969****	1.0219
6	0.8592	1.0193**	1.1235	1.1725	1.0721	1.0032*	1.0687
12	1.4659	1.0932**	1.0983	1.1176	1.1505	1.0532*	1.1502
24 month yields							
1	0.2447	1.1229	1.0534	1.0512	1.0299	1.0221**	1.0070*
6	0.8469	1.0191*	1.1240	1.1432	1.0695	1.0261**	1.0764
12	1.3501	**0.9953***	1.1371	1.1346	1.1853	1.0285**	1.1995
36 month yields							
1	0.2564	**0.9341***	1.0826	1.0353	1.0225	1.0498	1.0195**
6	0.8224	1.0169*	1.1261	1.1289	1.0761	1.0222**	1.0968
12	1.2355	**0.9984***	1.1751	1.1605	1.2072	1.0680**	1.2309
60 month yields							
1	0.2553	**0.9600***	1.0655	1.0285	1.0039**	1.0054	1.0181
6	0.7448	1.0095*	1.1068	1.1051	1.0920	1.0736**	1.1014
12	1.0372	1.0032*	1.2072	1.1930	1.2275	1.1884**	1.2388
84 month yields							
1	0.2459	1.1027	1.0553	1.0538	1.0192*	1.0249**	1.0380
6	0.6792	1.0192*	1.0789	1.0835	1.1123	1.0561**	1.0697
12	0.9132	1.0198*	1.1929	1.1879	1.2696	1.1600**	1.2022
120 month yields							
1	0.2369	1.0044*	1.0081	1.0145	1.0085**	1.0165	1.0144
6	0.6025	1.0268*	1.0709**	1.0842	1.0804	1.0930	1.0962
12	0.7825	1.0397*	1.2152**	1.2254	1.2224	1.2420	1.2451
D. Factors	–	3	3	4	*5*	*5*	*5*

Notes: This table presents the mean over T_E of the relative RMSE with respect to the Random Walk (RW) for the Diebold and Li model (DL), Svensson model (DSM), Bowsher and Meeks (BM), affine Gaussian, and two exponential loading models: weak segmented (NS4), and strong segmented (NS4S), all with AR factor dynamics. The second column presents the RMSE of the RW model in percentage points. Bold face values represent RMSEs lower than those of RW. The mean is obtained over 143 different out-of-sample forecasting exercises. D. factors indicate the number of dynamic factors in the model. * represents the lowest RMSE for each fixed horizon and maturity, and ** represents the second lowest.

latent factors, are nonstationary and cointegrated. For instance, in the Online Appendix, we provide results of unit root and cointegration tests of implied factors in the Diebold and Li model showing that, in the majority of samples analyzed in this article, its level and slope factors present unit root and in addition, its three factors (level, slope, and curvature) present at least one cointegration factor.

Table 3. RMSE mean—ECM dynamics

Horizons	RW	Affine	DL	DSM	BM ECM	NS4 ECM	NS4S ECM
3 month yields							
1	0.2167	**0.9066**	**0.9014**	**0.8767***	**0.9151**	**0.9165**	**0.8792****
6	0.8792	1.0076	**0.9102**	**0.9100****	**0.9218**	**0.9197**	**0.8944***
12	1.5394	**0.9904**	**0.9642**	**0.9362**	**0.9654**	**0.9211****	**0.8721***
6 month yields							
1	0.2115	**0.8817***	**0.9606**	**0.8906****	**0.9878**	**0.9642**	**0.9037**
6	0.8846	**0.9945**	**0.9560**	**0.9394****	**0.9706**	**0.9524**	**0.9232***
12	1.5426	1.0028	**0.9777**	**0.9454**	**0.9780**	**0.9282****	**0.8789***
12 month yields							
1	0.2141	**0.9418**	**0.9717**	**0.9224****	**0.9268**	**0.9243**	**0.8993***
6	0.8443	1.0193	1.0209	1.0055	1.0316	1.0020**	**0.9709***
12	1.4408	1.0932	1.0167	**0.9822**	1.0119	**0.9507****	**0.9003***
24 month yields							
1	0.2393	1.1229	**0.9891***	1.0099	1.0259	**0.9993**	**0.9943****
6	0.8305	1.0191*	1.0598	1.0524	1.0973	1.0493	1.0224**
12	1.3263	**0.9953**	1.0409	1.0087	1.0517	**0.9756****	**0.9242***
36 month yields							
1	0.2504	**0.9341***	1.0060	**0.9926**	1.0331	1.0418	**0.9950****
6	0.8064	1.0169*	1.0643	1.0559	1.1102	1.0581	1.0331**
12	1.2158	**0.9984**	1.0518	1.0206	1.0675	**0.9868****	**0.9378***
60 month yields							
1	0.2482	**0.9600***	1.0124	0.9960	1.0297	1.0007	**0.9867****
6	0.7312	1.0095*	1.0466	1.0412	1.0935	1.0313	1.0117**
12	1.0270	1.0032	1.0483	1.0268	1.0756	**0.9839****	**0.9423***
84 month yields							
1	0.2384	1.1027	1.0364*	1.0487	1.0520**	1.0576	1.0572
6	0.6686	1.0192*	1.0315	1.0403	1.1197	1.0386	1.0197**
12	0.9098	1.0198	1.0409	1.0353	1.1189	1.0015**	**0.9575***
120 month yields							
1	0.2293	1.0044*	1.0358	1.0377	1.0492	1.0320	1.0140**
6	0.6004	1.0268	1.0263	1.0360**	1.1031	1.0291	1.0013*
12	0.7927	1.0397	1.0566	1.0595	1.1385	1.0277**	**0.9813***
D. Factors	–	3	3	4	*5*	*5*	*5*

Notes: This table presents the mean over T_E of the relative RMSE with respect to the Random Walk (RW) for the Diebold and Li model (DL), Svensson model (DSM), Diebold and Li model with ECM factor dynamics (DL ECM), Svensson model with ECM factor dynamics (DSM ECM), Bowsher and Meeks (BM), and two exponential loading models: weak segmented (NS4) and strong segmented (NS4S). In the second column, we present the RMSE of the RW in percentage points. Bold face values represent RMSEs lower than those of RW. The mean is obtained over 143 different out-of-sample forecasting exercises. D. factors indicate the number of dynamic factors in the model. * represents the lowest RMSE for each fixed horizon and maturity, and ** represents the second lowest.

Table 3 is similar to Table 2 except that all models (but the affine) are estimated using ECM dynamics. These two tables reveal a first striking result: for each model, adopting ECM dynamics for the latent factors provides significantly better out-of-sample forecasts than using corresponding VAR/AR dynamics. RMSEs are smaller under ECM dynamics for

most maturities and forecasting horizons for all analyzed models. In particular, ECM dynamics drastically improve predictions for short-maturity yields.

For maturities of 3 and 6 months, all models have an impressive performance when compared with the Random Walk, for all forecasting horizons. Also, gains from ECM dynamics are particularly large for the Diebold and Li and the Dynamic Svensson models. Despite the strong performance of all models with ECM dynamics on the short-maturity segment, this performance is even stronger for the segmented exponential models (NS4 and NS4S): they outperform the Random Walk for all maturities and forecasting horizons at the short-maturity segment.

A combination of ECM dynamics with segmentation improves forecasting at the short end of the yield curve, a segment that has particular behavior due to liquidity issues (Duffee, 1996; Fontaine and Garcia, 2012). The strongest performance of the exponential segmented models appears to come from local factors driving short maturities that do not exist in traditional term structure models. The fast-decaying loadings of these factors, coupled with their impulse response analysis performed in Section 3.4, suggest the existence of local factors indeed driving term structure movements, as proposed by Vayanos and Vila (2009). Moreover, we calculate the correlation of all dynamic factors of the exponential segmented models with the funding liquidity factor of Fontaine and Garcia (2012) and find that our so-called local factors have a positive correlation with the funding liquidity factor on the order of 30%, whereas traditional term structure factors have half of this correlation.

Another important insight is that introducing segmentation in term structure models consistently improves long-horizon forecasts. In Table 3, for a 12 month forecasting horizon, the strong segmented model (NS4S) consistently outperforms the Random Walk on all maturities. Average RMSEs are around 5% smaller for longer maturities and 10% smaller for shorter maturities.

The exponential versions also have lower RMSEs than the polynomial model (BM). Given that a number of recent papers explore the structure of factor models with polynomial splines (Bowsher and Meeks, 2008; Koopman and van der Wel, 2013; Jungbaker, Koopman, and van der Wel, 2014), our novel results suggest that exponential spline models should be explored further in forecasting problems.

4.5.1 Statistical significance of forecasting performances

Are the differences in performances across models statistically significant? For each forecasting exercise, we adopt a Diebold and Mariano (1995) test with quadratic loss, and 5% significance level. We present comparisons in two tables, the first with all models against the Random Walk, and the second with individual model comparisons. The latter is particularly useful to better clarify the roles of ECM dynamics and segmentation.

In Table 4, candidate models (DL, DL ECM, DSM, DSM ECM, BM, NS4, and NS4S) are tested against the Random Walk.[27] This table shows that the exponential segmented

27 The first number in each entry of the table provides the percentage of forecasting exercises in which the candidate model achieves statistically significant smaller RMSE than the Random Walk. The second number on the same entry represents the percentage of forecasting

Table 4. Diebold and Mariano Test—RW

Horizons	Affine	DL	DSM	DL ECM	DSM ECM	BM	NS4 ECM	NS4S ECM
3 month yields								
1	0 / 0	0 / 72	0 / 64	28 / 0	37 / 0	28 / 0	29 / 0	42 / 0
6	0 / 0	0 / 44	0 / 34	0 / 1	16 / 1	12 / 0	15 / 0	15 / 2
12	0 / 0	0 / 0	0 / 0	0 / 4	12 / 5	5 / 0	15 / 4	15 / 4
6 month yields								
1	0 / 35	0 / 26	0 / 70	0 / 0	39 / 0	0 / 1	0 / 1	42 / 1
6	0 / 0	0 / 16	0 / 27	0 / 2	13 / 1	0 / 2	10 / 2	12 / 2
12	0 / 0	0 / 0	0 / 0	0 / 4	11 / 4	0 / 2	14 / 4	15 / 4
12 month yields								
1	0 / 0	27 / 56	0 / 62	0 / 0	20 / 0	25 / 0	30 / 0	39 / 0
6	0 / 0	0 / 37	1 / 47	0 / 12	0 / 1	0 / 0	10 / 2	8 / 0
12	0 / 12	0 / 0	0 / 0	0 / 11	5 / 0	0 / 0	15 / 3	19 / 0
24 month yields								
1	0 / 0	0 / 94	0 / 54	0 / 0	0 / 0	0 / 0	0 / 0	1 / 0
6	0 / 0	1 / 12	2 / 30	0 / 31	0 / 14	0 / 9	0 / 0	0 / 0
12	0 / 8	1 / 1	2 / 4	0 / 13	0 / 4	0 / 1	8 / 0	15 / 0
36 month yields								
1	0 / 0	0 / 54	0 / 3	0 / 0	0 / 0	0 / 0	0 / 6	1 / 0
6	0 / 0	0 / 11	2 / 5	0 / 41	0 / 13	0 / 19	0 / 6	0 / 0
12	0 / 4	0 / 8	0 / 5	0 / 25	0 / 10	0 / 4	7 / 1	11 / 2
60 month yields								
1	0 / 0	0 / 23	0 / 6	0 / 0	0 / 0	0 / 0	0 / 0	0 / 0
6	0 / 0	0 / 10	0 / 3	0 / 33	0 / 5	0 / 20	0 / 1	0 / 1
12	0 / 0	0 / 25	0 / 26	0 / 30	0 / 26	0 / 15	1 / 11	10 / 7
84 month yields								
1	0 / 0	0 / 15	0 / 15	0 / 0	0 / 0	0 / 6	0 / 0	0 / 0
6	0 / 0	0 / 5	0 / 4	0 / 30	0 / 11	0 / 21	0 / 4	0 / 2
12	0 / 0	0 / 25	0 / 25	0 / 25	0 / 25	0 / 26	0 / 17	5 / 8
120 month yields								
1	0 / 0	0 / 3	0 / 1	0 / 0	0 / 0	0 / 9	0 / 0	0 / 0
6	0 / 0	0 / 1	0 / 3	0 / 0	0 / 3	0 / 14	0 / 5	0 / 2
12	0 / 8	0 / 39	0 / 44	0 / 25	0 / 23	0 / 25	0 / 25	12 / 18

Notes: This table presents the Diebold and Mariano (1995) test with a 5% significance using quadratic loss for the Affine model, Diebold and Li model (DL), Svensson model (DSM), Diebold and Li model with ECM factor dynamics (DL ECM), Svensson model with ECM factor dynamics (DSM ECM), Bowsher and Meeks model (BM), and two exponential loading models: weak segmented (NS4), and strong segmented (NS4S). The first number in each cell represents the percentage of times that the model produces statistically significant better forecasts than the RW. The second number is the percentage of times that the RW produces statistically significant better forecasts than the model. Hundred minus the sum of percentages in a cell gives the number of statistical ties. A total of 143 out-of-sample forecasting exercises are considered.

models are stronger than the Random Walk in a robust way. The NS4 and NS4S models as a group present RMSE statistically smaller than the Random Walk in twenty-four out of forty-eight entries. On the other hand, the Random Walk is statistically better on only

exercises in which the Random Walk achieves statistically significant smaller RMSE than the candidate model.

fourteen entries, with ten ties. Moreover, the percentage of forecasting exercises in which the segmented models are better is significantly higher than the opposite.[28]

Considering the difference in performance between exponential versions of the segmented model and the Dynamic Svensson Model with ECM dynamics, the first panel of Table 5 presents Diebold and Mariano (1995) tests for individual forecasting exercises comparing those models. The aim of this test is to show that the superior forecasting performance of exponential segmented models for the 12 month horizon is neither due to the presence of a fourth latent factor nor due to the ECM dynamics. Data show that both segmented models strongly outperform the Dynamic Svensson for the 12 month forecasting horizon for all maturities, suggesting again the importance of allowing for local factors.

Finally, due to the relatively good performance of the Gaussian affine model (in terms of average RMSEs), we also adopt Diebold and Mariano tests to compare this model to the exponential segmented models. The second panel of Table 5 presents results comparing individual forecasting exercises between the Gaussian affine and exponential segmented models. Not surprisingly, due to the poor performance of the affine model in Diebold and Mariano tests against the Random Walk (Table 4), the segmented models provide better results, especially for the 12 month forecasting horizon.

4.5.2 A dynamic view on forecasting exercises

What can we say about model performance over time? To answer this question, Figure 6 gives us a dynamic view of RMSEs across different forecasting exercises. It presents the time series of RMSEs (relative to the Random Walk) over a 12 month forecasting horizon, for yields with some selected maturities.[29] For the 12 month forecasting horizon, from all models, exponential segmented models with ECM dynamics have the smallest RMSEs on a large number of forecasting exercises. In more than 60% of these exercises they outperform the Random Walk, achieving RMSEs smaller than one.

Figure 6 also allows us to observe a deterioration of model results with respect to the Random Walk, especially after the end of 2007. This is mainly due to a change in the behavior of the U.S. yield curve that was not directly identified by any of the models analyzed here, except perhaps by the affine model. Such change is due to the zero-lower-bound interest rate constraints reached in the period 2008–2012. In fact, analyzing Japanese yields, Kim and Singleton (2013) identify that during a zero-lower-bound period, interest rates have a very different behavior than during regular periods.[30] These peculiarities, along with the fact that all models considered have Gaussian shocks with constant volatilities, contribute to the models' bad performance in capturing term structure movements in a zero-lower-bound period. In contrast, since the Random Walk simply repeats previous values, it is much more sensible for a zero-lower-bound period.

28 For instance, the best result for the Random Walk is when it is statistically better than the NS4 model in 25% of the forecasting exercises, for the 120 month maturity and 12 month forecasting horizon. In contrast, the NS4S model is statistically better than the Random Walk in more than 35% of the forecasting exercises for at least three entries of the table.

29 Similar figures with 1 and 6 month forecasting horizons are available in the Online Appendix.

30 Interest rates present smaller volatilities, keeping substantial fluctuation of longer term yields, and have market prices of risk and yield volatilities more positively correlated across short and intermediate maturities.

Table 5. Diebold and Mariano test

	Dynamic Svensson		Affine	
Horizons	NS4	NS4S	NS4	NS4S
3 month yields				
1	0 / 33	0 / 0	0 / 0	0 / 0
6	0 / 0	0 / 0	0 / 0	0 / 0
12	15 / 0	29 / 0	8 / 0	16 / 3
6 month yields				
1	0 / 58	0 / 4	47 / 0	52 / 0
6	0 / 2	0 / 0	0 / 0	0 / 0
12	15 / 0	30 / 0	10 / 4	22 / 1
12 month yields				
1	0 / 0	10 / 0	54 / 0	54 / 0
6	1 / 2	0 / 0	0 / 0	0 / 0
12	17 / 1	31 / 0	12 / 2	24 / 0
24 month yields				
1	0 / 0	6 / 0	0 / 0	0 / 0
6	0 / 0	0 / 0	0 / 0	0 / 0
12	16 / 0	32 / 0	10 / 0	22 / 0
36 month yields				
1	0 / 13	0 / 0	0 / 0	0 / 0
6	0 / 0	0 / 0	0 / 0	0 / 0
12	16 / 0	31 / 0	8 / 2	35 / 0
60 month yields				
1	0 / 0	11 / 0	0 / 0	0 / 0
6	0 / 0	0 / 0	0 / 0	0 / 0
12	18 / 1	33 / 0	8 / 8	47 / 2
84 month yields				
1	0 / 4	6 / 1	0 / 0	0 / 0
6	6 / 0	0 / 0	0 / 0	0 / 0
12	26 / 1	34 / 0	2 / 13	26 / 7
120 month yields				
1	0 / 0	0 / 0	0 / 0	0 / 0
6	0 / 0	0 / 0	0 / 0	0 / 0
12	14 / 0	35 / 0	0 / 7	18 / 3

Notes: This table presents the Diebold and Mariano (1995) test with a 5% significance using quadratic loss for the two exponential loading models: weak segmented (NS4) and strong segmented (NS4S). The first panel presents the results for the comparison between the segmented models with the Dynamic Svensson Model. The second panel presents the same comparison with the affine model. The first number in each cell represents the percentage of times that the model produces statistically significant better forecasts than the compared model. The second number is the percentage of times that the compared model produces statistically significant better forecasts than the model. All 143 out-of-sample forecasting exercises are considered.

A surprising result is obtained by the affine Gaussian model, which achieves, for all maturities, its smallest RMSEs across the whole sample, precisely in the forecasting exercises performed with data from the zero-lower-bound period. We have no explicit explanation for this better performance. The affine Gaussian model is a multivariate Vasicek model

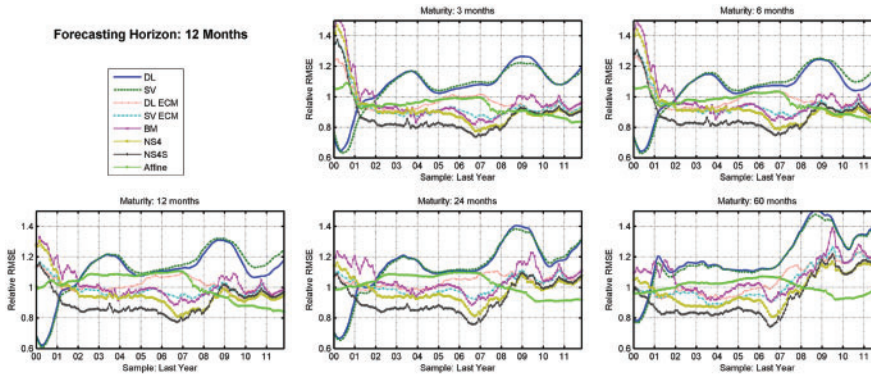

Figure 6. Time series of RMSE for 12 month forecasting horizon.

Notes: These graphs present the mean of the RMSE relative to the RW for a 12 month forecasting horizon for each analyzed model. The mean is computed for each forecasting exercise, defined by the date on which the last prediction is done (represented in the graphs by the *x*-axis). A forecasting exercise consists of a sequence of forecasts constructed such that when a new observation is available, the in-sample window is shifted in time one month ahead, the model is reestimated and a new forecast is computed. In each forecasting exercise, we perform 84 monthly out-of-sample forecasts. Models analyzed are Diebold and Li (DL), Svensson (DSM), Diebold and Li with ECM dynamics (DL ECM), Svensson with ECM dynamics (DSM ECM), Bowsher and Meeks (2008) (BM), affine Gaussian, weak segmented (NS4), and strong segmented (NS4S).

whose exponential loadings are determined by no-arbitrage conditions, and the only difference between this model and others like Diebold and Li and Svensson is the way the exponential loadings are chosen by no-arbitrage. No-arbitrage makes the decay parameters of exponential loadings vary across different forecasting exercises, perhaps allowing for a better adaptation of movements to low interest rates. Although Joslin, Singleton, and Zhu (2011) report that no-arbitrage does not improve forecasting power, the particular combination of no-arbitrage with a zero-lower-bound restriction might generate dynamic loadings that are more sensitive to the interest rate behavior along the zero-lower-bound period than the other models.

We conclude this section summarizing our main empirical findings. First, ECM dynamics for latent factors play a crucial role in improving forecasting results for all models considered. Second, segmentation strongly improves long-horizon forecasts for the whole term structure. In addition, exponential versions of splines improve forecasting results when compared with polynomial versions.

4 Conclusion

Motivated by the Vayanos and Vila (2009) formalized version of the preferred habitat theory of the term structure, we propose a family of segmented term structure models with local movements to forecast future yields. We show that the flexibility in capturing yields' idiosyncratic local shocks, possibly related to demand factors, helps these segmented models to better identify yield curve cross-section and dynamics.

Based on a series of out-of-sample forecasting exercises with U.S. Treasury yields, we identify that segmented models produce significantly smaller RMSEs than those produced

by the Random Walk and by some other established term structure benchmark models. In light of Diebold and Mariano tests, we also find that the segmented models consistently produce statistically significant better results than the competing models.

Due to the relatively bad forecasting performance of the pervasive AR process in latent factor dynamics, and to the well-known unit root behavior of yields, we adopt ECM as an alternative for factors' dynamics. In this context, for all models included in our study (Diebold and Li, Dynamic Svensson, Bowsher and Meeks, and two types of exponential segmented models), ECM dynamics consistently provide better statistically significant out-of-sample forecasting results, when compared with the corresponding AR dynamics. This should be welcomed as an interesting result for the fixed-income literature on its own.

Adopting ECM dynamics with exponential splines to segment the term structure of interest rates produces strong out-of-sample forecasts. These favorable forecasts are robust across different data sets (FRED, smoothed and unsmoothed Fama–Bliss) and across time, losing power only during the recent zero-lower-bound period.

Supplementary Data

Supplementary data are available online at http://jfin.oxfordjournals.org.

References

Adrian, T., E. Moench, and H. S. Shin. 2010. "Financial Intermediation, Asset Prices, and Macroeconomic Dynamics." Working paper, New York FED.

Almeida, C., R. Gomes, A. Leite, A. Simonsen, and J. Vicente. 2009. Does Curvature Enhance Forecasting?. *International Journal of Theoretical and Applied Finance* 12: 1171–1196.

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.

Carriero, A., G. Kapetanios, and M. Marcellino. 2012. Forecasting Government Bond Yields with Large Bayesian Vector Autoregressions. *Journal of Banking and Finance* 36**:** 2026–2047.

Cochrane, J., and M. Piazzesi. 2005. Bond Risk Premia. *American Economic Review* 95: 138–160.

D'Amico, S., and T. B. King. 2013. Flow and Stock Effects of Large-scale Treasury Purchases: Evidence on the Importance of Local Supply. *Journal of Financial Economics* 108: 425–448.

De Pooter, M. 2007. "Examining the Nelson-Siegel Class of Term Structure Models." Working paper, Tinbergen Institute Discussion Paper.

Diebold, F. X., and C. Li. 2006. Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics* 130**:** 337–364.

Diebold, F. X., and R. S. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13**:** 253–263.

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

Duffee, G. 1996. Idiosyncratic Variation of Treasury Bill Yields. *Journal of Finance* 51**:** 527–551.

Duffee, G. 2002. Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance* 57**:** 405–443.

Duffee, G. 2011. Information in (and not in) the Term Structure. *Review of Financial Studies* 24: 2895–2934.

Duffie, D. 2010. Presidential Address: Asset Price Dynamics with Slow-Moving Capital. *Journal of Finance* 55**:** 1237–1267.

Fontaine, J.-S., and R. Garcia. 2012. Bond Liquidity Premia. *Review of Financial Studies* 25: 1207–1254.

Giacomini, R., and B. Rossi. 2010. Forecast Comparisons in Unstable Environments. *Journal of Applied Econometrics* 25**:** 595–620.

Giacomini, R., and H. White. 2006. Tests of Conditional Predictive Ability. *Econometrica* 76: 1545–1578.

Greenwood, R., and D. Vayanos. 2010. Price Pressure in the Government Bond Market. *American Economic Review, P&P* 100: 585–590.

Greenwood, R., and D. Vayanos. 2014. Bond Supply and Excess Bond Returns. *Review of Financial Studies* 27**:** 663–713.

Gürkaynak, R. S., and J. Wright. 2012. Macroeconomics and the Term Structure. *Journal of Economic Literature* 50: 331–367.

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.

Hamilton, J., and J. C. Wu. 2012. "The Effectiveness of Alternative Monetary Policy Tools in a Zero Lower Bound Environment." *Journal of Money Credit and Banking* 44: 3–46.

Hansen, P. R., and A. Timmermann. 2012. "Choice of Sample Split in Out-of-Sample Forecast Evaluation." Working paper, European University Institute and UCSD.

Joslin, S., K. Singleton, and H. Zhou. 2011. A New Perspective on Gaussian Dynamic Term Structure Models. *Review of Financial Studies* 24**:** 926–970.

Jungbaker, B., S. J. Koopman, and M. van der Wel. 2014. Smooth Dynamic Factor Analysis with Application to the U.S. Term Structure of Interest Rates. *Journal of Applied Econometrics* 29: 65–90.

Kaminska, I., D. Vayanos, and G. Zinna. 2011. "Preferred-habitat Investors and the US Term Structure of Real Rates." Working paper, London School of Economics.

Kim, D., and K. Singleton. 2013. Term Structure Models and the Zero Bound: An Empirical Investigation of Japanese Yields. *Journal of Econometrics* 170**:** 32–49.

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.

Koopman, S. J., and M. van der Wel. 2013. Forecasting the U.S. Term Structure of Interest Rates using a Macroeconomic Smooth Dynamic Factor Model. *International Journal of Forecasting* 29: 676–694.

Krishnamurthy, A. 2002. The Bond/Old Bond Spread. *Journal of Financial Economics* 66**:** 463–506.

Krishnamurthy, A., and A. Vissing-Jorgensen. 2012. The Aggregate Demand for Treasury Debt. *Journal of Political Economy* 120**:** 233–267.

Li, C., and M. Wei. 2013. Term Structure Modeling with Supply Factors and the Federal Reserve's Large-Scale Asset Purchase Programs. *International Journal of Central Banking* 9: 3–39.

Litterman, R., and J. A. Scheinkman. 1991. Common Factors Affecting Bond Returns. *Journal of Fixed Income* 1**:** 54–61.

McCraken, M. W. 2007. Asymptotics for Out-of-Sample Tests of Granger Causality. *Journal of Econometrics* 140: 719–752.

Modigliani, F., and R. Sutch. 1966. Innovations in Interest-Rate Policy. *American Economic Review* 56: 178–197.

Nelson, C., and A. Siegel. 1987. Parsimonious Modeling of Yield Curves. *Journal of Business* 60: 473–489.

Rossi, B., and A. Inoue. 2012. Out-of-Sample Forecast Tests Robust to the Window Size Choice. *Journal of Business & Economic Statistics* 30**:** 432–453.

Svensson, L. 1994. "Estimating and Interpreting Forward Interest Rates: Sweden 1992 - 1994." Working paper, National Bureau of Economic Research.

Vayanos, D., and J.-L. Vila. 2009. "A Preferred-Habitat Model of the Term Structure of Interest Rates." Working paper, London School of Economics.

Zinna, G. 2016. Price Pressures on UK Real Rates: An Empirical Investigation. *Review of Finance* 20: 1587–1630.