A Hybrid Spline-Based Parametric Model for the Real Yield Curve

April 6, 2017

Abstract

In different markets, the short end of the real yield curve is influenced by seasonalities of the price index that imply a lack of smoothness in this segment. We propose a segmented model to deal with these seasonalities appearing in real yield curves. Borrowing from the flexibility of spline models, a B-spline function is used to fit the short end of the yield curve, while the medium and the long end are captured by a parsimonious parametric four-factor exponential model. We illustrate the benefits of the proposed term structure model by estimating real yield curves in one of the biggest government index-linked bond markets in the world. Our model is simultaneously able to fit the yield curve and to provide unbiased Value at Risk estimates for different portfolios of bonds negotiated in this market.

Keywords: Price index, Spline Models, Exponential Term Structure Models, Curve Fitting, Risk management.

JEL Code: C51, C58, G17.
1 Introduction

Several methods to extract the term structure of interest rates have been proposed over the past decades. Such effort is justified by the importance that this curve has for market participants, regulators, and studies in the fields of finance and macroeconometrics. Naturally, the differences between methods are related to the distinction in purposes for using the extracted yield curve.

A technical report produced by the Bank for International Settlements (BIS (2005)) presents the main term structure models adopted by major Central Banks, classifying them in two groups: Parametric and spline-based models. Parametric models fit the yield curve in a parsimonious way. They smooth idiosyncrasies\(^1\) of some bonds or curve regions in favor of a common global behavior. Those methods are typically used in macroeconomic studies,\(^2\) in which smoothness and the ability to capture a common trend are as important as model accuracy. As representatives of the parametric class, we have the Nelson and Siegel (1987) model\(^3\) and its extension proposed by Svensson (1994).

Unlike parametric models, the spline-based ones do not impose parsimonious functional forms along the entire length of maturity. Instead, they are constituted by several low-order polynomials which are smoothly linked over the range of maturities. Such constitution naturally implies: using a higher number of parameters than parametric models, and less smooth yield curves. On the other hand, spline-based models are more accurate and have higher ability to fit idiosyncrasies. As stated by Gurkaynak et al. (2007), a trader looking for pricing anomalies can be quite concerned about how a particular bond is priced relative to the others nearby. In this sense, less smoothness and a greater accuracy in curve regions with many idiosyncrasies may be desirable. Within the spline-based class, we can highlight the initial works of McCulloch (1971), McCulloch (1975) and Vasicek and Fong (1982) and the penalized spline\(^4\) models of Fisher et al. (1995), Waggoner (1997) and Jarrow et al. (2004).\(^5\)

\(^{1}\)Such as liquidity premium, inflation risk premium for index-linked bonds and hedging demand effects.

\(^{2}\)Among which we can mention Pooter et al. (2014), Gurkaynak et al. (2010) and Diebold et al. (2006). The Swiss Central Bank, although using a spline model in daily basis, also estimates a parametric one exclusively for macroeconomic studies purposes (BIS (2005)).

\(^{3}\)Dynamic versions of the Nelson and Siegel (1987) model are presented in Diebold and Li (2006) and Christensen et al. (2011).

\(^{4}\)For a comprehensive survey about the smooth spline technique and its properties, see Wahba (1990) and Hastie and Tibshirani (1990).

\(^{5}\)Jarrow et al. (2004) generalize the Fisher et al. (1995) model to corporate debt by modeling the corporate term structure as a term structure of government bond yields plus a parametric spread.
A drawback of the parametric approach is the instability of the solution in response to idiosyncratic shocks. As illustrated by Anderson and Sleath (2001), a small perturbation in the price of just one bond can significantly affect an estimate of the entire curve. On the other hand, the spline-based models avoid this undesirable characteristic. Because each polynomial that constitutes the spline can move with certain independence of the others, the shocks in any specific curve region can be accommodated within that region.\footnote{Almeida et al. (2017) propose segmented exponential-based spline models to forecast future yields. They show that these models’ superior ability in forecasting the short end of the yield curve is directly related to their success in capturing idiosyncratic shocks.}

When modeling the real yield curve, the instability of parametric models becomes an undesirable issue. Ejsing et al. (2007) show that the short end of the term structure is subject to the erratic and seasonal behavior of the price index. Such behavior does not affect its medium and long ends, since the seasonality effect diminishes over maturities. The erratic pattern allied to the high degree of smoothness and instability of the parametric models implies at least two main problems. First, these models are not able to adjust the idiosyncrasies of the short end of the curve. Second, when attempting to fit this short end, the model can distort estimates of the long end of the curve.

Trying to circumvent these problems, Ejsing et al. (2007) suggest a methodology in which the seasonality effect from bond prices is extracted before proceeding to curve estimation using the Nelson and Siegel model. The authors emphasize that their method is appropriate for medium term monetary policy analysis. However, the methodology is not suitable for pricing index-linked bonds, since pricing these instruments requires taking into account the inflation accrual, which is directly affected by seasonality effects.

Despite their instability issues, parametric models have advantages over spline-based ones when performing risk management tasks in interest rate markets. Their reduced number of parameters is crucial to simplify the estimation of Value at Risk (VaR) methodologies. Moreover, their functional forms have useful economic interpretations to build interest rate stress scenarios.\footnote{It is possible to build scenarios by applying shocks to the parameters associated with the level, slope, and curvature of the yield curve. These are the main movements of the term structure as pointed out by Litterman and Scheinkman (1991).}

This paper provides a unique approach capable of dealing with real yield curve estimation problems, while simultaneously meeting the main risk management demands appearing in fixed income markets. We propose a segmented model for the real yield
curve, considering splines to fit the short end, while medium and long ends are captured by a parametric Svensson model. The proposed model is useful for pricing purposes and monetary policy analysis, avoiding an additional step of seasonality component extraction proposed in other methodologies available to estimate the real yield curve. And, at the same time, the parsimonious parametric form for the medium and long ends of the yield curve enables the construction of economically interpretable risk controls via VaR and stress scenarios.

The main challenge of the proposed approach consists in solving a nonlinear optimization problem\(^8\) where two distinct models are jointly estimated. In the estimation process, restrictions must be imposed to guarantee the smoothness of the segmented yield curve. We develop a procedure in which the restricted problem is rewritten as an unconstrained problem that satisfies all the restrictions of the first one. The first step consists in specifying the full model as a spline function. After that, for the medium and long terms, we project the functional forms of the parametric model into the spline basis. Therefore, our method uses a single basis with the desired segmentation, allowing a direct and unrestricted estimation of the parameters associated with it.

We apply our model to the Brazilian government index-linked bond market, covering a period from January, 2009 to May, 2014. According to data extracted from Barclays Universal Government Inflation-Linked Bond Index and (UGILB), and Barclays Emerging Markets Government Inflation Linked Bond Index (EMGILB), in January of 2014, the Brazilian market was one the five biggest in the world and the largest among emerging countries, totaling US$ 220 billion.\(^9\) In addition to its importance, the presence of high inflation rates and strong seasonal patterns, makes its term structure one of the most challenging to be modeled. During the analyzed period, within each year, the annualized monthly inflation rate reaches values between 1% and 10%, implying huge variations on the short-term real interest rates. In the empirical section, we show that our model is not only able to fit the real yield curve as a whole, but also capable of providing unbiased estimates of VaR for portfolios built within the aforementioned market.

Although our focus is on the real yield curve, this model can also be useful to fit the nominal term structure. It is well known that money market rates can be a useful source of data to better fit the short end of the yield curve. Anderson and Sleath (2001) show

\(^8\)This comes from the non-linearity of the relationship between bond prices and the term structure.  
\(^9\)For comparison purposes, the TIPS market had US$ 700 billion outstanding in January 2014.
that using General Collateral (GC) repo rates in a nonparametric model can significantly improve estimates of the short end of the British term structure of interest rates. However, they do not recommended using the same procedure to improve Svensson’s estimates, because fitted rates at the very long end can be distorted. In such a case, our segmented approach could be applied to control the effects of including money market rates on Svensson’s estimates, avoiding the distortion of the long end of the yield curve.

The paper proceeds as follows. Section 2 presents the two classes of models as well as the proposed segmented design. This section also contains a numerical example, which illustrates how we build the segmented basis. Section 3 discusses our bond data and the Brazilian government index-linked bond market. Section 4 presents the results of the two empirical exercises. The first compares the in-sample and out-of-sample fitting abilities across different static term structure models. The second consists in VaR estimation for portfolios exposed to real term structure risk. Section 5 concludes.

2 The Model

This section introduces the case in which the term structure is expressed as a function of a general basis. Then, we define the Svensson and the B-spline basis representing respectively the parametric and the spline-based approaches. Merging these two approaches, we describe the methodology used to build the mixed basis, which defines our segmented model. Finally, a numerical example explaining how to construct such a basis is provided.

2.1 The general case

Consider a domain \([0,T]\) in which the term structure of interest rates \(z(\cdot)\) is defined. Let \(N\) be the number of bonds, \(p_i\) be the price of bond \(i\), \(c_i = (c_{i1}, c_{i2}, \ldots, c_{im_i})\) be the vector of cash flows with \(m_i\) payments for bond \(i\), and \(\tau_i = (\tau_{i1}, \tau_{i2}, \ldots, \tau_{im_i})\) be the vector of maturities of those payments. In this context, the relationship between \(p_i\) and \(z(\cdot)\) is given by:

\[
p_i = \sum_{j=1}^{m_i} c_{ij} \exp(-\tau_{ij} z(\tau_{ij})), i = 1, 2, \ldots, N.
\] (1)

In order to propose a model for the term structure, consider the existence of a certain
basis $\delta(.) = (\delta_1(.), \delta_2(.), \ldots, \delta_\kappa(.))$, with dimension $\kappa$, on which $z(.)$ is approximated by a linear combination. From the previous equation, we can express the price of a bond as the following function of $\delta$:

$$p_i = \sum_{j=1}^{m_i} c_{ij} \exp(-\tau_{ij} \delta(\tau_{ij}) \beta) + \epsilon_i = c_i \exp(-D_i \delta(\tau_i) \beta) + \epsilon_i,$$

where $\epsilon_i$ is an error term generated by the approximation of $z(.)$ by the basis, $D_i$ is a diagonal matrix in which the main diagonal is given by vector $\tau_i$ and

$$\delta(\tau_i) = (\delta(\tau_{i1})^T, \delta(\tau_{i2})^T, \ldots, \delta(\tau_{im_i})^T)^T,$$

is a $m_i \times \kappa$ matrix that results from applying $\delta$ to each element of $\tau_i$. Hence, the theoretical price of bond $i$ produced by this model is given by:

$$\pi_i = c_i \exp(-D_i \delta(\tau_i) \beta).$$  \hspace{1cm} (2)

Estimates of the parameters $\beta$ may be obtained by solving:

$$\min_{\beta(\alpha)} \left( \sum_{i=1}^{N} \left( \frac{p_i - \pi_i(\beta(\alpha))}{w_i} \right)^2 + \int_0^T \alpha(\tau) h''(\tau)^2 d\tau \right),$$ \hspace{1cm} (3)

where $w_i$ is the duration of bond $i$, $h(\tau) = \delta(\tau) \beta$ and $\alpha(.)$ is a penalty function such that $\alpha : [0, T] \rightarrow \mathbb{R}^+$. The first part of the objective function in (3) is a sum of quadratic pricing errors weighted by the inverse of each bond’s duration.\(^{10}\) It represents the component related to model fitting. The second part of the objective function penalizes excess variability in the estimated term structure. This is the component associated with model smoothness.

For the parametric and standard spline models, $\alpha(.)$ is simply null. In this case, the degree of smoothing is determined a priori by the chosen basis $\delta(.)$. On the other hand, on penalized spline models, $\alpha(.)$ can take any positive value. In this case, the solution of problem (3) must consider a trade-off between fitting and smoothness, with the importance of the latter determined by the magnitude of $\alpha(.)$. In the extreme case, when $\alpha \rightarrow \infty$, any minimal degree of curvature gets an infinite penalty. In such a case,

\(^{10}\)Weighting pricing errors by the inverse of bond durations, and using a first order Taylor approximation, these errors can be approximately converted into homoskedastic yield errors; see Vasicek and Fong (1982).
the solution will converge to the linear solution acceptable by the basis.\textsuperscript{11}

The functional form adopted for $\alpha(.)$ differs across different penalized spline models. Fisher et al. (1995) let $\alpha(.)$ be constant over the entire domain. The value of this constant is determined by the Generalized Cross Validation (GCV) procedure, explained later in the current paper. In Waggoner (1997), $\alpha(.)$ is a function that assumes different values over the range of maturities. Due to such variability of the penalty function, this method is denominated VRP, for Variable Roughness Penalty. The main argument for having a variable penalty function across maturities is the stylized fact that fluctuations in the long end of the curve are much less volatile than those appearing in the short end. Therefore, imposing the same degree of smoothing over the entire curve will not be appropriate. The parameters appearing in $\alpha(.)$ are estimated using an out-of-sample procedure.

The component associated with model smoothness can be written as:

$$
\int_0^T \alpha(\tau) h''(\tau)^2 d\tau = \beta^T \left( \int_0^T \alpha(\tau) \delta''(\tau)^T \delta''(\tau) d\tau \right) \beta = \beta^T H(\alpha) \beta,
$$

where $H$ is a matrix $\kappa \times \kappa$, in which element $H_{i,j}$ is defined as $\int_0^T \alpha(\tau) \delta''_i(\tau) \delta''_j(\tau) d\tau$. Using the last equation, we can rewrite the minimization problem (3), for a given $\alpha$, as follows:

$$
\min_{\beta(\alpha)} \left[ (P - \Pi(\beta(\alpha)))^T W (P - \Pi(\beta(\alpha))) + \beta(\alpha)^T H(\alpha) \beta(\alpha) \right],
$$

where $P$ is a $N \times 1$ vector of bond prices $\{p_i\}_{i=1,...,N}$, $\Pi(\beta)$ is the corresponding vector of modeled prices, $\{\pi_i(\beta)\}_{i=1,...,N}$, and $W$ is a $N \times N$ diagonal matrix with $W_{ii} = (1/w_i)^2$, $i = 1, ..., N$.

The minimization in (4) can be solved as a nonlinear least squares problem. Fisher et al. (1995) linearize $\Pi(\beta)$ around an initial guess $\beta^0$:

$$
\Pi(\beta) \approx \Pi(\beta^0) + \frac{\partial \Pi(\beta)}{\partial \beta^T} (\beta - \beta^0),
$$

where $\frac{\partial \Pi(\beta)}{\partial \beta^T} = -\pi_i(\beta)c_i D_i \delta(\tau_i)$. Defining $X(\beta_0) = \frac{\partial \Pi(\beta)}{\partial \beta^T} \big|_{\beta = \beta^0}$ and $Y(\beta^0) = P - \Pi(\beta^0) + X(\beta^0)\beta^0$, and rearranging (4) yields:

\textsuperscript{11}In the cubic spline case, the solution converges to the least-squares line.
\[
\min_{\beta(\alpha)} \left[ (Y(\beta^0) - X(\beta^0)\beta(\alpha))^\top W (Y(\beta^0) - X(\beta^0)\beta(\alpha)) + \beta(\alpha)^\top H(\alpha)\beta(\alpha) \right]
\]  

(6)

For a given \( \alpha \), the minimizer for (6) is:

\[
\beta^1(\alpha) = \left( X(\beta^0)^\top W X(\beta^0) + H(\alpha) \right)^{-1} X(\beta^0)^\top W Y(\beta^0),
\]

where \( \beta^1(\alpha) \) is an updated \( \beta^0 \). One can use \( \beta^1(\alpha) \) as an initial guess for the next iteration, obtaining \( \beta^2(\alpha) \), and continue iterating until convergence. The solution is the following fixed-point:

\[
\beta^*(\alpha) = \left( X(\beta^*(\alpha))^\top W X(\beta^*(\alpha)) + H(\alpha) \right)^{-1} X(\beta^*(\alpha))^\top W Y(\beta^*(\alpha)).
\]

(7)

The GCV procedure, as presented by Hastie and Tibshirani (1990), is originally defined only for linear problems. Fisher et al. (1995) generalize it to the non-linear case, using the approximation in (5). The GCV procedure consists in determining the value of \( \alpha \), which minimizes the following ratio:

\[
\gamma(\alpha) := \frac{((I - S(\alpha))Y(\beta^*(\alpha)))^\top((I - S(\alpha))Y(\beta^*(\alpha)))}{\left( N - tr(S(\alpha)) \right)^2},
\]

(8)

where \( S(\alpha) \) is the smoothing matrix defined as:

\[
S(\alpha) = X(\beta^*(\alpha))^\top WX(\beta^*(\alpha)) + H(\alpha)^{-1} X(\beta^*(\alpha))^\top W.
\]

(9)

Since \( S(\alpha)Y(\beta^*) \) is a vector of \( Y \)-fitted values,\(^{12}\) the numerator of the GCV ratio represents the model quadratic error, a well-known measure of goodness-of-fit. On the other hand, the smoothing matrix trace is a usual measure of the effective number of parameters (see Hastie and Tibshirani (1990)).\(^{13}\) As a consequence, the denominator of \( \gamma(\alpha) \) is the squared effective degrees of freedom. Therefore, the GCV ratio directly corresponds to a trade-off between goodness-of-fit and parsimony. To find the \( \alpha \) value,

---

\(^{12}\) When modeling the discount function, we end up in the linear case in which \( Y \) is the price vector, and \( S(\alpha)Y(\beta^*) \) is the vector of fitted prices.

\(^{13}\) Note that in the case of a least-squares regression, \( S(\alpha) \) is the projection matrix, and its trace equals the number of linearly independent predictors in the model.
we first create a grid of possible values. After that, for each value in the grid, we solve problem (6), obtaining $\beta^*(\alpha)$, and calculate the GCV ratio in (8). The optimal value, $\alpha^*$, is that one which minimizes the aforementioned ratio. The overall solution is given by $\beta^*(\alpha^*)$.

2.2 Svensson’s basis

The Svensson’s basis is the vector $\psi = \{\psi_k\}_{k=1}^4$, which is constituted by four functions: level, slope and two curvatures. For given strictly positive and distinct $\lambda_1$ and $\lambda_2$, the basis is defined as:

$$
\psi_k(\tau) = \begin{cases} 
1, & \text{if } k = 1 \\
S(\lambda_1, \tau) = (1 - \exp(-\lambda_1 \tau))/\lambda_1 \tau, & \text{if } k = 2 \\
C(\lambda_1, \tau) = S(\tau, \lambda_1) - \exp(-\lambda_1 \tau), & \text{if } k = 3 \\
C(\lambda_2, \tau) = S(\tau, \lambda_2) - \exp(-\lambda_2 \tau), & \text{if } k = 4 
\end{cases} 
$$

The Nelson and Siegel basis is obtained by eliminating the second curvature $\psi_4(\tau)$ of the Svensson’s basis. To specify and implement the parametric model, we have to substitute the generic basis $\delta(.)$, appearing in the general case, by the specific parametric basis $\psi(.)$.

2.3 B-spline basis

Let $\{s_k\}_{k=1}^K$ denote the set of knot points, with $s_k < s_{k+1}$, $s_1 = 0$, and $s_K = T$. The spline of degree $p$ is defined as:

$$
S(\tau) = \sum_{j=0}^{p} \theta_j \tau^j + \sum_{k=1}^{K-1} \eta_k |\tau - s_k|^3,
$$

for some constants $\{\theta\}_{j=1}^p$ and $\{\eta\}_{j=1}^{K-1}$ and for $\tau \in [0, T]$. As in most studies in the literature, we use a cubic spline, obtained with $p = 3$.

A stable numerical parameterization of a spline is provided by a B-spline basis. Let $\{d_k\}_{k=1}^{K+6}$ denote the augmented set of knot points, with $d_1 = d_2 = d_3 = s_1$, $d_{K+4} = d_{K+5} = d_{K+6} = s_K$ and $d_{k+3} = s_k$ for $1 \leq k \leq K$. A B-spline of degree $p$ is defined by the following recursion:
\[
\phi^p_k(\tau) = \frac{\phi_k^{-1}(\tau - d_k)}{(d_{k+p} - d_k)} + \frac{\phi_{k+1}^{-1}(\tau)(d_{k+p+1} - \tau)}{d_{k+p+1} - d_{k+1}},
\]

for \( k \in \{1, \ldots, K + 5 - p \} \) and \( \tau \in [0, T] \), with

\[
\phi_k^0(\tau) = \begin{cases} 
1 & \text{if } d_k \leq \tau < d_{k+1} \\
0 & \text{c.c.}
\end{cases}
\]

A B-spline basis of degree 3 is a vector of \( \kappa = K + 2 \) cubic B-splines given by \( \phi^3(\tau) := (\phi_1^3(\tau), \ldots, \phi_{\kappa}^3(\tau))_{1 \times \kappa} \). To simplify notation, let \( \phi_k(\tau) := \phi_k^3(\tau) \) and \( \phi(\tau) := \phi^3(\tau) \).

As shown by De Boor (2001), any cubic spline can be represented as a linear combination of the B-splines that constitute the basis \( \phi(.) \). That is: \( S(\tau) = \sum_{k=1}^{\kappa} \beta_k \phi_k(\tau) \), with \( \beta_k \in \mathbb{R} \ \forall k \in [1, \kappa] \).

To specify and implement a cubic B-spline model, we have to substitute, in the general case, the generic basis \( \delta(.) \) by the specific B-spline basis \( \phi(.) \).

### 2.4 Mixed basis

Among the existing properties of the B-spline basis,\(^{14}\) three of them are particularly useful to construct the mixed basis. (i) Partition of unity: for any \( \tau \in [0, T] \), \( \sum_{k=1}^{\kappa} \phi_k(\tau) = 1 \).

(ii) Local knots: the \( k \)th B-spline \( \phi_k \) depends only on the knots \( d_k, d_{k+1}, \ldots, d_{k+4} \). (iii) Local support: the interval \([d_k, d_{k+4}]\) is denominated support of \( \phi_k \). For any \( k \in (1, \kappa) \), if \( \tau \in (d_k, d_{k+4}) \), then \( \phi_k(\tau) > 0 \), otherwise, \( \phi_k(\tau) = 0 \).

Using the last two properties, if we split the domain \([0, T]\) into two intervals, \([0, \bar{\tau}]\) and \((\bar{\tau}, T]\), only a subset of the \( \kappa \) B-splines composing the basis will take non-zero values over the last interval. Formally, let \( \bar{\tau} \in (0, T) \) and \( d_{\tilde{k}} \) be the greatest knot in \((0, \bar{\tau}]\), which implies that \( \tilde{k} \in \{5, 6, \ldots, \kappa - 1\} \). Therefore, the set formed by B-splines that takes non-zero values in the range \((\bar{\tau}, T]\) is given by:

\[
\bar{\phi} := (\phi_{\kappa-3}, \phi_{\kappa-2}, \ldots, \phi_{\kappa}),
\]

containing \( \bar{\kappa} = \kappa - \tilde{k} + 4 \) elements. We denominate \( \bar{\phi} \) as the active B-spline basis, with dimension \( \bar{\kappa} \), over \((\bar{\tau}, T]\).

To construct the mixed basis, we start with a B-spline basis, \( \phi \), defined over the domain

---

\(^{14}\)See De Boor (2001) and Lyche and Morken (2011).
[0, T] and composed by \( \kappa \) elements. Then, we determine the range, \((\bar{\tau}, T]\), over which the term structure will be modeled using the Svensson’s basis. Finally, we approximate the smooth functions that constitute this basis using the active B-spline basis over the above-mentioned interval, \( \tilde{\phi} \). Therefore, this approximation uses just a subset of the \( \kappa \) B-splines, leaving the others free. A natural question that arises now is: how many knot points and, consequently, how many B-splines should be placed above the maturity \( \bar{\tau} \) to make the specified procedure to work properly?

A well-known result is that when a smooth function is approximated by a polynomial of degree \( p \), on an interval of length \( h \), the error is bounded by \( Ch^{p+1} \), where \( C \) is a constant that is independent of \( h \). Since a spline is formed by several polynomials that are linked smoothly together, Lyche and Morken (2011) show that it is possible to derive local error bounds for the spline approximation. Those bounds are functions of the spacing between the knot points and the degree of the spline. In the appendix, we present such bounds and the algorithm used for determining the knots location. For now, assume that \( \tilde{\phi} \) is formed with enough knot points such that the approximation error of the Svensson’s basis is within a certain tolerance.

Considering that the parametric approach described above will be active only for \( \tau > \bar{\tau} \), we redefine its basis to be:

\[
\tilde{\psi}_k(\tau) = \begin{cases} 
1, & \text{if } k = 1 \\
S(\lambda_1, \tau) = (1 - exp(-\lambda_1(\tau - \bar{\tau}))) / (\lambda_1(\tau - \bar{\tau})) , & \text{if } k = 2 \\
C(\lambda_1, \tau) = S(\tau, \lambda_1) - exp(-\lambda_1(\tau - \bar{\tau})) , & \text{if } k = 3 \\
C(\lambda_2, \tau) = S(\tau, \lambda_2) - exp(-\lambda_2(\tau - \bar{\tau})) , & \text{if } k = 4 
\end{cases}
\] (13)

The approximation of \( \tilde{\psi}_k(\cdot) \) by \( \tilde{\phi}(\cdot) \) over the interval \((\bar{\tau}, T]\) is given by:

\[
\tilde{\psi}_k(\tau) = \tilde{\phi}(\tau)\Gamma_k + \epsilon_k(\tau), k \in \{1, 2, 3, 4\}
\]

where \( \epsilon_k \) is the error term and \( \Gamma_k \) is a \( \bar{\kappa} \times 1 \) vector of parameters.

By the partition of unity property, making \( \Gamma_1 \) be a vector with all entries one, we get \( \tilde{\psi}_1(\tau) = \tilde{\phi}(\tau)\Gamma_1 \). Thus, \( \tilde{\psi}_1(\cdot) \) is perfectly replicated by \( \tilde{\phi} \). In turn, for each other function that constitute the basis \( \tilde{\psi} \), the parameter vector \( \Gamma_k \) is obtained by regression splines:
\[ \Gamma_k = (\bar{\Phi}^T \bar{\Phi})^{-1} \bar{\Phi}^T \bar{\Psi}_k, \]

where \( \bar{\Phi} \) and \( \bar{\Psi}_k \) result from applying \( \bar{\phi}(\cdot) \) and \( \bar{\psi}_k(\cdot) \) to each element in the interval \( (\bar{\tau}, T] \). For those cases, \( \epsilon_k \) is not null, and it depends on the spacing between the knot points.

Defining the \( k \times 4 \) matrix \( \Gamma := (\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4) \), the approximation of the Svensson’s basis \( \bar{\psi}(\cdot) \), defined over the interval \( (\bar{\tau}, T] \), is provided by:

\[ \hat{\psi}(\tau) = \bar{\phi}(\tau) \Gamma. \] (14)

Given (12) and (14), the mixed basis \( \Psi(\cdot) \), defined over the domain \([0, T]\), is given by:

\[ \Psi(\tau) = \begin{pmatrix} (\phi_1(\tau), \ldots, \phi_{k-4}(\tau))_{1 \times (k-4)} & \hat{\psi}(\tau) \end{pmatrix} = \begin{pmatrix} \phi_1(\tau), \ldots, \phi_{k-4}(\tau) \\ \hat{\psi}(\tau) \end{pmatrix} \begin{bmatrix} I_{(k-4)} & 0_{(k-4) \times 4} \\ 0_{k \times (k-4)} & \Gamma \end{bmatrix}_{k \times k} \]

\[ = \phi(\tau)A, \] (15)

where \( I_j \) is the identity matrix with dimension \( j \), \( 0_{i \times j} \) is a \( i \times j \) null matrix, and

\[ A = \begin{bmatrix} I_{(k-4)} & 0_{(k-4) \times 4} \\ 0_{k \times (k-4)} & \Gamma \end{bmatrix}_{k \times k}. \]

The \( \Psi \) basis has dimension \( \bar{k} \), with four functions destined to approximate the Svensson’s basis, and \( \bar{k} - 4 \) representing free B-splines.

By (15), the mixed basis, \( \Psi(\cdot) \), is a linear function of the B-spline basis, \( \phi(\cdot) \), determined by matrix \( A \). In short, this matrix determines the dimension of \( \Psi(\cdot) \), the number of free B-splines, and the starting point of the domain from which the mixed basis approximates the Svensson one. Therefore, matrix \( A \) restricts the B-spline basis in order to build the desired mixed basis.

Looking from another perspective, it is also possible to find a matrix \( M \) such that the product between \( M \) and the vector of parameters associated with the B-spline basis makes the models that use the B-spline basis and the mixed basis equivalent to each other. First, using \( \phi(\cdot) \) and \( \Psi(\cdot) \) in the pricing equation (2), we get:
\[ \pi^S_p(\beta^S_p) = c_i \exp\left( - D_i \phi(\tau_i) \beta^S_p \right) \]  
(16)

\[ \pi^M_i(\beta^M) = c_i \exp\left( - D_i \Psi(\tau_i) \beta^M \right) \]

\[ = c_i \exp\left( - D_i \phi(\tau_i) A \beta^M \right) \]

\[ = \pi^S_p(\beta^S_p = A \beta^M), \]  
(18)

where, \( \pi^M_i(\pi^S_p) \) is the modeled price for the bond \( i \) obtained using the mixed (B-spline) basis and the associated parameter vector \( \beta^M (\beta^S_p) \). By equation (18), the two bases considered will provide the same results when \( \beta^S_p = A \beta^M \), which is equivalent to:\(^{15}\)

\[ \beta^M = (A^\top A)^{-1} A^\top \beta^S_p. \]

(19)

Substituting (19) in (17):

\[ \pi^M_i(\beta^M(\beta^S_p)) = c_i \exp\left( - D_i \phi(\tau_i) M \beta^S_p \right), \]

(20)

where

\[ M = A (A^\top A)^{-1} A^\top \]

\[ = \begin{bmatrix} I_{(k-4)} & 0_{(k-4) \times \kappa} \\ 0_{\kappa \times (k-4)} & \Gamma (\Gamma^\top \Gamma)^{-1} \Gamma^\top \end{bmatrix}_{\kappa \times \kappa}. \]

Finally, comparing (20) to (16), it is possible to see that the only difference is the presence of matrix \( M \). Such matrix imposes the way in which the parameters in \( \beta^S_p \) must be linearly combined to achieve the equivalence between using the B-spline basis or the mixed one. Nevertheless, since the first \( \bar{k} - 4 \) elements of the mixed basis are free B-splines, \( M \) does not impose any restriction on the first \( \bar{k} - 4 \) parameters of \( \beta^S_p \).

Matrix \( A \) in (15) and matrix \( M \) in (20) represent linear transformations applied, respectively, to a subset of elements of \( \phi(\cdot) \), and a subset of associated parameters \( \beta^S_p \). Such transformations guarantee that the resulting mixed model approximates the Svensson’s basis are all linearly independent.

---

\(^{15}\)Here we are using the fact that the columns of matrix \( A \) are linearly independent, since the columns of matrix \( \Gamma \) are independent too. The independence of \( \Gamma \)’s columns comes from the fact that \( \lambda_1 \neq \lambda_2 \) guarantees that the exponential functions that compose the Svensson’s basis are all linearly independent.
son’s basis for maturities larger than a particular chosen point \( \bar{\tau} \) of the domain. Since the Svensson’s basis provides a higher degree of smoothness than the B-spline one, the above-mentioned transformations impose different levels of smoothness over the domain. Therefore, the model resulting from the use of the mixed basis \( \Psi(\cdot) \), is called VRP-SV. The term “VRP” comes from the non-uniformity of smoothness across maturities, and the term “SV” comes from the approximation of the Svensson’s model for maturities larger than \( \bar{\tau} \).

2.5 Numerical example

We provide here a graphical representation of how the mixed basis is built. Consider the domain \([0, 10]\) and a cubic spline characterized by the set of knot points \( s = \{0, 1, ..., 9, 10\} \), containing 11 elements. The B-spline basis capable of representing this spline is obtained by (11), using the augmented set of knot points \( \{d_k\}_{k=1}^{11+6} \), where \( d_1 = d_2 = d_3 = s_1 \), \( d_{11+4} = d_{11+5} = d_{11+6} = s_{11} \) and \( d_{k+3} = s_k \) for \( 1 \leq k \leq 11 \). This basis contains \( \kappa = 13 \) B-splines, represented in Figure 1a.

![B-spline basis \( \phi(\cdot) \)](image)

![Mixed basis \( \Psi(\cdot) \)](image)

Figure 1: Bases
Choosing arbitrarily $\tilde{\tau} = 2$ and noting that $\tilde{k} = 6$ and $d_{\tilde{k}} = 2$, we define the mixed basis. By (12), the active B-spline basis over the interval $(\tilde{\tau}, 10]$ is given by $\tilde{\phi} = (\phi_3, \phi_4, \ldots, \phi_{13})$. In Figure 1a, the active basis $\tilde{\phi}$ is represented by the set of B-splines plotted with a solid line. For given values of $\lambda_1$ and $\lambda_2$, the Svensson’s basis $\tilde{\psi}$, as defined in (13), is approximated by the basis $\tilde{\phi}$, as described in (15). As a result of this process, we obtain the mixed basis $\Psi(\cdot)$, plotted in Figure 1b. Note that this basis is formed by six functions, from which the first $\tilde{k} - 4$, i.e. two, functions are the B-splines that assume null value for maturities above $\tilde{\tau}$. The other four functions approximate those that constitute Svensson’s basis $\tilde{\psi}$, for maturities above $\tilde{\tau}$.

This approximation becomes more accurate with smaller spacing between the knot points. In fact, the number of knot points above $\tilde{\tau}$ used in this example is not sufficient to provide an accurate approximation. We used a small number of knot points to provide an easy graphical visualization. In the appendix, we show a histogram of the set of knot points obtained based on the knot location algorithm. Using that set, the approximation is guaranteed to work within a certain tolerance controlled by the researcher. For the empirical results appearing subsequently in the paper, we use those knot points to obtain our mixed basis.

Considering the B-spline basis built above and the mixed basis, defined using the proper set of knot points above $\tilde{\tau}$, we apply the cubic spline and VRP-SV models to a given series of simulated data. We consider four levels for the penalty parameter $\alpha$, as shown in panels (a) and (b) of Figure 2. In panels (c) and (d), we report the effective number of parameters as a function of $\alpha$, respectively for the VRP-SV and cubic spline models. Unlike the cubic spline model, for the VRP-SV model, when $\alpha \to \infty$, this number converges to one. This happens because the only linear solution acceptable by Svensson’s basis, and consequently by the mixed basis, is the one that gives nonzero weight only to the level function. Note that in panel (c) we also report the approximation error for the VRP-SV model, defined as:

$$\max_{\tau \in [0, T]} \left| (\Psi(\tau) - \tilde{\Psi}(\tau)) \beta(\alpha) \right|,$$

where $\tilde{\Psi}(\tau) = \left( (\phi_1(\tau), \ldots, \phi_{\tilde{k}-4}(\tau))_{1 \times (\tilde{k}-4)} \tilde{\psi}(\tau) \right)$. Such error is negatively correlated with alpha. In fact, when $\alpha$ grows, the solution converges to the level function, which is perfectly replicated by the B-spline basis.
3 Data

This section introduces and explains the Brazilian government index-linked bonds market. In a sequence, we explain the issue of scarcity of short-term data existing in this market and the kind of solution that we propose. The section ends providing a description of the data set used for curve estimation in the subsequent empirical section.

3.1 Brazilian government index-linked bonds

The main Brazilian Treasury inflation-protected security is the NTN-B bond. Interests on those bonds are paid twice a year, with an annualized rate of 6%, on its face value adjusted for changes in the IPCA, the leading consumer price index adopted on inflation targeting regimes of the Central Bank of Brazil. Since a NTN-B pays in the middle of a month, and since inflation accruals use the price index of the previous month, there is an
indexation lag of fifteen days when it is priced by market participants. Such lag exists in all index-linked bond markets across the world, and the lag in the Brazilian market is considerably smaller than in other countries. For instance, just to compare, in the TIPS and Gilts markets, there are indexation lags of 2.5 and 3 months, respectively. Due to the small lag within the NTN-B market, we abstract from this indexation issue when estimating the term structure. For those interested in a methodology to deal with this lag issue, see, for instance, Evans (1998).

The set of maturities of the NTN-B bonds available in our sample period is shown graphically in Figure 3a. Each bond is represented by a solid line, which associates each day in the period shown on the horizontal axis to the remaining years to maturity, shown in the vertical axis. Although there is not a large number of available bonds, there is a wide range of maturity, going from 0 to 40 years. For this reason, there is also a considerable gap in maturity between subsequent bonds. For yield curve estimation, such a gap in the short end is of particular concern, since it demands model extrapolations. It is possible to see that in some periods, no bond with time to maturity less than one year is available. For example, this is the case in 2013, when the shortest maturity available was of 1.25 years, as shown in Figure 3b.

Prices for the secondary market of government bonds are published daily by the Brazilian Financial and Capital Markets Association (ANBIMA). Since 1999, the Brazilian’s Treasury and the Central Bank delegate this task to the association. These prices are obtained by averaging prices collected from main market participants, who send to ANBIMA their estimated fair prices for those bonds, in a process similar to how the LIBOR rate is built.

3.2 The scarcity of short-term data

The lack of available data at the short end of the curve is more severe when trading prices of bonds close to mature are not available, or when those prices are affected by idiosyncratic premia, such as the liquidity premium. Unfortunately, this is the case for the NTN-B government bond market. The bottom of the maturities range is constituted by bonds that have been issued with medium or long duration but, that in the course of time, became bonds close to maturity. As reported by most Central Banks in BIS (2005), such bonds are not useful to obtain proper references for the short end of the curve, since
they are usually highly illiquid. In cases like that, the solution adopted by most Central Banks consists in eliminating from the yield curve estimation procedure, bonds with time to maturity less than one year and, to use instead, whenever available, bills and money market rates to populate the short end.

In the Brazilian market, there are not many established references available for the short end of the real yield curve. For this reason, in the term structure estimation process, we try to exclude the smallest possible number of short-term bonds. However, with the same spirit to the Central Banks cited above, we can not avoid excluding some highly illiquid short-term bonds. In fact, as shown in Figure 4\textit{a}, the strong seasonal pattern of the price index (IPCA) and its large variability for terms below six months represent important issues when marking to market real bonds, since these variations are usually barely reflected in prices of available illiquid short-term bonds.\textsuperscript{16} In line with that, many times it is possible to observe unrealistic behavior for bonds below six months, such as smooth price patterns or even almost flat yields over months. For this reason, we exclude bonds with time to maturity less than six months, and seek for alternative ways of anchoring the short end of the real yield curve.

\textsuperscript{16}In particular, when the observed prices do not come from observed trades, and are instead, collected prices.
Estimating the term structure without enough data at the short end may cause several problems. As stated by the Deutsche Bundesbank in BIS (2005), it generates unrealistic short end estimates, and sometimes even unrealistic estimates for the one-year zero-coupon yield, which has particular importance for policymakers. Moreover, it causes instability in the time series of model’s parameters, usually generating unrealistic volatility patterns. Those issues compromise model’s usefulness in forecasting exercises and in risk management procedures.

In the Brazilian fixed income market, there are two potential sources for short-term data: (i) Coupon STRIPS created from the coupons of NTN-B bonds; and (ii) Swaps DIxIGPM. The Brazilian STRIPS market is not well developed, that is, there is not much public information about the prices practiced there. The swap DIxIGPM is a contract in which the negotiated rate represents the expected real interest rate for the period between the trading date and the contract expiration date. Explicitly, celebrating the swap contract, one part pays inflation accrual, measured by the IGPM,\(^{17}\) over the

\(^{17}\)The General Market Price Index (IGP-M) is the second most used price index to measure inflation rates in Brazil. This index is used to readjust several contracts to account for inflation, such as the rental and the energy supply ones. Despite the discrepancy in the collecting period, the major difference between IGP-M and IPCA consists in the first one giving more weight to the wholesales prices than to the consumer prices. Since wholesales prices tend to be more volatile than the consumer ones, the inflation measure by IGP-M fluctuates more widely than the IPCA inflation. Moreover, the discrepancy between these two measures tends to be more pronounced in the short term, since the exchange rate pass-through into the IGP-M is higher and faster than that into the IPCA.
contract period and receives, in exchange, the accrual of the DI rate. The counterpart does the opposite. Swap rates can be considered proxies for default-free real interest rates since the Brazilian Securities, Commodities and Futures Exchange (BM&FBOVESPA) entity guarantees these contracts. However, due to the differences between the building methodologies of IPCA and IGPM, those indexes will differ, making real interest rates defined using each of those indexes to also differ. For this reason, using swap data to extract information to build the short end of the curve could generate inconsistencies in the estimated curve.

As an alternative, we construct synthetic short-term bonds, which are built using nominal yields and market expectations for inflation. ANBIMA publishes, every workday, estimates for the nominal yield curve, extracted from the Brazilian Treasury nominal bonds using the Svensson model. For inflation expectation, we use data from the Focus-Market Readout, a weekly survey of market forecasts for the main Brazilian macroeconomic variables, released by the Central Bank, on Mondays. The expectations are regularly offered by almost one hundred and twenty survey participant institutions, mostly banks, brokers, asset managers, consultancies, and other non-financial entities. The Central Bank also elaborates rankings including the best forecasters, classifying the surveyed institutions based on their accuracy considering expectations over the last six months. The forecasts of the best five institutions constitute the statistics of the Top five ranking, which is published in the Focus-Market Readout. For the IPCA, expectations of monthly inflation for each of the following fourteen months are collected in the survey. We opt to use the median of the Top five IPCA expectations for each of those months. Figure 4b shows a time series of forecasted inflation over our sample period. Comparing to Figure 4a, it is possible to see that the forecasted inflation resembles the actual inflation, with the same seasonal patterns.

A daily time series of the short-term synthetic bonds is built in the following way.

---

18 The DI rate is the average one-day interbank borrowing/lending rate, calculated by CETIP (Central of Custody and Financial Settlement of Securities), every workday.
19 The nominal bonds used are the NTN-F plain vanilla bonds and the LTN zero-coupon-bonds. The nominal yield curve does not face the lack of data problem for the short end. Because the nominal bonds have more liquidity than the real ones; there are bonds issued with a short time to maturity; the gap in maturity between subsequent bonds is small, and there are references to the short end that come from the money market rates.
20 In the Brazilian Market, using the inflation expectations provided by the Focus-Market Readout is practically a consensus among newspapers, analysts and policy makers. Furthermore, the use of surveys has offered better predictions than traditional models, as shown by Ang et al. (2007).
21 See Barbosa and Marques (2013) for a detailed description of the Focus-Market Readout.
First, on each day $t$, we calculate the terms to maturity between $t$ and each date in the upcoming months for which an inflation forecast exists. Then, we select yields from the nominal curve for each of these calculated terms. After that, each nominal yield is discounted by the corresponding accrued forecasted inflation for the respective term, obtaining a set of synthetic real yields. Using a flat-forward interpolation method, we obtain real yields for a predetermined set of maturities. Finally, for each of those maturities, prices for the synthetic zero-coupon bonds are obtained using the corresponding synthetic real yields obtained in previous steps. Figure 5 shows the swap DLxIGPM rates and the synthetic yields for maturities of 3, 6, and 12 months. Note that, despite the differences that exist between the two price indexes (IGPM and IPCA), there is still large similarity between swaps and synthetic yields, with both reflecting the volatility and seasonality patterns of these price indexes.

In the relationship between nominal and real yields, there are two components that we do not consider when building our synthetic bonds: the inflation risk and liquidity premiums. However, Vicente and Graminho (2015), following common procedure adopted in the literature,\textsuperscript{22} find that in the Brazilian market, for the period 2006-2013, the liquidity

\textsuperscript{22}For example D’Amico et al. (2014), Grishchenko and Huang (2013) and Gurkaynak et al. (2010).
premium is not significant. In addition, they show that the inflation risk premium is small for a one-year maturity, of the order of 0.2%. Therefore, there is not much harm in disregarding these two premiums, when we are focused in building the short end of the real yield curve.

In any case, determining the best reference for the short end of the real curve is not the main point of this paper. In what regards that, we just want to ensure that the expected behavior of the reference adopted arises from the seasonality and volatility of the inflation price index. In contrast, the main goal of our work is to present a model capable of fitting the seasonality of the short end without distorting a parsimonious estimation process, via a parametric model, for the medium and long ends of the curve.

### 3.3 Dataset

Our sample period covers January 2009 to May 2014, with daily frequency, totaling 1332 days. The prices of the NTN-B bonds come from ANBIMA, and we use the synthetic bonds with maturities of one day, three months and six months.

### 4 Empirical applications

In this section, we provide two empirical applications for the VRP-SV model. The first consists in modeling the Brazilian real yield curve, comparing in-sample and out-of-sample yield fitting results with those obtained by a fully parametric Svensson model and those by a pure spline-based model. In the second application, we use the time series of models’ parameters to estimate VaR for portfolios exposed to real term structure risk, and compare VaR accuracy obtained by the VRP-SV model to that obtained by Svensson’s model.

Using the X12-ARIMA methodology, we verify that the time series of the actual inflation presents seasonal components only for terms over one year. For the predicted inflation, seasonal components appear for terms over eleven months. We try to control the seasonality effects via parametric modeling of the yield curve by setting $\bar{\tau} = 1$. We define the short end as the interval $[0, \bar{\tau}]$, and the medium and long ends as $(\bar{\tau}, T]$, with $T = 45$ years. Thus, the term structure is modeled as a spline-based model in the short end and as a parametric model in the medium and long ends. The set of knots for
maturities smaller than $\bar{\tau}$ is given by $\{0, 0.25, 0.5\}$, and for maturities bigger than $\bar{\tau}$ is determined according to the knot location algorithm described in the appendix.

To choose the penalty parameter $\alpha$ for our model, we use the GCV procedure. To do so, we create a grid of possible values for $\alpha$, given by $\text{grid}_\alpha = \{0, 1e-8, 1e-7, ..., 1e7, 1e8\}$. For each value of this grid, we estimate the model for the first two years in our sample, obtaining the statistic $\gamma(\alpha)$, as described in (8), for each day. The chosen value, $\alpha^*$, is the one that minimizes the sum of daily statistics, and is given by $\alpha^*=1e-6$.

For comparison purposes, we estimate the Svensson model with and without the constructed synthetic references for the short end. We also estimate the VRP model, including short-term bonds synthetic references. For this model, we place knots in the maturity date of each bond, and the penalty function is given by a three-tiered step function, as in Waggoner (1997):

$$
\alpha(\tau) = \begin{cases} 
    s = 1e - 4 & \text{, if } \tau \leq 1 \\
    (s \times l)^{0.5} = 3.1623e - 2 & \text{, if } 1 < \tau \leq 10 \\
    l = 1e + 1 & \text{, if } \tau > 10 
\end{cases} \quad (22)
$$

The ad-hoc choice of the steps location comes from the volatility patterns observed for bond yields data in different maturities. In the short end ($\tau < 1$), due to the seasonal effect, there are huge differences among yields. In contrast, in the long end ($\tau > 10$), yields practically assume their asymptotic behavior. Finally, for $1 < \tau < 10$, there is no seasonal component but there are moderate variations across yields.\textsuperscript{23}

As in Waggoner (1997), the penalty function is determined by two parameters, $s$ and $l$, whose values are obtained by minimizing the out-of-sample error. Due to the limited number of bonds in the cross-section, we avoid splitting them into in-sample and out-of-sample groups, and instead, use the “leave one out” procedure. This consists in estimating the curve different times, and in each time, leaving out one specific bond and calculating the pricing error for this omitted bond. Such error is defined as the square of the ratio between the price residual and the duration of the bond. The average out-of-sample error for a specific date is, therefore, the average of all those duration-weighted

\textsuperscript{23}We have tested using other step locations, such as $\{[0, 1] [1, 5] [5, 45]\}$ and $\{[0, 5] [5, 10] [10, 45]\}$. However, fitting results were inferior.
pricing errors. We apply this procedure using the first two years of our sample, adopting the grid defined above for \( \alpha \), also for possible values to be taken by the parameters \( s \) and \( l \). The sum of the daily out-of-sample error in this period is minimized by the values of \( s \) and \( l \) reported in (22).

4.1 In-sample results

Table 1 shows the average absolute error of each model, when estimation is performed based on the whole cross-section of bonds, which we denominate in-sample results. SV refers to the Svensson model without inclusion of the three synthetically built bonds as references for the short end. In contrast, SV-Ref contemplates this inclusion. Both VRP and VRP-SV use the above-mentioned synthetic bonds in the estimation procedure. For comparison purposes, however, we do not consider such bonds in the error calculations, since the SV model is estimated without them. The error is expressed in terms of bond yields,\(^{24}\) and it is expressed in basis points.

<table>
<thead>
<tr>
<th>Maturity Range</th>
<th>( \leq 1 )</th>
<th>1-5</th>
<th>5-10</th>
<th>( &gt;10 )</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>0.34</td>
<td>2.02</td>
<td>2.79</td>
<td>2.06</td>
<td>2.11</td>
</tr>
<tr>
<td>SV-Ref</td>
<td>80.11</td>
<td>15.24</td>
<td>7.67</td>
<td>6.18</td>
<td>12.16</td>
</tr>
<tr>
<td>VRP</td>
<td>14.98</td>
<td>4.44</td>
<td>1.72</td>
<td>1.24</td>
<td>2.88</td>
</tr>
<tr>
<td>VRP-SV</td>
<td>6.11</td>
<td>3.33</td>
<td>3.46</td>
<td>2.30</td>
<td>2.99</td>
</tr>
</tbody>
</table>

The in-sample error for the SV model is about two bps on average, considering all maturities. This is a small value since the average bid-ask spread is about eight bps, and the average yield in this market is 5.4%. Note also that the error approaches zero at the short end since the model has considerable freedom to price a small number of bonds. In fact, in most cases, there is only one or none bond with maturity less than one year. However, by the results of SV-Ref, the error for the short end increases dramatically when such region is populated, achieving unacceptable high values (on average, 80 bps). Moreover, the error increases considerably for all range of maturities. Therefore, the presence of securities in the short end affects the estimates for the whole curve in the Svensson model.

\(^{24}\)For each bond, using the fitted price, we get the model implied bond yield. The absolute difference between this yield and the observed one gives the bond yield error.
The VRP model has average error slightly larger than SV and, as expected, the short-term securities do not impact the goodness of fit of medium and long ends. Finally, the VRP-SV model has an average error close to the one presented by the VRP model, but errors are more homogeneously distributed in the maturity range. Therefore, the proposed approach is capable of fitting the whole curve in a parsimonious way. Indeed, to describe the term structure above one year, both the Svensson model and the VRP-SV use the same number of parameters.

Figure 6 compares some zero-coupon yields estimated by each model. Note that the SV model presents completely non-realistic fitted yields for the short end, compromising even the one year rate. In contrast, when for the six-months yield, we compare estimates of the VRP and VRP-SV models to the synthetic six-months yield (not plotted in the picture), we see that both models deliver yields extremely close to the synthetic one.

Despite the unrealistic behavior of short-term yields implied by the SV model, for the
medium and long ends, the estimates provided by this model are very similar to those provided by the VRP and VRP-SV models. On the other hand, attempting to fit the short term references, the SV-Ref model presents fitted short-term yields that are far from the yields obtained by the other considered models, including some differences in yields also appearing in the long end of the curve. Lastly, the estimates of VRP and VRP-SV are close to each other across the whole range of maturities.

Regarding the error in approximation of Svensson’s basis defined in (21), the average and maximum values over the entire period are 1.31e-9 and 3.23e-9. Figure 7 shows the maximum error calculated for each maturity.

4.2 Out-of-sample results

Table 2 reports the average absolute error of each model for the out-of-sample exercise. The errors are obtained according to the leave one out procedure described before. All considerations about pricing errors described in the in-sample case are valid in this exercise too.

The SV model presents the largest average error, which comes from its terrible performance for maturities below five years. Once again, it is possible to see that the scarcity of data in the short end can also compromise model’s ability to fit yields above the one year term. As expected, the SV-Ref model also performs badly, since the entire fit is highly affected by the short-term references. Even so, this model provides a better result, on average, than the SV model.

The VRP model has the smallest average absolute errors and such model is out-
performed only within the range of one to five years, where the VRP-SV model shows smaller errors. In fact, the out-of-sample performances of those two models are again quite similar. Therefore, we conclude that the segmented VRP-SV approach can produce comparable results to those provided by the non-parsimonious VRP model.

Table 2: Out of sample errors

<table>
<thead>
<tr>
<th>Maturity Range</th>
<th>&lt; 1</th>
<th>1-5</th>
<th>5-10</th>
<th>&gt;10</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>514.06</td>
<td>25.97</td>
<td>3.81</td>
<td>2.71</td>
<td>30.53</td>
</tr>
<tr>
<td>SV-Ref</td>
<td>134.43</td>
<td>19.88</td>
<td>7.21</td>
<td>6.75</td>
<td>16.08</td>
</tr>
<tr>
<td>VRP</td>
<td>31.48</td>
<td>10.88</td>
<td>3.50</td>
<td>2.53</td>
<td>6.41</td>
</tr>
<tr>
<td>VRP-SV</td>
<td>35.21</td>
<td>10.77</td>
<td>4.27</td>
<td>2.90</td>
<td>6.83</td>
</tr>
</tbody>
</table>

4.3 VaR results

In this empirical application, we estimate a VaR model for portfolios exposed to real term structure risk. To do so, we estimate the variance of the yield curve parameters based on a multivariate GARCH process. We do not use the VRP model since its high number of parameters makes the estimation of a multivariate GARCH process extremely challenging. Thus, our competitors in this exercise are only the models SV, SV-Ref and VRP-SV.

We build three portfolios containing one unit of each NTN-B bond available within three different ranges of maturities. First, the general portfolio considers the entire range of maturities. In its turn, the short portfolio works with maturities lower than five years. Finally, the long portfolio covers the range over five years. Note that, since we use, in each day, only the available bonds in that day, the portfolios compositions are not constant over time.

In order to guarantee temporal consistency for models’ parameters $\beta$, the values of $\lambda_1$ and $\lambda_2$ must be fixed. By fixing those exponential decay parameters, we make the functional basis of those models constant over time. Thus, we split the data set in two parts. The first one contains the first 504 days, and it is used to determine $\lambda_1$ and $\lambda_2$, based on minimization of the global in-sample error within this period. The second part, comprising data on the 828 subsequent days, is used to estimate a time series of a one-day VaR, exactly as done by risk managers in most investment banks across the world.
The VaR is expressed as a daily change in the portfolio value, which is given by the sum of the prices of each bond belonging to it. To guarantee that the variation in the portfolio value only reflects market risk, we control, in bond prices, the effects of inflation accrual and coupon payment.

To forecast the volatility of the yield curve parameters, we estimate, using the daily changes of the vector of parameters $\beta$, $\Delta \beta$, a multivariate GARCH with constant conditional correlation (CCC-MGARCH).\textsuperscript{25} For each day $T$ in which the VaR is calculated, we estimate the GARCH model using a rolling window $[T - 504, T - 1]$, in business days, covering two years of data.

The VaR is tested based on three methodologies: Delta-Normal, Monte-Carlo\textsuperscript{26} and Bootstrap. We adopt the last one since it provides better accuracy results for all term structure models considered in this empirical exercise. The implementation of this method consists in creating a sample of one hundred thousand $\Delta \beta$ taken from the rolling-window using sampling with replacement. In this generated sample, each observation is normalized using the GARCH process:

\begin{equation}
\Delta \beta_{t, \text{Normalized}} = \frac{\Delta \beta_t}{\sqrt{h_t}} \sqrt{h_T},
\end{equation}

where $T$ is the date for which the VaR is built, $t \in [T - 504, T - 1]$ and $h_j$ is the variance forecasted by the GARCH that is estimated based on the rolling window $[j - 504, j - 1]$. After that, each standardized vector is added to the parameter vector of date $T - 1$, $\beta_{T - 1}$. This results in a new sample of one hundred thousand vectors, representing possible values for $\beta_T$. Finally, we use each one of those vectors to price the portfolio, and select the resulting price by estimating the VaR percentile specified.

For the usual VaR levels adopted in investment banks and hedge funds,\textsuperscript{27} Table 3 presents exception rates and p-values for tests of unconditional coverage (UC), independence (IND) and conditional coverage (CC), as proposed by Christoffersen (1998). The null hypotheses (H0) in the first test affirms that the exception rate is equal to the VaR level. Regarding the IND test, its H0 hypotheses states that all VaR violations are independently distributed over time. The H0 hypotheses for the CC test conjectures that

\textsuperscript{25}We use The Oxford MFE Toolbox developed by Kevin Sheppard, and available at http : //www.kevinsheppard.com/MFE_Toolbox.

\textsuperscript{26}Assuming that the daily changes in the vector of parameters follows a Brownian motion process.

\textsuperscript{27}The VaR level $\alpha$% is such that the confidence level is given by $(1 - \alpha)$%.
VaR violations are independently distributed and that the probability for one exception to happen is equal to the VaR level. Complementary to these tests for specific percentiles, we also report the p-value for the Kuiper test (K) that focus on the entire probability distribution, as shown by Crnkovic and Drachman (1996). Its H0 hypotheses states that the forecasted distribution of variations in portfolio value is equal to the realized distribution.

Examining first results for the general portfolio, the SV model presents exception rates close to the respective VaR levels. Consequently, all H0s for the UC and K tests are not rejected. However, it is possible to reject H0 for the CC and IND tests, for VaR levels above 5%. The VRP-SV performance is similar to that of the previous model, presenting slightly better results on the CC and IND tests. Regarding the SV-Ref, it is clear that such model can not predict the percentiles or the distribution as a whole. Up to this point, the proposed VRP-SV approach does not outperform the SV model. Moreover, it seems that scarcity of data in the short end does not cause damages to the risk management procedure of the SV model. This is a surprising result since we expect that non-realistic estimates for the short end of the curve would directly impact the volatility estimated for that region. To see if the conclusions obtained so far are robust to different portfolios, we analyze the VaR estimates for the short-term and long-term portfolios.

Regarding the short-term portfolio, the VRP-SV model is the only one which avoids rejecting the different H0 hypotheses. Furthermore, the SV and SV-Ref models present almost all exception rates extremely lower than the corresponding VaR levels. As a result, the forecasted distribution of variations in portfolio value is significantly different from the realized distribution. This can be visualized in Figure 8, which compares the realized distribution of daily variations for the short portfolio values with the forecasted distributions, using the three models. Therefore, the scarcity of short-term data issue below one-year maturities strongly compromises VaR estimates for portfolios composed of short- and medium-maturity bonds.

Finally, regarding the long-term portfolio, the proposed approach presents the same performance obtained in the general portfolio case. Once more the SV-Ref model shows complete inability to fit the data and its volatility. For the SV model, the results are again worse than those for the general portfolio, but now, the opposite to the short-term portfolio case happens: all exception rates are larger than the corresponding VaR levels. This effect is so strong that it is possible to reject the H0 for the UC test for the VaR
level 10% and also for the distribution test.

The P-P plots for daily variations in the values of general and long-term portfolios appear in the appendix, respectively in figures 17, and 18. Checking those pictures, we can see that, for both general and long-term portfolios, the VRP model does a better job in approximating the first half of their realized distributions when compared to the SV model, and they have a similar approximation for the second half.

In the appendix, we also report, as a visual complement to the information given in Table 3, figures of daily returns and VaR levels for the SV and the VRP-SV models. Figures 11, and 12 show, for the general portfolio, daily returns and VaR levels, respectively for the SV and the VRP-SV models. Figures 13, and 14, and Figures 15, and 16 show same series for the short-term and the long-term portfolios. By looking at those pictures we can see where exactly VaR exceptions occurred in time, and also how homogenous they were. For instance, we note that, for both models and all portfolios, no exception happened after January, 2014. This might be an indication that VaR based on historical information could be slightly overestimated during those last five months of the sample.

We want to draw our attention to figures 13, and 14, which are related to the short-term portfolio. Since in this empirical exercise we estimate VaR for a total of 828 daily observations, 1% exceptions would represent around 8 exceptions, while 5% would represent 41 exceptions. In this context, figure 13 clearly shows that the estimated VaR for the short-term portfolio, at 1% significance level, is strongly overestimated by the SV model, while this is not true for the VRP-SV model. In addition, both models overestimate the
5% VaR with worse results again for the SV model, that presents an exception rate of 3.6%, in contrast to the 4.2% rate of the VRP-SV model.

In summary, as a conclusion of this VaR exercise, the proposed VRP-SV not only outperforms the Svensson model but also it is the only model capable of providing reliable VaR estimates for portfolios consisting of short and medium-term bonds.

5 Conclusion

Relying on the main benefits provided by the two classes of methodologies mostly used by Central Banks to build the term structure of interest rates, we propose a segmented model for the real yield curve. Our main goal is to provide a parsimonious approach that can describe the curve as a whole, including the short end which is frequently characterized by a lack of smoothness issue. To do so, we consider splines to fit the short end, borrowing from its flexibility, while the medium and long ends are captured by a parametric Svensson model. To jointly estimate these two classes of models, we develop a procedure in which the spline basis spans the parametric one.

We illustrate in two empirical applications the benefits of the proposed term structure model, VRP-SV, for the Brazilian government index-linked bond market. The first one consists in modeling the Brazilian real yield curve. The VRP-SV and VRP models can fit the entire curve and provide similar outputs. In contrast, the Svensson model fails to fit the short end and is strongly outperformed by the other models.\textsuperscript{28} In the second application, we estimate VaR for portfolios exposed to real term structure risk. We avoid using the VRP method due to its high computational cost imposed by a large number of parameters. We find that the proposed VRP-SV not only outperforms the Svensson model, but also is the only model capable of providing reliable estimates for portfolios consisting of short- and medium-term bonds.

Our model represents a new unified technique, which meets the main demands of pricing, monetary policy analysis, and portfolio risk control. Although we have focused in the real yield curve, it can also be used to fit the nominal term structure, seeking to improve estimates of the short end of this curve when money market rates are available.

\textsuperscript{28}Those results are consistent with Yallup (2012), who shows that smoothing spline models offer superior results when compared to parametric models.
Table 3: VaR results.

<table>
<thead>
<tr>
<th>Model</th>
<th>VaR level</th>
<th>Exception rate</th>
<th>General Portfolio</th>
<th></th>
<th>Short Portfolio</th>
<th></th>
<th>Long Portfolio</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>IND UC CC K</td>
<td>Exception rate</td>
<td>IND UC CC K</td>
<td>Exception rate</td>
<td>IND UC CC K</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.0%</td>
<td>11.0% 0.01 0.35 0.02</td>
<td>7.0% 0.63 0.00 0.01</td>
<td>12.3% 0.00 0.03 0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>5.0% 0.02 0.95 0.05</td>
<td>3.6% 0.41 0.06 0.12</td>
<td>5.8% 0.01 0.30 0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV</td>
<td>2.5%</td>
<td>2.4% 0.32 0.88 0.60 0.29</td>
<td>1.2% 0.62 0.01 0.03 0.00</td>
<td>2.7% 0.27 0.77 0.53 0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0%</td>
<td>1.6% 0.52 0.13 0.26</td>
<td>0.5% 0.84 0.10 0.25</td>
<td>1.4% 0.55 0.22 0.40</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>0.6% 0.81 0.68 0.89</td>
<td>0.5% 0.84 0.94 0.98</td>
<td>0.6% 0.81 0.68 0.89</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.0%</td>
<td>3.1% 0.19 0.00 0.00</td>
<td>0.8% 0.73 0.00 0.00</td>
<td>4.2% 0.25 0.00 0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>1.2% 0.62 0.00 0.00</td>
<td>0.2% 0.92 0.00 0.00</td>
<td>1.6% 0.52 0.00 0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SV-Ref</td>
<td>2.5%</td>
<td>0.8% 0.73 0.00 0.00 0.00</td>
<td>0.2% 0.92 0.00 0.00 0.00</td>
<td>1.1% 0.66 0.00 0.01 0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0%</td>
<td>0.2% 0.92 0.01 0.03</td>
<td>0.2% 0.92 0.01 0.03</td>
<td>0.5% 0.84 0.10 0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>0.2% 0.92 0.24 0.50</td>
<td>0.0% 0.00 0.00 0.00</td>
<td>0.2% 0.92 0.24 0.50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>10.0%</td>
<td>10.3% 0.03 0.80 0.09</td>
<td>8.6% 0.42 0.16 0.27</td>
<td>10.0% 0.01 0.98 0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.0%</td>
<td>4.6% 0.84 0.58 0.84</td>
<td>4.2% 0.66 0.30 0.53</td>
<td>5.2% 0.61 0.80 0.85</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VRP-SV</td>
<td>2.5%</td>
<td>2.3% 0.34 0.70 0.59 1.00</td>
<td>1.9% 0.43 0.28 0.40 0.65</td>
<td>2.4% 0.50 0.88 0.79 1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.0%</td>
<td>1.2% 0.62 0.56 0.75</td>
<td>1.0% 0.69 0.92 0.92</td>
<td>1.3% 0.59 0.37 0.57</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.5%</td>
<td>0.6% 0.81 0.68 0.89</td>
<td>0.8% 0.73 0.20 0.41</td>
<td>0.7% 0.77 0.39 0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The exception rate and the p-values for the tests of unconditional coverage (UC), Independence (IND), conditional coverage (CC) and Kuiper (K). The p-value in bold indicates that it is not possible to reject the null hypothesis.
References


A Appendix

A.1 Local approximation

Let $d = \{d_k\}_{k=1}^{k+p+1}$ be the set of knot points with $d_1 = a$ and $d_{k+p+1} = b$. Define $S_{p,d}$ to be the space of all linear combinations of B-splines with degree $p$ and knot vector $d$. Lyche and Morken (2011) show that for a function $f \in C_{\Delta}^{p+1}[a,b]$\footnote{A function $f \in C_{\Delta}^{k}[a,b]$, if it has $k-1$ continuous derivatives on the interval $[a,b]$, and the $k$th derivative $D^k f$ is continuous everywhere except for a finite number of points in the interior $(a,b)$, given by $\Delta = (z_j)$. At the points of discontinuity $\Delta$ the limits from the left and right, given by $D^k f(z_j+)$ and $D^k f(z_j-)$, should exist so all the jumps are finite.}, the distance between $f$ and $S_{p,d}$ on the interval $[d_k,d_{k+1}]$ is bounded by:

$$\inf_{g \in S_{p,d}} ||f - g||_{\infty,[d_k,d_{k+1}]} \leq K_p h^{p+1} ||D^{p+1}f||_{\infty,[d_{k-p+1},d_{k+p}]}$$, \hspace{1cm} (24)

where $D^k f$ is the $k$th derivative of $f$, $||.||_{\infty,j}$ is the maximum norm on interval $j$, $h = d_{k-p+1} - d_{k+p}$, and the constant $K_p$ only depends on $p$, and is given by:

$$K_p = \frac{(2p(p-1))^{p} + p!}{2^{p+1}(p+1)!p!}$$.

By (24), with a sufficient number of knots, it is possible to make $h$ sufficiently small, such that the local error satisfies an acceptable tolerance level. Moreover, (24) also informs that if we want to obtain homogeneous approximation errors across all regions, we need to allocate more knot points in the regions where $|D^{p+1}f|$ assumes larger values. Therefore, the bound given by (24) not only ensures that the approximation is possible, but also provides useful information on how to efficiently allocate the knot points.

A.2 Knot location algorithm

Contrasting with a curve interpolation scheme, in an approximation scheme, error bounds are also turned into inputs jointly with the data that should be fitted. A priori, the number and distribution of knots required to achieve the desired accuracy satisfying the error bound, are unknown. Therefore, in most empirical applications, the approximation methods are iterative. According to Piegl and Tiller (1995), roughly speaking, these methods proceed in two possible ways: forward or backward incremental. The first one
starts with a minimum, or a small, number of knots, fit an approximating curve, check deviations of this curve from data, and increases the amount of knots if the error bound is not satisfied. In contrast, the backward method starts with a maximum, or a large number of knots, fit, check deviation, and reduces the number of knots, whenever possible.

More recently, some papers, starting with Li et al. (2005) and Park and Lee (2007), incorporate information about the shape of a given point set in the knot selection algorithm. Li et al. (2005) propose an incremental method where knots are placed to satisfy a heuristic rule that guarantees a good approximation to a given set of data points. This approach simplifies the general incremental methods, reducing time consumed since the problem of knot placement becomes how to best allocate knots in order to fulfill the above-mentioned rule. In that sense, our algorithm is quite similar to that of Li et al. (2005). We also adopt a forward incremental method using a heuristic rule. However, in our approach, the heuristic rule comes directly from the local bound (24). This allows us to set the desired tolerance, what can not be done in previously mentioned work.

For a given $\lambda$, consider the functions $\bar{\psi}_k(.)$, with $k \in \{1, 2, 3, 4\}$ and domain $(\bar{\tau}, T]$, as defined in (13). In this interval, allocate a reduced number of knot points equally spaced, with the first one allocated very close to $\bar{\tau}$ and the last one at $T$.\footnote{On our applications, we make the first knot above $\bar{\tau}$ to be $\bar{\tau} + 0.00794$, with this increment representing two business days expressed in years ($2/252$). In addition, when fitting the Brazilian real yield curve, we start with knot points equally spaced at every four years.} Then, letting $d$ be the set of those knots, for each interval $[d_{k-2}, d_{k+3}]$, starting from the smallest knots to
the largest ones, calculate:

$$\text{Sup}_k = \max_{j \in \{2,3\}} \left( K_3 h^4 \|D^4 \tilde{\psi}_j\|_{\infty,[d_{k-2},d_{k+3}]} \right).$$

(25)

If $\text{Sup}_k > \text{tol}$, where $\text{tol}$ is the acceptable tolerance, we calculate a new knot point $d = (d_{k+3} + d_{k-2})/2$. Then, we update the set of knots $d$ by adding $d$, and calculate $\text{Sup}_k$ once more time. We continue the iteration until $\text{Sup}_k \leq \text{tol}$. When the condition $\text{Sup}_k \leq \text{tol}$ is achieved, we go to the next interval, updating $k = k + 1$, calculate $\text{Sup}_k$, and restart the process. The whole process finishes when all intervals $[d_{k-2},d_{k+3}]$ in $d$ were covered.

Although it is possible to show that, $\forall \tau > \bar{\tau}$ and $\forall \lambda > 0$, $|\partial^4 \tilde{\psi}_2/\partial \tau^4| < 0$ and $|\partial^4 \tilde{\psi}_3/\partial \lambda^4| > 0$, the same is not true for $\bar{\psi}_3$. Figure 9 reports the absolute values for the fourth derivatives of $\tilde{\psi}_2$ and $\tilde{\psi}_3$. For those, we see that the largest values are found when $\tau$ approaches $\bar{\tau}$, and such values are even larger when $\lambda$ increases. Thus, according to (24), the number of necessary knot points is an increasing function of $\lambda$. To avoid the necessity of calculating the vector of knots and corresponding B-splines in each iteration when optimizing $\lambda$, we determine the set of knots that satisfies (24) for all possible predefined $\lambda$’s. To do so, we start defining a grid with possible values for $\lambda$, given by $\text{grid}_\lambda = [0.2, 0.25, 0.3, ..., 5.90, 5.95, 6.0]$. After that, for each $\lambda_j \in \text{grid}_\lambda$, we obtain the knot vector indexing it by $d^j$. The final knot vector $d$ is obtained as the union of all sets $d^j$, after elimination of any repeated sets.

Since $\text{Sup}_k$ is a superior bound for the approximation error, the bound might be well above the error, which is defined as:

$$\text{Error} = \max_{j \in \{2,3\}} |\tilde{\psi}_j - \tilde{\psi}_j|.$$  

(26)

Therefore, creating a knot allocation algorithm that makes (24) always below a desired tolerance level, automatically makes the approximation error below that tolerance level, but may cost an allocation of more knot points than necessary. To control for an unnecessarily large number of knots, we create a grid of possible values for $\text{tol}$, given by

\[\text{grid}_d = [0.2, 0.25, 0.3, ..., 5.90, 5.95, 6.0].\]

In a similar spirit to Diebold and Li (2006), the grid range is determined based in the maturity where the curvature function $\bar{\psi}_3$ assumes its maximum. With $\lambda = 0.2$, the maximum is reached at $\tau = \bar{\tau} + 9$, a large maturity in which a curvature effect is usually not expected to exist. For $\lambda$ values less than 0.2 the maximum occurs at maturities even larger. On the other hand, with $\lambda = 6$, $\bar{\psi}_3$ reaches its maximum at $\tau = \bar{\tau} + 0.3$, a value very close to $\bar{\tau}$. For this reason, $\bar{\psi}_3$ is also very close to $\bar{\psi}_2$ for $\lambda$ larger than six.
grid_{tol} = \{1e3, 1e2, ..., 1e-5, 1e-6\}, and for each grid entry, we perform the knot points allocation algorithm, and calculate the approximation error (26). We select the larger value in grid_{tol} and its associated knot vector d that makes the approximation error in (26) below a predefined tolerance level, which we take as 1e-6. When we perform such algorithm, the value in grid_{tol} chosen was 1e-2.

Once we obtain the vector of knot points d above \( \bar{\tau} \), we insert into that vector, the knots below \( \bar{\tau} \), completing the set of knots for the entire domain.

Figure 10 shows the histogram of the set of knot points above \( \bar{\tau} \) for the numerical example of section 2.5, in which \( \bar{\tau} = 2 \) and \( T = 10 \). As expected, the knots are more concentrated around maturities near \( \bar{\tau} \).

A.3 VaR Figures
Figure 11: VaR for general portfolio using SV model

Figure 12: VaR for general portfolio using VRP-SV model
Figure 13: VaR for short portfolio using SV model

Figure 14: VaR for short portfolio using VRP-SV model
Figure 15: VaR for Long Portfolio using SV model

Figure 16: VaR for Long Portfolio using VRP-SV model
Figure 17: P-P plot for daily variation in the general portfolio value

Figure 18: P-P plot for daily variation in the long portfolio value