Approximating Risk Premium on a Parametric Arbitrage-free Term Structure Model

Caio Almeida *
FGV/EPGE

Daniela Kubudi †
FGV/EPGE

March 28, 2014

Abstract

In this paper we approximate the risk factors of a polynomial arbitrage-free dynamic term structure model by running a sequential set of linear regressions independent across time. This approximation avoids full optimization in the estimation process allowing for a simple method to extract the risk premium embedded in interest rate instruments. Closed-form bond pricing formulas provide a clear interpretation of each source of aggregate risk as known term structure movements. Assuming, for illustrative purposes, the existence of three sources of aggregate risk (level, slope and curvature) in the economy, we test the validity of our approximation adopting a dataset of Brazilian zero coupon interest rate swaps. The new methodology generates accurate parameters, standard deviations and risk premium dynamics when compared to the exact dynamic model.

Keywords: Term structure of interest rates, parametric models, affine models, cross sectional estimation, time series analysis.

JEL Classification: C1,C5,G1

*Email: calmeida@fgv.br, Graduate School of Economics, FGV/EPGE, Rio de Janeiro, Brazil.
†Email: dkubudi@gmail.com, Address for correspondence: Praia de Botafogo 190, 11 andar, 22250-900, Botafogo, Rio de Janeiro. We thank Marcio Garcia (discussant) and participants at the XII Brazilian meeting of finance for useful comments. The first author acknowledges financial support from CNPq-Brazil.
1 Introduction

Parametric arbitrage-free term structure models are dynamic models with closed-form bond pricing formulas explicitly depending on the risk factors underlying the economy. Although theoretical results on such models were derived about fifteen years ago, corresponding empirical implementations and applications have only recently gained more attention\textsuperscript{1}. An important message behind these implementations is that parametric models are useful and simple to understand precisely because factor loadings are known before model estimation. The a priori knowledge of factor loadings brings some advantages when compared to traditional dynamic affine and quadratic models\textsuperscript{2}. In particular, closed-form formulas for bond prices guarantee that each source of risk in the economy has interpretation as a known movement of the term structure\textsuperscript{3}. Nonetheless, perhaps the most important advantage is that this a priori knowledge of factor loadings allows researchers to better disentangle the role of cross-sectional and time series information on the identification of model parameters.

In this paper, adopting a Legendre Dynamic Term Structure Model (Almeida, 2005)\textsuperscript{4}, we explore this last point to produce a simple and fast method to estimate the risk premium implicit in fixed income securities without having to fully optimize a dynamic term structure model.

We explore closed-form formulas of bond prices to separate the estimation procedure in two sequential steps: Cross-section and time-series. In the cross-section step, we approximate the model state vector by running, for each date, a linear multiple regression of observed zero-coupon yields on a set of predetermined term structure movements whose loadings are represented by Legendre polynomials. In the second step, we run a Vector autoregressive (VAR) regression of lag one on the time-series of the approximated state vector to estimate the parameters under the physical probability measure.

This separation of estimation in two stages allows us to come up with our approximation for risk premium. After running the two pairs of regressions to obtain respectively the approximated state vector, and estimates for model parameters under the physical measure, risk premium is backed out from a simple analytical formula that connects those two


\textsuperscript{2}For affine models see (Duffie and Kan, 1996), for quadratic see (Ahn, Dittmar and Gallant, 2002).

\textsuperscript{3}In contrast, the loadings of each source of risk in usual dynamic models, once estimated have to be associated with approximations for the level, slope, curvature, and other term structure movements.

\textsuperscript{4}The Legendre Dynamic Model is a separable term structure model that parameterizes the term structure of interest rates as a linear combination of Legendre polynomials, with the coefficients of the linear combination representing the model state vector.
probability measures as done in traditional dynamic models. As a byproduct, once all risk-neutral parameters are known, the model can be readily used in financial applications such as risk management procedures or pricing interest rate options.

Almeida (2005) showed that it is possible to have dynamics for the state vector of the Legendre model more general than that of affine models including stochastic volatility. However, due to the simplicity of the Gaussian models and also following the recent quest to better understand identification and estimation issues of Gaussian affine term structure models (Joslin, Singleton and Zhu, 2011; Hamilton and Wu, 2012), we propose our risk premium approximation in the context of a Gaussian Legendre dynamic model.

Our work is related to the above-cited recent papers that offer fast and efficient methods for the estimation of Gaussian arbitrage-free term structure models. Joslin, Singleton and Zhu (2011) suggest a two-step estimation procedure for Gaussian affine term structure models that consists on a combination of running OLS regressions under the physical probability measure combined with likelihood maximization methods to obtain risk-neutral parameters on a second step. Similarly Hamilton and Wu (2012) provide a new estimation methodology based on the minimization of chi-square distances. Although applicable to a more general class of Gaussian models than ours, both methods still rely on optimization processes to estimate part of the parameters while our approximation relies on running only two sets of linear regressions.

Adrian, Crump and Moench (ACM, 2013) have the closest work related to ours. They suggest using a three-step linear regression estimator to extract the risk premium of Gaussian term structure models with observable pricing factors. In contrast, in our method, since we build the pricing factors based on Legendre polynomials we don’t need observable pricing factors to exist. This has important consequences in especial when analyzing interest rate markets that contain coupon-bearing bonds. In such cases, by construction, the ACM approach needs external observable factors to serve as sources of risk since it is not designed to extract the risk factors from the analyzed market, except if it is a zero coupon market and principal component analysis can be applied. On the other hand, since we have a parametric arbitrage-free model, we can extract risk factors by using the model parametric nature and verify the prices of risk of such factors by using the model arbitrage-free nature, even in markets with coupon-bearing bonds.

Due to the parametric nature of the model, all risk neutral parameters, except for the diffusion of the state vector, are known a priori. Since the diffusion is a common parameter to both physical and risk-neutral measures, it is estimated with the VAR regression.

The focus of the current paper is not in empirical applications of the model but in showing that the proposed approximation is accurate.

In this paper, however, we concentrate on the zero-coupon case, keeping in mind that the methodology can be directly extended to a market with coupon bearing bonds. The extension should be based in
We apply this methodology to a database of Brazilian zero coupon swaps adopting an arbitrage-free Gaussian Legendre polynomial model with three risky stochastic factors. Two versions of the model are implemented: the exact dynamic model and our approximation based on regressions. Numerical results show that parameter estimates as well as their standard deviations for both versions are very close to each other. In particular, for both implementations, the same subset of parameters is significant at a 95% confidence interval. In addition, as can be noted in Figure 7, the time series of the risk premium implicit in bonds of different maturities for both versions of the model move on the top of each other for a large part of the sample, and are very similar on the remaining part of the sample. This reassures that the approximation works as a good substitute for the original more complex implementation of the dynamic model when we are interested in applications like option pricing and risk premium analysis.

The rest of the paper is organized as follows. Section 2 presents the cross-sectional and the dynamic Legendre models. A description of data, the results of the Gaussian dynamic model and its approximated version are presented in Section 3. Section 4 briefly introduces possible generalizations of the model approximation. Section 5 concludes. The two appendices present technical information including some important rotation matrices and details on the particular conditionally deterministic factors that appear on the Gaussian version of the model.

2 The Legendre Polynomial Model

2.1 The Static Model

Almeida, Duarte and Fernandes (1998) proposed modeling the term structure of interest rates as a benchmark curve plus a linear combination of Legendre polynomials:

$$R(\tau) = B(\tau) + \sum_{n \geq 0} c_n P_n(\frac{2\tau}{l} - 1)$$

where \(\tau\) denotes time to maturity, \(B(\tau)\) is a benchmark curve, \(P_n(.)\) is the Legendre polynomial of degree \(n\), and \(l\) is the longest maturity of a bond in the market considered.

The first four Legendre polynomials are respectively 1, \(x\), \(\frac{1}{2}(3x^2 - 1)\), and \(\frac{1}{2}(5x^3 - 3x)\), and usually they are sufficient to capture a large variability of interest rate curves in fixed income markets. Figure 1 depicts the first four Legendre polynomials. Note that there is a nice interpretation for each term structure movement driven by these polynomials. The constant polynomial is related to parallel shifts and may be interpreted as the loadings

Almeida, Duarte, and Fernandes (2003) who prove that we can build synthetic principal components with Legendre polynomials in markets with unobservable term structures.
of a level factor; the linear polynomial is related to changes in the slope; the quadratic polynomial is related to changes in the curvature; the cubic polynomial is related to more complex changes in the curvature.

Whenever there are yields available in the market, i.e., it is a zero coupon market, the model can be estimated based on a linear regression. In fact, assuming that we observe the term structure \( \tilde{R}(\tau) \) with measurement error and \( B(\tau) \) without measurement error, for maturities \( m_1, m_2, \ldots, m_k \), and defining:

\[
y = \begin{bmatrix}
\tilde{R}(m_1) - B(m_1) \\
\tilde{R}(m_2) - B(m_2) \\
\vdots \\
\tilde{R}(m_k) - B(m_k)
\end{bmatrix}, \quad \text{and } L = \begin{bmatrix}
P_0(m_1) & P_1(m_1) & \ldots & P_{N-1}(m_1) \\
P_0(m_2) & P_1(m_2) & \ldots & P_{N-1}(m_2) \\
\vdots & \vdots & \ddots & \vdots \\
P_0(m_k) & P_1(m_k) & \ldots & P_{N-1}(m_k)
\end{bmatrix}.
\]

the parametrization given in Equation (1) implies the following multiple linear regression:

\[
y = L\beta + \epsilon
\]

whose solution is known in closed-form:

\[
\hat{\beta} = (L' L)^{-1} L' y
\]

### 2.2 Legendre Forward Rates

The relation between the instantaneous forward rate curve and the term structure of interest rates is given by:

\[
r(t, \tau) = R(t, \tau) + \tau \frac{\partial R(t, \tau)}{\partial \tau}
\]

An application of Equation (4) to the Legendre parameterized term structure appearing in Equation (1) yields:

\[
r(\tau) = B(\tau) + \tau \frac{\partial B(\tau)}{\partial \tau} + \sum_{n \geq 0} c_n P_n(\frac{2\tau}{l} - 1) + \tau \left( \sum_{n \geq 1} c_n \frac{\partial P_n(\frac{2\tau}{l} - 1)}{\partial \tau} \right)
\]

Now, just to exemplify, suppose we are interested in obtaining the forward curve for the Legendre model with the first four Legendre polynomials, setting the benchmark curve to zero \( B(\tau) = 0 \). Defining \( x = \frac{2\tau}{l} - 1 \), and using the chain rule to get \( \frac{\partial P_n(\frac{2\tau}{l} - 1)}{\partial \tau} = \frac{2}{l} \frac{\partial P_n(x)}{\partial x} \), Equation (5) becomes:

\[
r(\tau) = c_0 + c_1 x + \frac{c_2}{2} (3x^2 - 1) + \frac{c_4}{2} (5x^3 - 3x) + \frac{2\tau}{l} \left[ c_1 + 3c_2 x + \frac{c_3}{2} (15x^2 - 3) \right]
\]

\(^8\)Of course the model also works to extract the term structure of interest rates in markets with coupon-bearing bonds as shown, for instance, in Almeida, Duarte and Fernandes (1998, 2003).
Observing Equation (6), note that the instantaneous forward rate can be rewritten as a polynomial in the maturity variable $\tau$:

$$r(\tau) = \sum_{n=0}^{3} L_n(c_0, c_1, ..., c_{N-1}) \tau^n$$

(7)

where each $L_n$ is a linear function of vector $c = [c_0, c_1, ..., c_{N-1}]'$. In this particular case, $L_0 = c_0 - c_1 + c_2 - c_3$, $L_1 = \frac{4}{7}c_1 - \frac{12}{7}c_2 - \frac{51}{7}c_3$, and so on.

When talking about the dynamic version of the model below, this polynomial form for the forward rate curve will be important to identify the state space variables.

2.3 The Legendre Dynamic Model

Our objective in this section is to frame the parametric Legendre model into a dynamic setting. To that end, assume the existence of a probability space $(\Omega, F, Q)$, with $Q$ being an equivalent martingale measure under which discounted bond prices are martingales. The existence of such measure guarantees absence of arbitrages in the market (see Duffie, 2001). Within this probability space we also assume that there is an $N$-dimensional Brownian Motion $W^*_t$ and a state vector $Y_t$ whose uncertainty is driven by $W^*_t$.

Under the Legendre model with a finite number of term structure movements $N$, bond prices are given by:

$$P(t, T) = e^{-\tau G(\tau) Y_t},$$

(8)

where $G(\tau)$ is a vector containing the first $N$ Legendre polynomials evaluated at maturity $\tau$:

$$G(\tau) = [P_0(\frac{2\tau}{T} - 1) P_1(\frac{2\tau}{T} - 1) ... P_{N-1}(\frac{2\tau}{T} - 1)]$$

(9)

Given a Stochastic Differential Equation (SDE) for the state vector $Y_t$,

$$dY_t = \mu^Q(Y_t)dt + \sigma(Y_t)dW^*_t$$

(10)

Almeida (2005) asked which restrictions $\mu^Q(Y_t)$ and $\sigma(Y_t)$ should satisfy in order to guarantee that discounted bond prices are martingales.

Writing bond prices as a function of the forward rate curve, applying the Heath, Jarrow and Morton (HJM, 1992) methodology and Ito’s formula, he derived the martingale drift restriction that the state vector should satisfy under $Q$.

Imposing the martingale restriction, he obtained a set of interesting results that we

---

9 Although the original statistical Legendre model is formulated with a series of infinite polynomials, to approach the problem in a dynamic setting similar to that of affine and quadratic factor models we limit the number of movements to be finite.
summarize here. First, he showed that arbitrage-free parametric Legendre models generalize the affine class of models proposed by Duffie and Kan (1996)\textsuperscript{10}. Duffie and Kan (1996) showed that the state vector dynamics of any affine model can be represented by SDEs whose drift and squared diffusion matrix are affine functions of the state vector under \( \mathcal{Q} \). In contrast, Almeida (2005) showed that in Legendre dynamic models, although the drift is also an affine function of the state vector, the diffusion matrix can be an arbitrary integrable function of the state vector.

But arbitrage-free parametric term structure models also have their drawbacks. In the particular case of the Legendre model, imposing the constraint that term structure movements should be represented by Legendre polynomials restricts, in a important dimension, the stochastic differential system satisfied by the dynamic factors \( Y_t \) under the risk neutral measure \( \mathcal{Q} \). Almeida (2005) showed that in a \( N \)-dimensional Brownian setting, the Dynamic Legendre Model with \( N \) factors will only have \( [\frac{N}{2}] \)\textsuperscript{11} truly stochastic factors, i.e., factors directly driven by the Brownian Motion vector having diffusion coefficients differing from zero. The other \( [\frac{N}{2}] \) factors are conditionally deterministic factors, i.e., they have null diffusion coefficients and their dynamics is driven exclusively by the other factors (and the time variable).

Nonetheless, Almeida and Vicente (2008) showed that for practical purposes, a way to bypass the issue of conditionally deterministic factors is to consider Legendre dynamic models with \( 2k \) polynomials driving the term structure whenever we believe that \( k \) stochastic factors are sufficient to drive the uncertainty of the yield curve.

In this paper, we restrict our attention to Gaussian Legendre models whose diffusion function is a constant\textsuperscript{12}. This restriction specializes Equation (10) to become:

\[
dY_t = \kappa^\mathcal{Q}(\theta^\mathcal{Q} - Y_t)dt + \Sigma dW_t^* \tag{11}
\]

Getting to the end of this section we learned that: i) in Legendre dynamic models the martingale risk-neutral restriction reduces the number stochastic factors in the state vector of the model, ii) the drift of the SDE driving the dynamics of the state vector should be affine, and iii) the diffusion should be constant (or a deterministic function of time) if we want the model to be Gaussian.

\textsuperscript{10}According to Duffie and Kan (1996), affine term structure models are those whose short term rate and bond prices are respectively affine and exponential-affine functions of the dynamic factors \( Y_t \). Note that this is exactly what happens with the Legendre model. The main question is how different should these two classes of models be? The answer lays in how different the dynamics of the state vector are within each of these classes.

\textsuperscript{11}where function \( [\cdot] \) represents the integer part of the argument.

\textsuperscript{12}Note that diffusions that are deterministic functions of time also generate a Gaussian model.
In the next section, we present the specific Legendre model that we analyze in this paper: A Gaussian model with six term structure movements, three stochastic (level, slope and curvature)\textsuperscript{13}, and three conditionally deterministic movements (those related to the Legendre polynomials of degree 3, 4 and 5). We provide the main steps for the theoretical derivation of this model.

2.3.1 The Gaussian Model with Six Factors

From Equations (8) and (9), it is straightforward to extract the term structure of interest rates as a function of the state vector under the Dynamic Legendre Model with $N$ factors ($N = 6$ here):

$$R(t, \tau) = G(\tau)'Y_t = \sum_{n=1}^{N} Y_{t,n}P_{n-1}(\frac{2\tau}{T} - 1)$$ (12)

Using the forward rate Equation (4) with $R(t, \tau)$ given in (12), and collecting terms in powers of $\tau$ as in (7) we obtain an auxiliary state vector $\tilde{Y}$, related to the forward rates by the following equation:

$$r(t, \tau) = \sum_{n=1}^{N} \tilde{Y}_{t,n}\tau^{n-1}$$ (13)

This auxiliary state vector $\tilde{Y}$ is used only to simplify the calculation of the Ito’s formula when imposing the HJM martingale condition (details appear in Almeida (2005), or Almeida and Vicente (2008)). Note, however, that recovering the original state space vector $Y$ and its SDE from $\tilde{Y}$ and its SDE should be a simple task.

To obtain $Y$ from $\tilde{Y}$, we only have to solve the linear system:

$$\tilde{Y}_{t,j} = L_j(Y_t), j = 1, 2, ..., N.$$ (14)

where each $L_j, j = 1, 2, ..., N$ comes from (7).

Similarly, there is a simple relationship between the SDE followed by $Y$ and $\tilde{Y}$. Once we impose the martingale restriction to one of them the corresponding restrictions on the other are automatically determined. We will start with restrictions imposed to the SDE of $\tilde{Y}$ and obtain the restrictions imposed to the dynamics of $Y$.

\textsuperscript{13}Litterman and Scheinkman (1991) and Heidari and Wu (2003) applying principal component analysis showed that three factors are statistically adequate to capture the variability of the term structures of U.S. Treasury bonds and swap yields, respectively. More recent papers advocate in favor of a larger number of factors (say four or five) if the goal is to forecast future yields (Duffee, 2011) or future bond returns (Cochrane and Piazzesi, 2005).
By the fact that $Y$ has Gaussian affine dynamics and since $\tilde{Y}$ is a linear transformation of $Y$, $\tilde{Y}$ will also present Gaussian affine dynamics:

$$d\tilde{Y}_t = \tilde{\mu}^Q(\tilde{Y}_t)dt + \tilde{\Sigma}dW^*_t$$

where $\tilde{\Sigma}$ is a $N \times N$ matrix\(^{14}\) and $\tilde{\mu}^Q(\tilde{Y}_t)$ is an affine function of $\tilde{Y}$.

Almeida (2005) proved that discounted bond prices are $Q$-martingales whenever the following restriction holds for the drift $\tilde{\mu}^Q(\tilde{Y}_t)$:

$$\sum_{j=2}^{N} (j-1)\tilde{Y}_{t,j}\tau^{j-2} = \sum_{j=1}^{N} (\tilde{\mu}^Q(\tilde{Y}_t))_j\tau^{j-1} - \sum_{j=1}^{N} \sum_{k=1}^{N} H_{0,jk}\tau^{j+k-1}$$

with $\tilde{\Sigma}\tilde{\Sigma}^T = \tilde{H}_0$.

By matching $\tau$ coefficients in (16), we obtain an explicit expression for the drift of $\tilde{Y}$:

$$\tilde{\mu}^Q(\tilde{Y}_t)_{2j-1} = (2j-1)\tilde{Y}_{t,2j} + \frac{\tilde{H}_{0,2j-1}}{j-1} + \frac{\tilde{H}_{0,2j-1}}{j-1} +, \ 1 \leq j \leq \left[ \frac{N}{2} \right]$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_{2j} = 2j\tilde{Y}_{t,2j+1} + \frac{\tilde{H}_{0,2j}}{j} + \frac{\tilde{H}_{0,2j+1}}{j-1}, \ 1 \leq j \leq \left[ \frac{N-1}{2} \right]$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_N = \left\{ \begin{array}{ll}
\frac{4(N-2)}{(N-1)(N-3)} & \text{if } N \text{ odd} \\
\frac{2}{N} & \text{if } N \text{ even}
\end{array} \right.$$ \hspace{1cm} (17)

where $\tilde{H}_{0,2j-1}/j$, $\tilde{H}_{0,2j-1}/j$, and $\tilde{H}_{0,2j+1}/j-1$ are zero when $j = 1$.

Further simplifying this restriction we obtain:

$$\tilde{\mu}^Q(\tilde{Y}_t)_1 = \tilde{Y}_{t,2}$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_2 = 2\tilde{Y}_{t,3} + \tilde{H}_{0,11}$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_3 = 3\tilde{Y}_{t,4} + \frac{\tilde{H}_{0,12}}{2} + \tilde{H}_{0,21}$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_4 = 4\tilde{Y}_{t,5} + \frac{\tilde{H}_{0,22}}{2} + \tilde{H}_{0,31}$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_5 = 5\tilde{Y}_{t,6} + \frac{\tilde{H}_{0,23}}{3} + \tilde{H}_{0,32}$$

$$\tilde{\mu}^Q(\tilde{Y}_t)_6 = \frac{\tilde{H}_{0,33}}{3}$$

(18)

Now we are interested in obtaining what the restrictions on the SDE of $\tilde{Y}$ impose on the SDE of the original state vector $Y$. To obtain the original state space dynamics we have to solve for (14). After some algebraic manipulations expressing $r(t, \tau)$ as a polynomial of degree five in $\tau$ and calculating each coefficient for powers of $\tau$ we explicitly obtain the

\(^{14}\)For identification purposes we follow Dai and Singleton (2002) and choose to have a diagonal matrix $\Sigma$. Note, however that the diffusion coefficient $\Sigma$ of $\tilde{Y}$ will be upper triangular since it is a linear transformation of $\Sigma$ by an upper triangular matrix $L^{-1}$ (see Equation (22) and the appendix).
linear transformation of (14)\textsuperscript{15}: \[ \tilde{Y}_t = LY_t \] (19)

where $L$ is the upper triangular matrix presented in (32) in the Appendix I.

Note that we are ready to express the drift and diffusion restrictions for the original state vector $Y$, by using (19). Let us first rewrite the drift $\tilde{\mu}^Q$ in matrix notation as an affine function of $\tilde{Y}$:

\[ \tilde{\mu}^Q(\tilde{Y}_t) = M + U\tilde{Y}_t \] (20)

where $U$ and $M$ are defined in (33) and (34) in the Appendix I.

Since $Y = L^{-1}\tilde{Y}$, the drift and diffusion of $Y$ will be:

\[ \mu^Q(Y_t) = L^{-1}\mu^Q(\tilde{Y}_t) = L^{-1}M + L^{-1}ULY_t \] (21)

\[ \sigma(Y_t) = \Sigma = L^{-1}\tilde{\Sigma} \] (22)

Equation (21) provides the functional relationship between $\kappa^Q, \theta^Q$ and the matrices $L, M, U$:

\[ \kappa^Q = L^{-1}UL \]
\[ \kappa^Q\theta^Q = L^{-1}M \] (23)

Equation (23) represents the two restrictions that the risk neutral drift in the Legendre model should satisfy to avoid arbitrages. The first makes $\kappa^Q$ predetermined before model estimation (see Eq. (37)). This is a restriction coming from the polynomial nature of the model. The second is a HJM-type restriction relating drift and volatility: The product $\kappa^Q\theta^Q$ should be a specific function of the volatility process $\Sigma \Sigma^T$, since $M$ is a function of $\Sigma$ (see Eq. (38)).

An important point to be noted is that $Y_4, Y_5,$ and $Y_6$ are deterministic factors under the Gaussian case. This is a consequence of the fact that their dynamics do not depend on the Brownian motion vector, and their drifts do not depend on the first three components of the state vector. In this case, we can write down explicitly their dynamics as functions of time and volatility parameters $\Sigma$ (see Appendix II).

2.3.2 Market Prices of Risk and Risk Premia

The element that connects factors dynamics under the risk neutral (or pricing) measure $Q$ and the physical measure, which we denominate $P$, is the market price of risk, or in

\textsuperscript{15}Noticing that the Legendre polynomials of degrees four and five are $P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$ and $P_5(x) = \frac{1}{8}(63x^5 - 70x^3 + 15x)$ and adopting $x = \frac{2\tau}{\tau} - 1$, we obtain the forward curve $r(t, \tau) = Y_{t,1} + Y_{t,2}x + \frac{Y_{t,3}}{3}(2x^2 - 1) + \frac{Y_{t,4}}{5}(5x^3 - 3x) + \frac{Y_{t,5}}{3}(35x^4 - 30x^2 + 3) + \frac{Y_{t,6}}{8}(63x^5 - 70x^3 + 15x) + \frac{2\tau}{T}\left[Y_{t,2} + 3Y_{t,3}x + \frac{Y_{t,4}}{2}(15x^2 - 3)\right] + \frac{2\tau}{T}\left[\frac{Y_{t,5}}{8}(140x^3 - 60x) + \frac{Y_{t,6}}{8}(315x^4 - 210x^2 + 15)\right]$. 

9
Duffee (2002) and Dai and Singleton (2002) tested different specifications of affine processes using U.S. treasury bonds data and concluded that a flexible time varying risk premia, with prices of risk potentially depending on all risk factors driving interest rates uncertainty, is the most appropriate to capture interest rates dynamics.

Based on their results and since we are restricting the Legendre model to the affine class, we adopt the essentially affine risk premia of Duffee (2002) implying the following form for the market prices of risk:

$$\Lambda_t = \lambda_0 + \lambda_Y Y_t,$$  \hspace{1cm} (24)

where $\lambda_0$ is a $N \times 1$ vector, $\lambda_Y$ is a $N \times N$ matrix.

The market prices of risk allow us to relate the Brownian Motion vectors under the probability measures $P$ and $Q$:

$$W^*_t = W_t + \int_0^t \Lambda_s ds,$$ \hspace{1cm} (25)

where, by Girsanov’s theorem, $W_t$ is an $N$-dimensional Brownian Motion under $P$.

Moreover, while the risk neutral dynamics of the state vector appears in (11), when we restrict market prices of risk to be essentially affine, its physical dynamics should be expressed by:

$$dY_t = \mu(Y_t)dt + \Sigma dW_t,$$ \hspace{1cm} (26)

where $\mu(Y_t) = \kappa(\theta - Y_t)$.

A substitution of Eq. (24) in (25), another of (25) in (11) and, matching (11) with (26) yields:

$$\kappa = \kappa^Q + \Sigma \lambda_Y,$$ \hspace{1cm} (27)

$$\kappa \theta = \kappa^Q \theta^Q + \Sigma \begin{bmatrix} \lambda^1_0 \\ \vdots \\ \lambda^N_0 \end{bmatrix}$$ \hspace{1cm} (28)

### 2.3.3 Exact Estimation of the Model

Understanding the role of each parameter driving factors’ dynamics under probability measures $P$ and $Q$ is fundamental, during the estimation process. Parameters under the pricing measure $Q$ are used to fit model implied yields to observed yields, while parameters under

---

16 The state price deflator is a strictly positive process with the property that it turns deflated price processes into martingales under the physical measure. Absence of arbitrage implies the existence of at least one state price deflator (Duffie, 2001).
the physical measure $P$ determine factors’ conditional probability distributions used as inputs when the estimation technique is Maximum Likelihood (Duffie and Singleton, 1997; DS, 2002), or Quasi Maximum Likelihood (Duffee, 2002; Duarte, 2004). Those techniques are implemented either with direct inversion of the state vector from a subset of yields like in Chen and Scott (1993) or with the use of a linear Kalman filter that assumes that all yields are priced with error like in Duan and Simonato (1999).

In order to better understand the estimation process, we first have to identify which parameters should be estimated and what is their relative importance during estimation. Note, from Eq. (23), that the risk-neutral drift of the Gaussian model is known a priori as a function of matrices $L, U, M$, having no free parameters to be estimated except for its dependence in $\Sigma$ through matrix $M$. Therefore, the parameters of interest will be those appearing on the drift under the physical probability measure $(\kappa, \theta)$, the volatility parameters $(\Sigma_{11}, \Sigma_{22}, \Sigma_{33})$, and the initial values for the deterministic factors $(Y_{0,4}, Y_{0,5}, Y_{0,6})$.

It is important to emphasize that, in principle, the volatility parameters are the only reason that prevent us from separating the estimation process into two stages. As seen before, separation in two stages would be useful since it would avoid having to solve an optimization process back and forth between the two probability measures. The reason for failure of separation is that the latent factors’ volatility parameters are necessary to characterize the deterministic factors that on their turn are necessary to determine the latent factors, and only a full optimization of the model dealing simultaneously with the parameters under the risk neutral and physical probabilities would solve this problem. We will see how to relax this condition in the model approximation proposed below.

In this paper, to estimate the exact dynamic model with full optimization, we follow Chen and Scott (1993) adopting a Maximum Likelihood Estimation procedure with inversion of the state vector directly from a set of observed yields.

Once we are able to choose the three yields to be used in the inversion of the state vector, for a set of fixed volatility parameters and initial values of deterministic factors we first calculate the dynamics of the deterministic factors (see Appendix II), and then extract the contribution that they have on the values of the three chosen yields for inversion. After this step, inversion of the state vector is immediately obtained by solving a linear system where the three chosen yields minus the contribution of the determinist factors are the dependent variables and the three Legendre polynomials are the independent variables.

\[17\] In a recent paper, Duffee and Stanton (2012) provide empirical evidence against the Efficient Method of Moments in estimation of term structure models for small sample data.

\[18\] Following Dai and Singleton (2002) we set the vector $\theta$ equal to zero for econometric identification purposes. We experienced the estimation with free values for $\theta$ and observed that it does not affect any of the important results of the dynamic model, including the implied state vector and risk-premium dynamics.
The remaining yields are priced with i.i.d Gaussian errors with unknown variances that are also estimated parameters under the maximization of the likelihood function (see Chen and Scott, 1993 for more details).

2.3.4 Estimation by Approximation of the State Vector

The Six-Legendre Polynomial Gaussian Dynamic Model presents three risk factors: level, slope and curvature and three deterministic factors. The only reason for the existence of the deterministic factors is model internal consistency with the HJM conditions. Since these factors are not able to capture features of the stochastic behavior of the term structure they should be small by nature. In our empirical analysis, reinforcing the empirical results obtained by Almeida and Vicente (2008) with U.S. zero coupon data, we show that the three deterministic factors have negligible influence on the implied values of the dynamic level, slope, and curvature factors of the Brazilian analyzed zero coupon curve.

Due to the estimation cost of a full optimization of the dynamic model, we propose an approximation to simplify estimation. First, we determine the vector of the three risky factors by running the linear regression in Equation (2) using only the three-yields chosen for inversion in the exact estimation of the dynamic model described in section 2.3.3. With the time series of the approximated latent factors in hands, we estimate the equations for their dynamics under the physical probability measure via a discrete version of the autoregressive process appearing in Equation (26).

Now, by Equation (23) that determines $\kappa^Q, \theta^Q$ as a function of matrices $L, U, M$, we observe that the risk neutral parameters depend only on the volatility parameters $\Sigma$. Thus, once we have estimated the parameters under the physical measure, which includes $\Sigma$, the dynamics under the risk neutral measure $Q$ is completely determined. Knowledge of the parameters under the physical measure then allow us to easily extract the market prices of risk by using Equations (27) and (28).

it is important to emphasize that making use of a pair of linear regressions, one to approximate the state vector and one to fit an autoregressive process for the extracted state vector, we are able to approximate the risk premia embedded on the original set of zero coupon yields.

The approximated state vector could have been obtained either by running the regression in Equation (2) adopting all observed yields or adopting only the yields used to invert the state vector in the exact estimation of the dynamic model. We decided to use only

---

19 We discuss our choice of adopting only the three yields in the regression two paragraphs below.

20 To be more precise, to approximate the state vector we run a sequence of $T$ cross-sectional linear regressions where $T$ is the number of dates in the dataset adopted.
the yields chosen to invert the state vector in the exact estimation of the dynamic model to isolate the effects of the deterministic Legendre factors on the difference between the approximated state vector and the original state vector, from effects of pricing errors, which would appear if all yields were included. In fact, if we were adopting all yields in the regression, an extra term distinguishing the approximated state vector from the exact state would come from having the remaining yields not chosen to invert the state directly affecting the approximate state. The use of such regression would be justified in an approximation of the state vector if the original exact model was estimated with a linear Kalman filter instead of with the Chen and Scott (1993) method.\footnote{In any case, we experimented with both regressions (using only the three yields versus all yields) and identified that the influence of the additional yields has a magnitude three times bigger than the influence of the deterministic factors. Results of the approximation adopting all yields in the regression are available upon request.}

Next, we estimate both the dynamic model and our suggested approximation and show that the risk premia parameters and processes are very similar under the two distinct methods of estimation.

3 Empirical Results

3.1 The Brazilian Swap Market

The One-Day Inter Bank Deposit Future Contract (ID-Future) with maturity $T$ is a future contract whose underlying asset is the accumulated daily ID rate\footnote{The ID rate is the average one-day interbank borrowing/lending rate, calculated by CETIP (Central of Custody and Financial Settlement of Securities) every workday. The ID rate is expressed in effective rate per annum, based on 252 business-days.} capitalized between the trading time $t$ ($t \leq T$) and $T$. The contract size corresponds to R$ 100,000.00 (one hundred thousand Brazilian Reals) discounted by the accumulated rate negotiated between the buyer and the seller of the contract.

This contract is very similar to a zero coupon bond, except that it pays margin adjustments every day. Each daily cash flow is the difference between the settlement price\footnote{The settlement price at time $t$ of an ID-Future with maturity $T$ is equal to R$ 100,000.00 discounted by its closing yield quotation.} on the current day and the settlement price on the day before corrected by the ID rate of the day before.

A swap that changes the floating rate of an ID-future for a fixed rate is quoted as a zero coupon bond. At time $t$, the yield to maturity $sw$, of a swap with time to maturity $\tau = T - t$ is given by $P_t = 100,000e^{-sw\tau}$, where $P_t$ is the settlement price of the corresponding ID-future with time to maturity equal to $\tau$.

The Brazilian Mercantile and Futures Exchange (BM&F) is the entity that offers the
ID-Future and the swaps. The number of authorized contract-maturity months is fixed by BM&F (on average, there are about twenty authorized contract-maturity months for each day). Contract-maturity months are the first four months subsequent to the month in which a trade has been made and, after that, the months that initiate each following quarter. Expiration date is the first business day of the contract-maturity month. BM&F maintains a historical database with interest rates on swaps with different fixed maturities synthetically constructed from the prices of ID-futures. The database can be accessed at "www.bmf.com.br".

In this paper, data consists of a set of historical series of Brazilian interest rates swaps obtained from BM&F with a daily frequency, for maturities of 30, 60, 90, 120, 180, 360 and 720 days. Figure 2 presents their historical evolution during the sample period from January 1, 2005 to December 31, 2011. Note that at the beginning of the sample (2005) yields had high values around 18%. Then, with the gradual improvement of macroeconomic conditions they had decreased to a level around 8% before the financial crisis, but ended up reaching values around 11% in the end of our sample (2011).

Applying principal component analysis (PCA) to the current sample we identify that three principal components (PCs) explain around 86% of the term structure movements. Since Almeida, Duarte and Fernandes (2003) showed that Legendre factors extracted by running cross-sectional regressions (section 2.1) will be specific rotations of PCs of yields, the PCA analysis suggests that the static Legendre model with three-factors will capture the very same percent of yields' variability captured by the first three PCs.

3.2 Legendre Dynamic Model: Exact Estimation

We start by running for each day the cross-sectional linear regressions of Eq. (2) in section 2.1, using the first three Legendre polynomials (constant, linear and quadratic) to capture different movements of the term structure of interest rates. These regressions will help us to choose the best yields (in the mean square error sense) to be used in the inversion of the state vector on the estimation of the exact dynamic Gaussian Legendre model.

Figure 3 presents the time series of the three Legendre term structure factors and Table 4 presents their descriptive statistics. The factor attached to the constant Legendre polynomial represents the level factor. It has mean and standard deviation of respectively

24The inclusion of a fourth stochastic Legendre factor in the regressions would increase the percent of yields’ variance captured by Legendre factors but with the cost of having a bunch of extra Legendre forward rate calculations. Since the empirical section works more as an illustration for our risk premia approximation than a detailed empirical application, we chose not to include a fourth factor to avoid these extra calculations. It is worth mentioning that these calculations, which are only used to define the linear transformations that appear in Appendix I, are performed only one time before model estimation therefore not directly affecting the optimization steps of the estimation process.
12.94% and 2.70%. The factor attached to the Legendre polynomial of degree 1 is the slope factor, with low values associated with flat term structures and high values with steep term structures indicating expectations of future risks in lending money for short and medium term maturities. It presents mean and standard deviation much smaller than the level factor, of respectively 0.18% and 0.86%. The factor attached to the Legendre polynomial of degree 2 is the curvature factor and it basically indicates the degree of concavity of the term structure. Negative values of the curvature factor are associated to a concave term structure. Its mean and standard deviations are the smallest among the three factors, being respectively of 0.06% and 0.35%.

By the fact that these factors are built based on our parametric term structure model rather than a principal component analysis procedure, their time series should be potentially correlated. Table 1 presents their correlation structure. Note that similarly to results obtained with the use of other parametric term structure models such as the exponential model of Diebold and Li (2006), the three factors have indeed non-negligible correlations. In particular, the negative correlation between level and slope indicates that increases on the level are linked to further increases of short-term rates. Similarly, the negative correlation between slope and curvature suggests that increases in short-term rates due to the slope come together with increases of longer-term rates due to the curvature factor.

We use the time series of the Legendre static factors to help us to identify which swap rates should be priced without error when using the Chen and Scott (1993) methodology. From a qualitative point of view we would like to choose one short-term rate, one medium-term rate and one long-term rate to invert from, since it will better represent the dynamic properties of the yield curve as a whole.

The statistical properties of the residuals obtained with the static fitting procedure shown in Table 2 indicate that the residuals for the yields with 60 and 720 days to maturity present the smallest bias and standard deviations among all alternatives. We adopt these two yields (60, 720) together with the 360-day yield (since this is usually a liquid intermediate maturity in the swap market) to invert the state vector when estimating the model by Maximum Likelihood. Following Chen and Scott (1993), the remaining yields with maturities of 30, 90, 120 and 180 days are priced assuming i.i.d Gaussian errors.

Figure 4 presents the time series of the three Legendre term structure factors obtained with the exact estimation of the dynamic model. Note that the time series of all three factors are very similar to those of the corresponding static factors in figure 3. Their similarities are more formally confirmed when we compare the basic statistical properties (i.e., mean and standard deviation) of these dynamic factors with the static ones. The level factor has mean and standard deviation of respectively 12.95% and 2.68%, the slope factor
presents mean and standard deviation of respectively 0.17% and 0.87%, and the mean and standard deviation of the curvature factor are respectively 0.04% and 0.32%. Of course, as is usual with factor analysis, the differences between static and dynamic factors increase for higher order factors (such as the curvature) but are still small.

Table 3 presents parameter’s values, their standard deviations calculated by the Outer Product Method (BHHH), and the ratio \( \frac{\text{value}}{\text{std value}} \), which allows for the asymptotic analysis of parameter’s significance. Bold ratios indicate significant parameters at a 95% confidence level. Note that most parameters are statistically significant, except for some parameters related to the slope factor (mean reversion and one risk premium parameter). First, we can observe that the speeds of mean reversion of the three factors are very distinct with the level being the slowest mean-reverting factor and curvature the fastest, with speed 30 times bigger than the level speed.

By observing Figure 4 we can see how these factors have distinct cycles. Indeed, while curvature oscillates a lot around its long term mean, in contrast, the level doesn’t complete even one cycle during the whole sample period. The diffusion coefficients associated to each Brownian Motion follow the same relative order of magnitude of the standard deviations of the dynamic factors of figure 4, with level having higher diffusion volatility coefficient than slope (0.014 versus 0.012), and slope having higher diffusion volatility coefficient than the curvature (0.012 versus 0.009).

Due to our choice of econometric identification structure that follows Duffee (2002) and Dai and Singleton (2002), the risk premium of the three factors is estimated based on an upper triangular matrix \((\lambda_Y)\), with the level risk premia allowing for a feedback of the three factors, the slope premium allowing for feedback of slope and curvature, and curvature premium being affected by only the curvature factor itself. Given this structure, we note that the three market prices of risk parameters linked to the level factor \((\lambda_Y(1, 1), \lambda_Y(1, 2), \lambda_Y(1, 3))\) are statistically significant. This is an indication that the three factors have influence on the level risk premium. In contrast, only the curvature factor is statistically significantly responsible for the slope risk premia (see parameter \(\lambda_Y(2, 3)\)). The curvature risk premium is also significantly affected by the curvature factor (see \(\lambda_Y(3, 3)\)).

Table 5 presents statistical properties of the residuals for the maturities 30, 90, 120, and 180 days. The residuals have means and standard deviation values comparable to results that appeared in the literature using the same estimation method, but for U.S. treasury data (Dai and Singleton (2000)) as well as for Russian Brady Bonds data (Duffie, Pedersen and Singleton (2003)).

All that was presented up to this point is standard in the affine term structure literature. A point that differs though is the existence of deterministic factors in this Gaussian
Legendre model. For this reason, let us temporarily concentrate our analysis on the three deterministic factors appearing in Figure 5 to understand how they affect the dynamic model. Note that they are all small when compared to the magnitude of the three factors driving the uncertainty of the term structure. $Y_4$ is the one that varies more, with variation between -13 and 7 bps and mean value of 0.95 bps. $Y_5$ and $Y_6$ range between -1 and 3 bps, with means of 0.2 and -0.03 bps, respectively. Our intention is to identify how much these deterministic factors affect the truly stochastic factors. Hopefully, if they don’t affect much, this will justify our approximation that just neglects the existence of these factors in the estimation process.

For a fixed set of model parameters $(\kappa, \theta, \Sigma, Y_{0,4}, Y_{0,5}, Y_{0,6})$, at each time $t$, the implied stochastic factors, which are the first three variables in the state vector $Y_{t,1}, Y_{t,2}, Y_{t,3}$, are extracted using the following linear system:

$$Sw_t^{\text{exact}} - \left[P_3(x_{\text{exact}})P_4(x_{\text{exact}})P_5(x_{\text{exact}})\right] = \left[P_0(x_{\text{exact}})P_1(x_{\text{exact}})P_2(x_{\text{exact}})\right] = \left[Y_{t,1}, Y_{t,2}, Y_{t,3}\right]$$

(29)

where $Sw_t^{\text{exact}}$ denotes the vector of yields priced exactly, $P_i(x_{\text{exact}})$ the Legendre polynomial of degree i, $x_{\text{exact}}$ is a vector of transformed maturities $x_{\text{exact}} = \frac{\tau_{\text{exact}}}{2} - 1$, $\tau_{\text{exact}}$ the maturities of the exactly priced yields, and $Y_{t,4}, Y_{t,5}, Y_{t,6}$ the deterministic factors that are known functions of time and of the subset of parameters $(\Sigma, Y_{0,4}, Y_{0,5}, Y_{0,6})$ (see Appendix II).

The influence that the deterministic factors have on the three truly stochastic factors is closely related to the system above. The question is how much the vector of linear combinations $P_3(x_{\text{exact}})Y_{t,4} + P_4(x_{\text{exact}})Y_{t,5} + P_5(x_{\text{exact}})Y_{t,6}$ affect the yields priced exactly and consequently the values of the stochastic Legendre factors. By the facts that the deterministic factors are small and that the Legendre polynomials form an orthonormal basis, we conjecture that these linear combinations will be small. Figure 6 confirms the intuition that their influence is negligible.

This picture presents how much the deterministic factors (obtained with the parameters that maximize the Likelihood function) affect the three stochastic factors (level, slope and curvature). It confirms that their influence is negligible. In fact, for most of the sample period, the deterministic factors affect the level factor in less than 1 bp, the slope factor in less than 5 bps and the curvature in less than 2 bps. This means that the model restricts the initial values $Y_{0,4}, Y_{0,5}$ and $Y_{0,6}$, and the parameters in $\Sigma$ to guarantee that the influence of the deterministic factors is minimal.
3.3 Legendre Dynamic Model: Proposed Approximation

The previous results concerning the deterministic factors indicate that we may discard them when estimating the state vector. The elimination of the deterministic factors allows us to propose a simple approximation of the state vector of the dynamic model by only solving for each time $t$ the following linear system:

$$
S u_t^{\text{exact}} = [P_0(x_{\text{exact}})P_1(x_{\text{exact}})P_2(x_{\text{exact}})] \begin{bmatrix} Y_{t,1} \\ Y_{t,2} \\ Y_{t,3} \end{bmatrix}, \quad t = 1, \ldots, T_{\text{final}}.
$$

(30)

The most interesting feature of this approximation is that the linear systems appearing in Eq. (30) do not depend on any model parameters and therefore can be solved without an optimization process. Once we solve the sequence of linear systems we end up with a time series for each stochastic factor. Given that, estimation of a vector autoregressive process of lag one for these three factors will deliver values of the parameters under the physical probability $P^{25}$.

Table 6 shows the statistical properties of the difference between the state vector estimated under the dynamic model and the state vector obtained with our approximation: For the three factors, the differences have means smaller than one basis point, mean absolute values smaller than 2 bps, and standard deviations smaller than 2.5 bps, confirming the fact that the deterministic factors have very small influence on the estimation process of the risky factors (the truly stochastic level, slope and curvature). On average they affect the truly stochastic factors in less than 3% of their magnitude, at most. Also comparing the properties of these differences with the properties of the residuals of the yields priced with error we observe that the magnitude of the residuals is three to four times higher than the magnitude of these differences suggesting one more time that the deterministic factors do not affect much the estimation of the three truly stochastic factors in the state vector.

For our proposed approximation, parameters estimates and statistical properties of the residuals of yields priced with error appear in Tables 7 and 8. Standard deviations are calculated based on small numerical perturbations of the autoregressive parameters ($\kappa$ and $\Sigma$), followed by an application of the delta method (Casella and Berger, 2002) for the risk-premium parameters, justified by the fact that these parameters are functions of $\kappa$ and $\Sigma$. Note that both implementations of the model agree on the statistical significance of

---

25If we want to restrict the autoregressive coefficient matrix to have some zeros like is usually done in the affine term structure literature, than instead of running a linear regression we should maximize a Gaussian likelihood but with a known predetermined state vector. This has the same computational cost of a linear regression and is much simpler than maximizing a likelihood with respect to parameters and simultaneously determining the implied states.
all parameters at a 95% confidence level. Moreover, the estimated parameters under the approximation are very close to the ones obtained under the exact dynamic model (compare tables 3 and 7). The same happens with the statistical properties of the residuals. Results overall indicate that our approximation of the state vector generates an accurate proxy for model parameters and their statistical significances.

### 3.4 Implied Risk Premia

Both implementations of the model agree on the statistical significance of the time varying risk premium parameters appearing in $\lambda_Y$: The only entry that is not significant at a 95% confidence interval is $\lambda_Y(2, 2)$. This number measures how much the slope factor influences its own premium.

We note that the three risky factors have significant effects on the level risk premium, as we can see with parameters $\lambda_Y(1, 1)$, $\lambda_Y(1, 2)$ and $\lambda_Y(1, 3)$, which are all significant at a 95% confidence. Once $\lambda_Y(1, 2)$ is 26 times higher than $\lambda_Y(1, 1)$ and the mean absolute level factor is only 15 times higher than the mean absolute slope factor, we can infer that slope is an important component of level risk premia. For instance, during the crisis, when the slope was negative, the level risk premia was higher due to the slope contribution. This may be associated with the fact that monetary policy shocks are usually estimated to have a slope effect on the yield curve, affecting short term rates and, in this particular case as shown by the significance of parameter $\lambda_Y(1, 2)$, affecting the risk premia associated with the level factor$^{26}$.

Although the slope is an important component of the level risk-premium it does not significantly affect its own risk premium, i.e., the parameter $\lambda_Y(2, 2)$ is not significant under both implementations of the model. In contrast, the curvature factor has significant effect on the risk premia charged on all fundamental sources of uncertainty on the term structure$^{27}$ directly indicating that this movement plays an important role in the term structure dynamics.

For our Gaussian Legendre model, the time $t$ instantaneous expected excess return of a zero-coupon bond with maturity $\tau$ is given by:

$$
e^i_{t, \tau} = [P_0(\tau)P_1(\tau)P_2(\tau)]\Sigma \Lambda_Y Y_t$$

Equation (31) indicates that the instantaneous expected excess return is a linear combination of the factors of the model with weights on specific factors coming from a combination

$^{26}$For an interesting macroeconomic interpretation of latent factors in an essentially affine dynamic term structure model see Dewachter and Lyrio (2006).

$^{27}$That is, parameters $\lambda_Y(1, 3)$, $\lambda_Y(2, 3)$ and $\lambda_Y(3, 3)$ are all significant at a 95% confidence.
of parameters from matrices $\lambda Y$, and $\Sigma$ and also from Legendre polynomial terms that are maturity-dependent.

Figure 7 shows the risk premium obtained for two different bond maturities under both implementations of the model (30-day and 360-day maturities). We observe that the time series of the risk premium implicit in these two bonds for both versions of the model move on the top of each other for a large part of the sample, and are highly correlated on the remaining parts of the sample. This in part reassures that the approximation works as a good substitute for the original more complex implementation of the dynamic model.

Table 9 gives more precise information about the differences in risk premium between the two implementations of the model for all yields adopted in the estimation process. Mean values of the differences in risk-premiums range from 21 bps for the shortest-term yield (30-day maturity) to 43 bps for the longest-term yield (2-year maturity), values that represent something around 10% to 30% of the absolute value of the original risk-premium obtained under the exactly estimated dynamic model. Nonetheless, note that standard deviations of these differences are also high ranging from 36 to 65 bps indicating that although the similarities between the two implementations are not perfect they still motivate a substitution of the dynamic model by the much simpler approximation. Indeed, the correlation of pairs of risk-premium time series for the exact and the approximated models is higher than 99% for all maturities: 0.9995, 0.9996, 0.9997, 0.9998, 0.9999, 0.9997, and 0.9989, respectively for the 30-, 60-, 90-, 120-, 180-, 360-, and 720-day maturities.

4 Model Extensions

4.1 Risk Premia Approximation in Interest Rate Markets with Coupon-Bearing Bonds

Almeida, Duarte and Fernandes (2003) used the static version of the Legendre model to generate zero-coupon yield curves in markets with coupon-bearing bonds (Brady bonds and Global bonds). The idea is very similar to the static version of the model described in section 2.1 but instead of running linear regressions one should run nonlinear regressions that try to minimize the quadratic distance between observed prices of bonds and their corresponding discounted cash-flows. They show that once a sequence of regressions is performed, a sequence of synthetic zero-coupon term structures driven by the Legendre polynomial factors is obtained. Moreover, they show that it is possible to identify the principal components of this panel of synthetic zero-coupon term structures through specific rotations of the Legendre polynomial time-series factors.

\[\text{Results for the other maturities are qualitatively very similar and are available from the authors upon request.}\]
We suggest that, apart from the error terms in the original prices of coupon-bearing bonds\textsuperscript{29}, once these synthetic zero-coupon term structures are extracted from the original set of coupon-bearing bonds we can proceed with the same methodology proposed in this paper and, assuming that the Legendre factors follow a dynamic Gaussian model, obtain the parameters under the physical probability measure $P$ by running a vector autoregression on the vector of time series of the Legendre factors obtained with the nonlinear regressions. In this case, instead of running a set of linear regressions and one vector autoregressive estimation, we would have a set of nonlinear cross-sectional regressions and the same vector autoregressive estimation.

Treating the synthetic zero-coupon term structures as our initial data, with knowledge of the physical parameters, risk premia can be extracted exactly as done in this paper.

\textbf{4.2 Separation of the Cross Sectional and Time Series Roles Under More General Model Specifications}

Restricting the loadings to be polynomials makes the parameter estimation process to be focused on factors dynamics under the physical probability measure $P$, rather than under $Q$. The only elements which prevent the estimation process to be completely separated under measures $P$ and $Q$ are the conditionally deterministic factors\textsuperscript{30}. However, in Section 3, we offered empirical evidence for the Gaussian case, that the only role for the conditionally deterministic factors is to create specific restrictions on the vector space in terms of choices for the volatility parameters, not substantially affecting the implied values for the truly stochastic Legendre factors. Based on this observation, we proposed extracting an approximation for the state vector by simply neglecting the conditionally deterministic factors and running cross-sectional regressions with the term structure parameterized by the truly stochastic Legendre polynomials factors. The state-vector approximation allowed us to come up with a simple method to extract interest rate risk premia in a Gaussian polynomial model.

Although we only showed that the approximation of the state vector is valid under the Gaussian model, the parametric structure of the model suggests that it should be robust under other versions of the dynamic model, i.e., under other probability distributions for the state variables\textsuperscript{31}. The reason is simple. The difference between the state vector obtained with the independent regressions (linear systems) and the state vector obtained under any

\textsuperscript{29}That might become large if the number of bonds is too big.

\textsuperscript{30}This can be noted through Equations (44)-(46) that present the dependence of these conditionally deterministic factors on parameters that define the volatility structure of the whole set of factors.

\textsuperscript{31}For instance, instead of having state variables with a Gaussian distribution it could be a mixed distribution coming from a combination of Gaussian and Feller stochastic processes.
version of the exact dynamic model is controlled by the conditionally deterministic factors. The conditionally deterministic factors, on their turn, will be negligible as long as we choose to implement the dynamic model with a number of truly stochastic factors equal to the number of principal components (PCs) needed to capture most of the variability of the term structure movements. That is, there should be a correspondence between non-negligible principal components and truly stochastic Legendre factors (See Almeida, Duarte and Fernandes, 2003), meaning that the conditionally deterministic factors may be identified with negligible principal components of the term structure.

This opens space to use the state vector approximation to extract more information about the state-vector dynamics, even when the model is not Gaussian. For instance, we can obtain qualitative information regarding the dynamics of the state vector. To that end, after the extraction of the approximated time series of the state vector, we may adjust ARMA + GARCH models to these time series to test if specifications involving stochastic volatility would be appropriate. We could also use a more sophisticated non-parametric test (Hong and Li (2005)) to verify the adequacy of introducing stochastic volatility in the dynamics of the state vector.

5 Conclusion

In this paper, we analyzed a Gaussian version of the Legendre polynomial arbitrage-free term structure model. Its parametric nature, which predetermines the loading functions of the term structure, allows for an approximation of the state vector of the dynamic model by solving a sequence of independent linear systems (across time). This approximation was used to provide a fast method to extract risk premium from an interest rate market with zero-coupon yields. In addition, we briefly describe how this method can be extended to consider: i) factor dynamics more general than the Gaussian; ii) markets with coupon bearing bonds.
References


6 Appendix I - Analytical and Numerical Matrices for the Six-Factor Legendre Dynamic Model

In this appendix, we present some of the matrices used on the specialized Gaussian version of the model. It contains the main matrices that allow for the relation between a power polynomial model and the Legendre polynomial model.

\[
L = \begin{bmatrix}
  1 & -1 & 1 & -1 & 1 & -1 \\
  0 & \frac{4}{l} & \frac{-12}{l^2} & \frac{24}{l^3} & \frac{-40}{l^4} & \frac{60}{l^5} \\
  0 & 0 & \frac{18}{l^2} & \frac{-90}{l^3} & \frac{270}{l^4} & \frac{-630}{l^5} \\
  0 & 0 & 0 & \frac{80}{l^3} & \frac{-560}{l^4} & \frac{2450}{l^5} \\
  0 & 0 & 0 & 0 & \frac{350}{l^4} & \frac{-3150}{l^5} \\
  0 & 0 & 0 & 0 & 0 & \frac{1512}{l^5}
\end{bmatrix}
\]  

(32)

\[
M = \begin{bmatrix}
  0 \\
  \tilde{H}_{0,11} \\
  \tilde{H}_{0,12} + \tilde{H}_{0,21} \\
  \frac{\tilde{H}_{0,22}}{2} + \tilde{H}_{0,31} \\
  \frac{\tilde{H}_{0,23}}{3} + \frac{\tilde{H}_{0,32}}{2} \\
  \frac{\tilde{H}_{0,33}}{3}
\end{bmatrix}
\]  

(33)

\[
U = \begin{bmatrix}
  0 & 1 & 0 & 0 & 0 & 0 \\
  0 & 0 & 2 & 0 & 0 & 0 \\
  0 & 0 & 0 & 3 & 0 & 0 \\
  0 & 0 & 0 & 0 & 4 & 0 \\
  0 & 0 & 0 & 0 & 0 & 5 \\
  0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]  

(34)

\[
L = \begin{bmatrix}
  1 & -1 & 1 & -1 & 1 & -1 \\
  0 & 2 & -6 & 12 & -20 & 30 \\
  0 & 0 & 4.50 & -22.50 & 67.50 & -157.50 \\
  0 & 0 & 0 & 10 & -70 & 306.50 \\
  0 & 0 & 0 & 0 & 21.88 & -196.88 \\
  0 & 0 & 0 & 0 & 0 & 47.25
\end{bmatrix}
\]  

(35)

\[
L^{-1} = \begin{bmatrix}
  1 & 0.5000 & 0.4444 & 0.5000 & 0.6400 & 0.6111 \\
  0 & 0.5000 & 0.6667 & 0.9000 & 1.2800 & 1.4048 \\
  0 & 0 & 0.2222 & 0.5000 & 0.9143 & 1.3095 \\
  0 & 0 & 0 & 0.1000 & 0.3200 & 0.6852 \\
  0 & 0 & 0 & 0 & 0.0457 & 0.1905 \\
  0 & 0 & 0 & 0 & 0 & 0.0212
\end{bmatrix}
\]  

(36)
7 Appendix II - Deterministic Factors in the Gaussian Six-Factor Legendre Dynamic Model

The factors $Y_4$, $Y_5$, and $Y_6$ are deterministic under the Gaussian case. As indicated before in the body of the article, we adopt the same econometric identification as Dai and Singleton (2002) where $\Sigma$ is a diagonal matrix. From Equation (22) we directly conclude that:

$$
\sigma(Y_t)\sigma(Y_t)^T = L^{-1}\tilde{H}_0(L^{-1})^T = H_0
$$

or equivalently:

$$
\tilde{H}_0 = L\Sigma^2((L^{-1})^T)^{-1} = L\Sigma^2L^T
$$

We use matrix $L$ defined by Equation (35) which uses the fact that the largest maturity of observed Brazilian swaps in our data base is 2 years.

Equation (40) combined with Equation (38) inform that the contribution to the drift of each of the state variables $Y_4$, $Y_5$, and $Y_6$ coming from $L^{-1}M$ can be expressed as a linear combination of squared diagonal elements of matrix $\Sigma$, as shown in what follows. Explicitly
substituting the value of \( L \) presented in subsection 2.3 we obtain:

\[ \tilde{H}_0 = \begin{bmatrix} \Sigma_{11}^2 + \Sigma_{22}^2 + \Sigma_{33}^2 & -2\Sigma_{22}^2 - 6\Sigma_{33}^2 & 4.5\Sigma_{33}^2 & 0 & 0 & 0 \\ -2\Sigma_{22}^2 - 6\Sigma_{33}^2 & 4\Sigma_{22}^2 + 36\Sigma_{33}^2 & -27\Sigma_{33}^2 & 0 & 0 & 0 \\ 4.5\Sigma_{33}^2 & -27\Sigma_{33}^2 & 20.25\Sigma_{33}^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \] (41)

Using Equations (38) and (41):

\[
\begin{align*}
(L^{-1}M)_4 &= 0.2\Sigma_{22}^2 + 2.6751\Sigma_{33}^2 \\
(L^{-1}M)_5 &= 0.2577\Sigma_{33}^2 \\
(L^{-1}M)_6 &= 0.1431\Sigma_{33}^2
\end{align*}
\] (42)

From Equations (21), (37) and (42) we obtain the drift of the state variables \( Y_4 \), \( Y_5 \), and \( Y_6 \):

\[
\begin{align*}
\mu^Q(Y_t)_4 &= 0.2\Sigma_{22}^2 + 2.6751\Sigma_{33}^2 + 8.75Y_{t,5} - 3.15Y_{t,6} \\
\mu^Q(Y_t)_5 &= 0.2577\Sigma_{33}^2 + 10.8Y_{t,6} \\
\mu^Q(Y_t)_6 &= 0.1431\Sigma_{33}^2
\end{align*}
\] (43)

Explicitly solving the simple ODE’s implied for these factors, we get:

\[
Y_{t,4} = Y_{0,4} + (0.2\Sigma_{22}^2 + 2.6751\Sigma_{33}^2 + 8.75Y_{0,5} - 3.15Y_{0,6})t + (1.8041\Sigma_{33}^2 + 94.5Y_{0,6})\frac{t^2}{2} + 13.5231\Sigma_{33}^2\frac{t^3}{6}
\] (44)

\[
Y_{t,5} = Y_{0,5} + (0.2577\Sigma_{33}^2 + 10.8Y_{0,6})t + 1.5455\Sigma_{33}^2\frac{t^2}{2}
\] (45)

\[
Y_{t,6} = Y_{0,6} + 0.1431\Sigma_{33}^2 t
\] (46)

Note that the dynamics of the state variables \( Y_{t,4} \), \( Y_{t,5} \) and \( Y_{t,6} \) are deterministic and completely determined by the parameters \( \Sigma_{22} \) and \( \Sigma_{33} \), and the initial conditions \( Y_{0,4} \), \( Y_{0,5} \) and \( Y_{0,6} \), which are also treated as parameters of the model.
<table>
<thead>
<tr>
<th>Factor</th>
<th>Level</th>
<th>Slope</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Slope</td>
<td>-0.54</td>
<td>1.00</td>
</tr>
<tr>
<td>Curvature</td>
<td>-0.13</td>
<td>-0.59</td>
</tr>
</tbody>
</table>

Table 1: Correlation Matrix of the Legendre Static Factors.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Mean (bp)</th>
<th>Mean Abs (bp)</th>
<th>Standard Deviation (bp)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>1.80</td>
<td>6.13</td>
<td>7.64</td>
<td>-0.25</td>
<td>3.88</td>
</tr>
<tr>
<td>60</td>
<td>0.17</td>
<td>2.35</td>
<td>3.23</td>
<td>-0.40</td>
<td>11.82</td>
</tr>
<tr>
<td>90</td>
<td>-1.15</td>
<td>3.53</td>
<td>4.57</td>
<td>-0.05</td>
<td>4.61</td>
</tr>
<tr>
<td>120</td>
<td>-0.05</td>
<td>5.20</td>
<td>6.74</td>
<td>-0.28</td>
<td>5.09</td>
</tr>
<tr>
<td>180</td>
<td>-2.49</td>
<td>5.87</td>
<td>6.95</td>
<td>0.50</td>
<td>3.61</td>
</tr>
<tr>
<td>360</td>
<td>2.06</td>
<td>6.30</td>
<td>7.55</td>
<td>-0.56</td>
<td>3.38</td>
</tr>
<tr>
<td>720</td>
<td>-0.34</td>
<td>1.09</td>
<td>1.31</td>
<td>0.53</td>
<td>3.40</td>
</tr>
</tbody>
</table>

Table 2: Statistical Properties of the Residuals of the Static Model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error</th>
<th>ratio (\frac{\text{abs(Value)}}{\text{Std Err}})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\kappa_{11})</td>
<td>0.141</td>
<td>0.046</td>
<td>3.05</td>
</tr>
<tr>
<td>(\kappa_{12})</td>
<td>-1.608</td>
<td>1.468</td>
<td>1.10</td>
</tr>
<tr>
<td>(\kappa_{22})</td>
<td>0.557</td>
<td>1.355</td>
<td>0.41</td>
</tr>
<tr>
<td>(\kappa_{13})</td>
<td>10.303</td>
<td>3.781</td>
<td>2.72</td>
</tr>
<tr>
<td>(\kappa_{23})</td>
<td>-4.710</td>
<td>3.274</td>
<td>1.44</td>
</tr>
<tr>
<td>(\kappa_{33})</td>
<td>3.247</td>
<td>1.340</td>
<td>2.42</td>
</tr>
<tr>
<td>(\Sigma_{11})</td>
<td>0.014</td>
<td>0.0001</td>
<td>107.15</td>
</tr>
<tr>
<td>(\Sigma_{22})</td>
<td>0.012</td>
<td>0.0002</td>
<td>77.30</td>
</tr>
<tr>
<td>(\Sigma_{33})</td>
<td>0.009</td>
<td>0.0000</td>
<td>214.68</td>
</tr>
<tr>
<td>(\lambda_0(1))</td>
<td>0.005</td>
<td>0.0002</td>
<td>21.81</td>
</tr>
<tr>
<td>(\lambda_0(2))</td>
<td>0.018</td>
<td>0.0005</td>
<td>35.09</td>
</tr>
<tr>
<td>(\lambda_0(3))</td>
<td>0.016</td>
<td>0.0002</td>
<td>97.96</td>
</tr>
<tr>
<td>(\lambda_Y(1,1))</td>
<td>10.30</td>
<td>5.261</td>
<td>1.96</td>
</tr>
<tr>
<td>(\lambda_Y(1,2))</td>
<td>-263.94</td>
<td>106.700</td>
<td>2.47</td>
</tr>
<tr>
<td>(\lambda_Y(2,2))</td>
<td>47.07</td>
<td>113.755</td>
<td>0.41</td>
</tr>
<tr>
<td>(\lambda_Y(1,3))</td>
<td>863.39</td>
<td>263.647</td>
<td>3.27</td>
</tr>
<tr>
<td>(\lambda_Y(2,3))</td>
<td>-778.77</td>
<td>276.014</td>
<td>2.82</td>
</tr>
<tr>
<td>(\lambda_Y(3,3))</td>
<td>374.43</td>
<td>118.159</td>
<td>3.17</td>
</tr>
<tr>
<td>(Y_{0,4})</td>
<td>-0.0014</td>
<td>0.0001</td>
<td>10.62</td>
</tr>
<tr>
<td>(Y_{0,5})</td>
<td>0.0003</td>
<td>0.00001</td>
<td>20.74</td>
</tr>
<tr>
<td>(Y_{0,6})</td>
<td>-0.00003</td>
<td>0.000001</td>
<td>51.93</td>
</tr>
</tbody>
</table>

Table 3: Parameters and Standard Errors Estimated under the Exact Dynamic Model.
<table>
<thead>
<tr>
<th>Factors</th>
<th>Mean (bp)</th>
<th>Mean Abs (bp)</th>
<th>Standard Deviation (bp)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>1295</td>
<td>1295</td>
<td>268</td>
<td>0.88</td>
<td>2.77</td>
</tr>
<tr>
<td>Slope</td>
<td>17</td>
<td>76</td>
<td>87</td>
<td>0.12</td>
<td>1.95</td>
</tr>
<tr>
<td>Curvature</td>
<td>4</td>
<td>27</td>
<td>32</td>
<td>-0.25</td>
<td>2.16</td>
</tr>
</tbody>
</table>

Table 4: Statistical Properties of the State Factors under the Exact Dynamic Model.

<table>
<thead>
<tr>
<th>Residuals</th>
<th>Mean (bp)</th>
<th>Mean Abs (bp)</th>
<th>Standard Deviation (bp)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\epsilon}_{30}$</td>
<td>1.75</td>
<td>6.80</td>
<td>8.87</td>
<td>0.36</td>
<td>5.05</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{90}$</td>
<td>-1.89</td>
<td>5.49</td>
<td>6.81</td>
<td>0.37</td>
<td>4.60</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{120}$</td>
<td>-1.04</td>
<td>8.54</td>
<td>10.68</td>
<td>0.16</td>
<td>3.78</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{180}$</td>
<td>-3.83</td>
<td>10.25</td>
<td>12.12</td>
<td>0.53</td>
<td>3.30</td>
</tr>
</tbody>
</table>

Table 5: Statistical Properties of the Residuals of the Dynamic Gaussian Model.

<table>
<thead>
<tr>
<th>Factors</th>
<th>Mean (bp)</th>
<th>Abs. Mean (bp)</th>
<th>Standard Deviation (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Level Diff$</td>
<td>-0.202</td>
<td>0.463</td>
<td>0.519</td>
</tr>
<tr>
<td>$Slope Diff$</td>
<td>-0.673</td>
<td>1.899</td>
<td>2.172</td>
</tr>
<tr>
<td>$Curvat. Diff$</td>
<td>-0.248</td>
<td>0.934</td>
<td>1.144</td>
</tr>
</tbody>
</table>

Table 6: Statistical Properties of the Differences Between State Factors in the Approximated and Exact Dynamic Models.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Standard Error</th>
<th>ratio $\frac{\text{abs(Value)}}{\text{Std Err.}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_{11}$</td>
<td>0.112</td>
<td>0.046</td>
<td>2.43</td>
</tr>
<tr>
<td>$\kappa_{12}$</td>
<td>-1.709</td>
<td>1.431</td>
<td>1.19</td>
</tr>
<tr>
<td>$\kappa_{13}$</td>
<td>-0.124</td>
<td>1.329</td>
<td>0.09</td>
</tr>
<tr>
<td>$\kappa_{22}$</td>
<td>9.264</td>
<td>3.667</td>
<td>2.52</td>
</tr>
<tr>
<td>$\kappa_{23}$</td>
<td>-4.644</td>
<td>3.178</td>
<td>1.46</td>
</tr>
<tr>
<td>$\kappa_{33}$</td>
<td>3.799</td>
<td>1.457</td>
<td>2.60</td>
</tr>
<tr>
<td>$\Sigma_{11}$</td>
<td>0.014</td>
<td>0.0001</td>
<td>107.80</td>
</tr>
<tr>
<td>$\Sigma_{22}$</td>
<td>0.012</td>
<td>0.0002</td>
<td>77.73</td>
</tr>
<tr>
<td>$\Sigma_{33}$</td>
<td>0.009</td>
<td>0.0000</td>
<td>192.20</td>
</tr>
<tr>
<td>$\lambda_{0}(1)$</td>
<td>-0.009</td>
<td>0.0001</td>
<td>101.69</td>
</tr>
<tr>
<td>$\lambda_{0}(2)$</td>
<td>-0.007</td>
<td>0.0002</td>
<td>38.73</td>
</tr>
<tr>
<td>$\lambda_{0}(3)$</td>
<td>0.005</td>
<td>0.0002</td>
<td>25.15</td>
</tr>
<tr>
<td>$\lambda_{Y}(1,1)$</td>
<td>8.170</td>
<td>3.440</td>
<td>2.37</td>
</tr>
<tr>
<td>$\lambda_{Y}(1,2)$</td>
<td>-271.424</td>
<td>107.239</td>
<td>2.53</td>
</tr>
<tr>
<td>$\lambda_{Y}(2,2)$</td>
<td>-10.474</td>
<td>112.405</td>
<td>0.09</td>
</tr>
<tr>
<td>$\lambda_{Y}(1,3)$</td>
<td>787.700</td>
<td>275.708</td>
<td>2.86</td>
</tr>
<tr>
<td>$\lambda_{Y}(2,3)$</td>
<td>-772.455</td>
<td>278.452</td>
<td>2.77</td>
</tr>
<tr>
<td>$\lambda_{Y}(3,3)$</td>
<td>416.518</td>
<td>161.991</td>
<td>2.57</td>
</tr>
</tbody>
</table>

Table 7: Parameters and Standard Errors Estimated under the Approximated Model.

<table>
<thead>
<tr>
<th>Residuals</th>
<th>Mean (bp)</th>
<th>Mean Abs (bp)</th>
<th>Standard Deviation (bp)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\epsilon}_{30}$</td>
<td>2.01</td>
<td>7.06</td>
<td>9.07</td>
<td>0.16</td>
<td>4.81</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{90}$</td>
<td>-1.66</td>
<td>5.14</td>
<td>6.54</td>
<td>0.17</td>
<td>4.70</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{120}$</td>
<td>-0.88</td>
<td>8.34</td>
<td>10.48</td>
<td>0.11</td>
<td>3.84</td>
</tr>
<tr>
<td>$\hat{\epsilon}_{180}$</td>
<td>-3.84</td>
<td>10.33</td>
<td>12.20</td>
<td>0.56</td>
<td>3.35</td>
</tr>
</tbody>
</table>

Table 8: Statistical Properties of the Residuals of the Approximated Gaussian Model.

<table>
<thead>
<tr>
<th>Maturities (Days)</th>
<th>Mean (bp)</th>
<th>Mean Abs (bp)</th>
<th>Standard Deviation (bp)</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>21.26</td>
<td>55.21</td>
<td>64.67</td>
<td>0.14</td>
<td>2.16</td>
</tr>
<tr>
<td>60</td>
<td>22.75</td>
<td>54.31</td>
<td>62.37</td>
<td>0.07</td>
<td>2.13</td>
</tr>
<tr>
<td>90</td>
<td>24.19</td>
<td>53.37</td>
<td>60.05</td>
<td>0.01</td>
<td>2.11</td>
</tr>
<tr>
<td>120</td>
<td>25.59</td>
<td>52.38</td>
<td>57.67</td>
<td>-0.04</td>
<td>2.09</td>
</tr>
<tr>
<td>180</td>
<td>28.24</td>
<td>50.22</td>
<td>52.73</td>
<td>-0.14</td>
<td>2.09</td>
</tr>
<tr>
<td>360</td>
<td>35.07</td>
<td>43.02</td>
<td>36.18</td>
<td>-0.27</td>
<td>2.15</td>
</tr>
<tr>
<td>720</td>
<td>43.67</td>
<td>47.01</td>
<td>36.61</td>
<td>0.16</td>
<td>2.27</td>
</tr>
</tbody>
</table>

Table 9: Risk Premium: Differences between the Exact Dynamic Model and the Approximated Model.
Figure 1: Legendre Polynomials.

Figure 2: Temporal Evolution of the Brazilian Term Structure.
Figure 3: Static Model: Temporal Evolution of the Legendre Coefficients for the Brazilian Swaps.

Figure 4: Dynamic Model: Temporal Evolution of the Legendre Coefficients for the Brazilian Swaps.
Figure 5: Conditionally Deterministic Factors in the Multi-factor Gaussian Model.

Figure 6: Difference Between Legendre Dynamic and Static Factors for the Gaussian Model.
Figure 7: Comparing Risk Premium for Two Different Maturities.