

UNIVERSIDAD DEL ROSARIO

FACULTAD DE ECONOMÍA

**MODELO DE CASCADA SIGMA VARIANTE PARA LA  
ESTRUCTURA A TÉRMINO DE TASAS DE INTERÉS EN  
COLOMBIA**

TESIS PRESENTADA POR LAURA NATALIA LÓPEZ RODRÍGUEZ  
PARA OPTAR POR EL TÍTULO DE MAESTRÍA EN FINANZAS CUANTITATIVAS

DIRECTOR: RAFAEL SERRANO

JUNIO 2019

---

## Resumen

El presente estudio contrasta dos modelos de cascada recursiva para determinar las dinámicas de la estructura a término de las tasas de interés en Colombia, estos modelos poseen factores heterogéneamente persistentes que reversionan en media al factor inmediatamente anterior. Bajo ciertas especificaciones, el número de parámetros es invariante al número de factores de las cascadas. El modelo estándar posee un supuesto en su volatilidad que la hace siempre constante, mientras que el modelo sigma variante (SV), evaluado en este documento, relaja este supuesto dándole una estructura que depende de la dimensión de la cascada. Ambos modelos presentan fluctuaciones en los valores de los parámetros ante aumentos en la dimensión de la cascada. En el ajuste en la muestra el modelo estándar posee más estabilidad en los resultados, pero el modelo SV en ciertas dimensiones (número de factores) tiene mejor ajuste. En el pronóstico por fuera de la muestra, ambos modelos poseen un bajo desempeño comparado con una caminata aleatoria, sin embargo, el modelo SV parece ser más consistente en resultados en la muestra y por fuera de la muestra que el estándar. En general, puede ser más valioso elegir un modelo de al menos 3 factores para el ajuste y pronóstico de la curva yield colombiana, aunque se tiene que ser cuidadoso en su elección.

*Clasificación JEL* : E43, E47, G12, C51.

**Palabras claves:** Estructura a término de tasas de interés, modelo de cascada, dimensión invariante, volatilidad cambiante.

## Abstract

The present study contrasts two recursive cascade models to determine the dynamics of the term structure of interest rates in Colombia, these models have heterogeneously persistent factors that mean-revert to the immediately preceding factor. Under certain specifications, the number of parameters is invariant to the number of factors in the cascades. The standard model has an assumption in its volatility that makes it always constant, while the sigma variant model (SV) relaxes this assumption giving it a structure that depends on the cascade's size. Both models show fluctuations in the values of the parameters when increases the cascade's size. In the in-sample fit, the standard model shows more stability in the results, but the SV model in certain dimensions (number of factors) has a better fit. In the out-of-sample forecast, both models have a low performance compared to a random walk, but the SV model seems to be more consistent in sample and out-of-sample than the standard model. In general, it may be more valuable to choose a model with at least three factors for the adjustment and forecast of the Colombian yield curve, although it must be a meticulous choice.

**Keywords:** Term structure of interest rates, cascade model, dimension-invariance, changing volatility.

# Índice general

<b>1</b>	<b>Introducción</b>	<b>5</b>
<b>2</b>	<b>Modelo de Cascada</b>	<b>8</b>
2.1	Una cascada multifrecuente . . . . .	8
2.2	Risk Premia . . . . .	12
2.3	Modelo de Dimensión Invariante . . . . .	16
<b>3</b>	<b>Datos y Estimación</b>	<b>18</b>
3.1	Datos . . . . .	18
3.2	Modelo . . . . .	19
3.3	Estimación . . . . .	20
3.3.1	Filtro de Kalman . . . . .	20
3.3.2	Optimización . . . . .	22
<b>4</b>	<b>Resultados Empíricos</b>	<b>24</b>
4.1	Estimación en la muestra . . . . .	24
4.1.1	Ajuste en la muestra . . . . .	27
4.1.2	Pronóstico por fuera de la muestra . . . . .	29
<b>5</b>	<b>Análisis Gráfico</b>	<b>33</b>
<b>6</b>	<b>Conclusiones</b>	<b>39</b>

# Índice de figuras

3.1	Algoritmo de Estimación y Optimización . . . . .	23
5.1	Ajuste en la Muestra Modelo Estándar . . . . .	34
5.2	Ajuste en la Muestra Modelo SV . . . . .	35
5.3	Ajuste por Fuera de la Muestra Modelo Estándar . . . . .	36
5.4	Ajuste por Fuera de la Muestra Modelo SV . . . . .	37
5.5	Ajuste Curva Yield Modelo Estándar . . . . .	38
5.6	Ajuste Curva Yield Modelo SV . . . . .	38
5.7	Ajuste Curva Yield Modelo Estándar . . . . .	38
5.8	Ajuste Curva Yield Modelo SV . . . . .	38

# Índice de tablas

3.1	Resumen Estadístico de Tasas Yield . . . . .	19
4.1	Parámetros Estimados y Errores Estándar Modelo Estándar . . . . .	25
4.2	Parámetros Estimados y Errores Estándar Modelo SV . . . . .	26
4.3	Raíz de Errores de Valoración Cuadráticos Medios (RMSE) Modelo Estándar . . . . .	28
4.4	Raíz Cuadrada de Errores de Valoración Cuadráticos Medios (RMSE) Modelo SV . . . . .	28
4.5	RMSE (Puntos Básicos) Fuera de la Muestra Modelo Estándar . . . . .	30
4.6	RMSE (Puntos Básicos) Fuera de la Muestra Modelo SV . . . . .	31

# Capítulo 1

## Introducción

La estructura a plazos de tasas de interés o yield curve, tiene un rol importante en la economía ya que permite a los inversionistas tomar decisiones de inversión, de valoración de activos financieros y de manejo de riesgo financiero, a los consumidores decisiones de consumo y ahorro, y a los bancos centrales de expectativas de inflación, actividad económica real y otros factores macroeconómicos. Uno de los problemas más estudiados en la literatura es el de cómo modelar el comportamiento de la estructura a plazos a través del tiempo, asegurando que tenga un buen ajuste empírico y además que esta sea libre de arbitraje.

Al no poder ser observadas con antelación, las tasas de interés son variables aleatorias. Por lo cual, la formación de los modelos se ha centrado en reducir la incertidumbre para lograr un buen ajuste y caracterización de la estructura temporal. El desarrollo de modelos de estructura a plazos dinámicos (DTSM por su siglas en inglés) libres de arbitraje, como los modelos afines de Duffie y Kan(1996) [9] han incitado la investigación en esta área como por ejemplo el análisis multifactorial de Dai y Singleton(2000) [7].

Generalmente se usan modelos de factores basados en el modelo de Vasicek(1997) [18] y el CIR(1985) [6], los cuales son convenientes por su facilidad de trazabilidad y sus soluciones de forma cerrada para distintos tipos de derivados de tasa de interés. La gran mayoría de los estudios se han centrado en el estudio teórico y empírico de un número pequeño de factores, típicamente de uno a tres, por razones de identificación de todos los parámetros obtenidos en los modelos y de tractabilidad.

Varios autores han hablado de las limitaciones de un modelo con solo un factor, donde encontramos por ejemplo a Duan y Simonato(1999) [8], Bolder(2001) [2], Zeytun y Gup-

ta(2007) [21]. Existen diferentes conclusiones en los estudios sobre modelos de dos factores, Nowman(2010) [13] dice que se ajusta bien a la curva, mientras que Walther(2003) [20] sugiere que agregaría valor al ajuste el tener más factores en los modelos. Chatterjee(2005) [5] concluye que un modelo de tres factores es adecuado para explicar la variabilidad en los yields por su similitud con el nivel, la pendiente y la curvatura.

Calvet, Laurent y Fisher(2018) [4] sustentan que los DTSM de baja dimensionalidad, es decir, de menos de 3 factores, tienen limitaciones empíricas, como su incapacidad de ajustarse a la curva yield en un punto en el tiempo, generando errores de valoración altos. Además, tienen un bajo desempeño en pronóstico comparado con una caminata aleatoria. Por ello, proponen un modelo de cascada multifactorial que, bajo ciertos supuestos, mantiene invariante su dimensión con respecto al número de factores. Esto elimina el *trade-off* entre la riqueza del espacio de estados y la fuerza de la identificación.

Las conclusiones a las que llegan es que los modelos con más de 7 factores tienen mejor ajuste a la curva yield que los de baja dimensionalidad, al igual que menor raíz cuadrada del error cuadrático medio (RMSE). En el caso de pronóstico, por fuera de la muestra, el modelo de alta dimensionalidad tiene mayor estabilidad en resultados que la caminata aleatoria.

En Colombia, pocos han sido los estudios que han utilizado DTSM afines con el propósito de modelar el comportamiento de las tasas de interés. Rodríguez(2016) [14] muestra que un modelo de un solo factor tiene bajo desempeño de ajuste en la muestra y en pronóstico por fuera de la muestra, mientras que en Velásquez et al(2016) [19] hacen un análisis exhaustivo de modelos afines hasta con tres factores con diferentes representaciones de los modelos base, concluyendo que el modelo de tres factores se ajusta bien a la curva con los datos de mercado de bonos. El presente estudio será el primero en utilizar un modelo multifactorial para más de 3 factores en Colombia con análisis de ajuste en la muestra y de pronóstico por fuera de la muestra.

El objetivo de este trabajo de grado es proponer un modelo de cascada donde se supone que el riesgo es dependiente del tamaño de la cascada y comparar dicho modelo con la cascada de dimensión invariante propuesta en Calvet et Al(2018) [4] ajustado para el caso colombiano. Además, se analizan los resultados encontrados con respecto a las curvas yields tanto en ajuste

como en pronóstico. Se adaptó la cascada a los datos de las curvas yield, en contraposición a las tasas LIBOR y swap usadas por Calvet et Al, debido a las limitaciones en las series de datos que tiene el país. Luego se realizó una estimación del modelo en espacio-estado por medio del Filtro de Kalman y el método de Máxima Verosimilitud para obtener los vectores de estado y las matrices de parámetros. Finalmente se hizo un análisis empírico usando el error cuadrático medio (RMSE) para el ajuste y para el pronóstico comparado con una caminata aleatoria.

El resto del documento está organizado de la siguiente manera: El capítulo 2 define el modelo de cascada invariante de manera continua, presentando el modelo estándar y el modelo sigma variante (SV) con sus supuestos claves. El capítulo 3 describe los datos que se utilizan en el estudio además de explicar los métodos utilizados para la estimación. El capítulo 4 exhibe los resultados del ejercicio empírico. Finalmente, se presentan las conclusiones y se discuten posibles extensiones del modelo para futuras investigaciones.

# Capítulo 2

## Modelo de Cascada

Para enmarcar este modelo se especifica una economía tiempo-continua con un horizonte infinito, donde la variable de tiempo se denota  $t \in [0, \infty)$ , con un espacio de probabilidad  $\{\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}_t)_{t \geq 0}\}$ . Se asume que  $\mathbb{P}$  es la medida física y que cumple las condiciones usuales de continuidad por derecha y de completitud.

### 2.1 Una cascada multifrecuente

Se modela las dinámicas de la tasa de interés instantánea sobre la medida  $\mathbb{P}$  por medio de una **cascada** de  $n$ -difusiones. Se considera la constante  $\theta \in \mathbb{R}$  y un proceso de Wiener  $\mathbf{W}_t = (\mathbf{W}_{1,t}, \dots, \mathbf{W}_{n,t})^T$   $n$ -dimensional con componentes independientes.

**Supuesto 1. Estructura de los factores**  $x_{0,t}$  es la tasa de largo plazo igual a  $\theta$  constante, a la cual el proceso revierte en media. Se define la cascada de difusión para los factores:

$$dx_{j,t} = \kappa_j(x_{j-1,t} - x_{j,t})dt + \sigma_j d\mathbf{W}_{j,t} \quad (2.1)$$

para todo  $j \in \{1, \dots, n\}$ , donde  $\kappa_1, \dots, \kappa_n, \sigma_1, \dots, \sigma_n$  son constantes estrictamente positivas. Además, las velocidades de ajuste son estrictamente crecientes en  $j$ :  $\kappa_1 < \kappa_2 < \dots < \kappa_n$ .

Se puede observar que el primer factor  $x_{1,t}$  es un proceso de Ornstein-Uhlenbeck con reversión a la media a  $\theta$ . Más generalmente para cada  $j \geq 1$  el  $j$ -ésimo factor de  $x_{j,t}$  revierte en media hacia el  $(j - 1)$ -ésimo factor  $x_{j-1,t}$ . Los factores son menos persistentes cuando  $j$  incrementa.

**Supuesto 2. Tasa Corta** La tasa de interés instantánea coincide con el último factor:  $r_t = x_{n,t}$  para todo  $t \geq 0$ .

El vector de estados  $\mathbf{X}_t = (x_{1,t}, \dots, x_{n,t})^T \in \mathbb{R}^n$  satisface:

$$d\mathbf{X}_t = \boldsymbol{\kappa}(\boldsymbol{\theta} - \mathbf{X}_t)dt + \boldsymbol{\Sigma}^{1/2}d\mathbf{W}_t, \quad (2.2)$$

Donde  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)^T$ ,  $\boldsymbol{\Sigma}$  es una matriz diagonal con elementos constantes  $\sigma_1^2, \dots, \sigma_n^2$ , y  $\boldsymbol{\kappa}$  es una matriz con elementos en la diagonal  $\kappa_{j,j} = \kappa_j$  y elementos fuera de la diagonal  $\kappa_{j,j-1} = -\kappa_j$ . Por tanto,  $\boldsymbol{\kappa}$  genera una matriz con forma de cascada, que le da el nombre al modelo.

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & 0 & \dots & 0 \\ 0 & 0 & \sigma_3^2 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \sigma_n^2 \end{bmatrix} \quad \boldsymbol{\kappa} = \begin{bmatrix} \kappa_1 & 0 & 0 & \dots & 0 \\ -\kappa_2 & \kappa_2 & 0 & \dots & 0 \\ 0 & -\kappa_3 & \kappa_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & -\kappa_n & \kappa_n \end{bmatrix}$$

Cada ecuación separa los factores de acuerdo a sus tasas de reversión a la media, es decir que tan rápido, ante un shock, vuelven a su estado estacionario o de largo plazo.

Se denota la convolución de  $a$  y  $b$  como  $a * b$  usando un asterisco o estrella. Se define como la integral del producto de estas dos funciones donde una de ellas se revierte y desplaza, por lo cual es una transformación integral particular [15]:

$$\begin{aligned} (a * b)(t) &\triangleq \int_{-\infty}^{\infty} a(t - \tau)b(\tau)d\tau \\ &= \int_{-\infty}^{\infty} a(\tau)b(t - \tau)d\tau \end{aligned}$$

La función  $a_j(\tau)$  se conoce como función de respuesta y se define como la convolución del producto de las funciones de densidad de probabilidad exponencial:

$$a_j(\tau) = (\mathbf{K}_j * \dots * \mathbf{K}_n)(\tau)/\kappa_j \quad (2.3)$$

Donde  $\mathbf{K}_i(\tau) = \kappa_i e^{-\kappa_i \tau}$  si  $\tau > 0$  y  $\mathbf{K}_i(\tau) = 0$  si  $\tau < 0$ . La función  $a_j(\tau)$  toma valores positivos para cada  $\tau \in (0, \infty)$  y satisface  $\int_0^{+\infty} a_j(\tau)d\tau = 1/\kappa_j$  para todo  $j$  (ver [Calvet et al(2018) [4]])

Para un  $j$  dado, la función de respuesta  $a_j(\tau)$  cuantifica cómo una unidad de shock al factor  $j$  en el tiempo  $t - \tau$  impacta a la tasa de interés instantánea  $r_t = x_{n,t}$ . La ecuación (2.3) implica que la función de respuesta depende de tanto  $j$  como de  $n$ , pero se omite el índice  $n$ .

Trabajando backwards, la última función de respuesta  $a_n(\tau) = e^{-\kappa_n \tau}$  inicia en  $a_n(0) = 1$ , consecuencia del supuesto de que la tasa corta es igual al  $n$ -ésimo factor. La función  $a_n(\tau)$  decrece exponencialmente con el horizonte de tiempo a una tasa controlada por  $\kappa_n$ . La función  $a_{n-1}(\tau)$  crece y alcanza un máximo en  $\bar{\tau}_{n-1} = (\ln(\kappa_n) - \ln(\kappa_{n-1})) / (\kappa_n - \kappa_{n-1})$  y decrece hacia cero. Las funciones de respuesta satisfacen más generalmente:

**Proposición 1. Funciones de respuesta** Para todo  $j < n$  las funciones de respuesta  $a_j(\tau)$  tienen forma de campana y sus horizontes máximos de respuesta son monótonamente decrecientes con  $j$ . La función satisface la expresión de forma cerrada:

$$a_j(\tau) = \sum_{i=j}^n \alpha_{i,j} \kappa_i e^{-\kappa_i \tau}, \quad \text{donde } \alpha_{i,j} = \frac{\kappa_{j+1} \cdots \kappa_n}{\kappa_i \prod_{k=j, k \neq i}^n (\kappa_k - \kappa_i)} \quad (2.4)$$

Además

$$0 \leq \sum_{j=1}^n a_j(\tau) \leq 1 \quad \text{para todo } \tau \geq 0 \quad (2.5)$$

*Demostración.* Las soluciones analíticas y demostraciones de la convolución de funciones de densidad exponenciales están dados, entre otros lugares, en Akkouchi(2008) [1]. Se iniciará con un lema sobre el comportamiento de las funciones y luego como se llega a las ecuaciones (2.4) y (2.5).

**Lemma 1.** Si  $\tau > 0$  es un óptimo local de la función de respuesta  $a_{j,n}$ , entonces  $a''_{j,n}(\tau) = \kappa_{j+1} a'_{j+1,n}(\tau)$

*Demostración.* Se sabe por (2.3) que

$$a_{j,n}(\tau) = \frac{\kappa_{j+1}}{\kappa_j} \int_0^\tau \mathbf{K}_j(\tau - s) a_{j+1,n}(s) ds$$

Diferenciando esta relación con respecto a  $\tau$  utilizando la regla de Leibniz se obtiene  $a'_{j,n}(\tau) = \kappa_{j+1} a_{j+1,n}(\tau) - \kappa_j a_{j,n}(\tau)$ . Esto implica que  $a''_{j,n}(\tau) = \kappa_{j+1} a'_{j+1,n}(\tau) - \kappa_j a'_{j,n}(\tau)$ . Por tanto, un óptimo local de  $a_{j,n}$  satisface  $a''_{j,n}(\tau) = \kappa_{j+1} a'_{j+1,n}(\tau)$   $\square$

**Un solo pico.** Por inducción hacia se atrás se muestra que para todo  $j = n - 1, \dots, 1$ , la función  $a_{j,n}(\tau)$  tiene solo un pico y alcanza su máximo en  $\bar{\tau}_{j,n}$ . Además,  $\bar{\tau}_{1,n} \geq \dots \geq \bar{\tau}_{n,n}$ :

- Esta propiedad se mantiene para  $j = n - 1$ , como se menciona en la proposición.
- Suponiendo ahora que la propiedad se mantiene para  $j + 1 \leq n$ . Sea  $\bar{\tau}_{j,n}$  el máximo local más pequeño de  $a_{j,n}$ . Se sabe que  $a''_{j,n}(\bar{\tau}_{j,n}) \leq 0$ , y que  $a''_{j,n}(\bar{\tau}_{j,n}) = \kappa_{j+1} a'_{j+1,n}(\bar{\tau}_{j,n})$ . Por tanto,  $a'_{j+1,n}(\bar{\tau}_{j,n}) \leq 0$ , esto implica que  $\bar{\tau}_{j,n} \geq \bar{\tau}_{j+1,n}$ . Si la función es no monotónica, existe un mínimo local  $\tau > \bar{\tau}_{j,n}$ . Ya que  $\tau > \bar{\tau}_{j,n} \geq \bar{\tau}_{j+1,n}$ , se sabe que  $a''_{j,n}(\tau) = \kappa_{j+1} a'_{j+1,n}(\tau) < 0$ , que es una contradicción. Se puede concluir que  $a_{j,n}$  tiene solo un pico y alcanza su máximo en  $\bar{\tau}_{j,n} \geq \bar{\tau}_{j+1,n}$ .

**para las ecuaciones (2.4) y (2.5).** La desigualdad (2.5) puede ser demostrada para todos los  $n$  utilizando una recursión hacia adelante. Iniciando en  $n = 1$ , la inecuación se mantiene ya que  $a_{1,1} = e^{-\kappa_1 t} \leq 1$  para todo  $t \geq 0$ . Ahora se asumen que la desigualdad se mantiene para  $n - 1$  factores. Se infiere que :

$$\sum_{j=1}^n a_{j,n}(\tau) = e^{-\kappa_n t} + \int_0^t \kappa_n e^{-\kappa_n(t-s)} \sum_{j=1}^{n-1} a_{j,n-1}(s) ds$$

La desigualdad  $\sum_{j=1}^{n-1} a_{j,n-1}(s) \leq 1$  para todos los  $s$ , implicando que  $\sum_{j=1}^n a_{j,n}(\tau) \leq 1$ . □

**Proposición 2. Representación de factores para la tasa corta** La tasa de interés instantánea  $r_t$  es una función afín del vector inicial de estado  $X_0$  y las integrales de los conductores de innovación de cada factor entre 0 y  $t$ .

$$r_t = \theta_r + \sum_{j=1}^n a_j(t)(x_{j,0} - \theta_r) + \sum_{j=1}^n \sigma_j \int_0^t a_j(t-s) dW_{j,s} \quad (2.6)$$

*Demostración.* Se prueba la proposición por inducción. Considere el caso de un solo factor ( $n = 1$ ). Se infiere por el lema de Itô que  $d(e^{\kappa_1 t} x_{1,t}) = e^{\kappa_1 t} (\kappa_1 \theta_r dt + \sigma_1 dW_{1,t})$ . Integrando ambos lados y luego dividiendo por  $e^{\kappa_1 t}$ , se obtiene la ecuación (2.6) para  $n = 1$ .

Asumiendo que la propiedad (2.6) se mantiene para  $(n - 1)$  factores,

$$\mathbf{x}_{n-1,t} = \boldsymbol{\theta}_r + \sum_{j=1}^{n-1} (x_{j,0} - \boldsymbol{\theta}_r) a_{j,n-1}(t) + \sum_{j=1}^{n-1} \boldsymbol{\sigma}_j \int_0^t a_{j,n-1}(t-s) d\mathbf{W}_{j,s}$$

El lema de Itô implica que  $d(e^{\kappa_n t} x_{n,t}) = \kappa_n e^{\kappa_n t} \mathbf{x}_{n-1,t} dt + \boldsymbol{\sigma}_n e^{\kappa_n t} dW_{n,t}$ . Integrando a ambos lados y dividiendo por  $e^{\kappa_n t}$ , se obtiene:

$$\mathbf{x}_{n,t} = e^{-\kappa_n t} \mathbf{x}_{n,0} + \int_0^t \kappa_n e^{-\kappa_n(t-s)} \mathbf{x}_{n-1,s} ds + \boldsymbol{\sigma}_n \int_0^t e^{-\kappa_n(t-s)} dW_{n,s}$$

Sustituyendo  $\mathbf{x}_{n-1,s}$ ,

$$\begin{aligned} \int_0^t \kappa_n e^{-\kappa_n(t-s)} \mathbf{x}_{n-1,s} ds &= \boldsymbol{\theta}_r (1 - e^{-\kappa_n t}) + \sum_{j=1}^{n-1} (x_{j,0} - \boldsymbol{\theta}_r) a_{j,n}(t) \\ &\quad + \sum_{j=1}^{n-1} \boldsymbol{\sigma}_j \int_0^t \kappa_n e^{-\kappa_n(t-s)} \left[ \int_0^s a_{j,n-1}(s-u) d\mathbf{W}_{j,u} \right] ds \end{aligned}$$

Aplicando el teorema de Fubini para integrales estocásticas en los últimos términos, se puede concluir que la propiedad se mantiene para la estructura de  $n$  factores.  $\square$

## 2.2 Risk Premia

Asumiendo que **no existen** oportunidades de arbitraje, es decir que existe una medida neutral al riesgo  $\mathbb{Q}$ , se selecciona un precio de riesgo de mercado no afín a los factores haciendo el risk premia constante  $\gamma_j$  y los precios de los bonos disponibles analíticamente.

**Supuesto 3. Medida neutral al riesgo.** Se define  $\mathbb{Q}$  con la derivada de Radón-Nikódyém:

$$\left. \frac{d\mathbb{Q}}{d\mathbb{P}} \right|_t = \prod_{j=1}^n \exp \left( - \int_0^t \gamma_j dW_{j,s} - \frac{1}{2} \int_0^t \gamma_j^2 ds \right) \quad (2.7)$$

Con respecto a la medida física  $\mathbb{P}$ .

Realizando el *cambio de medida* de  $\mathbb{P}$  a  $\mathbb{Q}$  a través del teorema de Girsanov, las dinámicas de los factores serán:

$$dx_{j,t} = \gamma_j \boldsymbol{\sigma}_j dt + \kappa_j (x_{j-1,t} - x_{j,t}) dt + \boldsymbol{\sigma}_j d\mathbf{W}_{j,t}^{\mathbb{Q}} \quad (2.8)$$

Por tanto, las dinámicas de los factores se mantienen afines bajo la medida neutral al riesgo. De manera matricial, tenemos:

$$d\mathbf{X}_t = \boldsymbol{\kappa}(\boldsymbol{\theta}^{\mathbb{Q}} - \mathbf{X}_t)dt + \boldsymbol{\Sigma}^{1/2}d\mathbf{W}_t^{\mathbb{Q}}, \quad (2.9)$$

Donde  $\boldsymbol{\theta}^{\mathbb{Q}}$  representa a la *media* de largo plazo del vector de estados bajo la medida neutral al riesgo y está dado por

$$\theta_1^{\mathbb{Q}} = \theta_r - \frac{\gamma_1 \sigma_1}{k_1}, \theta_2^{\mathbb{Q}} = -\frac{\gamma_2 \sigma_2}{k_2}, \dots, \theta_n^{\mathbb{Q}} = -\frac{\gamma_n \sigma_n}{k_n} \quad (2.10)$$

Sea  $P(X_t, \tau) = \mathbb{E}_t^{\mathbb{Q}}[\exp(-\int_t^{t+\tau} r_s ds)]$  el valor en el tiempo  $t$  de un bono cero cupón, con una unidad monetaria como valor par y fecha de vencimiento  $t + \tau$ .

**Proposición 3. Precio Bonos Cero Cupón** El precio de un bono cero cupón que inicia en  $t$  y vencimiento en  $\tau$  es dado por un modelo afín a  $X_t$ :

$$P(\mathbf{X}_t, \tau) = \exp[-b(\boldsymbol{\tau})^T \mathbf{X}_t - c(\boldsymbol{\tau})] \quad (2.11)$$

Los coeficientes  $b(\boldsymbol{\tau})$  y  $c(\boldsymbol{\tau})$  satisfacen el sistema de ecuaciones diferenciales ordinarias

$$b'(\boldsymbol{\tau}) = \mathbf{e}_n - \boldsymbol{\kappa}^T b(\boldsymbol{\tau}) \quad (2.12)$$

$$c'(\boldsymbol{\tau}) = b(\boldsymbol{\tau})^T \boldsymbol{\kappa} \boldsymbol{\theta}^{\mathbb{Q}} - \frac{1}{2} b(\boldsymbol{\tau})^T \boldsymbol{\Sigma} b(\boldsymbol{\tau}) \quad (2.13)$$

Con condiciones iniciales  $b(\boldsymbol{\tau}) = 0$  y  $c(\boldsymbol{\tau}) = 0$ , donde  $\mathbf{e}_n$  denota un vector con el valor de uno en la  $n$ -ésima posición, y cero en las demás.

*Demostración.* Sea  $P(\mathbf{X}_t, \tau) = \exp[-b(\boldsymbol{\tau}, n)^T \mathbf{X}_t - c_n(\boldsymbol{\tau})]$  el precio de un bono cero cupón bajo la cascada de  $n$  factores. Bajo  $\mathbb{Q}$ , el retorno esperado del bono es igual a la tasa de interés  $r_t = x_{n,t}$ , entonces:

$$-b(\boldsymbol{\tau}, n)^T \boldsymbol{\kappa} (\boldsymbol{\theta}^{\mathbb{Q}} - \mathbf{X}_t) + \frac{1}{2} b(\boldsymbol{\tau}, n)^T \boldsymbol{\Sigma} b(\boldsymbol{\tau}, n) + b'(\boldsymbol{\tau}, n)^T \mathbf{X}_t + c'_n(\boldsymbol{\tau}) = x_{n,t}$$

Donde  $b'(\boldsymbol{\tau}, n)$  y  $c'_n(\boldsymbol{\tau})$  denotan, respectivamente, las derivadas de  $b(\boldsymbol{\tau}, n)$  y  $c_n(\boldsymbol{\tau})$  con respecto

a  $\tau$ . Se infiere que:

$$\begin{aligned} b(\tau, n)^T \kappa \mathbf{X}_t + b'(\tau, n)^T \mathbf{X}_t &= \mathbf{e}_n^T \mathbf{X}_t \\ -b(\tau, n)^T \kappa \boldsymbol{\theta}^Q + \frac{1}{2} b(\tau)^T \Sigma b(\tau) + c'_n(\tau) &= 0 \end{aligned}$$

Las ecuaciones anteriores implican las ecuaciones (2.12) y (2.13) □

**Proposición 4. Valoración bajo premia constante** Las cargas de precios  $b_j(\tau)$  satisfacen

$$b_j(\tau) = \int_0^\tau a_j(\tau') d\tau' = \sum_{i=j}^n \alpha_{i,j} (1 - e^{-\kappa_i \tau}) \quad (2.14)$$

para todo  $j \in \{1, \dots, n\}$ , y el intercepto  $c(\tau)$  es dado por:

$$\begin{aligned} c(\tau) &= \theta_r \kappa_1 \sum_{i=1}^n \alpha_{i,1} \left( \tau - \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} \right) - \sum_{j=1}^n \gamma_j \sigma_j \sum_{i=j}^n \alpha_{i,j} \left( \tau - \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} \right) - \\ &\sum_{j=1}^n \frac{\sigma_j^2}{2} \sum_{i=j}^n \sum_{k=j}^n \alpha_{i,j} \alpha_{k,j} \left[ \tau - \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} - \frac{1 - e^{-\kappa_k \tau}}{\kappa_k} + \frac{1 - e^{-(\kappa_i + \kappa_k) \tau}}{\kappa_i + \kappa_k} \right] \end{aligned}$$

Además, los niveles de largo plazo del vector de estados bajo la medida  $\mathbb{Q}$  son

$$\theta^Q = \theta_r - \gamma_1 \sigma_1 / \kappa_1, \theta_2^Q = \theta_r - \gamma_1 \sigma_1 - \gamma_2 \sigma_2 / \kappa_2, \dots, \theta_n^Q = \theta_r - \sum_{i=1}^n \gamma_i \sigma_i / \kappa_i$$

La proposición muestra que las cargas de precios  $b_j(\tau)$  y son integrales de la función de respuesta  $a_j(\tau)$ . El intercepto  $c(\tau)$  tiene tres componentes:

1. El primero es proporcional a la media de largo plazo  $\theta$  del vector de estados.
2. El segundo depende del risk premia  $\gamma_j$ .
3. el último es un ajuste por convexidad, que corresponde a la forma cuadrática de la ecuación (2.13).

*Demostración.* Por la ecuación (2.12), la función  $b(\tau, n) = [b_{1,n}(\tau), \dots, b_{n,n}(\tau)]^T$  satisface:

$$\begin{aligned} b'_{j,n}(\tau) &= -\kappa_j b_{j,n}(\tau) + \kappa_{j+1} b_{j+1,n}(\tau), \quad j = 1, \dots, n-1. \\ b'_{n,n}(\tau) &= 1 - \kappa_n b_{n,n}(\tau) \end{aligned}$$

La segunda ecuación implica que  $b_{n,n}(\tau) = (1 - e^{-\kappa_n \tau}) / \kappa_n = \int_0^\tau a_{n,n}(s) ds$ . Se infiere por la primera ecuación que para todos los  $j \in \{1, \dots, n-1\}$ ,

$$\frac{d}{d\tau} [e^{\kappa_j \tau} b_{j,n}(\tau)] = e^{\kappa_j \tau} \kappa_{j+1} b_{j+1,n}(\tau)$$

y así,  $b_{j,n} = (\kappa_{j+1} / \kappa_j) \mathbf{K}_j * b_{j+1,n}$ . Mediante una recursión simple, la función  $b_{j,n}$  es igual a  $(\kappa_n / \kappa_j) \mathbf{K}_j * \dots * \mathbf{K}_{n-1} * b_{n,n}$ , y su derivada es

$$b'_{j,n} = \kappa_j^{-1} \mathbf{K}_j * \dots * \mathbf{K}_{n-1} * \mathbf{K}_n = a_{j,n}$$

Se concluye que (2.14) se mantiene para todo los  $j \in \{1, \dots, n-1\}$

Para el intercepto, por (2.13), la función  $c_n(\tau)$  satisface

$$c_n(\tau) = \theta \kappa_1 \int_0^\tau b_{1,n}(\tau) d\tau - \sum_{j=1}^n \gamma_j \sigma_j \int_0^\tau b_{j,n}(s) ds - \frac{1}{2} \sum_{j=1}^n \sigma_j^2 \int_0^\tau b_{j,n}^2(s) ds$$

Así, por (2.14)

$$\begin{aligned} \int_0^\tau b_{j,n}(s) ds &= \sum_{j=1}^n \left( \tau - \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} \right) \\ \int_0^\tau b_{j,n}^2(s) ds &= \sum_{i=j}^n \sum_{k=j}^n \alpha_{i,j} \alpha_{k,j} \left[ \tau - \frac{1 - e^{-\kappa_i \tau}}{\kappa_i} - \frac{1 - e^{-\kappa_k \tau}}{\kappa_k} + \frac{1 - e^{-(\kappa_i + \kappa_k) \tau}}{\kappa_i + \kappa_k} \right] \end{aligned}$$

la expresión para  $c_n(\tau)$  se mantiene.

**Los niveles de largo plazo del vector de estados bajo la medida  $\mathbb{Q}$ .** Se verifica fácilmente que  $\kappa^{-1}$  es una matriz triangular inferior con elementos  $(\kappa^{-1})_{i,j} = (-1)^{i+j} \kappa_j^{-1}$  si  $j \leq i$  se infiere que  $\theta^{\mathbb{Q}} = (\theta_r - \gamma_1 \sigma_1^2 / k_1, \theta_r - \gamma_1 \sigma_1^2 - \gamma_2 \sigma_2^2 / k_2, \dots, \theta_r - \sum_{i=1}^n \gamma_i \sigma_i / k_i)^T$   $\square$

El **yield de largo plazo**  $y_\infty = \lim_{\tau \rightarrow +\infty} [b(\tau)^T \mathbf{X}_t + c(\tau)] / \tau$  es dado por:

$$y_\infty = \theta_r - \sum_{j=1}^n \frac{\sigma_j^2}{\kappa_j^2} \left( \frac{\gamma_j}{\sigma_j} \kappa_j + \frac{1}{2} \right), \quad (2.15)$$

*Demostración.* Se infiere de la proposición 4 que

$$y_\infty = \lim_{\tau \rightarrow +\infty} \frac{c(\tau)}{\tau} = \theta \kappa_1 \sum_{i=1}^n \alpha_{i,1,n} - \sum_{j=1}^n \gamma_j \sigma_j \sum_{i=j}^n \alpha_{i,j,n} - \sum_{j=1}^n \frac{\sigma_j^2}{2} \sum_{i=j}^n \sum_{k=j}^n \alpha_{i,j,n} \alpha_{k,j,n}$$

Integrando  $a_{j,n}(\tau) = \mathbf{K}_j * \dots * \mathbf{K}_n(\tau) / \kappa_j = \sum_{i=j}^n \alpha_{i,j,n} \mathbf{K}_i(\tau)$  se obtiene:

$$\int_0^{+\infty} a_{j,n}(\tau) d\tau = \frac{1}{\kappa_j} = \sum_{i=j}^n \alpha_{i,j,n}.$$

Por lo cual, la ecuación (2.15) se mantiene. □

### 2.3 Modelo de Dimensión Invariante

En esta sección se desarrollan cascadas parsimoniosas para las cuales la dimensión del espacio de parámetros es invariante al número de factores basándose en las siguientes condiciones impuestas al modelo.

**Supuesto 4. Velocidades de ajuste** La secuencia de velocidades de ajuste sigue una progresión geométrica:

$$\kappa_j = kb^{j-1}, j = 1, 2, \dots, n. \quad (2.16)$$

donde  $b > 1$  y  $k > 0$  son números reales contantes.

La velocidad de ajuste del primer factor  $k$  y el parámetro de espacio  $b$  controlan la velocidad de ajuste de todos los demás factores. Esta forma se debe a como el mercado ve la información que se distribuye a lo largo de la curva de la estructura a plazos.

Para obtener una cascada estrictamente especificada, se hacen unas suposiciones finales que son consistentes con la intuición de que cada factor carga aproximadamente una cantidad igual de información y riesgo:

**Supuesto 5.a. Volatilidades** Volatilidad de los factores es idéntica:  $\sigma_j = \sigma$  para todo  $j$ .

Se propone una variación al **supuesto 5.a** presentado en Calvet et al(2018) [4] para permitir que las dinámicas de la volatilidad estocástica del proceso de reversión a la media sean

diferentes a un valor constante

**Supuesto 5.b. Volatilidades** Volatilidad de los factores siguen una progresión geométrica ajustada por un nuevo parámetro  $s$ :  $\sigma_j = \sigma b^{(j-1)s}$  para todo  $j$ .

El supuesto 5.b, es el cambio en la especificación de la volatilidad del Modelo de cascada Estándar que lo hace dependiente. Al modelo resultante se le conocerá como el Modelo Sigma Variante (SV).

**Supuesto 6. Risk Premia** Los precios de riesgo del mercado en todos los factores de riesgo son idénticos:  $\gamma_j = \gamma$  para todo  $j$ .

Los supuestos 1-6 definen la versión base del modelo de estructura a plazos, por lo que se tendrán dos especificaciones de la cascada: El modelo de cascada estándar presentado por Calvet et Al [4] ajustado para curvas yield, y el modelo de cascada sigma variante, que es el propuesto en este trabajo de tesis con la variación única del supuesto 5.b.

El modelo estándar está completamente especificado por 5 parámetros  $k, b, \sigma, \theta_r$  y  $\gamma$ , mientras que el sigma variante contiene 6, agregándole el parámetro  $s$  en la volatilidad. Bajo la medida riesgo-ajustada, los niveles de largo plazo del primer factor son  $\theta^Q = \theta_r - \gamma\sigma/k$ . Entonces, hay una relación afín 1 a 1 entre  $\theta^Q$  y el parámetro del risk premia  $\gamma$ . Por ser  $\theta^Q$  directamente comparable con  $\theta_r$ , se define el vector de parámetros del modelo estándar como:

$$(k, b, \sigma, \theta_r, \gamma) \in (0, \infty) \times (1, \infty) \times (0, \infty) \times \mathbb{R}^2$$

Y los del modelo sigma variante:

$$(k, b, \sigma, \theta_r, \gamma, s) \in (0, \infty) \times (1, \infty) \times (0, \infty) \times \mathbb{R}^2 \times (0, \infty)$$

# Capítulo 3

## Datos y Estimación

### 3.1 Datos

El presente trabajo utiliza datos de la curva cero cupón soberana de Colombia, extraídos de Bloomberg, con vencimientos de 1, 2, 4, 6, 7, 8, 10 y 15 años. Los datos son semanales con corte los viernes, con el dato de la tasa media entre bid y ask desde el 23 de septiembre de 2005 hasta el 28 de diciembre del 2018. Existen 704 observaciones semanales de cada una de las 8 series de tiempo, para un total de 5632 observaciones.

Para el análisis de robustez, se separó la muestra en 2 submuestras, la primera para la estimación, la cual cubre el periodo del 23 de septiembre del 2005 hasta el 30 de diciembre del 2016, para un total de 589 observaciones semanales para cada una de las series de tiempo. La otra submuestra corresponde al periodo del 6 de enero del 2017 hasta el 28 de diciembre del 2018, con un total de 115 observaciones.

Se hizo una transformación de las tasas yield con capitalización compuesta a capitalización continua. Para los vacíos en ciertas fechas para ciertos vencimientos se utilizó el método de interpolación lineal con backward fill, ya que al hacer un análisis gráfico comparativo entre este y los métodos spline, polinomial (con diferentes órdenes), era el que tenía mejor comportamiento. Los otros métodos formaban gráficas a través del tiempo que no eran intuitivamente creíbles.

La Tabla 3.1 muestra el resumen estadístico de cada serie con respecto a los plazos de la estructura a plazos. En estas se incluye la media, desviación estándar, asimetría, kurtosis y la autocorrelación semanal de orden 1. Consistente con la literatura, el promedio de la estructura a plazos es creciente, mientras que la desviación estándar disminuye con la madurez. La asimetría y kurtosis tienen valores pequeños y generalmente positivos. Todas las series de las tasas de

interés son altamente persistentes, sus autocorrelaciones de primer orden van desde 0.9892 a 0.9954, siendo la de dos años la más persistente de todas.

Tabla 3.1: Resumen Estadístico de Tasas Yield

Madurez	Media	Desv. Std	Skew.	Kurt.	Autocorrelación
					1
1 año	5.695	1.849	1.007	-0.037	0.9937
2 años	6.201	1.799	1.126	0.239	0.9958
4 años	6.859	1.770	0.909	0.137	0.9904
6 años	7.331	1.721	0.823	0.112	0.9964
7 años	7.607	1.525	0.907	0.418	0.9975
8 años	7.567	1.665	0.448	-0.752	0.9982
10 años	7.776	1.599	0.807	0.112	0.9959
15 años	7.953	1.429	0.681	0.029	0.9938
Promedio	7.077	1.669	0.8385	0.032	0.9952

La Tabla 3.1 reporta un resumen estadístico de observaciones semanales (con los mid-quotes de los viernes) de las tasas yield con vencimientos de 1,2,4,6,7,8,10 y 15 años. Cada serie contiene 704 observaciones semanales desde 23 de septiembre del 2015 al 28 de diciembre del 2018. Las entradas reportan media (Media), desviación estándar (Desv. Std), asimetría (Skew), exceso de kurtosis (Kurt) y la autocorrelación semanal de primer orden.

### 3.2 Modelo

Las tasas yields relacionan los precios de los bonos cero cupón por medio de una transformación logarítmica de la función (2.11):

$$y(t, \tau) = -\frac{1}{\tau} \ln [P(\mathbf{X}_t, \tau)]$$

Para la estimación se modifica el DTSM de dimensión invariante para que posea una forma de espacio-estado y así estimar los parámetros usando un método de **Máxima Verosimilitud**. Se supone que los precios de los bonos son observados en intervalos de tiempo  $\Delta t$ . La ecuación de estado propagación (o de transición) es una discretización de la dinámica de la ecuación (2.2):

$$\mathbf{X}_t = \mathbf{A} + \Phi \mathbf{X}_{t-\Delta t} + \sqrt{\Sigma_X} \epsilon_t, \quad (3.1)$$

Donde  $\Delta t = 1/52$ ,  $\Phi = \exp(-\kappa \Delta t)$ ,  $\mathbf{I}_n$  denota una matriz identidad  $n$ -dimensional,  $\mathbf{A} = (\mathbf{I}_n - \Phi) \boldsymbol{\theta}$ ,  $\{\epsilon_t\}$  es i.i.d  $\mathcal{N}(0, \mathbf{I}_n)$ , y  $\Sigma_X = \sigma^2 \Delta t \mathbf{I}_n$  para el modelo estándar y  $\Sigma_X = \sigma_j^2 \Delta t \mathbf{I}_n$

con  $\sigma_j = \sigma b^{(j-1)s}$  para el modelo sigma variante.

Las ecuaciones de medición están construidas de observaciones de las tasas yield continuas, pero agregándole un error de medida gaussiano que permita discrepancias entre tasas observadas y tasas teóricas para evitar problemas de subidentificación por características del mercado Chatterjee(2005) [5]:

$$\mathbf{y}(t, \boldsymbol{\tau}) = \frac{\mathbf{b}(\boldsymbol{\tau})^T \mathbf{X}_t + \mathbf{c}(\boldsymbol{\tau})}{\boldsymbol{\tau}} + \mathbf{e}_t \quad (3.2)$$

Donde  $\mathbf{y}(t, \boldsymbol{\tau})$  denota el vector de datos,  $\mathbf{b}(\boldsymbol{\tau})$  y  $\mathbf{c}(\boldsymbol{\tau})$  son los coeficientes que satisfacen la relación del modelo con los yield a través de los factores  $\mathbf{X}_t$ , y  $\mathbf{e}_t$  denota un vector de errores de medición normalmente distribuidos i.i.d.  $\mathcal{N}(0, \sigma_e^2)$ .

Como tanto la ecuación de propagación o transición y la ecuación de medición son modelos Gaussianos con comportamiento lineal, se puede usar el Filtro de Kalman para obtener la trayectoria de los factores  $\mathbf{X}_t$ , además de unos estimados de *maxima verosimilitud* para encontrar el vector de parámetros  $(k, b, \sigma, \theta_r, \sigma_e^2)$  para el modelo estándar y  $(k, b, \sigma, \theta_r, \sigma_e^2, s)$  para el sigma variante, donde  $\sigma_e^2$  cuantifica el error de ajuste en las tasas yield observadas y los demás parámetros gobiernan las dinámicas de la estructura a plazos.

### 3.3 Estimación

#### 3.3.1 Filtro de Kalman

Una vez que el modelo se encuentra en la forma de espacio-estado, el filtro de Kalman [10] puede ser utilizado para obtener pronósticos óptimos de la media y la matriz de covarianza de un vector de estado normalmente distribuido, basado en la información a través del tiempo. se define como un filtro ya que actualiza el conocimiento sobre el vector de estado cada vez que nueva información de la variable observada está disponible, en este caso  $\mathbf{y}(t, \boldsymbol{\tau})$  Megner(2009) [12]. En esta especificación del modelo no son observables las variables de estado  $\mathbf{X}_t$ , los parámetros de la ecuación de transición y la ecuación de medida. Por ello es importante usar este método de filtrado recursivo.

El modelo base de la ecuación de transición y medición para el Filtro de Kalman es:

$$\mathbf{X}_t = \mathbf{A} + \mathbf{F}\mathbf{X}_{t-\Delta t} + \mathbf{D}\boldsymbol{\varepsilon}_t \quad (3.3)$$

$$\mathbf{y}_t = \mathbf{C} + \mathbf{H}\mathbf{X}_t + \mathbf{e}_t \quad (3.4)$$

Con  $\mathbf{X}_t$ ,  $\mathbf{A}$ ,  $\mathbf{F} = \Phi$ ,  $\mathbf{D} = \sqrt{\Sigma_X}$  y  $\boldsymbol{\varepsilon}_t$ <sup>1</sup>. Además,  $\mathbf{y}_t$ ,  $\mathbf{C} = \frac{c}{\tau}$ ,  $\mathbf{H} = \frac{b}{\tau}$  y  $\mathbf{e}_t$ <sup>2</sup>. Los errores  $\boldsymbol{\varepsilon}_t$  y  $\mathbf{e}_t$  se distribuyen normales con media cero y varianzas  $\mathbf{Q}$  y  $\mathbf{R}$  respectivamente.

La distribución condicional de  $\mathbf{X}_t$  puede ser caracterizada por su media  $\hat{\mathbf{x}}_{t|t}$  y varianza condicional  $\mathbf{P}_{t|t}$ . El filtro de Kalman provee una manera directa de computar  $\hat{\mathbf{x}}_{t|t} = \mathbb{E}[\mathbf{X}_t | \mathcal{F}_t]$  recursivamente dado cualquier valor de partida para  $\hat{\mathbf{x}}_{1|0}$  y  $\mathbf{P}_{1|0}$ , donde el gorro sobre una variable significa su valor esperado condicional a la filtración en  $t$ ,  $\mathcal{F}_t$  [11].

$$K_{t-1} = \mathbf{P}_{t-1|t-2} \mathbf{H}^T \Sigma_e^{-1} \quad \text{Kalman Gain} \quad (3.5)$$

$$\Sigma_e = \mathbf{H} \mathbf{P}_{t-1|t-2} \mathbf{H}^T + \mathbf{R} \quad \text{Covarianza Error de Medición} \quad (3.6)$$

$$\mathbf{P}_{t|t-1} = \mathbf{F} \mathbf{P}_{t-1|t-2} \mathbf{L}_{t-1}^T + \mathbf{D} \mathbf{Q} \mathbf{D}^T \quad \text{Actualización Covarianza de Estado} \quad (3.7)$$

$$\mathbf{L}_{t-1} = \mathbf{F} (\mathbf{I} - K_{t-1} \mathbf{H}) \quad \text{Corrección Covarianza de Estado} \quad (3.8)$$

$$\hat{\mathbf{y}}_{t|t} = \mathbf{C} + \mathbf{H} \hat{\mathbf{x}}_{t|t}^T \quad \text{Actualización Vector Observaciones} \quad (3.9)$$

$$\mathbf{e}_t = \mathbf{y}_t - \hat{\mathbf{y}}_{t|t} = \mathbf{H} \mathbf{f}_{t-1} + \mathbf{e}_{t-1} \quad \text{Error de Estimación Medición} \quad (3.10)$$

$$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t-1} + K_t \mathbf{e}_t \quad \text{Actualización Vector Estado} \quad (3.11)$$

$$\hat{\mathbf{x}}_{t+1|t} = \mathbf{A} + \mathbf{F} \hat{\mathbf{x}}_{t|t} \quad \text{Estimación Vector Estado} \quad (3.12)$$

$$\mathbf{f}_{t+1} = \mathbf{x}_{t+1} - \hat{\mathbf{x}}_{t+1|t} = \mathbf{F} \mathbf{f}_t + \mathbf{D} \boldsymbol{\varepsilon}_t \quad \text{Error de Estimación Estado} \quad (3.13)$$

Así, el Filtro de Kalman es esencialmente un conjunto de ecuaciones matemáticas que implementan un estimador predictor-corrector que es óptimo en el sentido de que minimiza la covarianza estimada del error. Cuando se cumple alguna de las siguientes condiciones, se puede afirmar que es el mejor filtro de cualquier forma concebible: el sistema es lineal en el vector de

<sup>1</sup>con dimensión  $n \times 1$ ,  $n \times 1$ ,  $n \times n$ ,  $n \times n$  y  $n \times 1$  respectivamente para la ecuación de transición donde  $n$  es la dimensión de la cascada

<sup>2</sup>con dimensión,  $N \times 1$ ,  $N \times 1$ ,  $N \times n$  y  $N \times 1$  respectivamente, donde  $N$  es el número de plazos de la estructura a plazos

factores  $X_t$  y/o los errores de transición y medición son ruido blanco con distribución gaussiana. Ambos se cumplen en el modelo presente.

El algoritmo usa las ecuaciones de Kalman (3.5)-(3.13), primero con los valores iniciales asumidos  $\hat{x}_{1|0}$  y  $P_{1|0}$ , que se eligen como la convergencia cuando  $t \rightarrow \infty$  de la solución analítica del modelo de Vasicek [18] siendo  $\hat{x}_{1|0} = \theta_r$ , la tasa de largo plazo y  $P_{1|0} = \frac{\sigma^2}{2k}I$ . Luego hace una repetición de las ecuaciones, tantas veces como número de factores se tengan, pero con los valores ajustados generados por el filtro. Esto para hallar el trayecto de los factores, la matriz de varianza condicional  $\Sigma_e$  y el error de las variables observadas  $e_t$ .

Con las matrices generadas por el filtro, se pueden estimar las matrices de parámetros desconocidas por medio del método de máxima verosimilitud. El modelo debe ser paramétrico y totalmente especificado por medio de la función de densidad de probabilidad conjunta. Para el número de observaciones en el tiempo en la muestra  $t \leq T$ , se utiliza la suma de las funciones de verosimilitud en forma logarítmica. Esto disminuye las dificultades computacionales y es equivalente a su forma no logarítmica.

Para este modelo, como se supuso que los errores tenían una distribución normal, la función de verosimilitud se comportará como un kernel en forma logarítmica, sumándolo para cada observación  $T$ . Como las observaciones de una serie de tiempo no son generalmente independientes, se usa es la función conjunta y no el producto de las densidades individuales:

$$\log L = -\frac{NT}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \det |\Sigma_e| - \frac{1}{2} \sum_{t=1}^T e_t' [\Sigma_e]^{-1} e_t \quad (3.14)$$

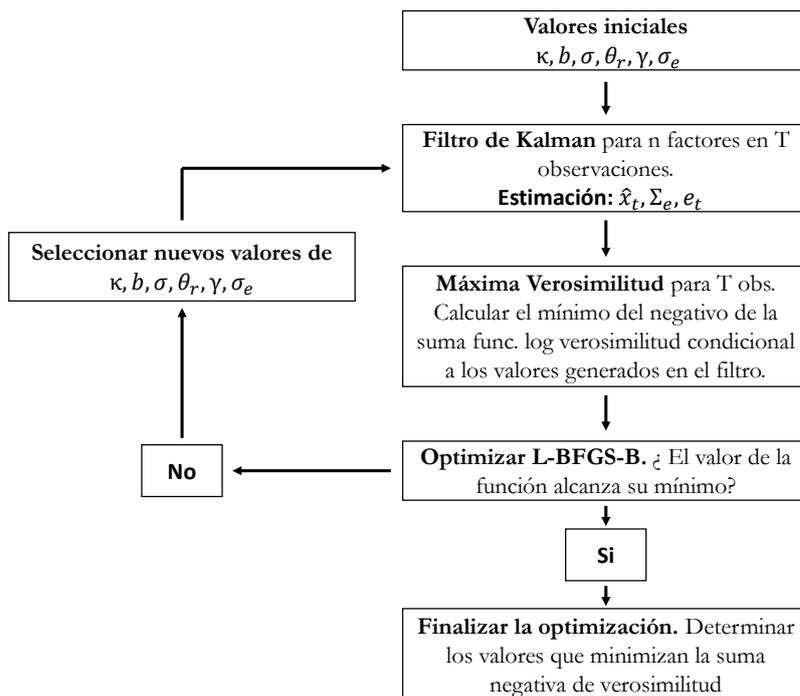
### 3.3.2 Optimización

El siguiente paso es la optimización a través de un método Broyden–Fletcher–Goldfarb–Shanno de memoria limitada con fronteras o L-BFGS-B. Este aproxima el algoritmo de BFGS usando una pequeña cantidad de memoria computacional, ya que solo utiliza la función y el gradiente, no la matriz Hessiana y además permite la limitación de los parámetros, Byrd(1995) [3]. Este algoritmo quasi Newtoniano es útil para optimizar la función de log verosimilitud ya que minimiza el valor negativo de la función  $\log L$  limitando el espacio de posibilidades de los parámetros deseados  $(k, b, \sigma, \theta_r, \gamma, \sigma_e^2)$  para el modelo estándar y  $(k, b, \sigma, \theta_r, \gamma, \sigma_e^2, s)$  para el SV.

Donde con  $\gamma$  se obtiene, bajo la medida riesgo ajustada, el **nivel** de largo plazo del primer factor  $\theta_r^Q = \theta_r - \gamma\sigma/k$ .

El algoritmo de estimación y optimización se puede observar en la Figura 3.1, donde se muestran los pasos a seguir para obtener los parámetros del modelo y las trayectorias de los factores. Este algoritmo es usado para ambas versiones del modelo, tanto estándar como SV.

Figura 3.1: Algoritmo de Estimación y Optimización



El algoritmo depende de unos valores o parámetros iniciales, que se filtran eficientemente con el filtro de Kalman (3.13) obteniendo las trayectorias de los  $n$  factores, los errores de medición de la estructura a plazos y la varianza del error para cada observación. Con esto dos últimos se construye la función de verosimilitud que se suma para todas las observaciones según la ecuación (3.14). Luego se hace el proceso de minimización del negativo de la función de log verosimilitud con el algoritmo L-BFGS-B hasta encontrar los parámetros con los que se llega al menor valor. Este algoritmo se hace para 8 factores en los dos modelos.

# Capítulo 4

## Resultados Empíricos

En este capítulo se presentan los resultados obtenidos de la aplicación de los modelos expuestos en el capítulo 2 usando el proceso de estimación explicado en el capítulo 3 para la estructura a plazos de las tasas yield colombianas.

### 4.1 Estimación en la muestra

La Tabla 4.1 reporta los resultados de estimación del modelo estándar donde  $n$  es el número de factores que se usa en la cascada. Se puede observar que no existe una convergencia clara en los componentes cuando se aumentan los factores, se podría decir que son fluctuantes al tamaño de la cascada. Esto se puede deber a la interpolación lineal que se le hizo a los datos y los cambios realizados a la forma funcional del modelo.

La velocidad de ajuste del componente de menor frecuencia  $k$  se encuentra entre 0,1975 (6)<sup>1</sup> y 1,7900 (4), implicando una duración ( $1/k$ ) entre 5 años y 7 meses, donde todos sus valores son altamente significativos. El coeficiente de escala  $b$  fluctúa entre 1,12128 (6) y 2,01 (5), con valores significativos, representando que la mayor frecuencia tiene una duración entre 20 días y un mes. La tasa media de largo plazo estadística  $\theta_r$  parece más estable y con un pseudo decrecimiento con valores entre 0,0379 (8) y 0,0753 (2), aunque pierde su significancia estadística en los factores  $n = 5, 6$  y  $7$

Por otro lado, el risk premia es negativo en todos sus valores, excepto en el modelo de 6 factores, y altamente significativo en todos los factores, obteniendo un nivel de largo plazo  $\theta^Q$  con valores entre 0,0381 (8) y 0,0761 (2), vislumbrando que la estructura a plazos media tiene

---

<sup>1</sup>Los número entre paréntesis representan el número de factores usados en el modelo para obtener el valor correspondiente

una pendiente creciente. Por último, los errores de medición del modelo también son fluctuantes y no muestran un decrecimiento claro, son inestables en su significancia, ya que pasan de alta a ninguna significancia en la parte de media de la cascada. Esto puede ser un resultado de la linealización de la ecuación de medición para adaptarlo a las limitaciones de los datos en Colombia o la interpolación lineal de los datos.

Tabla 4.1: Parámetros Estimados y Errores Estándar  
Modelo Estándar

Parámetros								
$n$	$k$	$b$	$\sigma$	$\theta$	$\gamma$	$\sigma_e^{2+}$	$\theta^{Q_r}$	$\mathcal{L}$
1	1.0572*** (0.0233)	(—)	0.0110 (0.1076)	0.0702* (0.0412)	-0.4315*** (0.0023)	0.0311*** (0.0044)	0.0702	-12148.49
2	0.4977*** (0.0029)	1.4088*** (0.0028)	0.0312*** (0.0053)	0.0753*** (0.0110)	-0.4341*** (0.00001)	0.03741*** (0.0050)	0.0761	-12120.65
3	0.5795*** (0.0011)	1.4043*** (0.0010)	0.0106*** (0.0007)	0.0455*** (0.0013)	-0.4317*** (0.000005)	0.0302*** (0.0026)	0.0456	-12755.38
4	1.7900*** (0.1007)	1.4053*** (0.0485)	0.0305 (0.0548)	0.0610* (0.0336)	-0.4351*** (0.0129)	0.0612 (0.7079)	0.3164	-13107.42
5	1.3255*** (0.0909)	2.0069*** (0.1139)	0.02314 (0.0343)	0.0644 (0.0721)	-0.4345*** (0.0005)	0.0249*** (0.3373)	0.0646	-12796.83
6	0.1975*** (0.0522)	1.1218*** (0.0442)	0.0187*** (0.0476)	0.0466 (0.6501)	0.1290*** (0.0004)	0.0739 (6.4866)	0.0464	-9287.79
7	0.3771*** (0.0798)	1.8281** (0.7829)	0.0151 (0.0211)	0.0572 (0.0973)	-0.4105*** (0.0025)	0.0099 (0.483)	0.0574	-14163.49
8	0.4547*** (0.0031)	1.6104*** (0.0035)	0.0126*** (0.0025)	0.0379*** (0.0034)	-0.4309*** (0.00002)	0.0370*** (0.0098)	0.0381	-11532.20

La Tabla 4.1 reporta los estimados de máxima verosimilitud y los errores estándar (en paréntesis) de los parámetros del modelo estándar trabajados en Calvet(2018) [4]. Cada fila representa un set de parámetros estimados con la cascada con  $n = 1$  a  $n = 8$  componentes. La columna  $\mathcal{L}$  reporta el valor de log verosimilitud promedio semanal maximizada para cada modelo. Asintóticamente, el estadístico tiene una distribución normal estándar. Nivel de significancia: \* 10 %, \*\* 5 % y \*\*\* 1 %.

+ Tanto los errores de medición como sus desviaciones se encuentran en puntos básicos.

Para encontrar los errores estándar se utilizó el método de bootstrap con un remuestreo de los errores estandarizados expuesto en Stoffer(1984) [16] y Stoffer(1991) [17], teniendo como resultado que los únicos factores que logran hacer significantes todos los parámetros del modelo son  $n = 2, 3$  y  $8$ .

Es importante recalcar que este tipo de modelos dinámicos estocásticos cuentan con un número considerable de máximos locales, lo cual hace que los resultados de estimación en la optimización sean bastante sensibles a los parámetros iniciales que se elijan para el modelo. Los parámetros que resultan dependen bastante de la capacidad computacional que se posea, pareciera que el algoritmo de optimización hiciera menos iteraciones cuando no tiene suficiente capacidad, haciendo que sean similares al punto inicial (estancado en un máximo local) o que

se desvíen en el proceso de optimización, aunque se tenga el mismo punto inicial. Esto podría ser un determinante de la falta de convergencia de los parámetros del modelo y su falta de significancia. Por lo tanto, no se le da tanta importancia a la significancia estadística en el análisis de los modelos, pero sí se presentará y comentará.

Por su parte, la Tabla 4.2 reporta los resultados de estimación del modelo SV. Aunque no se observa una mejora contundente en la convergencia de los componentes o parámetros del modelo, sí existe un decrecimiento claro en los errores de medición cuando se aumentan los factores, por lo menos hasta el factor 7, pese a que estos no sean significantivos sino hasta el cuarto factor. Asimismo, posee una mejor significancia estadística en algunos parámetros a través de los modelos. Parece que hay una mejora en utilizar el modelo SV sobre el modelo estándar, sobre todo en estabilidad de los resultados. Sin embargo, como se mencionó, la forma del modelo es muy cambiante y no genera certidumbre sobre su comportamiento.

Tabla 4.2: Parámetros Estimados y Errores Estándar  
Modelo SV

Parámetros									
$n$	$k$	$b$	$\sigma$	$\theta$	$\gamma$	$\sigma_e^{2+}$	$s$	$\theta_r^Q$	$\mathcal{L}$
1	0.2596*** (0.0256)	(—)	0.0257 (0.0253)	0.0601*** (0.0118)	-0.1406*** (0.0012)	0.7948*** (0.0030)	(—)	0.0605	-6862.94
2	0.1234*** (0.0007)	1.1523*** (.0001)	0.0122*** (0.0019)	0.0286 (0.1924)	0.4575*** (0.00002)	0.3780*** (0.0025)	0.1890*** (0.00001)	0.0280	-9876.86
3	1.0768*** (0.0069)	2.3100*** (0.0040)	0.0261*** (0.0066)	0.0317*** (0.0055)	-0.4314*** (0.00007)	0.1110*** (0.0023)	0.4989*** (0.0041)	0.0314	-10508.42
4	0.3271*** (0.0202)	1.4317*** (0.1635)	0.0024 (0.1041)	0.0805*** (0.0155)	-0.4463*** (0.0577)	0.0130*** (0.0075)	0.5027*** (0.0519)	0.0805	-15171.54
5	0.3373*** (0.0466)	1.4027*** (0.0601)	0.0108*** (0.0057)	0.0526* (0.0307)	-0.4346*** (0.0494)	0.0103 (0.0750)	0.5007*** (0.0770)	0.0528	-13998.85
6	1.4325*** (0.0827)	1.3914*** (0.0571)	0.0441 (0.0271)	0.0302 (0.0245)	-0.4326*** (0.0787)	0.0099 (0.7834)	0.5094*** (0.1013)	0.0309	-14192.72
7	0.1422*** (0.0016)	1.1927*** (0.0009)	0.0014 (0.0038)	0.0963*** (0.0001)	0.2779*** (0.00001)	0.0046 (0.2547)	0.4909*** (0.0002)	0.0963	-15587.02
8	0.2492*** (0.0060)	1.2806*** (0.0267)	0.0291* (0.0152)	0.0647** (0.0274)	-0.1018*** (0.0306)	0.0094 (1.5207)	0.3817*** (0.0215)	0.0650	-13291.66

La Tabla 4.2 reporta los estimados de máxima verosimilitud y los errores estándar (en paréntesis) de los parámetros del modelo  $\sigma_j$  variante. Cada fila representa un set de parámetros estimados con la cascada con  $n = 1$  a  $n = 8$  componentes, donde el error de medición está en puntos básicos. La columna  $\mathcal{L}$  reporta el valor de log verosimilitud promedio semanal maximizada para cada modelo. Asintóticamente, el estadístico tiene una distribución normal estándar. Nivel de significancia: \* 10 %, \*\* 5 % y \*\*\* 1 %.

+ Tanto los errores de medición como sus desviaciones se encuentran en puntos básicos.

En este modelo, la velocidad de ajuste del componente de menor frecuencia  $k$  está entre 0,1234 (2) y 1,4325 (6), implicando una duración de  $(1/k)$  entre años 9.7 y 8 meses. El coeficiente de escala  $b$  fluctúa en entre 1,1523 (2) y 2,3100 (3) con duraciones entre 20 días y un

poco más de un mes. Ambos parámetros son altamente significantes para todos los modelos. La tasa media de largo plazo estadística  $\theta_r$  tiene valores menores que el modelo estándar pero no posee una tendencia clara. Sus valores se encuentran en un rango entre 0,0286 (2) y 0,0063 (7), donde solo los modelos de 2 y 6 factores no muestran significancia en sus parámetros.

El risk premia generalmente es negativo, con excepciones en el modelo de 2 y 7 factores, y con todos sus valores son significantes. Así, se obtiene un nivel de largo plazo  $\theta^Q$  con valores entre 0,0280 (2) y 0,0963 (7), donde se genera una pendiente positiva para todos los factores. Por último, la posible mejora en los errores del modelo se puede deber al parámetro de ajuste de la volatilidad del modelo  $s$ . Primero, porque es altamente significativo en todos los modelos, robándole el poder a la volatilidad  $\sigma$ . Segundo, porque está disminuyendo el impacto estocástico de la varianza en el modelo al tener valores entre 0,1890 (2) y 0,5094 (6), es decir, menores a 1.

#### 4.1.1 Ajuste en la muestra

Para el análisis del ajuste de los modelos en la muestra se utilizará la raíz cuadrada del error cuadrático medio (RMSE por sus siglas en inglés), para determinar qué tanto se alejan los valores generados por el modelo de los reales. El RMSE tiene la siguiente fórmula:

$$RMSE = \sqrt{\mathbb{E} (\mathbf{y}_{t,\tau} - \hat{\mathbf{y}}_{t,\tau})' (\mathbf{y}_{t,\tau} - \hat{\mathbf{y}}_{t,\tau})} \quad (4.1)$$

La Tabla 4.3 reporta los errores de ajuste en la muestra cuando se agregan factores en el modelo estándar. Se observa cómo el error va disminuyendo cuando se van agregando factores hasta el factor 5 con un valor de 0,2794. Luego hay un aumento para los factores más grandes, donde existe un muy mal ajuste en el modelo de 6 factores con desviaciones de casi 2%. Esto puede deberse a que en el valor de su tasa de largo plazo hay alta probabilidad de que genere una curva invertida. Disulmbrando que contrario a los valores surgidos en los parámetros, el error de medición necesariamente no coincide con el ajuste de los modelos.

Por su parte, la Tabla 4.4 reporta los errores de ajuste en la muestra cuando se agregan factores en el modelo SV. Se puede notar que no necesariamente el modelo con menor error de medición en la Tabla 4.2 es el que posee mejor ajuste en la muestra, reafirmando lo ya discutido

de que hay una disociación entre lo que expresan los parámetros y el verdadero ajuste en la muestra de los modelos.

Tabla 4.3: Raíz de Errores de Valoración Cuadráticos Medios (RMSE) Modelo Estándar

Madurez	$n$							
	1	2	3	4	5	6	7	8
1 año	0.5037	0.8222	0.6478	0.6681	0.4894	2.7089	0.9226	1.3220
2 años	0.4569	0.6496	0.3949	0.2561	0.1153	2.5289	0.7779	1.0354
4 años	0.4260	0.2942	0.0763	0.1182	0.1850	2.2392	0.5338	0.5584
6 años	0.4245	0.0864	0.1884	0.2196	0.2692	1.9981	0.3657	0.2788
7 años	0.3843	0.0955	0.2585	0.2126	0.2545	1.8555	0.3460	0.2165
8 años	0.4140	0.1902	0.3140	0.2680	0.3036	1.7671	0.3029	0.2039
10 años	0.4125	0.2751	0.3747	0.2756	0.3093	1.5476	0.1834	0.14283
15 años	0.3907	0.4614	0.4546	0.2840	0.3091	1.0422	0.1572	0.2479
Promedio	0.4266	0.3593	0.3384	0.2878	0.2794	1.9609	0.4487	0.5014

La Tabla 4.3 reporta la raíz cuadrada de los errores de valoración cuadráticos medios de las tasas de rendimiento yield de la cascada de Calvet(2018) [4] con  $n = 1$  a  $n = 8$  componentes. La última fila reporta el promedio sobre las 8 series de tasa de interés.

Tabla 4.4: Raíz Cuadrada de Errores de Valoración Cuadráticos Medios (RMSE) Modelo SV

Madurez	$n$							
	1	2	3	4	5	6	7	8
1 año	0.2778	0.7127	0.1975	0.3307	0.6080	0.6390	0.7046	0.4186
2 años	0.2212	0.4690	0.1971	0.2139	0.5500	0.5233	0.5642	0.3800
4 años	0.2411	0.1197	0.4391	0.0661	0.4355	0.2320	0.3726	0.3575
6 años	0.2971	0.2862	0.5649	0.1109	0.3486	0.1797	0.2875	0.3765
7 años	0.3047	0.4390	0.6081	0.1057	0.3566	0.1483	0.2548	0.4222
8 años	0.3363	0.5540	0.6375	0.1662	0.3007	0.2135	0.2408	0.4062
10 años	0.3549	0.7759	0.6786	0.1678	0.2154	0.2002	0.1837	0.3558
15 años	0.3717	1.2278	0.7155	0.2016	0.1918	0.2241	0.1502	0.2936
Promedio	0.3006	0.5730	0.5048	0.1704	0.3758	0.2950	0.3448	0.3762

La Tabla 4.4 reporta la raíz cuadrada de los errores de valoración cuadráticos medios (en puntos básicos) de las tasas de rendimiento yield de la cascada  $\sigma_j$  variante con  $n = 1$  a  $n = 8$  componentes. La última fila reporta el promedio sobre las 8 series de tasa de interés.

En este caso, no existe un decrecimiento del error promedio en la muestra cuando se aumentan los factores como en el modelo estándar. Es más, los resultados reflejan que no se deberían escoger los modelos de 2 y 3 factores cuando el modelo tiene una volatilidad que varía en el tiempo con la forma funcional presentada en el supuesto 5.b. Este mal ajuste para estos dos factores puede deberse a que existe un significativo error en la madurez de 15 años para el factor 3 y tanto de 1 como de 15 años para el modelo de factor 2, creando valores extremos que afectan

la media. En general, el tener una volatilidad cambiante en el tiempo genera más fluctuaciones en el ajuste de los modelos comparado con una volatilidad constante.

La Tabla 4.4, también muestra que el modelo que tiene mejor ajuste es el que utiliza 4 factores, con error medio de solo 0,1706. Esto, comparado con el error de 0,2794 del modelo estándar de 5 factores, demuestra que es mejor en el ajuste *en la muestra*. Por lo cual, existe evidencia de que el modelo sigma variante sí posee una mejora en estimación en la muestra sobre el modelo estándar, aunque se debe ser muy cuidadoso en cuantos números de factores se utilizan.

#### 4.1.2 Pronóstico por fuera de la muestra

Para el pronóstico por fuera de la muestra se analizará el desempeño del modelo estándar y el modelo SV para factores de 1 a 8, comparado con una caminata aleatoria por medio de una función de pérdida de RMSE:

$$RMSE = \sqrt{\mathbb{E} (\mathbf{y}_{t+h,\tau} - \hat{\mathbf{y}}_{t+h,\tau})' (\mathbf{y}_{t+h,\tau} - \hat{\mathbf{y}}_{t+h,\tau})} \quad (4.2)$$

Donde  $h$  representa la longitud de pasos hacia adelante que se toman para la evaluación de pronóstico, es decir el horizonte de pronóstico. En este caso se harán pronósticos cada 1 y 4 semanas con una reestimación de los parámetros cada 6. Esto por la alta demanda computacional del modelo.

La caminata aleatoria (RW) se elige como modelo benchmark por su simplicidad teórica y facilidad de estimación, cuya fórmula es:

$$\hat{\mathbf{y}}_{t+h,\tau} = \mathbf{y}_{t,\tau} \quad (4.3)$$

La Tabla 4.5 muestra el desempeño de pronóstico por fuera de la muestra del modelo estándar para 1 y 4 semanas, con reestimación cada 6 semanas. La primera columna de cada pronóstico reporta el RMSE de la caminata aleatoria, mientras que las demás columnas reportan el RMSE de la cascada con 1-8 frecuencias. En ambos horizontes hay comportamiento similar. Se puede observar que la caminata aleatoria muestra el mejor desempeño en el pronóstico. Para

el yield a 1 y 15 años, tiene un error de predicción de 0,8119 y 1,1219 puntos básicos para un horizonte de 1 semana con promedio de 0,8572 y, 3,1380 y 2,2490 para un horizonte de 4 semanas con promedio de 2,6934 , respectivamente.

Tabla 4.5: RMSE (Puntos Básicos) Fuera de la Muestra  
Modelo Estándar

	Madurez (años)								Promedio
	1	2	4	6	7	8	10	15	
<b>Pronóstico 1 Semana</b>									
RW	0.8119	1.1001	0.8062	0.4097	0.4199	1.1391	1.1925	1.1219	0.8572
<b>Cascada</b>									
$n = 1$	50.0708	40.8546	30.1365	21.7134	18.7932	21.8067	18.7226	15.2931	27.1801
$n = 2$	113.8480	98.3167	74.6923	57.3738	51.0289	50.8753	43.9347	35.1057	65.6469
$n = 3$	66.3987	55.5995	41.3273	30.8611	27.2226	29.5263	25.7817	21.4080	37.3123
$n = 4$	97.9679	64.1557	39.5245	26.9895	22.9334	25.0857	20.9140	15.9833	39.1932
$n = 5$	82.4970	53.5531	33.6517	22.7980	19.2099	21.6646	17.9714	13.6820	33.1285
$n = 6$	139.1800	133.8105	124.3817	113.6382	109.0188	109.1236	101.3739	83.3652	114.2365
$n = 7$	166.1269	140.3636	104.5433	80.4709	71.8436	69.7512	59.7771	46.2397	92.3958
$n = 8$	161.3521	137.7070	100.2842	75.0047	66.1986	64.0633	54.1253	41.0136	87.4686
<b>Pronóstico 4 Semanas</b>									
RW	3.1380	3.8857	2.6221	2.5791	2.2035	2.5233	2.3478	2.2490	2.6934
<b>Cascada</b>									
$n = 1$	24.9899	20.3511	15.0426	10.8606	9.3656	10.8850	9.4053	7.6595	13.5699
$n = 2$	56.8268	49.0314	37.2470	28.5880	25.3788	25.3080	21.8663	17.4364	32.7103
$n = 3$	32.6134	27.3813	20.4755	15.3710	13.5398	14.7156	12.9078	10.7428	18.4684
$n = 4$	48.4612	31.9930	19.7493	13.5030	11.4439	12.5299	10.4842	8.0096	19.5718
$n = 5$	41.1227	26.6817	16.8045	11.4076	9.5777	10.8125	9.0068	6.8547	16.5335
$n = 6$	66.2249	63.5917	59.0507	53.9605	51.6998	51.8423	48.2093	39.7739	54.2942
$n = 7$	82.7235	69.8719	52.1042	40.0995	35.7622	34.7238	29.7890	23.0379	46.0140
$n = 8$	79.7636	68.1500	49.7537	37.2862	32.8887	31.8485	26.9648	20.4698	43.3907

La Tabla 4.5 reporta la raíz cuadrada de los errores de valoración cuadráticos medios (RMSE) por fuera de la muestra de cada serie de tasa de interés para 1 y 4 semanas con reestimación de parámetros cada 6 semanas, del modelo estándar con  $n = 1$  a  $n = 8$  componentes. En cada panel se reporta los resultados de una caminata aleatoria (RW) y el desempeño de la cascada para  $n = 1$  a  $n = 8$  frecuencias. Se utilizan los parámetros generados por la reestimación, se pronostica desde el 1 de enero del 2017 hasta el 28 de diciembre del 2018 y se encuentra el error medio en la serie.

No hay evidencia clara de que las cascadas de baja dimensión (1-3) tengan un desempeño peor de que las de alta dimensión (6-8). Pareciera que existe una fluctuación en el desempeño de la cascada similar al visto en el análisis en la muestra, donde el modelo del modelo de 1 factor tiene un desempeño mejor que todos los demás para vencimientos menores a 7 años. Sin embargo, para la parte media y larga de la cascada el modelo de 5 factores tiene un mejor comportamiento. Aun así, la cascada con  $n = 1$  tiene en promedio un ajuste por fuera de la muestra de 27,1801 puntos básicos para un horizonte de 1 semana y 13,5699 para 4 semanas, mientras

que el  $n = 5$  tiene 33,1285 y 16,5335 para 1 y 4 semanas respectivamente. Esto muestra que el modelo estándar no genera un mejor pronóstico al aumentar el número de factores. Existe evidencia para decir que no es consistente con sus resultados, ya que propone el modelo de 5 factores para ajuste en la muestra y el de 1 factor para por fuera de la muestra, los cuales están considerablemente separados entre sí.

Tabla 4.6: RMSE (Puntos Básicos) Fuera de la Muestra  
Modelo SV

	Madurez (años)								Promedio
	1	2	4	6	7	8	10	15	
Pronóstico 1 Semana									
RW	0.8119	1.1001	0.8062	0.4097	0.4199	1.1391	1.1925	1.1219	0.8572
Cascada									
$n = 1$	48.1732	44.9362	37.9085	30.6922	27.9127	30.6793	27.6851	24.0275	34.0949
$n = 2$	179.0069	168.4306	149.8036	131.2246	123.0020	120.0676	107.1137	81.4439	132.5116
$n = 3$	57.1067	42.6106	30.5068	22.2517	19.3976	22.3339	19.4689	16.2654	28.7427
$n = 4$	58.0041	48.5964	37.0742	28.2334	25.1618	28.1364	25.0162	21.5376	33.9700
$n = 5$	123.1515	118.7957	107.2306	92.9443	86.4143	85.2952	75.8871	60.6507	93.7962
$n = 6$	155.0287	126.4439	80.0947	55.9316	48.4779	48.0555	40.2810	30.5007	73.1020
$n = 7$	87.5492	82.8867	74.6509	66.0652	62.3569	63.3738	58.3494	48.9972	68.0287
$n = 8$	88.2565	83.6355	75.4394	66.9595	63.2627	64.2591	59.2202	49.7731	68.8508
Pronóstico 4 Semanas									
RW	3.1380	3.8857	2.6221	2.5791	2.2035	2.5233	2.3478	2.2490	2.6934
Cascada									
$n = 1$	24.4034	22.3924	18.9395	15.3478	13.9223	15.3289	13.8696	12.0429	17.0309
$n = 2$	89.2273	83.9012	74.6637	65.4187	61.2830	59.8508	53.4040	40.5544	66.0379
$n = 3$	28.4098	21.2365	15.2719	11.1313	9.6603	11.1760	9.7724	8.1389	14.3496
$n = 4$	28.8383	24.1455	18.4552	14.0789	12.5095	13.9966	12.4768	10.7235	16.9030
$n = 5$	61.2230	59.0352	53.3265	46.2422	42.9525	42.3996	37.7398	30.1424	46.6326
$n = 6$	77.2383	62.9606	39.9157	27.8969	24.1453	23.9383	20.1002	15.2136	36.4261
$n = 7$	43.2106	40.8666	36.8082	32.5894	30.7097	31.2156	28.7540	24.1452	33.5374
$n = 8$	43.5365	41.1816	37.1069	32.8936	30.9977	31.4712	28.9769	24.2955	33.8075

La Tabla 4.6 reporta la raíz cuadrada de los errores de valoración cuadráticos medios (RMSE) por fuera de la muestra de cada serie de tasa de interés para 1 y 4 semanas con reestimación de parámetros cada 6 semanas, de la cascada  $\sigma_j$  variante con  $n = 1$  a  $n = 8$  componentes. En cada panel se reporta los resultados de una caminata aleatoria (RW) y el desempeño de la cascada para  $n = 1$  a  $n = 8$  frecuencias. Se utilizan los parámetros generados por la reestimación, se pronostica desde el 1 de enero del 2017 hasta el 28 de diciembre del 2018 y se encuentra el error medio en la serie.

La Tabla 4.6 reporta el desempeño de pronóstico por fuera de la muestra del modelo SV para 1 y 4 semanas, con reestimación cada 6 semanas, con un formato similar a la Tabla 4.5. En este caso, al igual que en el modelo estándar, la caminata aleatoria predice mejor a la cascada en cualquier frecuencia. Se puede observar que cuando  $\sigma$  depende del tiempo los errores son más grandes, tienen peor ajuste, y siguen el mismo comportamiento fluctuante a través de la cascada

ya encontrado en el análisis en la muestra. Aunque en general los valores de ajuste son mayores que en el modelo estándar, existe una mejora entre elegir  $n = 1$  con un promedio de 34,0949 y 17,0309 puntos básicos de error para horizontes de 1 y 4 semanas, y  $n = 4$  con 33,970 y 16,9030 puntos básicos.

Aun así, el mejor modelo es la cascada de 3 factores con un promedio de 28,7427 y 14,3496 puntos básicos de error para horizontes de 1 y 4 semanas. Esto muestra a pesar de que el modelo de 4 factores es el mejor modelo en la muestra, no es el mejor para el pronóstico. El modelo que lo supera no tiene una dimensión muy lejana, es solo un factor menos. Contrario a lo que sucede en el modelo estándar, parece que el modelo SV es más consistente con sus resultados en la muestra y por fuera de la muestra. Sin embargo, se puede ver que el modelo estándar es ligeramente superior en pronóstico que el modelo SV.

Analizando la información de error medio en ajuste y en pronóstico en la muestra y por fuera de la muestra, los resultados son más consistentes en el modelo SV que en el modelo estándar, por lo mencionado anteriormente. Así, la diferencia de desempeño de pronóstico que hace ligeramente mejor el modelo estándar por sobre el SV puede verse disminuida por la mejor estabilidad del modelo SV con  $n = 3$  y  $n = 4$  factores.

Los resultados con diferentes horizontes de tiempo muestran que al pasar de un paso cada 1 semana a pasos cada 4 semanas se disminuye el error de pronóstico a casi la mitad de su valor para cada modelo. Por el contrario, la caminata aleatoria aumenta su error. Este resultado puede deberse a que el horizonte de 4 semanas está mucho más cerca a periodo de reestimación de parámetros de 6 semanas que el horizonte de 1 semana, disminuyendo el ruido de parámetros sin actualizar. Por lo cual, podrían obtenerse menores errores si la reestimación se hiciera más recurrentemente.

En general, los dos modelos poseen un pobre desempeño de pronóstico para vencimientos de corto plazo, es decir, la parte de la curva yield de corto plazo. Esto se puede deber a la falta de liquidez que presenta el mercado de Colombia para plazos cortos. Ya que la cascada predice los niveles futuros de la tasa corta a través de los valores de las tasas de plazos más largos, el modelo no predice bien esta parte de la curva por no tener en consideración la iliquidez del mercado.

# Capítulo 5

## Análisis Gráfico

Se expone en las Figuras 5.1-5.4 cómo se compara el modelo de cascada estándar y SV a través del tiempo para las diferentes frecuencias  $n = 1$  hasta  $n = 8$ , en cada plazo o vencimiento de la curva yield colombiana con el verdadero valor de las tasas. Las Figuras 5.1 y 5.2 muestran el ajuste en la muestra, mientras que las Figuras 5.3 y 5.4 el ajuste por fuera de la muestra.

Por último, las Figuras 5.5-5.8 muestran el comportamiento de la curva yield o estructura a plazos en un punto específico del tiempo. En este caso, se eligieron el 19 de junio del 2009 y el 11 de octubre del 2013 para el análisis de ajuste en la muestra por ser puntos donde hay un cambio de tendencia en la serie. Por otro lado, se eligieron las fechas del 8 de septiembre del 2017 y el 9 de noviembre del 2018 por ser los puntos donde más se aleja y se acerca la predicción del modelo a la curva, respectivamente. Se observa que los modelos de frecuencias particulares con mal ajuste y/o pronóstico tienden a ser los que invierten totalmente la curva, como el caso del ajuste en la muestra de la cascada SV de 2 factores y la del modelo estándar de 6.

Figura 5.1: Ajuste en la Muestra Modelo Estándar

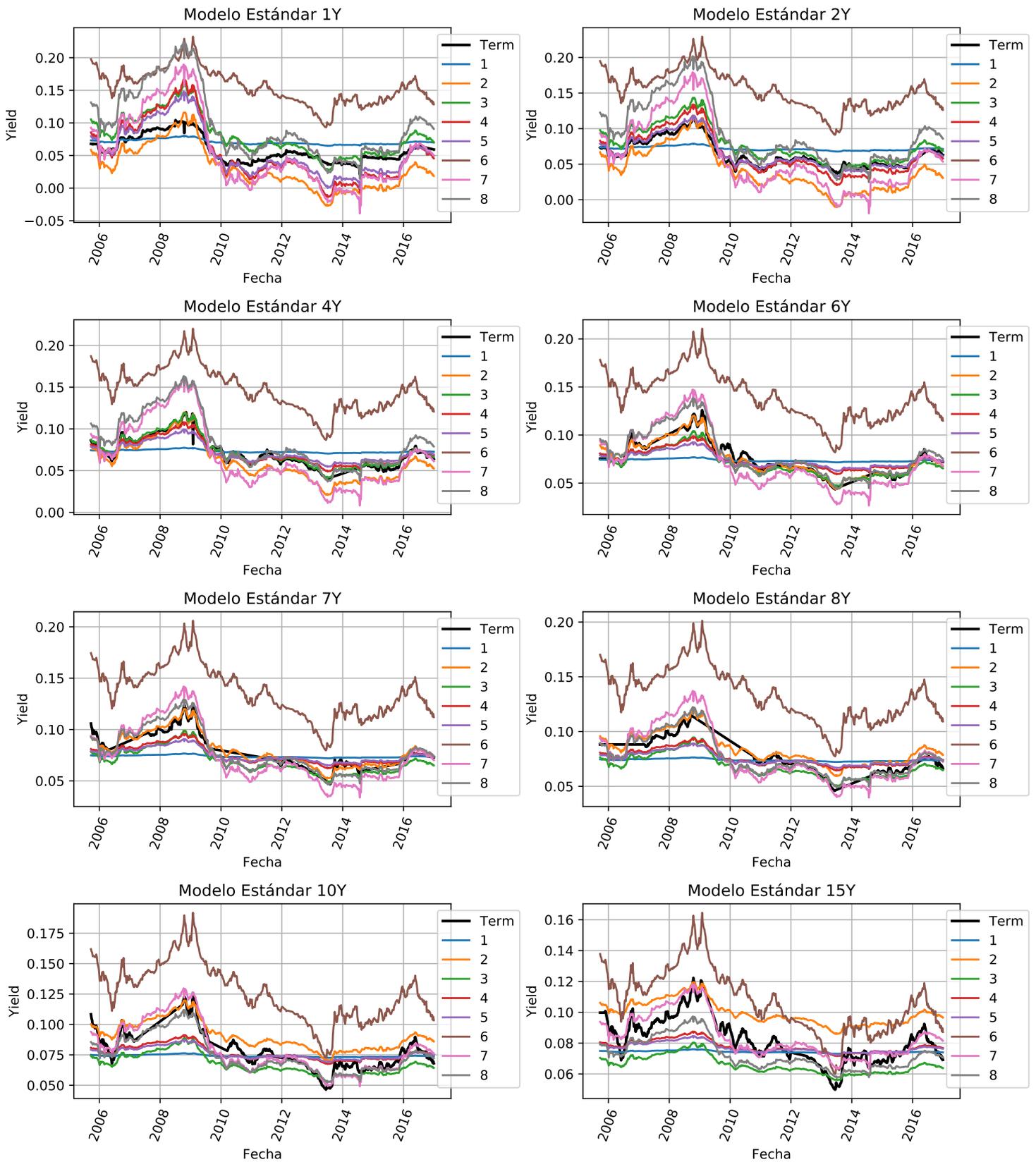


Figura 5.2: Ajuste en la Muestra Modelo SV

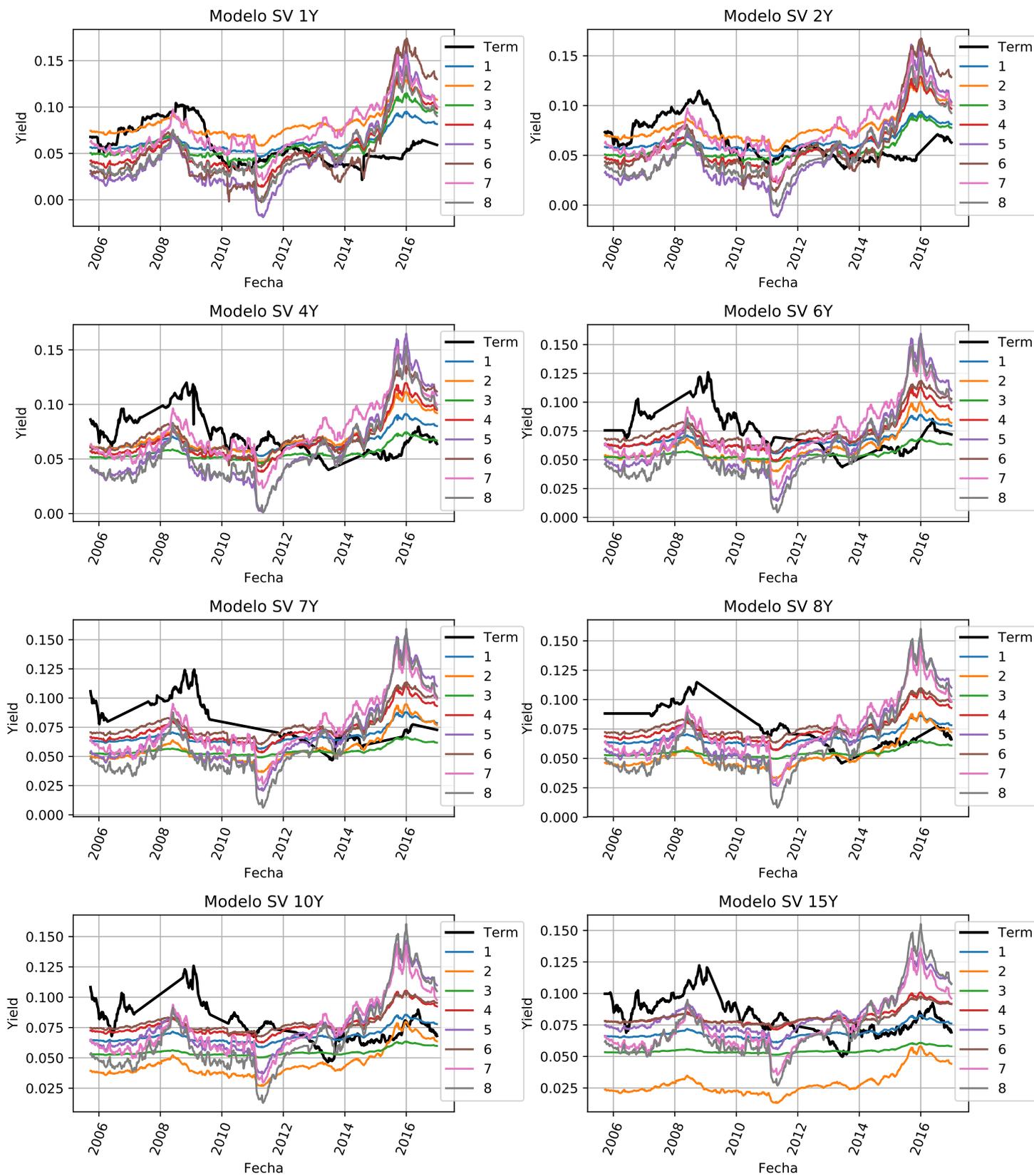


Figura 5.3: Ajuste por Fuera de la Muestra Modelo Estándar

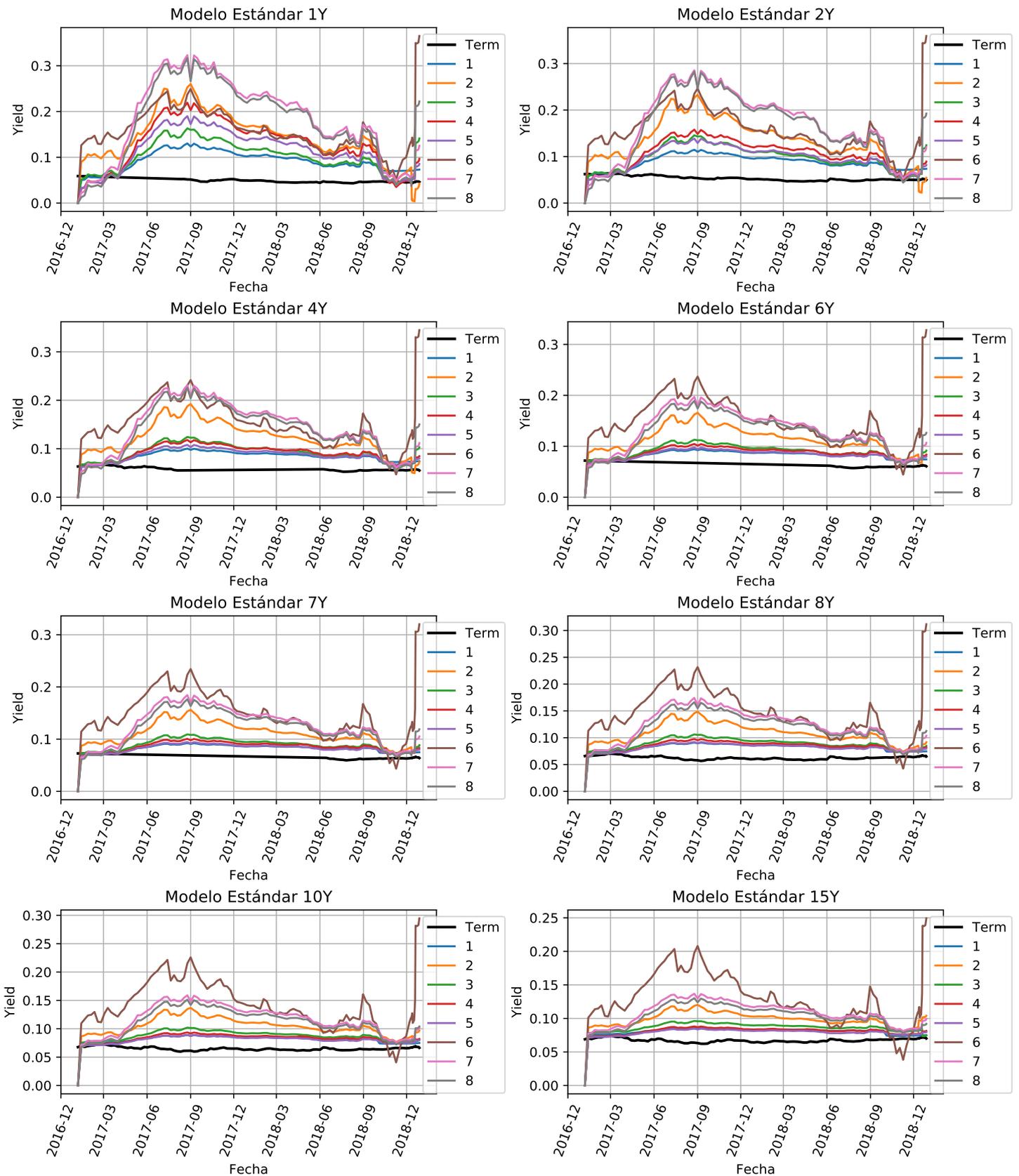


Figura 5.4: Ajuste por Fuera de la Muestra Modelo SV

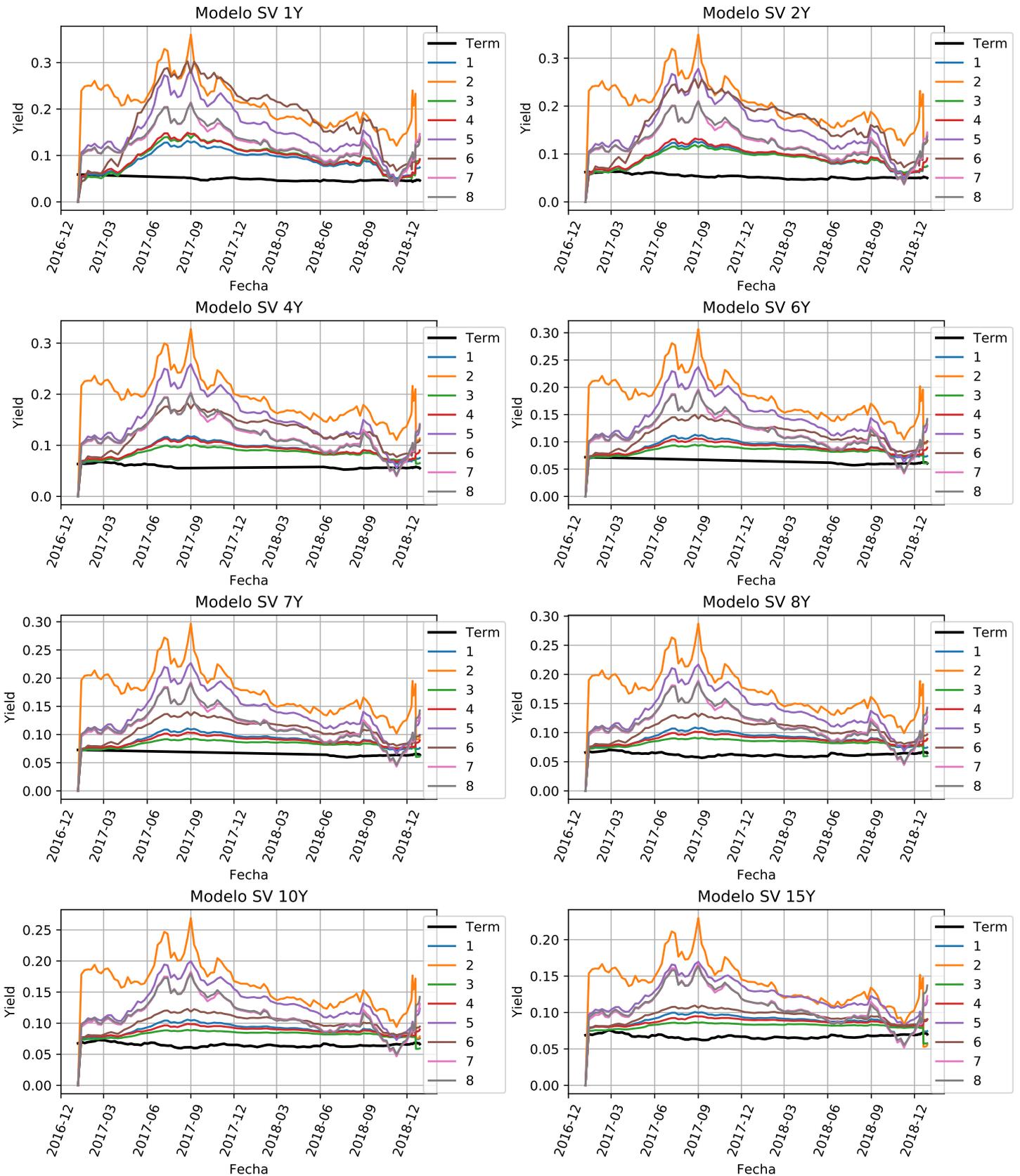


Figura 5.5: Ajuste Curva Yield Modelo Estándar

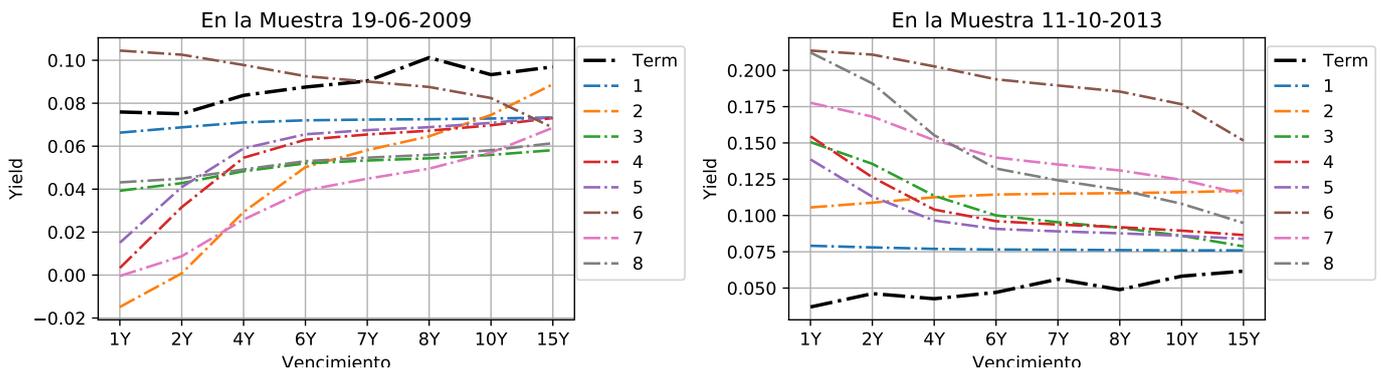


Figura 5.6: Ajuste Curva Yield Modelo SV

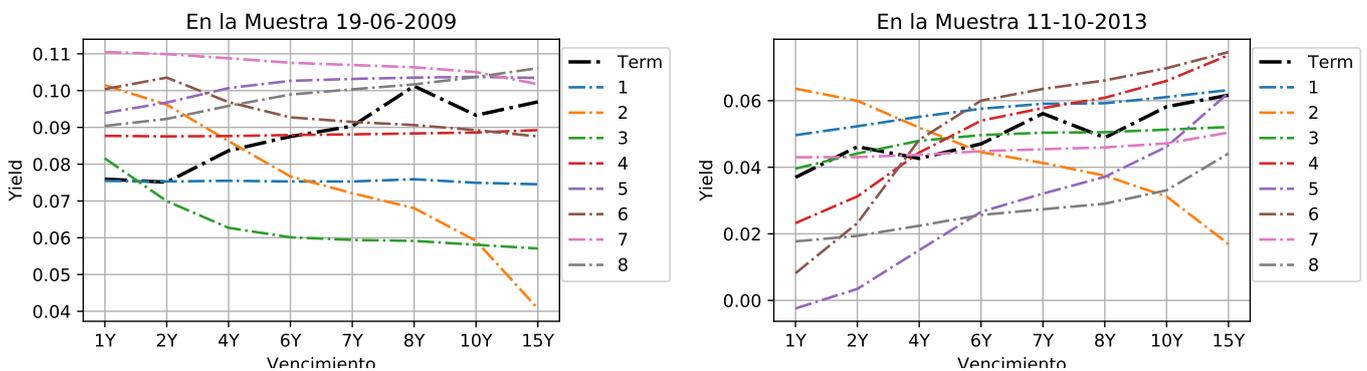


Figura 5.7: Ajuste Curva Yield Modelo Estándar

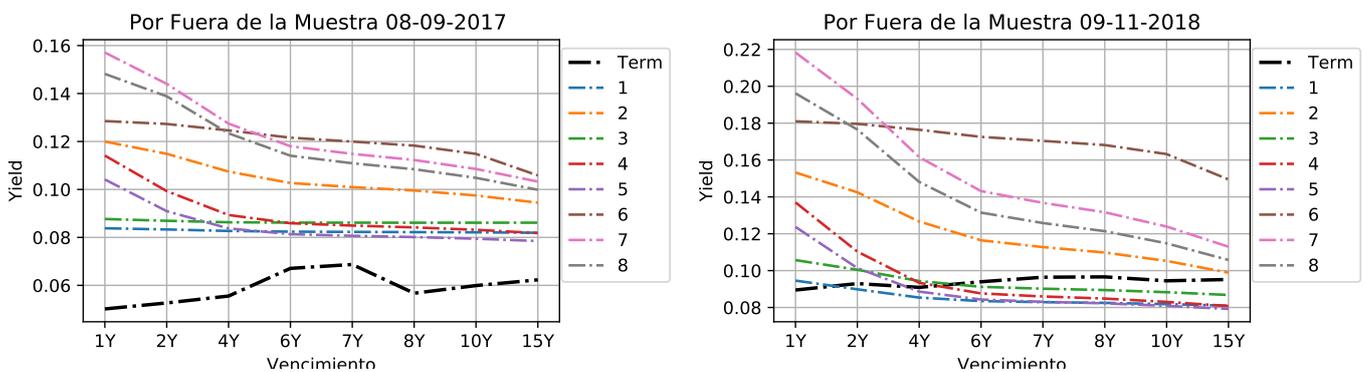
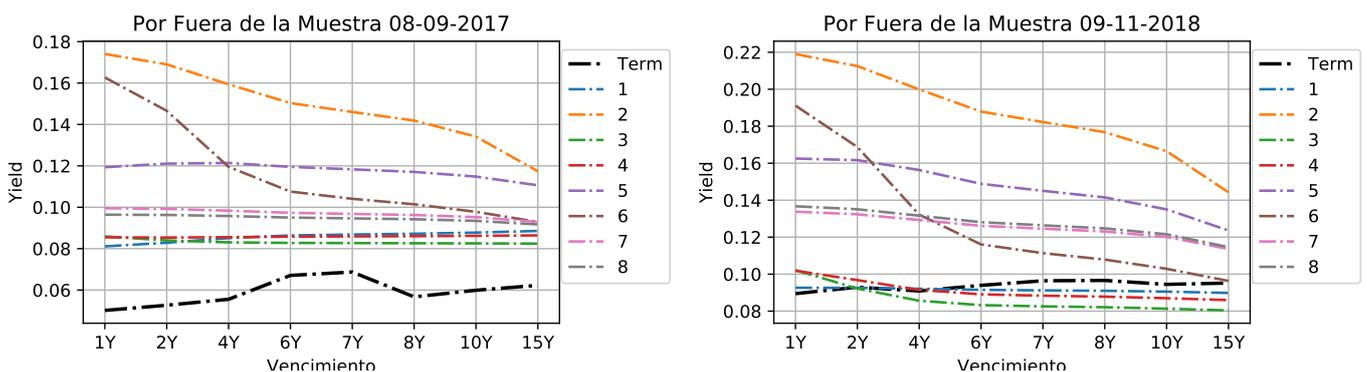


Figura 5.8: Ajuste Curva Yield Modelo SV



# Capítulo 6

## Conclusiones

Se ha desarrollado un análisis comparativo entre dos modelos basados en un modelo de estructura a término dinámicos con forma afín que intenta acomodar un número arbitrario de factores de duraciones heterogéneas creando una cascada de  $n$  difusiones. En este, cada variable de estado revierte en media hacia su antecesora, tratando de capturar la idea de que cualquier fluctuación económica ocurre en un contexto de ciclos de poca frecuencia. La base de este modelo impone supuestos en la distribución del riesgo, las velocidades de ajuste y la premia de riesgo, generando una estructura de dimensión invariante especificada por solo 5 parámetros, independiente de la dimensión de la cascada.

Los dos modelos comparados en este estudio se diferencian en el cambio del supuesto sobre la forma de la volatilidad o riesgo del modelo, donde es constante para el modelo estándar, generando la misma cantidad de parámetros y dependiente de la dimensión de la cascada para el nuevo modelo planteado, el Sigma Variante con 6 parámetros. En la estimación se encuentra que los valores de los parámetros resultantes para diferentes difusiones no convergen al aumentar los factores, algo deseable, si no que fluctúan a lo largo de la dimensión de la cascada para ambos modelos.

En primera instancia, el ajuste en la muestra del modelo estándar sigue un comportamiento deseado de disminución del error, pero solo hasta la parte media de la cascada, mientras que el modelo SV tiene un error mucho más fluctuante. En ambos casos, el mejor ajuste lo tiene un modelo con al menos 3 factores. Por su parte, en el error de pronóstico, el modelo estándar es mucho más fluctuante que el SV. Este último muestra más estabilidad de resultados tanto en la

muestra como por fuera de la muestra, aunque ambos modelos siguen teniendo un desempeño de pronóstico peor que el de una caminata aleatoria.

Así, hay indicios de que en este tipo de modelos afines es recomendable usar al menos 3 factores para su ajuste en los datos de estructura a plazos en Colombia, siendo cuidadosos en la elección de este número por la inestabilidad de los modelos y sus fluctuaciones. Por su parte, los modelos afines con dimensión invariante no parecen tener un ajuste en pronóstico adecuado para las dinámicas de las estructuras a término colombianas.

En investigaciones futuras se propone mejorar la estimación a través de la programación directa de un algoritmo de optimización con fronteras para evaluar si se puede llegar a una convergencia los parámetros ante aumentos de la dimensión de la cascada. También es posible buscar otras estructuras de dependencia para el riesgo y adicionar una forma al risk premia para ver si hay mejor ajuste en los datos de Colombia. Se puede modificar el modelo para utilizar datos de tasas IBR y Swap colombianos y así observar ciertos comportamientos de la estructura a plazos no lineales del mercado. La estructura lineal del modelo tiene la limitación de que permite que las tasas de interés sean negativas, por lo cual sería una mejora para el caso colombiano el adaptar la cascada a una especificación CIR o cuadrática para obtener solo tasas positivas.

# Bibliografía

- [1] AKKOUCHI, M. On the convolution of exponential distributions. *J. Chungcheong Math. Soc* 21, 4 (2008), 501–510.
- [2] BOLDER, D. J. Affine term-structure models: Theory and implementation. *Available at SSRN 1082826* (2001).
- [3] BYRD, R. H., LU, P., NOCEDAL, J., AND ZHU, C. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 5 (1995), 1190–1208.
- [4] CALVET, L. E., FISHER, A. J., AND WU, L. Staying on top of the curve: A cascade model of term structure dynamics. *Journal of Financial and Quantitative Analysis* 53, 2 (2018), 937–963.
- [5] CHATTERJEE, S. Application of the kalman filter for estimating continuous time term structure models: The case of uk and germany.
- [6] COX, J. C., INGERSOLL, JONATHAN E, J., AND ROSS, S. A. A Theory of the Term Structure of Interest Rates. *Econometrica* 53, 2 (March 1985), 385–407.
- [7] DAI, Q., AND SINGLETON, K. J. Specification analysis of affine term structure models. *The Journal of Finance* 55, 5 (2000), 1943–1978.
- [8] DUAN, J.-C., AND SIMONATO, J.-G. Estimating and testing exponential-affine term structure models by kalman filter. *Review of Quantitative Finance and Accounting* 13, 2 (1999), 111–135.
- [9] DUFFIE, D., AND KAN, R. A yield-factor model of interest rates. *Mathematical finance* 6, 4 (1996), 379–406.
- [10] KALMAN, R. E. A new approach to linear filtering and prediction problems. *Journal of basic Engineering* 82, 1 (1960), 35–45.
- [11] KARAPANAGIOTIDIS, P. Dynamic state-space models.
- [12] MERGNER, S. *Applications of State Space Models in Finance: An Empirical Analysis of the Time-Varying Relationship Between Macroeconomics, Fundamentals and Pan-European Industry Portfolios*. Universitätsverlag Göttingen, 2009.
- [13] NOWMAN, K. B. Modelling the uk and euro yield curves using the generalized vasicek model: Empirical results from panel data for one and two factor models. *International Review of Financial Analysis* 19, 5 (2010), 334–341.

- 
- [14] RODRIGUEZ REVILLA, C. A., ET AL. Estimating and forecasting the term structure of interest rates: Us and colombia analysis. Master's thesis, Universidad del Rosario, 2016.
- [15] SPEICHER, R., AND WOROUDI, R. Boolean convolution. In *OF FIELDS INST. COMMUN* (1993), Citeseer.
- [16] STOFFER, D. Bootstrapping the kalman filter. Tech. rep., PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS, 1984.
- [17] STOFFER, D. S., AND WALL, K. D. Bootstrapping state-space models: Gaussian maximum likelihood estimation and the kalman filter. *Journal of the american statistical association* 86, 416 (1991), 1024–1033.
- [18] VASICEK, O. An equilibrium characterization of the term structure. *Journal of financial economics* 5, 2 (1977), 177–188.
- [19] VELÁSQUEZ-GIRALDO, M., AND RESTREPO-TOBÓN, D. Affine term structure models: Forecasting the yield curve for colombia. *Lecturas de Economía*, 85 (2016), 53–90.
- [20] WALTHER, B., MAY, A., AND FISCHER, T. Fitting yield curve models using the kalman filter. In *PAMM: Proceedings in Applied Mathematics and Mechanics* (2003), vol. 3, Wiley Online Library, pp. 507–508.
- [21] ZEYTUN, S., AND GUPTA, A. A comparative study of the vasicek and the cir model of the short rate.