

FUNDAÇÃO GETÚLIO VARGAS  
ESCOLA DE ECONOMIA DE SÃO PAULO

BRUNO MÜLLER BERNZ

MODELO NELSON-SIEGEL DINÂMICO DA ESTRUTURA A  
TERMO DA TAXA DE JUROS COM FATORES EXÓGENOS  
MACROECONÔMICOS  
UMA APLICAÇÃO AO MERCADO BRASILEIRO

SÃO PAULO  
2014

**BRUNO MÜLLER BERNZ**

**MODELO NELSON-SIEGEL DINÂMICO DA ESTRUTURA A  
TERMO DA TAXA DE JUROS COM FATORES EXÓGENOS  
MACROECONÔMICOS  
UMA APLICAÇÃO AO MERCADO BRASILEIRO**

Dissertação apresentada ao Programa de  
Mestrado Profissional em Economia da  
Fundação Getúlio Vargas/EESP, como  
parte dos requisitos para a obtenção do  
título de Mestre em Economia.

Área de concentração: Finanças Quanti-  
tativas

Orientador:  
Prof. Dr. Juan Carlos Ruivola Terán

**SÃO PAULO  
2014**

Bernz, Bruno Müller

Modelo Nelson-Siegel Dinâmico da Estrutura a Termo da Taxa de Juros  
com Fatores Exógenos Macroeconômicos - Uma aplicação ao mercado brasileiro  
/ Bruno Müller Bernz - 2014.

67 f.

Orientador: Juan Carlos Ruivola Terán.

Dissertação (MPFE) - Escola de Economia de São Paulo.

1. Modelos econométricos. 2. Taxas de juros. 3. Métodos de espaço de  
estados. 4. Macroeconomia. 5. Kalman, Filtragem de. Terán, Juan Carlos  
Ruivola. II. Dissertação (MPFE) - Escola de Economia de São Paulo. III.  
Título.

CDU 336

BRUNO MÜLLER BERNZ

MODELO NELSON-SIEGEL DINÂMICO DA ESTRUTURA A  
TERMO DA TAXA DE JUROS COM FATORES EXÓGENOS  
MACROECONÔMICOS  
UMA APLICAÇÃO AO MERCADO BRASILEIRO

Dissertação apresentada ao Programa de  
Mestrado Profissional em Economia da  
Fundação Getúlio Vargas/EESP, como  
parte dos requisitos para a obtenção do  
título de Mestre em Economia.

Área de concentração: Finanças Quanti-  
tativas

**Data de aprovação:**

— / — / —

**Banca Examinadora:**

---

Prof. Dr. Juan Carlos Ruivola Terán  
(Orientador)  
FGV - EESP

---

Prof. Dr. Roberto Barbosa Cintra  
FGV - EESP

---

Prof. Dr. Marcelo Ferman  
Membro Externo

# AGRADECIMENTOS

Ao professor Juan Carlos Ruivola Terán, pela orientação no trabalho.

Ao professor Roberto Barbosa Cintra, pelos conselhos, por todo o conhecimento compartilhado e pelas inúmeras contribuições para a conclusão deste trabalho.

Aos meus pais, Pedro e Cleusa, e minhas irmãs, Iara e Livia, pela ajuda e incentivo durante toda a minha vida.

À Suellen e ao pequeno Guilherme, pela paciência, carinho e apoio incondicionais.

# RESUMO

A tradicional representação da estrutura a termo das taxas de juros em três fatores latentes (nível, inclinação e curvatura) teve sua formulação original desenvolvida por Charles R. Nelson e Andrew F. Siegel em 1987. Desde então, diversas aplicações vêm sendo desenvolvidas por acadêmicos e profissionais de mercado tendo como base esta classe de modelos, sobretudo com a intenção de antecipar movimentos nas curvas de juros. Ao mesmo tempo, estudos recentes como os de Diebold, Piazzesi e Rudebusch (2010), Diebold, Rudebusch e Aruoba (2006), Pooter, Ravazzallo e van Dijk (2010) e Li, Niu e Zeng (2012) sugerem que a incorporação de informação macroeconômica aos modelos da ETTJ pode proporcionar um maior poder preditivo. Neste trabalho, a versão dinâmica do modelo Nelson-Siegel, conforme proposta por Diebold e Li (2006), foi comparada a um modelo análogo, em que são incluídas variáveis exógenas macroeconômicas. Em paralelo, foram testados dois métodos diferentes para a estimação dos parâmetros: a tradicional abordagem em dois passos (Two-Step DNS), e a estimação com o Filtro de Kalman Estendido, que permite que os parâmetros sejam estimados recursivamente, a cada vez que uma nova informação é adicionada ao sistema. Em relação aos modelos testados, os resultados encontrados mostram-se pouco conclusivos, apontando uma melhora apenas marginal nas estimativas dentro e fora da amostra quando as variáveis exógenas são incluídas. Já a utilização do Filtro de Kalman Estendido mostrou resultados mais consistentes quando comparados ao método em dois passos para praticamente todos os horizontes de tempo estudados.

**Palavras-Chave:** Modelo Nelson-Siegel, Estrutura a Termo da Taxa de Juros, Macro-Finanças, Modelos de Espaço de Estados, Brasil.

# ABSTRACT

The traditional representation of the term structure of interest rates in three latent factors (level, slope and curvature) had its original formulation developed by Charles R. Nelson and Andrew F. Siegel in 1987. Since then, several applications have been developed by academics and practitioners based on this class of models, mainly with the intention of anticipating movements in yield curves. At the same time, recently published papers as Diebold, Piazzesi and Rudebusch (2010), Diebold, Rudebusch and Aruoba (2006), Pooter, Ravazallo and van Dijk (2010) and Li, Niu and Zeng (2012) suggest that incorporating macroeconomic information to interest rates models can provide higher predictive power. In this study, the dynamic version of the Nelson-Siegel model, as proposed by Diebold and Li (2006), was compared to a similar model in which exogenous macroeconomic variables are included, for Brazilian data. In parallel, two different methods of parameter estimation were tested: the traditional two-step approach, and another, with the use of an Extended Kalman Filter, which allows parameters to be estimated recursively, every time a new information is added to the system. Regarding the models tested, the results were shown to be inconclusive, indicating only a marginal improvement in the estimates both in-sample and out-of-sample when exogenous variables are included. Nonetheless, the use of the Extended Kalman Filter showed more consistent results compared to the two-step method for almost all horizons studied.

**Keywords:** Nelson-Siegel model, Term-Strucutre of Interest Rates, Macro-Finance, State-Space Models, Brazil.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>13</b>
1.1	Objetivo . . . . .	16
1.2	Estrutura do trabalho . . . . .	16
<b>2</b>	<b>Revisão da Literatura</b>	<b>17</b>
2.1	Modelo Nelson-Siegel da ETTJ . . . . .	17
2.2	Modelagem da ETTJ com fatores macroeconômicos . . . . .	18
<b>3</b>	<b>Modelo</b>	<b>20</b>
3.1	Conceitos Preliminares . . . . .	20
3.1.1	Análise de Componentes Principais (PCA) . . . . .	20
3.1.2	A Abordagem de Litterman e Sheinkman . . . . .	21
3.2	O Modelo Nelson-Siegel . . . . .	25
3.2.1	A Abordagem de Diebold e Li . . . . .	25
3.2.2	Nelson-Siegel Dinâmico . . . . .	28
3.2.3	Incorporando Variáveis Macroeconômicas ao Modelo . . . . .	30
3.3	Modelo de Espaço de Estados . . . . .	31
3.3.1	A Representação em Espaços de Estados . . . . .	31
3.3.2	Filtro de Kalman . . . . .	32
3.3.3	Filtro de Kalman Estendido . . . . .	33
3.3.4	Estimação de Parâmetros . . . . .	35
<b>4</b>	<b>Metodologia</b>	<b>38</b>
4.1	Modelos Comparados . . . . .	38



4.2	Calibragem do parâmetro $\lambda$	41
4.3	Estimação dos Parâmetros	41
4.3.1	2-Step DNS	41
4.3.2	EKF DNS com $\theta_t = \theta_{t-1} + \eta_t$	42
4.4	Análise dentro da amostra	44
4.5	Análise fora da amostra	45
4.5.1	<i>Benchmarks</i>	46
4.5.2	Medidas de Desempenho	47
<b>5</b>	<b>Resultados</b>	<b>49</b>
5.1	Dados Utilizados	49
5.1.1	Taxas de Juros	49
5.1.2	Séries das Variáveis Macro	50
5.2	Resultados Dentro da Amostra	52
5.3	Resultados Fora da Amostra	54
<b>6</b>	<b>Conclusões</b>	<b>63</b>
6.1	Conclusões	63
	<b>Referências Bibliográficas</b>	<b>65</b>

# Lista de Figuras

1.1	Séries temporais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses dos juros no Brasil no período de jan/2004 a dez/2013 . . .	13
1.2	Série tridimensional da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013 . . . . .	14
3.1	Evolução Temporal dos fatores Nível, Inclinação e Curvatura para a Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013 . . . . .	22
3.2	Poder explicativo da variância de cada um dos componentes principais aplicados à Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013 . . . . .	22
3.3	Autovetores dos fatores Nível, Inclinação e Curvatura da Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013 . . . . .	23
3.4	Nível empírico ( $72m$ ) e primeiro componente principal (nível) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013 . . . . .	23
3.5	Inclinação empírica ( $72m - 1m$ ) e segundo componente principal (inclinação) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013	24
3.6	Curvatura empírica ( $3m + 72m - 2 \times 18m$ ) e terceiro componente principal (curvatura) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013 . . . . .	24
3.7	Séries temporais da taxa de juros curto prazo brasileira (DI CETIP) e série de $\beta_1 + \beta_2$ . . . . .	26
3.8	Séries temporais da taxa de 72 meses e de $\beta_1$ . . . . .	26
3.9	Possíveis formas assumidas pela ETTJ Nelson-Siegel em datas específicas .	27

3.10	Peso dos fatores no modelo Nelson-Siegel . . . . .	27
3.11	Séries temporais do <i>spread</i> ( $72m - 1m$ ) e de $\beta_2$ . . . . .	28
3.12	Séries temporais do <i>butterfly spread</i> ( $2 \times 18m - 3m - 72m$ ) e de $\beta_3$ . . . . .	28
3.13	Dinâmica dos fatores DNS para Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013 . . . . .	29
4.1	Soma dos erros quadráticos médios dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 para diferentes valores do parâmetro $\lambda$ . . . . .	41
4.2	Estimação dos parâmetros do modelo DNS para os dados mensais dos vér- tices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 usando Filtro de Kalman Estendido com $\theta_t = \theta_{t-1} + \varepsilon_t$ . . . . .	43
4.3	Estimação dos coeficientes autorregressivos do modelo DNS com variáveis exógenas para os dados mensais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 usando Filtro de Kalman Estendido com $\theta_t = \theta_{t-1} + \varepsilon_t$ . . . . .	44
5.1	Séries temporais mensais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses dos juros no Brasil no período de jan/2004 a dez/2013	49
5.2	Séries temporais mensais dos primeiros componentes de Atividade e Infla- ção no período de jan/2004 a dez/2013 . . . . .	52
5.3	Série dos erros de estimação dentro da amostra para os vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013, com parâmetros estimados pelo método 2-STEP DNS .	53
5.4	Série dos erros de previsão do Filtro de Kalman dentro da amostra para os vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013, com parâmetros estimados usando Filtro de Kalman Estendido com $\theta_t = \theta_{t-1} + \eta$ . . . . .	54

5.5	Erro quadrático médio acumulado para previsões 3 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	55
5.6	Erro quadrático relativo acumulado para previsões 3 meses a frente para os vértices de 6, 12, 18 e 24 meses . . . . .	56
5.7	Erro quadrático médio acumulado para previsões 6 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	57
5.8	Erro quadrático relativo acumulado para previsões 6 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	58
5.9	Erro quadrático acumulado para previsões 9 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	59
5.10	Erro quadrático relativo acumulado para previsões 9 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	60
5.11	Erro quadrático acumulado para previsões 12 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	61
5.12	Erro quadrático relativo acumulado para previsões 12 meses à frente para os vértices de 6, 12, 18 e 24 meses . . . . .	62

# Lista de Tabelas

3.1	Dimensões das matrizes do modelo de espaço de estados. . . . .	32
5.1	Estatísticas descritivas das séries mensais das taxas de juros para os vértices estudados . . . . .	50
5.2	Estatísticas descritivas das séries mensais macroeconômicas utilizadas . . .	51
5.3	Análise dos Componentes Principais para as variáveis macroeconômicas de atividade. . . . .	51
5.4	Análise dos Componentes Principais para as variáveis macroeconômicas de inflação. . . . .	52
5.5	Raiz do Erro Quadrático Médio para as previsões dentro da amostra para 1 mês a frente para os modelos analisados. . . . .	53
5.6	Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 3 meses à frente fora da amostra para vértices selecionados . . . . .	55
5.7	Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 6 meses à frente fora da amostra para vértices selecionados . . . . .	57
5.8	Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 9 meses à frente fora da amostra para vértices selecionados . . . . .	58
5.9	Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 12 meses à frente fora da amostra para vértices selecionados . . . . .	61

# Capítulo 1

## Introdução

A taxa de juros é, sem dúvidas, uma das mais importantes variáveis econômicas de qualquer país. Seja no mercado financeiro ou na economia real, os agentes buscam, direta ou indiretamente, antecipar os movimentos das curvas a fim de obterem resultados favoráveis.

Na prática, as taxas de juros são derivadas dos preços de títulos negociados no mercado financeiro. Isso implica em um grande número de taxas sendo negociadas em um mesmo momento (*cross-section*), cada uma correspondendo a um determinado prazo de vencimento e uma qualidade de crédito. Uma curva de juros, portanto, pode ser entendida como um vetor em que cada valor diz respeito à taxa para um prazo específico. Isso implica que, ao modelar a dinâmica da estrutura a termo das taxas de juros, além de uma série temporal existe uma dimensão transversal.

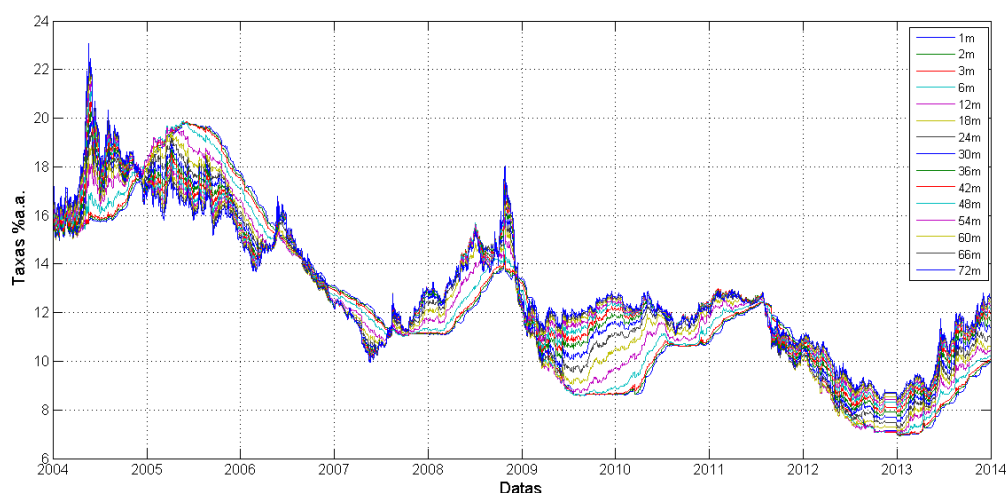


Figura 1.1: Séries temporais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses dos juros no Brasil no período de jan/2004 a dez/2013

Seria possível supor, como coloca Piazzesi (2010), que a dinâmica da estrutura a

termo poderia ser descrita por um modelo vetorial autorregressivo (VAR). No entanto, o comportamento das taxas nos diferentes vértices é estreitamente correlacionado, dando origem a uma série de restrições teóricas. Matematicamente, essas restrições aparecem na forma de equações cruzadas que "amarram" os vértices e elevam o grau de complexidade dos modelos.

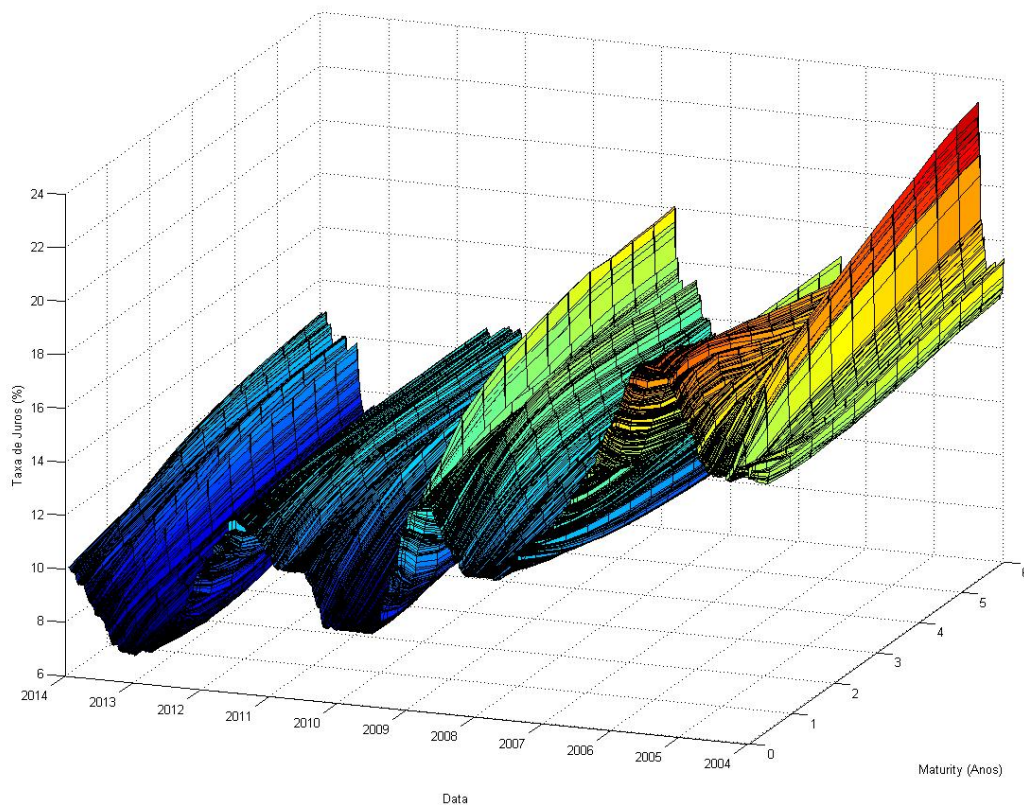


Figura 1.2: Série tridimensional da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013

São muitas as aplicações das teorias que descrevem o comportamento da estrutura a termo da taxa de juros (ETTJ). Conforme Diebold et al. (2006) mencionam, macroeconomistas, economistas financeiros e profissionais de mercado têm se debruçado sobre este campo resultando em abordagens tão diversas quanto os contextos aos quais são aplicadas.

Piazzesi (2010) cita quatro razões para a importância de compreender os movimentos da curva de juros. A primeira é como ferramenta para previsão, já que as taxas longas podem ser interpretadas como o valor esperado das médias das taxas curtas futuras após um ajuste pelo risco. Isso significa que a curva de juros atual contém informação sobre o futuro da economia que podem ser utilizadas para tomada de decisão em investimentos e risco. O segundo aspecto diz respeito à política monetária. Como, de forma geral, os bancos centrais conseguem controlar a taxa de curto prazo, e o que determina a "demanda

agregada" são os juros de longo prazo, um modelo pode ajudar a compreender como ocorrem as dinâmicas de transmissão dos movimentos da extremidade curta para o resto da curva. A terceira razão está associada às políticas de endividamento, já que ao emitir um título, o governo pode decidir sobre o seu vencimento para influenciar determinados movimentos das curvas. O último motivo está relacionado ao apreçamento de ativos e derivativos.

Em razão dos possíveis direcionamentos, constata-se que a progressão literária no tema não ocorreu de forma linear. Duas vertentes populares são os modelos de não-arbitragem e os de equilíbrio. A primeira tem como objetivo principal o perfeito ajuste (*fitting*) da estrutura a termo das taxas, importante para a precificação de derivativos. São exemplos os modelos de Hull e White (1990) e de Heath, Jarow and Morton (1992). A segunda abordagem tem como foco a modelagem da dinâmica da taxa de juros instantânea, em geral, utilizando-se de modelos afins (*affine models*) através dos quais as taxas para diferentes vencimentos podem ser estimadas utilizando hipóteses sobre o prêmio de risco. Destacam-se os modelos de Vasicek (1977), Cox, Ingersoll e Ross (1985) e de Duffie e Kan (1996). (Diebold e Li, 2006)

Apesar do apelo teórico, os modelos citados acima têm desempenhos fracos para previsão fora da amostra (Diebold e Li, 2006). Diante deste fato, acadêmicos e profissionais do mercado têm direcionado as atenções para os modelos da chamada classe "Nelson-Siegel" de estrutura a termo da taxa de juros. Essa classe é fundamentada no trabalho de Charles R. Nelson e Andrew F. Siegel (1987), que utiliza componentes exponenciais (fatores) para derivar todos os pontos da curva de juros com uma estrutura tridimensional paramétrica. Os três parâmetros são interpretados como nível, inclinação e curvatura e são análogos aos fatores ortogonais aplicados à curva de juros de Litterman e Scheinkman (1991), porém com imposição de estrutura.

Independentemente da vertente teórica, durante um período relativamente longo foi possível perceber um descompasso entre as literaturas de finanças e macroeconômica em relação à modelagem da ETTJ. Conforme Diebold e Rudebusch (2013), esta é uma lacuna que vem sendo preenchida nos últimos anos, com diversos pesquisadores incorporando variáveis macroeconômicas a modelos tradicionais, como o Nelson-Siegel dinâmico e os modelos afins de não-arbitragem<sup>1</sup>.

Neste contexto, trabalhos como os de Ang e Piazzesi (2004), Diebold, Piazzesi e Rudebusch (2010), Diebold, Rudebusch e Aruoba (2006), Pooter, Ravazallo e van Dijk (2010) e Li, Niu e Zeng (2012) têm buscado identificar e explorar as relações e dinâmicas conjuntas entre as variáveis macroeconômicas e a curva de juros.

A justificativa comumente encontrada para adotar essa linha de pesquisa é a de que os dados econômicos contém informação sobre o futuro das curvas de juros que ainda não estão totalmente refletidos nos fatores latentes. É possível, portanto, supor que introdu-

---

<sup>1</sup> Affine Term-Structure Models



zindo variáveis, por exemplo, de inflação e atividade real, que certamente direcionam as decisões de política monetária, seja possível melhorar as estimativas dos modelo.

## 1.1 Objetivo

O objetivo deste trabalho é comparar, testando diferentes métodos para estimação dos parâmetros, o poder preditivo do modelo Nelson-Siegel Dinâmico (DNS) com uma extensão, em que são adicionadas variáveis macroeconômicas. A ideia central consiste em aplicar os modelos propostos para os dados brasileiros e verificar seus desempenhos dentro e fora da amostra.

## 1.2 Estrutura do trabalho

Este trabalho está organizado da seguinte forma:

Após a parte introdutória, um capítulo será dedicado à revisão da literatura relacionada ao modelo Nelson-Siegel e aos estudos que analisaram a dinâmica conjunta da estrutura a termo das taxas de juros e das variáveis macroeconômicas. O objetivo é situar o trabalho desenvolvido no contexto literário e justificar determinadas escolhas metodológicas e de modelagem.

No capítulo seguinte, serão apresentados os referenciais teóricos. Nesta seção serão introduzidas as ideias e modelos que fundamentam a abordagem adotada. Serão discutidos brevemente alguns conceitos relacionados à curva de juros, a teoria de componentes principais e o trabalho de Litterman e Scheinkman (1991). Em um segundo momento, será descrita a abordagem estática do modelo Nelson-Siegel para que então seja introduzida a dinâmica, e em seguida, incorporadas as variáveis exógenas macroeconômicas. Uma seção também será dedicada à representação em espaço de estados e a algumas abordagens para a estimação dos estados e parâmetros, onde serão explorados os conceitos de filtro de Kalman (KF) e Filtro de Kalman Estendido (EKF).

Após, serão descritos os dados e os resultados apresentados e analisados à luz da literatura e dos objetivos traçados previamente.

O capítulo final será dedicado à discussão das conclusões, suas possíveis implicações e introduzidas ideias para trabalhos futuros dentro da mesma linha de pesquisa.

# Capítulo 2

## Revisão da Literatura

O interesse central desta dissertação consiste em comparar o poder preditivo da versão dinâmica do modelo Nelson-Siegel para a estrutura a termo das taxas de juros a uma extensão deste modelo, em que são incluídas variáveis macroeconômicas. Para tanto, na primeira seção deste capítulo, será apresentada a literatura clássica relacionada ao modelo DNS (*Dynamic Nelson-Siegel*), focando na sua evolução ao longo dos últimos anos a partir da sua formulação original. Para completar o enquadramento literário, serão mencionados alguns estudos que relacionam a dinâmica da ETTJ com variáveis macroeconômicas.

### 2.1 Modelo Nelson-Siegel da ETTJ

A tradicional representação da ETTJ em três fatores latentes teve sua formulação original desenvolvida por Charles R. Nelson e Andrew F. Siegel (1987). No seminal trabalho intitulado *Parsimonious Modeling of Yield Curves*, os autores introduziram uma classe paramétrica de modelos para a curva de juros capaz de caracterizar as diversas formas assumidas ao longo do tempo, aplicando-as aos títulos do tesouro americano.<sup>1</sup>

A fim de melhorar o ajuste, Svensson(1995) propôs a adição de um quarto fator, similar ao terceiro, porém com outro parâmetro de decaimento, o que permitiria mais de um ponto de mínimo ou máximo local à curva.

Francis X. Diebold e Canlin Li (2006) através de uma pequena modificação na forma funcional, interpretaram a série temporal dos parâmetros do modelo como fatores correspondentes ao nível, inclinação e curvatura, associando a estrutura Nelson-Siegel à análise de componentes principais para os vértices da estrutura a termo, conforme trabalho de Litterman e Sheinkman (1991). Aplicando modelos autorregressivos para identificar a dinâmica dos fatores, consolidaram o modelo que ficou conhecido como Diebold e Li ou Nelson-Siegel Dinâmico.

Em Diebold, Li e Ji (2006), foram exploradas as relações da estrutura Nelson-Siegel

---

<sup>1</sup>U.S. Treasury bills

com outras classes de modelos. Desta forma, foram realizados exercícios para testar a capacidade preditiva e desenvolvidas aplicações em gestão de risco, comparando as estimativas realizadas utilizando os três fatores com a abordagem tradicional da *Duration de Macaulay*. Em Diebold, Li e Yue (2008), os autores ampliaram o escopo da abordagem ao aplicá-la simultaneamente a vários países, permitindo fatores de nível, inclinação e curvatura globais e fatores idiossincráticos em uma estrutura hierárquica.

Complementarmente, nos últimos anos, muitos autores têm incorporado modificações na estrutura clássica com o objetivo de melhorar a performance do poder preditivo. É possível citar o trabalho de Koopman et al. (2010), que relaxaram algumas restrições a fim de permitir que os parâmetros de decaimento da curvatura e volatilidade assumidos como constantes, fossem variantes no tempo.

No Brasil, é importante citar o trabalho de Caldeira, Laurini e Portugal (2010), que assim como Koopman et al.(2010) permitiram generalizações aos pesos dos fatores e à dinâmica da volatilidade, e Caldeira, Moura e Santos (2012), que aplicaram a classe para a otimização de portfólios de títulos obtendo bons resultados ajustados ao risco.

O último grande passo dado dentro da evolução literária desta classe de modelos refere-se à imposição de não-arbitragem. Destaca-se o trabalho de Christensen, Diebold e Rudebusch(2011), que criaram uma classe generalizada de modelos de não-arbitragem com a estrutura Nelson-Siegel, obtendo bons resultados na previsão de taxas.

## 2.2 Modelagem da ETTJ com fatores macroeconômicos

Assim como na modelagem da estrutura a termo, a literatura que relaciona a curva de juros com variáveis macroeconômicas tem se desenvolvido por caminhos distintos.

Os primeiros autores a estudar essas relações tinham como objetivo compreender o efeito unidirecional das curvas de juros sobre a economia. São exemplos os artigos de Estrella e Hardouvelis (1991) e Mishkin e Estrella (1998), que aplicaram modelos para prever a atividade real com base na estrutura a termo.

Em direção contrária estão trabalhos como os de Ang e Piazzasi (2003) que incorporaram variáveis macroeconômicas em modelos multifatoriais da estrutura a termo livre de arbitragem. Neste artigo, os autores utilizam fatores representando inflação e atividade real como variáveis explicativas dos movimentos nas curvas de juros. Eles encontraram que choques na atividade tem impacto significativo sobre a curvatura (parte intermediária da curva de juros), enquanto choques na inflação geram efeitos no nível de toda a curva.

Uma terceira vertente tem estudado os efeitos bidirecionais, como no trabalho de Diebold, Rudebusch e Aruoba (2006). No artigo, os autores combinam a estrutura Nelson-Siegel com uma modelagem autorregressiva para as variáveis econômicas. Os resultados indicaram que ambos os efeitos são importantes, porém há um efeito preponderante das variáveis macro sobre as curvas de juros.

A última linha de pesquisa tem incorporado uma estrutura macroeconômica mais complexa aos modelos. É importante citar o artigo de Rudebusch e Wu (2008), que combina uma modelagem afim livre de arbitragem com especificações macroeconômicas para a política monetária, inflação e produto.

Conforme Diebold e Rudebusch (2013), há uma grande quantidade de estudos que relacionam os fatores *nível*, *inclinação* e *curvatura* com os fundamentos macroeconômicos. Sob o ponto de vista da teoria econômica, há uma grande ênfase na correlação entre o nível da curva de juros e o nível da inflação esperada. E inflação esperada é frequentemente associada à inflação corrente. Koopman et al.(2010) chegaram a conclusões similares, sugerindo que a expectativa de inflação de longo prazo rege o fator nível. Além disso, Diebold, Rudebusch e Aruoba (2006) usam o mesmo racional para incluir a inflação como variável no modelo citado anteriormente.

Também segundo Diebold e Rudebusch (2013), de forma geral, uma interpretação da teoria econômica permite relacionar a atividade real com o fator inclinação, porém, as relações de causalidade variam. Por um lado, quando os bancos centrais utilizam uma política contracionista, a taxa de curto prazo sobe e segue um período de recessão. Desta forma, segundo, por exemplo, Estrella e Mishkin (1998), curvas de juros invertidas precedem períodos de baixa atividade. Por outro lado, uma boa perspectiva sobre o crescimento da atividade de curto prazo também tende a elevar as taxas de curto prazo, conforme, por exemplo, Kurmann e Otrok (2012).

Finalmente, a literatura relata o fator curvatura como o de mais difícil interpretação. Alguns autores o associam com a volatilidade do fator nível, porém, os resultados não foram replicados em períodos mais recentes. No entanto, por ser o menos importante dos três fatores, a sua interpretação econômica tende a ser menos relevante.

Neste trabalho, as variáveis macro serão incorporadas ao modelo para a ETTJ segundo a interpretação de Diebold e Rudebusch (2013), baseando-se nas evidências de que o nível da curva é afetado pela inflação e a inclinação responde aos choques na atividade.

# Capítulo 3

## Modelo

### 3.1 Conceitos Preliminares

#### 3.1.1 Análise de Componentes Principais (PCA)

Análise de componentes principais refere-se a uma técnica matemática de redução de dimensionalidade para um conjunto de observações com grau elevado de correlação. A ideia principal consiste em aplicar uma transformação ortogonal sobre as variáveis originais a fim de obter um novo conjunto de variáveis não correlacionadas, os componentes principais.

Rachev et al.(2007) define intuitivamente o conceito aplicando-o a um portfólio de ativos. Considere um conjunto de  $n$  séries estacionárias representadas por retornos de ativos  $X_i$ . Considere agora uma combinação linear dessa série, ou seja, um portfólio de ativos. Cada portfólio  $P$  corresponde a um vetor  $n$ -dimensional de pesos  $\omega_P$  e é caracterizado por uma variância  $\sigma_P^2$ . Considere então um portfólio normalizado (cuja a soma dos quadrados dos pesos seja igual a um) cuja variância seja a maior possível.

Se for assumido que os retornos tenham média nula, com distribuição normal conjunta e matriz de covariâncias  $\sigma$ , pode-se mostrar que os retornos de cada um dos portfólios será normalmente distribuído com variância:

$$\sigma_P^\top = \omega_P^\top \sigma \omega_P, \quad (3.1)$$

e o portfólio de máxima variância será dado pela seguinte função de maximização:

$$\arg \max_{\|\omega\|=1} \{\omega_P^\top \sigma \omega_P\} \quad (3.2)$$

em que a solução é o autovetor  $\omega_1$  associado ao maior autovalor  $\lambda_1$  da matriz de covariâncias  $\sigma$ .

Agora, considere o conjunto de todos os portfólios normalizados ortogonais a  $\omega_1$ :

$$\omega_1^\top \omega_P = \omega_P^\top \omega_P = 0 \quad (3.3)$$

Repetindo-se o processo anterior  $n$  vezes, é possível chegar a  $n$  portfólios  $P_i$  de máxima variância. Os pesos desses portfólios são os autovetores ortonormais da matriz de covariâncias  $\sigma$ . A cada portfólio, então, existe uma serie temporal correspondente que é dada por uma combinação linear das séries originais dos retornos dos ativos  $X_i$ .

Importante notar que, por construção, todos os portfólios de máxima variância são mutuamente não-correlacionados. Portanto, para recuperar o retorno das séries originais  $X_j$  basta fazer:

$$X_j = \sum_{i=1}^n \alpha_{j,i} P_i \quad (3.4)$$

Ocorre que, quando as séries originais possuem alto grau de correlação, apenas um pequeno número  $p$  de componentes principais possuem variância estatisticamente diferente de zero. Portanto, é possível reduzir a dimensionalidade do problema aproximando o retorno de qualquer uma das series originais como uma combinação linear dos componentes de maior variância  $F$ , mais um termo de ruído  $\varepsilon$ . Segue que:

$$X_j = \sum_{i=1}^p \alpha_{j,i} F_i + \sum_{i=p+1}^n \alpha_{j,i} P_i = \sum_{i=1}^n \alpha_{j,i} F_i + \varepsilon_j \quad (3.5)$$

Portanto, para implementar a análise de componentes principais, deve-se calcular os autovalores e autovetores da matriz de covariância e escolher os autovalores significativamente diferentes de zero, ou os maiores. Os autovetores correspondentes equivalem aos pesos dos portfólios que formam os fatores.

### 3.1.2 A Abordagem de Litterman e Sheinkman

Litterman e Sheinkman (1991) aplicaram o conceito de componentes principais aos *treasuries* norte-americanos para estudar os fatores comuns que afetam os retornos dos títulos de renda fixa de diversas maturidades. Os autores encontraram que os três primeiros componentes são responsáveis por aproximadamente 97% da variação nos retornos naquele mercado. Além disso, analisando o comportamento e as características de cada componente, os interpretaram como *nível*, *inclinação* e *curvatura* da taxa de juros.

A figura 3.1 mostra as séries temporais dos três primeiros componentes para ETTJ brasileira quando aplicados diretamente sobre as taxas de juros. Visualmente, e conforme esperado, é possível perceber que o primeiro componente (nível) responde pela maior parte da variabilidade das taxas. O segundo componente (inclinação), tem uma variância consideravelmente menor do que o primeiro, e o terceiro componente quase não varia.

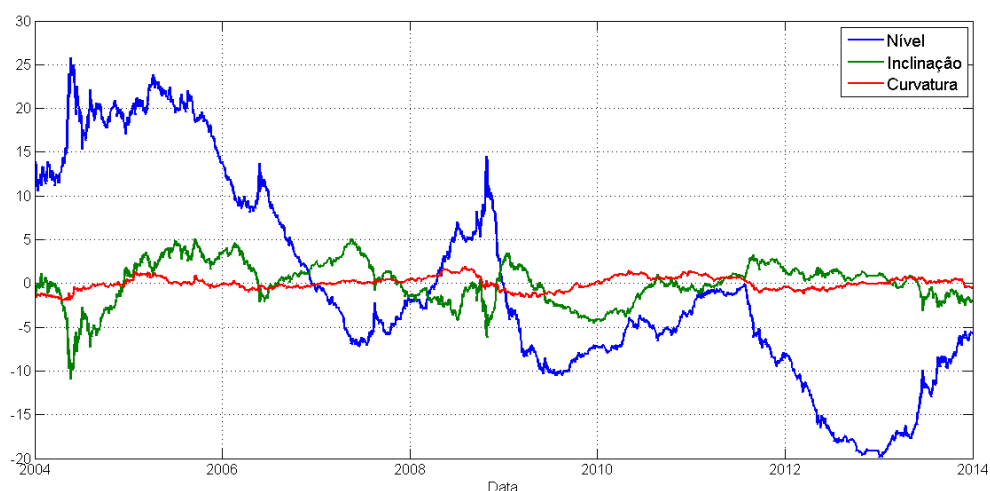


Figura 3.1: Evolução Temporal dos fatores Nível, Inclinação e Curvatura para a Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013

O poder explicativo da variância (autovalores normalizados) de cada um dos componentes principais pode ser melhor visualizado na figura 3.2. A exemplo do trabalho original de Litterman e Sheinkman (1991), os três primeiros componentes explicam quase 100% da variância das curvas de juros brasileiras.

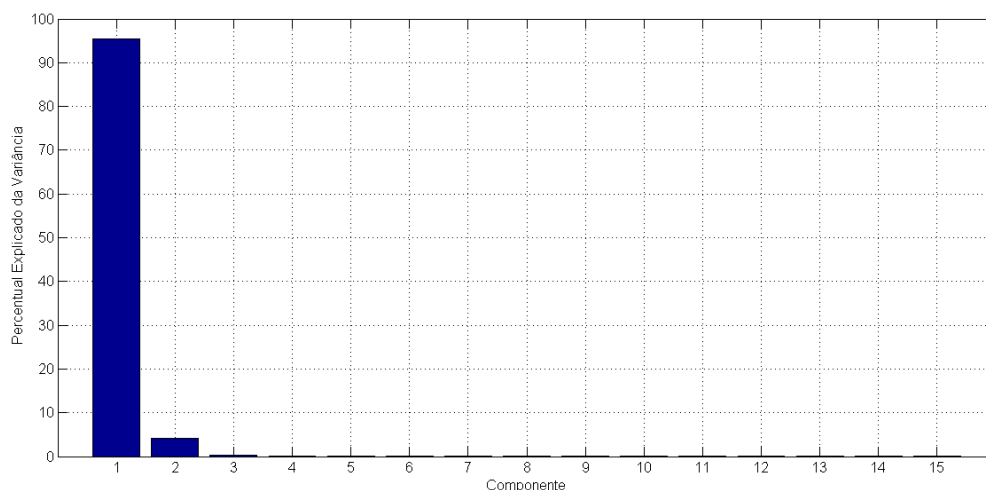


Figura 3.2: Poder explicativo da variância de cada um dos componentes principais aplicados à Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013

Compreender a interpretação dos fatores como nível, inclinação e curvatura também é um exercício simples.

Observando a figura 3.3, em que estão plotados os gráficos dos autovetores dos três primeiros componentes, verifica-se que os coeficientes do primeiro são praticamente constantes ao longo do tempo. Assim, este fator afeta toda a curva e é o único a afetar as

maturidades mais longas. Pode-se, então, concluir que trata-se do nível da ETTJ. Para confirmar, na figura 3.4 estão as séries do primeiro componente junto com a série das taxas para o vértice de  $72meses$ , o nível empírico, ambas normalizadas.

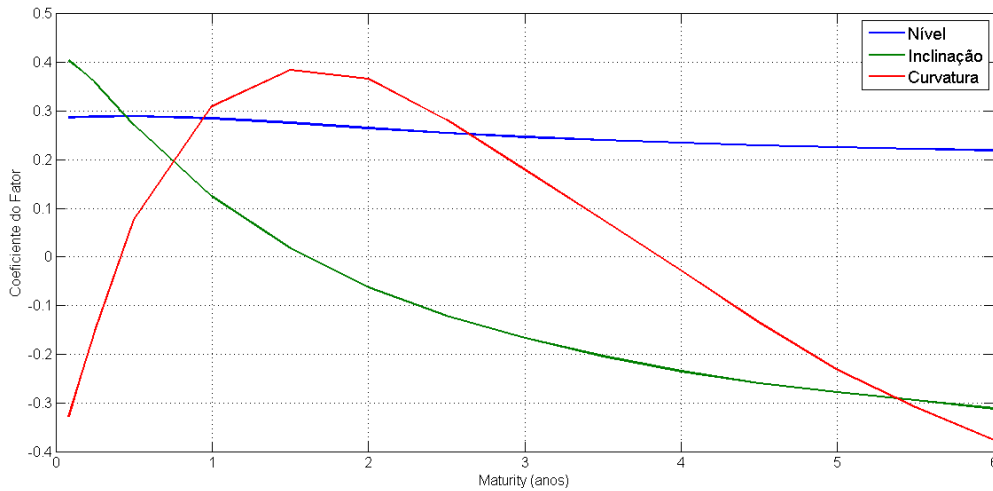


Figura 3.3: Autovetores dos fatores Nível, Inclinação e Curvatura da Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013

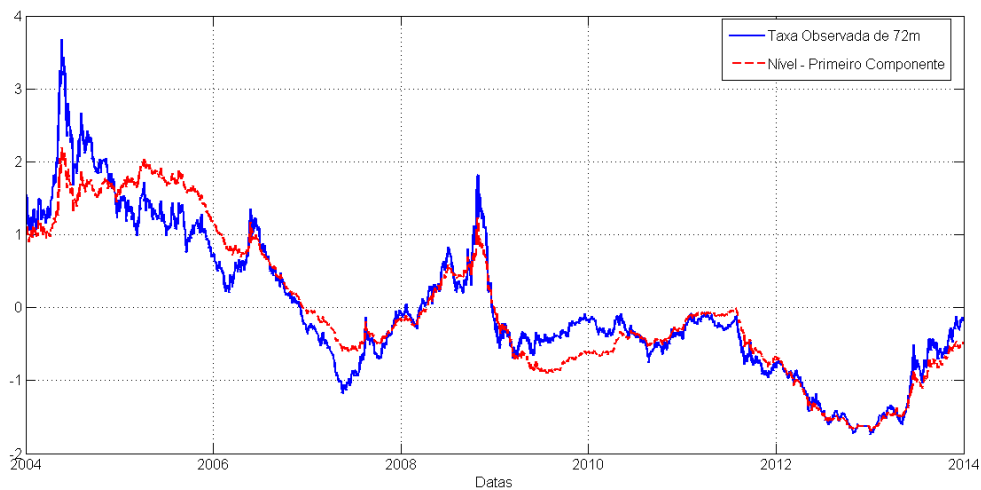


Figura 3.4: Nível empírico ( $72m$ ) e primeiro componente principal (nível) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013

Da mesma forma, pode-se construir a série da inclinação empírica através de um *spread*  $72m - 1m$  e compará-la com a o segundo componente principal, conforme figura 3.5. Este fator afeta a parte curta da curva de juros, juntamente com o nível, conforme é possível inferir a partir da análise do seu autovetor na figura 3.3 que forma o máximo no primeiro vértice e decai exponencialmente.



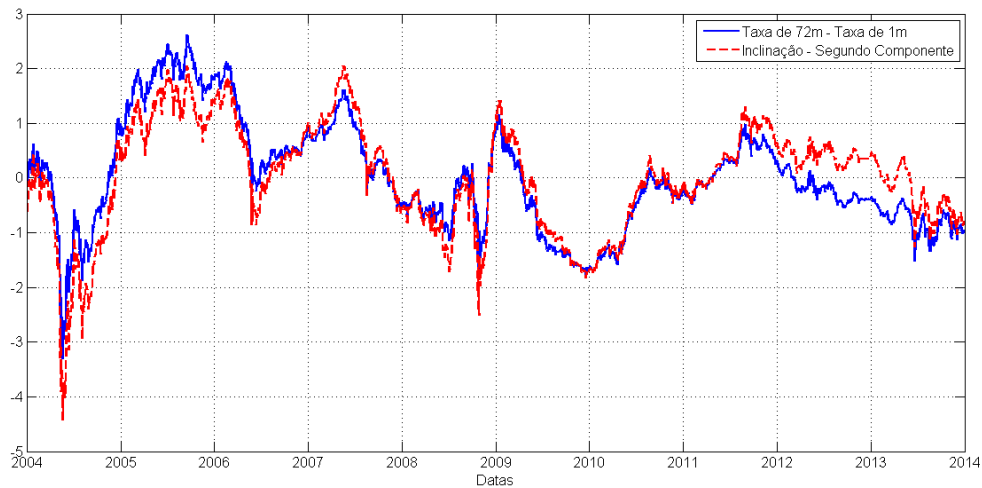


Figura 3.5: Inclinação empírica ( $72m - 1m$ ) e segundo componente principal (inclinação) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013

Finalmente, verifica-se na figura 3.6 que o terceiro componente corresponde à curvatura da curva de juros ao comparar a respectiva série temporal com a da curvatura empírica, dada pelo *butterfly spread* nos vértices  $3m + 72m - 2 \times 18m$ . O fator curvatura impacta os vértices de médio prazo, já que o seu autovetor começa com valores baixos para os vértices mais curtos, em seguida forma um ponto de máximo, e cai novamente na parte longa da curva.

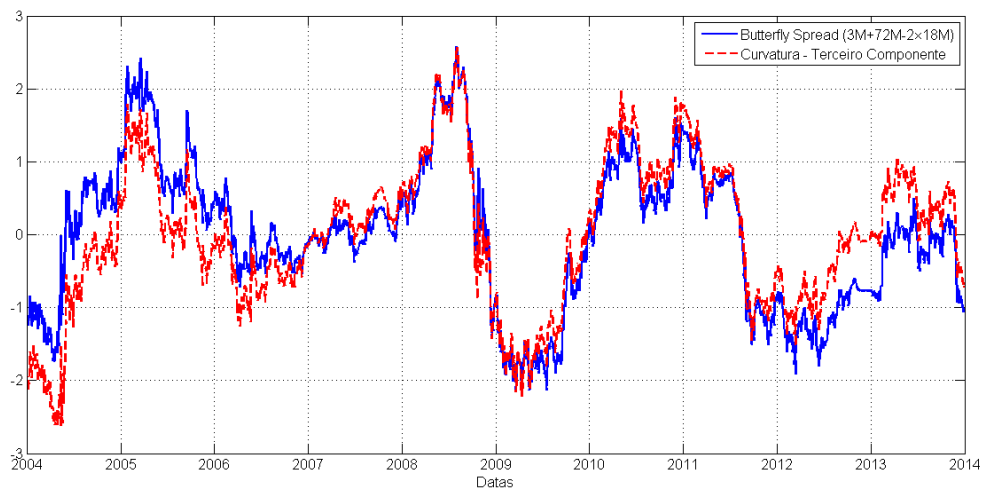


Figura 3.6: Curvatura empírica ( $3m + 72m - 2 \times 18m$ ) e terceiro componente principal (curvatura) da estrutura a termo da taxa de juros no Brasil de jan/2004 a dez/2013

## 3.2 O Modelo Nelson-Siegel

### 3.2.1 A Abordagem de Diebold e Li

O trabalho seminal de Charles R. Nelson e Andrew F. Siegel publicado em 1987 introduziu uma classe de modelos paramétricos para a estrutura a termo da taxa de juros ancorado no princípio da parcimônia.

A ideia era construir um modelo simples, porém flexível o suficiente para que se ajustasse às diferentes formas da curva de juros americana.

Desta forma, a formulação geral da classe Nelson-Siegel consiste em aplicar uma forma funcional para ajustar uma curva suavizada a um conjunto de taxas observáveis.

A interpretação de Diebold e Li (2006) para a ETTJ Nelson-Siegel é definida por uma equação baseada em três fatores exponenciais:

$$y(\tau) = \beta_1 + \beta_2\left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_3\left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) \quad (3.6)$$

em que  $y_t(\tau)$  é a taxa de juros para o correspondente prazo  $\tau$ , e com  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  e  $\lambda$  parâmetros do modelo.

Essa estrutura que parece arbitrária em primeiro momento, no entanto, possui algumas características com grande apelo empírico e teórico, conforme discutido em Diebold e Rudebusch (2013).

Em primeiro lugar, as correspondentes curvas de desconto obedecem  $P(0) = 1$  e  $\lim_{\tau \rightarrow \infty} P(\tau) = 0$ , ou seja, os preços dos correspondentes títulos zero cupom são 1 no vencimento e zero para um prazo tendendo ao infinito. Além disso,  $\lim_{\tau \rightarrow 0} y(\tau) = \beta_1 + \beta_2 = r$ , em que  $r$  é a taxa instantânea de curto prazo, e  $\lim_{\tau \rightarrow \infty} y(\tau) = \beta_1$ , uma constante. Um exercício de visualização dessas restrições pode ser feito facilmente, como mostram as figuras 3.7 e 3.8, assumindo-se como *proxies* de  $r$  e da taxa de longo prazo, respectivamente, a taxa DI CETIP de 1 dia e a taxa do vértice de 72 meses.

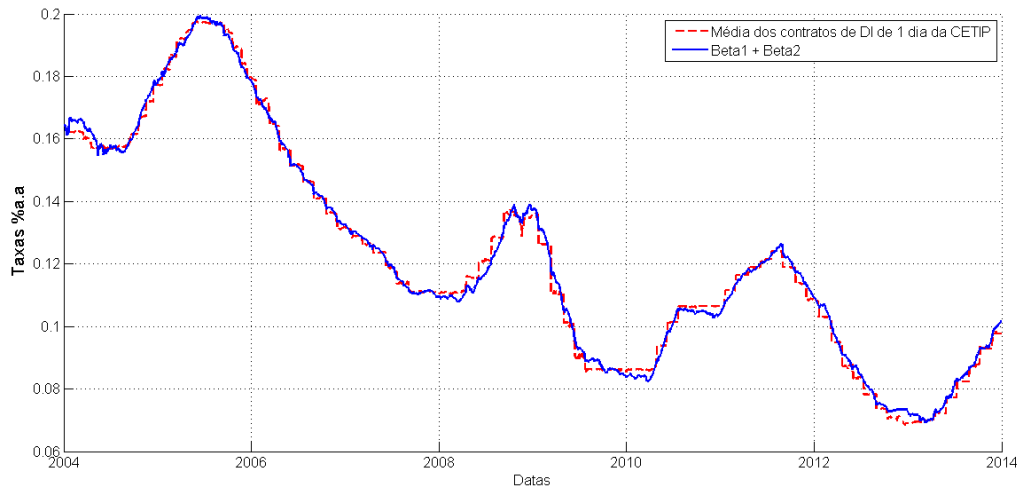


Figura 3.7: Séries temporais da taxa de juros curto prazo brasileira (DI CETIP) e série de  $\beta_1 + \beta_2$

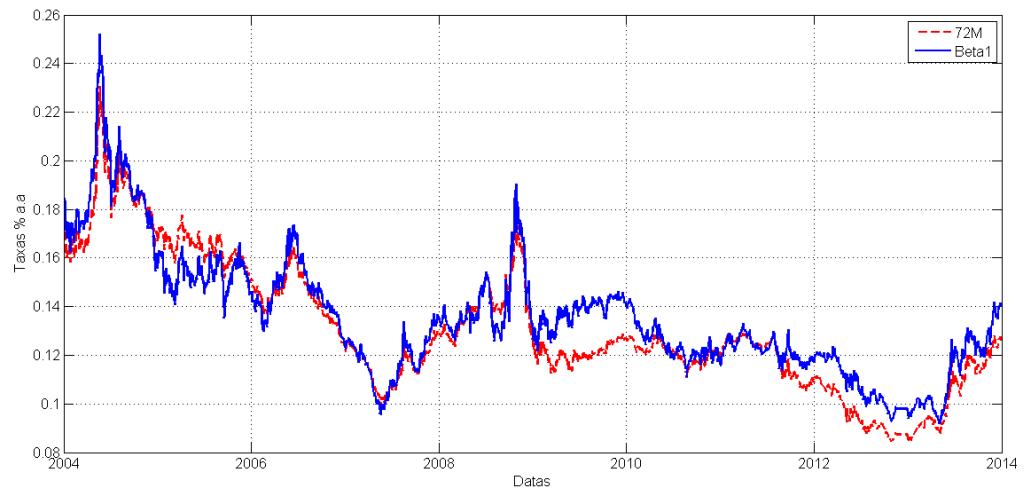


Figura 3.8: Séries temporais da taxa de 72 meses e de  $\beta_1$

Em segundo, vale notar que a estrutura parcimoniosa empregada, além de tratabilidade matemática, protege o modelo contra *overfitting*, importante para aplicações cujo foco é previsão.

Um terceiro ponto remete à flexibilidade do modelo, já que as curvas observadas podem assumir uma grande variedade de formas. É fácil verificar que a curva Nelson-Siegel, dependendo dos valores de  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  e  $\lambda$ , adapta-se facilmente às taxas observadas, conforme figura 3.9, abaixo.

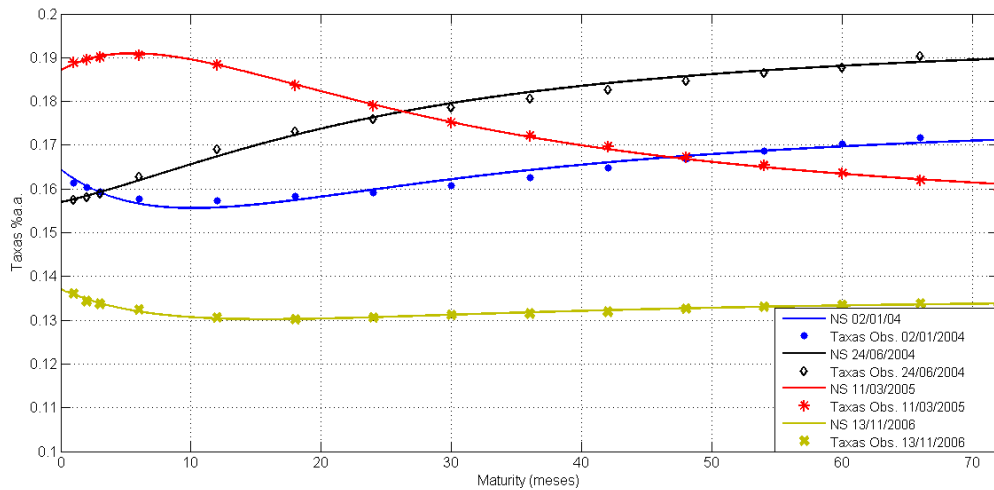


Figura 3.9: Possíveis formas assumidas pela ETTJ Nelson-Siegel em datas específicas

Além disso, a interpretação dos três fatores do modelo Nelson-Siegel assemelha-se muito com a abordagem de componentes principais de Litterman e Sheinkman (1991), conforme será apresentado nos gráficos seguintes.

A primeira semelhança a ser notada encontra-se entre os pesos  $(1, (\frac{1-e^{-\lambda\tau}}{\lambda\tau}), (\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}))$  que multiplicam os parâmetros  $\beta_1, \beta_2, \beta_3$  (figura 3.10) e os autovetores dos três primeiros componentes principais, conforme apresentado na seção anterior (figura 3.3).

Desta forma, analogamente a análise por componentes principais, o primeiro fator afeta a curva toda (nível), o segundo fator impacta os vértices de curto prazo (inclinação) e o terceiro fator afeta a parte intermediária da curva (curvatura). No modelo, o parâmetro  $\lambda$  é o fator de decaimento exponencial e determina o prazo em que o peso da curvatura é máximo.

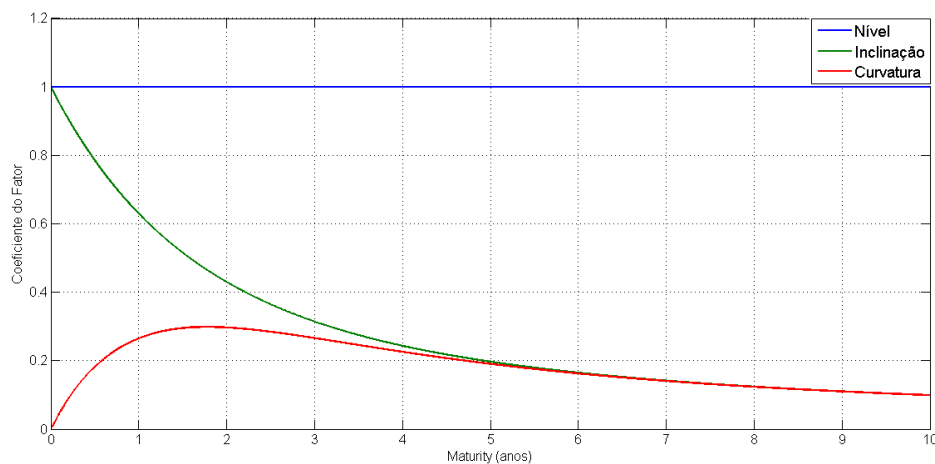


Figura 3.10: Peso dos fatores no modelo Nelson-Siegel

Para completar a análise de relação entre a abordagem de Litterman e Sheinkman (1991) e a interpretação de Diebold e Li (2006) para o modelo Nelson-Siegel, as figuras 3.8, 3.11 e 3.12, mostram as séries dos fatores plotados contra o nível, a inclinação e a curvatura empíricos, da mesma forma como feito na seção anterior.

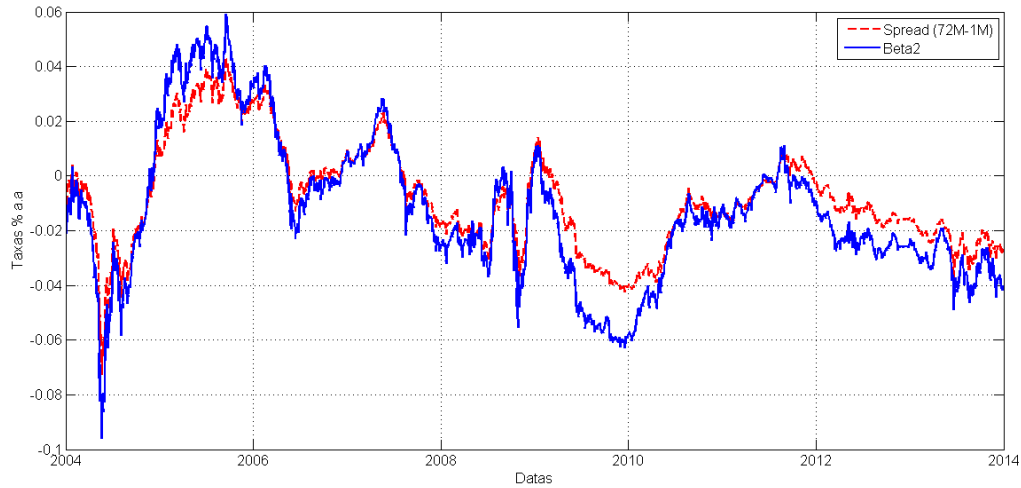


Figura 3.11: Séries temporais do *spread* ( $72m - 1m$ ) e de  $\beta_2$

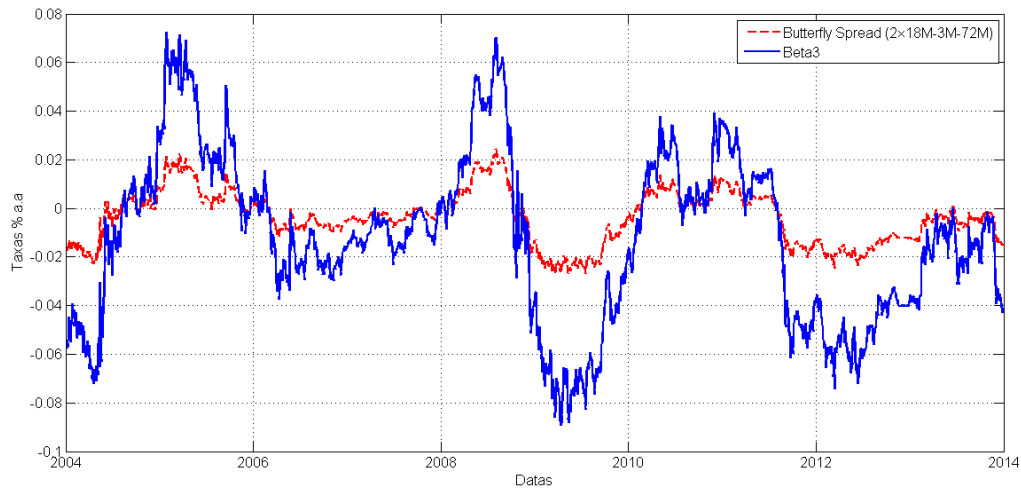


Figura 3.12: Séries temporais do *butterfly spread* ( $2 \times 18m - 3m - 72m$ ) e de  $\beta_3$

### 3.2.2 Nelson-Siegel Dinâmico

A versão estática do modelo Nelson-Siegel apresentada na seção anterior, pode ser interpretada como uma projeção de  $y(\tau)$  sobre as variáveis  $(1, (\frac{1-e^{-\lambda\tau}}{\lambda\tau}), (\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}))$ , com parâmetros  $\beta_1$ ,  $\beta_2$  e  $\beta_3$ , e  $\lambda$  conhecidos, em um determinado  $t$  fixo.

Assumindo que para cada  $\tau$  fixo exista uma serie temporal, é possível escrever:

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + \beta_{3t}\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right) \quad (3.7)$$

Alternativamente, tem-se agora uma série temporal da projeção linear de  $y_t$ , nas variáveis  $\beta_{1t}$ ,  $\beta_{2t}$  e  $\beta_{3t}$  com parâmetros  $(1, (\frac{1-e^{-\lambda\tau}}{\lambda\tau}), (\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}))$ .

Portanto, do ponto de vista *espacial*, os  $\beta$ s são parâmetros, mas do ponto de vista *temporal*, os  $\beta$ s são variáveis. Combinando as perspectivas espacial e temporal, obtém-se o modelo Nelson-Siegel Dinâmico (DNS). (DIEBOLD e RUDEBUSCH, 2013).

O DNS é, portanto, um modelo de fatores dinâmicos em que um grande número de variáveis (vértices de taxas de juros) é determinado por um pequeno número de variáveis de estado (nível, inclinação e curvatura) latentes.

Além disso, em decorrência da semelhança entre os pesos  $(1, (\frac{1-e^{-\lambda\tau}}{\lambda\tau}), (\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}))$  que multiplicam os parâmetros  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  (figura 3.10) e os autovetores dos três primeiros componentes principais (figura 3.3), é fácil inferir que  $\beta_{1t}$ ,  $\beta_{2t}$  e  $\beta_{3t}$  governam, respectivamente, nível, inclinação e curvatura da taxa de juros.

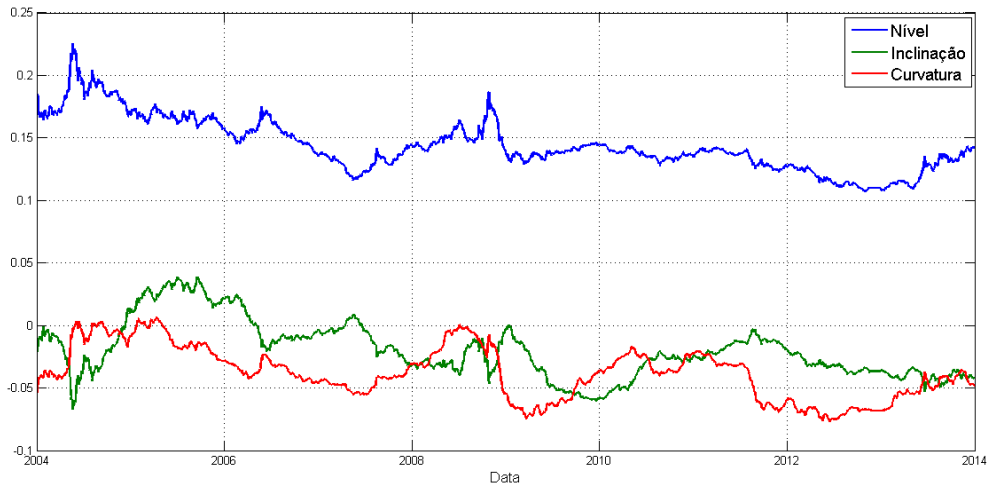


Figura 3.13: Dinâmica dos fatores DNS para Estrutura a Termo da Taxa de Juros no Brasil de jan/2004 a dez/2013

Seguindo a notação de Diebold et al. (2006), é possível reescrever o modelo DNS enfatizando a interpretação dos fatores (nível, inclinação e curvatura):

$$y_t(\tau) = l_t + s_t\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau}\right) + c_t\left(\frac{1 - e^{-\lambda_t\tau}}{\lambda_t\tau} - e^{-\lambda_t\tau}\right), \quad (3.8)$$

$$t = 1, \dots, T, \tau = 1, \dots, N.$$

Incluindo um termo de erro estocástico, chega-se a equação de medida, que relaciona estocasticamente o conjunto de taxas observáveis com os três fatores latentes do modelo:

$$y_t = \Lambda f_t + \varepsilon_t, \quad (3.9)$$

com:

$$y_t = \begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix}, f_t = \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix}, \varepsilon_t = \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_N) \end{pmatrix}, \text{ e } \Lambda = \begin{pmatrix} 1 & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}\right) & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1}\right) \\ 1 & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}\right) & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2}\right) \\ \vdots & \vdots & \vdots \\ 1 & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}\right) & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N}\right) \end{pmatrix}$$

em que  $\Lambda$  é a matriz de parâmetros.

A equação que governa a dinâmica dos estados (equação de transição) é dada por um processo vetorial autorregressivo de primeira ordem:

$$(f_t - \mu) = A(f_{t-1} - \mu) + \eta_t, \quad (3.10)$$

com:

$$f_t = \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix}, \eta_t = \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}, \mu = \begin{pmatrix} \mu^l \\ \mu^s \\ \mu^c \end{pmatrix}, \text{ e } A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

em que  $f_t$  e  $l_t$  são variáveis e  $\mu$  e  $A$  são parâmetros.

Para completar o sistema, é ainda necessária a especificação da estrutura de covariâncias das equações (3.9) e (3.13). A abordagem convencional consiste em assumir os erros das equações de medida e transição como ruído branco ortogonais entre si e ao estado inicial:

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim WN \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right), \quad (3.11)$$

$$E(f_0 \eta_t') = 0, \quad E(f_0 \varepsilon_t') = 0 \quad (3.12)$$

### 3.2.3 Incorporando Variáveis Macroeconômicas ao Modelo

Sob o ponto de vista macroeconômico, o juros de curto prazo é um dos instrumento de política monetária utilizados pelos bancos centrais para atuar em uma determinada economia sobre os objetivos de estabilização. No caso brasileiro em particular, a taxa SELIC é individualmente o principal instrumento de política monetária no sistema de metas de inflação. Por outro lado, na teoria de finanças, essa variável é tratada como elemento fundamental para modelar o restante da curva, já que os demais prazos podem ser interpretados como o valor futuro do juros de curto prazo ajustado ao risco. (Diebold, Piazzesi e Rudebusch, 2005).

Além disso, os vértices de diferentes prazos podem reagir de maneiras distintas à política monetária e aos choques na economia. Como consequência, é natural supor que uma

modelagem que considere a dinâmica conjunta das variáveis financeiras e macroeconômicas possa proporcionar um maior poder explicativo e capacidade de previsão.

Em linha com essa ideia, diversos trabalhos têm sido desenvolvidos nos últimos anos. Os primeiros a incorporar variáveis macroeconômicas a um modelo de fatores latentes para a ETTJ foram Ang e Piazzesi (2003). Através de um modelo vetorial autorregressivo e usando restrições de não-arbitragem, os autores concluíram que os fundamentos econômicos contribuíram significativamente para melhorar o poder de previsão dos modelos empregados.

Conforme Diebold e Rudebusch (2013), há uma grande quantidade de estudos que relacionam os fatores *nível, inclinação e curvatura* com os fundamentos macroeconômicos. Desta forma, conforme detalhado na seção 2.1 da Revisão da Literatura, as variáveis macro serão incorporadas ao modelo para a ETTJ segundo a interpretação de Diebold e Rudebusch (2013), baseando-se nas evidências de que o nível da curva é afetado pela inflação e a inclinação responde aos choques na atividade.

Assim, a fim de adicionar variáveis macroeconômicas ao modelo, é introduzida uma modificação na equação da dinâmica dos estados do modelo original, permitindo variáveis  $n$  exógenas além do componente autorregressivo. Desta forma, temos:

$$(f_t - \mu) = A(f_{t-1} - \mu) + BX_{t-1} + \eta_t, \quad (3.13)$$

onde  $X$  é um vetor contendo os valores das  $n$  variáveis macro disponíveis no tempo  $t - 1$  e  $B$  é a matriz de coeficientes com dimensão  $3 \times n$ .

A equação de medida e a especificação da matriz de covariância dos ruídos permanece a mesma da formulação original.

## 3.3 Modelo de Espaço de Estados

### 3.3.1 A Representação em Espaços de Estados

Segundo Dubin e Koopman (2012), um modelo de espaço de estados é uma abordagem para a análise de séries temporais em que o sistema em estudo é determinado por uma série de vetores não observáveis  $\alpha_1, \dots, \alpha_n$  guiados por um processo estocástico e associados a uma série de observações  $y_1, \dots, y_n$ .

Conforme Lutkepohl (2005), as relações entre os  $\alpha_t$ s e os  $y_t$ s são especificadas pela *equação de observação* ou de *medida*:

$$y_t = Z_t \alpha_t + \varepsilon_t \quad (3.14)$$

em que  $Z_t$  é uma matriz de parâmetros constantes ou contemporânea a  $y_t$  e  $\varepsilon$  é o erro de observação, assumido como um ruído.



Complementarmente, o vetor de estados é definido como:

$$\alpha_{t+1} = T_t \alpha_t + \eta_t \quad (3.15)$$

que é tipicamente chamada de *equação de transição* já que descreve a transição do estado, de um período  $t$  a um período  $t + 1$ . A matriz  $T_t$  é uma matriz de coeficientes e  $\eta_t$  equivale ao erro do processo.

O principal propósito da análise de espaços de estados é inferir as propriedades dos fatores latentes  $\alpha_t$ s a partir do conhecimento das observações  $y_t$ s (medidas), e a partir daí, ser utilizada para previsão, extração de sinal ou estimação de parâmetros.

### 3.3.2 Filtro de Kalman

Segundo Martin et al. (2013), o Filtro de Kalman é um algoritmo utilizado para previsão de  $m$  fatores latentes, com base nas suas médias condicionais e na atualização dessas previsões de forma sistemática, a medida em que novas medidas das  $p$  variáveis observáveis tornam-se disponíveis. É, portanto, tipicamente associado aos modelos de espaço de estados.

A representação em espaço de estados do modelo DNS faz com que o Filtro de Kalman seja uma escolha natural para estimações ótimas dos estados (nível, inclinação e curvatura) e das taxas de juros observadas  $h$ -passos à frente, conhecidos os parâmetros do modelo.

Considere o modelo de espaço de estados:

$$y_t = Z_t \alpha_t + \epsilon_t, \quad \epsilon_t \sim N(0, H_t) \quad (3.16)$$

$$\alpha_{t+1} = T_t \alpha_t + \eta_t, \quad \eta_t \sim N(0, Q_t) \quad (3.17)$$

para  $t = 1, \dots, n$ , em que  $\eta_t$ s e  $\epsilon_t$ s são mutuamente independentes e também independentes do estado inicial  $\alpha_1 \sim N(a_1, P_1)$ , com  $a_1$  e  $P_1$  conhecidos e dimensões dos vetores e matrizes dados conforme tabela 3.1.

Vetor		Matriz	
$y_t$	$p \times 1$	$Z$	$p \times m$
$\alpha_t$	$m \times 1$	$T$	$m \times m$
$\epsilon_t$	$p \times 1$	$H$	$p \times p$
$\eta_t$	$r \times 1$	$Q$	$r \times r$
$a_1$	$m \times 1$	$P_1$	$m \times m$

Tabela 3.1: Dimensões das matrizes do modelo de espaço de estados.  
Fonte: Durbin e Koopman(2012), p.82.

Seja  $Y_{t-1}$  o vetor de observações  $(y_1, \dots, y_{t-1})'$  para  $t = 2, 3, \dots$ , com  $Y_0$  indicando que não exista observação anterior a  $t = 1$ . Começando em  $t = 1$  e calculando recursiva-

mente as distribuições de  $\alpha_t$  e  $y_t$ , é possível provar que  $p(y_t|\alpha_1, \dots, \alpha_t, Y_{t-1}) = p(y_t|\alpha_t)$  e  $p(\alpha_{t+1}|\alpha_1, \dots, \alpha_t, Y_{t-1}) = p(\alpha_{t+1}|\alpha_t)$ .

Assuma que a distribuição condicional de  $\alpha_t$  dado  $Y_{t-1}$  seja  $N(\alpha_t, P_t)$ , em que  $\alpha_t$  e  $P_t$  são conhecidos. Assuma também que a distribuição condicional de  $\alpha_t$  dado  $Y_t$  seja  $N(a_{t|t}, P_{t|t})$  e a distribuição de  $\alpha_{t+1}$  dado  $Y_t$  é  $N(a_{t+1}, P_{t+1})$ .

Sob essas hipóteses, é possível estimar  $a_{t|t}$ ,  $P_{t|t}$ ,  $a_{t+1}$ ,  $P_{t+1}$  quando  $y_t$  torna-se disponível, em que  $a_{t+1}$  (a média condicional de  $a_{t+1}$ ) representa o estimador ótimo em  $t+1$  para o vetor de estados.

O Filtro de Kalman, portanto, é um algoritmo que calcula recursivamente as estimações de  $a_{t+1}$  e  $P_{t+1}$  a partir de  $a_t$  e  $P_t$  através de uma sequência de equações conforme segue:

$$\nu_t = y_t - E(y_t|Y_{t-1}) = y_t - Za_t \quad (3.18)$$

$$F_t = \text{Var}(\nu_t|Y_{t-1}) = ZP_tZ' + H \quad (3.19)$$

$$a_{t|t} = a_t + P_tZ'F_t^{-1}\nu_t \quad (3.20)$$

$$P_{t|t} = \text{Var}(\alpha_t|Y_t) = P_t - P_tZ'F_t^{-1}ZP_t \quad (3.21)$$

$$\alpha_{t+1} = Ta_{t|t} = Ta_t + K_t\nu_t \quad (3.22)$$

$$P_{t+1} = TP_tT' + Q = TP_t(T - K_tZ_t)' + Q \quad (3.23)$$

para  $t = 1, \dots, n$ , onde a matriz  $K_t = TP_tZ'F_t^{-1}$  é chamada de *ganho de Kalman* e a matriz  $F_t$  é assumida como não singular, o que às vezes pode gerar problemas práticos como mau condicionamento numérico e divergência.

O conjunto de equações 3.22 e 3.23 é conhecido como *passo de previsão* e as equações 3.20 e 3.21 são chamadas de *passo de atualização*. Quando calculadas forma iterativa, permitem atualizar o conhecimento do sistema a cada vez que uma nova observação é disponibilizada.

Maiores detalhes sobre as derivações das fórmulas podem ser encontrados em Durbin e Koopman (2012), Lutkepohl (2005), Goodwin e Sin (2009) ou Grewal e Andrews (2001).

### 3.3.3 Filtro de Kalman Estendido

Segundo Fulop(2011), é frequente que o modelo em finanças seja normal, mas que as equações de medida e estado sejam não-lineares.

Nesses casos, o Filtro de Kalman conforme apresentado na seção anterior não pode mais ser aplicado diretamente. Segundo Goodwin e Sin (2009), é possível utilizar uma

aproximação linear para superar essa limitação, em que destacam-se as abordagens conhecidas como Filtro de Kalman Estendido e Filtro de Kalman *Unscented*.

Desta forma, o Filtro de Kalman Estendido, que será utilizado neste trabalho, baseia-se na ideia de linearização das equações de estado e medida e subsequente aplicação do Filtro de Kalman ao modelo linearizado (Durbin e Koopman, 2012).

Considere agora o modelo não-linear em que os ruídos sejam não necessariamente Gaussianos:

$$y_t = Z_t(\alpha_t) + \epsilon_t \quad (3.24)$$

$$\alpha_{t+1} = T_t(\alpha_t) + \eta_t \quad (3.25)$$

para  $t = 1, \dots, n$ , com  $Z_t(\alpha_t)$  e  $T_t(\alpha_t)$  funções diferenciáveis de  $\alpha_t$ , e os ruídos  $\epsilon_t$  e  $\eta_t$  sejam serialmente e mutuamente ortogonais com média zero e matriz de covariâncias  $H$  e  $Q$ , respectivamente. Além disso, o vetor de estados inicial  $\alpha_1$  é aleatório, com média  $a_1$  e matriz de covariâncias  $P_1$ .

Considere ainda, como na seção anterior, que  $a_{t|t}$  seja o estimador filtrado do estado  $\alpha_t$ , com variância  $P_{t|t}$ , e  $a_{t+1}$  seja o estimador de  $\alpha_{t+1}$  um passo a frente e  $P_{t+1}$  sua variância.

Então, é possível definir:

$$\dot{Z}_t = \left. \frac{\partial Z_t(\alpha_t)}{\partial \alpha'_t} \right|_{\alpha_t = a_t}, \quad \dot{T}_t = \left. \frac{\partial T_t(\alpha_t)}{\partial \alpha'_t} \right|_{\alpha_t = a_{t|t}} \quad (3.26)$$

e expandir as matrizes  $Z_t(a_t)$  e  $T_t(a_t)$  em séries de Taylor, nos pontos  $a_t$  e  $a_{t|t}$ :

$$Z_t(\alpha_t) = Z_t(a_t) + \dot{Z}_t(\alpha_t - a_t) + \dots$$

$$T_t(\alpha_t) = T_t(a_t) + \dot{T}_t(\alpha_t - a_t) + \dots$$

Desconsiderando os termos de ordem maior e substituindo as expansões em séries de Taylor nas equações 3.24 e 3.25, é possível chegar em um novo conjunto de fórmulas para o algoritmo do Filtro de Kalman aplicado ao modelo linearizado:

$$\nu_t = y_t - E(y_t|Y_{t-1}) = y_t - Z a_t \quad (3.27)$$

$$F_t = Var(\nu_t|Y_{t-1}) = \dot{Z} P_t \dot{Z}' + H \quad (3.28)$$

$$a_{t|t} = a_t + P_t \dot{Z}' F_t^{-1} \nu_t \quad (3.29)$$

$$P_{t|t} = Var(\alpha_t|Y_t) = P_t - P_t \dot{Z}' F_t^{-1} \dot{Z} P_t \quad (3.30)$$

$$\alpha_{t+1} = Ta_{t|t} = Ta_t + K_t \nu_t \quad (3.31)$$

$$P_{t+1} = \dot{T}P\dot{T}' + Q = \dot{T}P_t(\dot{T} - K_t\dot{T})' + Q \quad (3.32)$$

Onde o ganho de Kalman é definido como  $K_t = \dot{T}P_t\dot{Z}'F_t^{-1}$ .

Essa sequência recursiva de equações, juntamente com as especificações dos valores iniciais de  $a_1$  e  $P_1$  é chamado de Filtro de Kalman Estendido. É importante destacar que em decorrência do processo de linearização, o Filtro de Kalman Estendido produz estimativas sub-ótimas.

Maiores detalhes sobre a derivação das fórmulas e análises podem ser encontrados em Durbin e Koopman (2012), Goodwin e Sin (2009) e Grewal e Andrews (2001).

Uma aplicação do Filtro de Kalman Estendido para o modelo Nelson-Siegel pode ser encontrado em Koopman et al.(2011), em que os autores interpretam o parâmetro de decaimento exponencial ( $\lambda$ ) como um quarto elemento no vetor de estados, modelando-o conjuntamente no processo autorregressivo.

### 3.3.4 Estimação de Parâmetros

De forma geral, ao aplicar o Filtro de Kalman para um modelo de espaços de estados, supõe-se que as matrizes de parâmetros das equações de estado e medida sejam conhecidos. Na prática, porém, esta hipótese não é verdadeira e estas matrizes precisam ser estimadas.

Inúmeros procedimentos estão disponíveis para a estimação de parâmetros em um modelo de espaço de estados. As alternativas vão desde simples modelos em dois passos até modelos de análise Bayesiana usando método de Monte-Carlo com cadeias Markovianas (Diebold e Rudebusch, 2013).

Particularmente, para as aplicações envolvendo o modelo dinâmico de Nelson-Siegel, existem duas abordagens populares.

Na primeira, tradicionalmente conhecida como *two-step DNS*, ou DNS em dois passos, introduzida em Diebold e Li (2006), aplica-se o método dos mínimos quadrados ordinários (MQO) em cada seção transversal, a fim de obter as séries temporais dos fatores latentes e os erros das medidas. Em um segundo momento, também por MQO, estima-se os parâmetros da dinâmica através de um modelo vetorial-autorregressivo sobre as séries temporais dos fatores latentes, obtendo-se simultaneamente os erros do estado. Segundo Diebold e Rudebusch (2013), o custo da estimação em dois passos é uma possível sub otimalidade estatística, já que os erros de estimação do primeiro passo não são considerados na segunda estimação.

O segundo método é conhecido como *one-step DNS*. Nele, todos os parâmetros são estimados simultaneamente. Conforme Diebold e Rudebusch (2013), a estimação em um passo pode ser obtida de diversas maneiras. O procedimento clássico consiste em

maximizar a função de máxima-verossimilhança usando o Filtro de Kalman, que entrega as inovações necessárias para o cálculo das pseudo-verossimilhança Gaussiana. É possível, portanto, achar uma configuração de parâmetros que maximize tal função. Em geral, são aplicados métodos convencionais de cálculo numérico para a otimização, como por exemplo o método de gradientes.

Diebold e Rudebusch (2013), no entanto, destacam que apesar da superioridade em princípios do método em um passo quando utilizado com métodos de gradientes, os resultados são, em geral, intratáveis ou não confiáveis. A razão para esta afirmação é devida ao número de parâmetros estimados. Por exemplo, em um sistema com 15 vértices, é possível que até 139 parâmetros tenham que ser estimados, gerando uma quantidade grande de máximos locais para a função a ser otimizada, além de diversas combinações de parâmetros que geram o mesmo máximo, um problema de identificação.

Este problema é ainda maior quando variáveis macroeconômicas são incluídas nos modelos. Laurini (2011), aponta que a estimação por verossimilhança é notoriamente problemática nesses modelos da estrutura a termo da taxa de juros. O problema é relacionado à existência de múltiplos máximos locais com vetores de parâmetros distintos, que implicam em diferentes modelos identificados para um mesmo sistema com interpretações econômicas diferentes.

Desta forma, uma alternativa aos métodos apresentados será apresentado em seguida.

### **Estimação de Parâmetros com o Filtro de Kalman Estendido**

Quando um problema de estimação de parâmetros é resolvido simultaneamente à estimação dos estados, o modelo linear torna-se não-linear. Desta forma, uma alternativa possível à estimação por máxima-verossimilhança é a aplicação do Filtro de Kalman Estendido com o vetor de estados aumentado pela inclusão dos parâmetros desconhecidos. Desta forma, os parâmetros são estimados recursivamente, dada a informação contida no sistema até o instante da estimação.

Segundo Ljung e Söderström (1987), o uso do Filtro de Kalman Estendido para a identificação de parâmetros parece ter sido sugerido pela primeira vez por Koop e Oxford (1963) e Cox (1964), sendo então utilizado em diversas aplicações.

Nestes casos em que é aplicado para estimação de parâmetros, segundo Goodwin e Sin (2009), o Filtro de Kalman Estendido pode ser interpretado como estimador de estados adaptativo. Isso significa que toda a informação necessária para a arquitetura do filtro ou preditor é obtida diretamente através da análise dos dados do próprio sistema. Segundo os autores, esta é uma alternativa atrativa nos casos em que se dispõe de pouco conhecimento do sistema ou quando os parâmetros variam no tempo.

Considere o seguinte modelo:

$$y_t = Z(\theta)\alpha_t + \varepsilon_t \quad (3.33)$$

$$\alpha_{t+1} = T(\theta)\alpha_t + \eta_t \quad (3.34)$$

em que  $\varepsilon_t$  e  $\eta_t$  são vetores aleatórios independentes com média zero e matrizes de covariâncias  $H_t$  e  $Q_t$ . O vetor de parâmetros  $\theta$  é desconhecido. Suponha que sua distribuição *a priori* tenha média  $\theta_0$  e matriz de covariâncias  $P_0$ .

Para determinar o estimador recursivo para  $\theta$ , considere o vetor de estados estendido:

$$A_t = \begin{pmatrix} \alpha_t \\ \theta \end{pmatrix} \quad (3.35)$$

e o problema de estimação dos estados para o vetor estendido, que obedece as seguintes equações:

$$y_t = \bar{Z}(A_t) + \varepsilon_t \quad (3.36)$$

$$A_{t+1} = \bar{T}(A_t) + \bar{\eta}_t \quad (3.37)$$

em que

$$\bar{T}(A_t) = \begin{pmatrix} T(\alpha_t) \\ \theta \end{pmatrix}, \quad (3.38)$$

$$\bar{\eta}_t = \begin{pmatrix} \eta_t \\ 0 \end{pmatrix}, \quad (3.39)$$

$$\bar{Z}(A_t) = Z(\theta)\alpha_t. \quad (3.40)$$

Desta forma, o problema da identificação recursiva dos parâmetros pode ser formulado como um problema de filtragem ou de estimação dos estados não-linear. Consequentemente, é possível aplicar o algoritmo do Filtro de Kalman Estendido.

# Capítulo 4

## Metodologia

Em resumo, foram comparados os modelos Nelson-Siegel Dinâmico com e sem variáveis exógenas para os dados brasileiros, e para cada um deles, foram aplicadas duas alternativas para a estimação dos parâmetros. Desta forma, a análise concentrou-se em quatro modelos finais:

1. Nelson-Siegel Dinâmico com estimação em dois passos (2-STEP DNS);
2. Nelson-Siegel Dinâmico com variáveis exógenas e estimação em dois passos (2-STEP DNSX);
3. Nelson-Siegel Dinâmico com estimação com Filtro de Kalman Estendido (EKF DNS);
4. Nelson-Siegel Dinâmico com variáveis exógenas e estimação com Filtro de Kalman Estendido (EKF DNSX);

A implementação foi feita em MATLAB, com o uso do *toolbox* EKF/UKF. A análise concentrou-se nas estimativas das taxas e seus respectivos erros. Neste capítulo serão descritos em detalhes as especificações dos modelos e os métodos adotados para a comparação dos modelos.

### 4.1 Modelos Comparados

#### Nelson-Siegel Dinâmico

O modelo Nelson-Siegel Dinâmico pode ser representado a partir da seguinte formulação em espaço de estados, para um dado  $\lambda$  fixo:

$$y_t(\tau) = \Lambda(\tau)f_t + \varepsilon_t(\tau), \quad (4.1)$$

$$(f_t - \mu) = A(f_{t-1} - \mu) + \eta_t, \quad (4.2)$$

Em que a equação de medida é dada por 4.1, onde  $f_t$  é o vetor  $(3 \times 1)$  contendo os fatores latentes no instante  $t$ ,  $\Lambda$  é a matriz  $(1, (\frac{1-e^{-\lambda\tau}}{\lambda\tau}), (\frac{1-e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}))$  com dimensão  $(15 \times 3)$ , contendo os pesos dos fatores para cada prazo  $\tau$ ,  $\lambda$  é um parâmetro escalar que determina o prazo onde o peso para o fator *curvatura* é máximo.

A equação 4.2 rege a dinâmica do modelo, em que  $A$  é uma matriz  $(3 \times 3)$  contendo os coeficientes autorregressivos e  $\mu$  é o vetor  $(3 \times 1)$  que representa a média dos fatores.

Neste estudo, para reduzir a dimensionalidade do sistema, é assumido que a matriz  $A$ , que contém os coeficientes autorregressivos, é diagonal. Apesar de restritiva, é uma hipótese plausível, já que os três fatores de Nelson-Siegel têm interpretação análoga aos componentes principais de Litterman e Sheinkman (1991), ortogonais pos construção.

Além disso, é assumido que os ruídos das equações de transição e medida sejam normalmente distribuídos e mutuamente não-correlacionados.

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim WN \left( \begin{pmatrix} 0 \\ 0 \end{pmatrix} \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right), \quad (4.3)$$

em que  $Q$  e  $H$  são matrizes diagonais de dimensão  $(15 \times 15)$  e  $(3 \times 3)$ .

Assim, considerando os quinze vértices para a ETTJ, foi utilizada a seguinte formulação matricial para o modelo com os parâmetros fixos e lambda calibrado:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_{15}) \end{pmatrix} = \begin{pmatrix} 1 & (\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}) & (\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1}) \\ 1 & (\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}) & (\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2}) \\ \vdots & \vdots & \vdots \\ 1 & (\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}) & (\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N}) \end{pmatrix} \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_{15}) \end{pmatrix}, \quad (4.4)$$

$$\begin{pmatrix} l_t - \mu_l \\ s_t - \mu_s \\ c_t - \mu_c \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} l_{t-1} - \mu_l \\ s_{t-1} - \mu_s \\ c_{t-1} - \mu_c \end{pmatrix} + \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}, \quad (4.5)$$

E, para o modelo com os parâmetros variando no tempo e lambda calibrado:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_{15}) \end{pmatrix} = \begin{pmatrix} 1 & (\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}) & (\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1}) \\ 1 & (\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}) & (\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2}) \\ \vdots & \vdots & \vdots \\ 1 & (\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}) & (\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N}) \end{pmatrix} \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_{15}) \end{pmatrix}, \quad (4.6)$$

$$\begin{pmatrix} l_t - \mu_t^l \\ s_t - \mu_t^s \\ c_t - \mu_t^c \end{pmatrix} = \begin{pmatrix} a_{11,t} & 0 & 0 \\ 0 & a_{22,t} & 0 \\ 0 & 0 & a_{33,t} \end{pmatrix} \begin{pmatrix} l_{t-1} - \mu_{t-1}^l \\ s_{t-1} - \mu_{t-1}^s \\ c_{t-1} - \mu_{t-1}^c \end{pmatrix} + \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}, \quad (4.7)$$



### Nelson-Siegel Dinâmico com variáveis exógenas macroeconômicas

O último modelo avaliado neste estudo consiste em introduzir à dinâmica da estrutura Nelson-Siegel as variáveis exógenas de inflação e atividade real com o intuito de obter melhores estimativas das taxas futuras. Desta forma, as equações de medida e transição do modelo de espaço de estados são expressas da seguinte forma:

$$y_t(\tau) = \Lambda(\tau)f_t + BX_{t-1} + \varepsilon_t(\tau), \quad (4.8)$$

$$(f_t - \mu) = A(f_{t-1} - \mu) + \eta_t, \quad (4.9)$$

em que os vetores e matrizes possuem interpretação e dimensões equivalentes ao modelo anterior, e  $X$  é um vetor contendo os valores das variáveis de inflação e atividade disponíveis no tempo  $t - 1$  e  $B$  é a matriz de coeficientes com dimensão  $(3 \times 2)$ .

Para a especificação da matriz  $B$ , foram utilizadas as conclusões de Diebold e Rudebusch (2013) de que o nível é correlacionado com a inflação e a inclinação responde a choques na atividade.

Desta forma, o modelo final com parâmetros fixos e lambda calibrado é dado por:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_{15}) \end{pmatrix} = \begin{pmatrix} 1 & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}\right) & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1}\right) \\ 1 & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}\right) & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2}\right) \\ \vdots & \vdots & \vdots \\ 1 & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}\right) & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N}\right) \end{pmatrix} \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_{15}) \end{pmatrix}, \quad (4.10)$$

$$\begin{pmatrix} l_t - \mu_l \\ s_t - \mu_s \\ c_t - \mu_c \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix} \begin{pmatrix} l_{t-1} - \mu_l \\ s_{t-1} - \mu_s \\ c_{t-1} - \mu_c \end{pmatrix} + \begin{pmatrix} b_{11} & 0 \\ 0 & b_{22} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Inf_{t-1} \\ Ativ_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}, \quad (4.11)$$

E o modelo com parâmetros variando no tempo e lambda calibrado ficou:

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_{15}) \end{pmatrix} = \begin{pmatrix} 1 & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1}\right) & \left(\frac{1-e^{-\lambda\tau_1}}{\lambda\tau_1} - e^{-\lambda\tau_1}\right) \\ 1 & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2}\right) & \left(\frac{1-e^{-\lambda\tau_2}}{\lambda\tau_2} - e^{-\lambda\tau_2}\right) \\ \vdots & \vdots & \vdots \\ 1 & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N}\right) & \left(\frac{1-e^{-\lambda\tau_N}}{\lambda\tau_N} - e^{-\lambda\tau_N}\right) \end{pmatrix} \begin{pmatrix} l_t \\ s_t \\ c_t \end{pmatrix} + \begin{pmatrix} \varepsilon_t(\tau_1) \\ \varepsilon_t(\tau_2) \\ \vdots \\ \varepsilon_t(\tau_{15}) \end{pmatrix}, \quad (4.12)$$

$$\begin{pmatrix} l_t - \mu_t^l \\ s_t - \mu_t^s \\ c_t - \mu_t^c \end{pmatrix} = \begin{pmatrix} a_{11,t} & 0 & 0 \\ 0 & a_{22,t} & 0 \\ 0 & 0 & a_{33,t} \end{pmatrix} \begin{pmatrix} l_{t-1} - \mu_t^l \\ s_{t-1} - \mu_t^s \\ c_{t-1} - \mu_t^c \end{pmatrix} + \begin{pmatrix} b_{11,t} & 0 \\ 0 & b_{22,t} \\ 0 & 0 \end{pmatrix} \begin{pmatrix} Inf_{t-1} \\ Ativ_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t^l \\ \eta_t^s \\ \eta_t^c \end{pmatrix}, \quad (4.13)$$

## 4.2 Calibragem do parâmetro $\lambda$

Durante a análise, o parâmetro  $\lambda$  foi considerado como uma constante, a exemplo de Diebold e Li (2006). Para calibrá-lo, foi adotado procedimento similar a Caldeira et al.(2012), em que o  $\lambda$  é escolhido de forma a minimizar a diferença entre as taxas de juros obtidas no modelo Nelson-Siegel e as taxas observadas. Desta forma, foi encontrado o valor ( $\lambda = 0.08525$ ), que equivale a um máximo para a maturidade de 18 meses na curvatura. A figura 4.1 ilustra o procedimento.

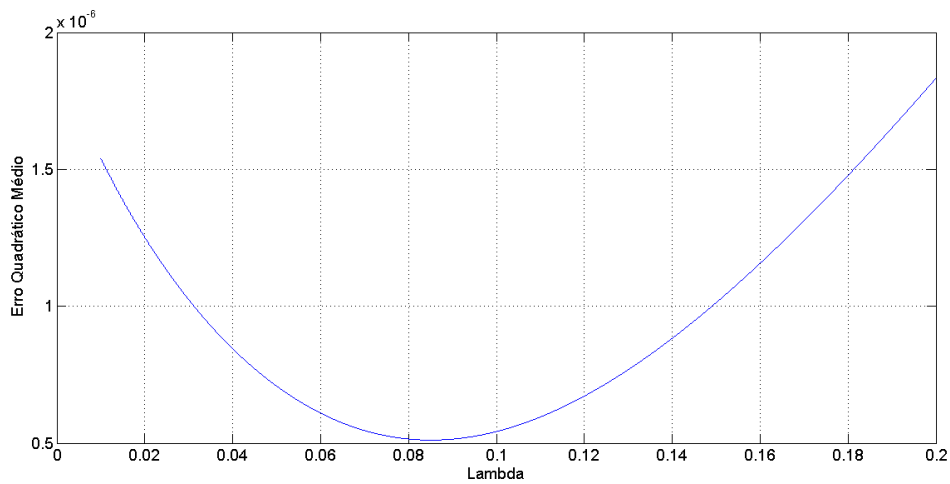


Figura 4.1: Soma dos erros quadráticos médios dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 para diferentes valores do parâmetro  $\lambda$

## 4.3 Estimação dos Parâmetros

### 4.3.1 2-Step DNS

O procedimento para estimação em dois passos dos parâmetros do modelo DNS foi introduzido em Diebold e Li (2006). Para um dado  $\lambda$  calibrado, o primeiro passo consiste em aplicar o método dos mínimos quadrados ordinários (MQO) em cada seção transversal, a fim de obter as séries temporais dos fatores latentes e os erros das medidas. Em um segundo momento, também por MQO, estima-se os parâmetros da dinâmica através de um modelo vetorial-autorregressivo sobre as séries temporais dos fatores latentes, obtendo-se simultaneamente os erros do estado.

### 4.3.2 EKF DNS com $\theta_t = \theta_{t-1} + \eta_t$

Tratando os modelo em espaço de estados apresentados nas seções anteriores como sistemas não-lineares, é possível resolver simultaneamente os problemas de estimação de parâmetros e dos estados.

Dada a complexidade envolvida em um processo de filtragem não-linear, uma alternativa consiste em utilizar o Filtro de Kalman Estendido, conforme descrito na seção. 3.3.4.

Usando esta técnica, é possível permitir ao vetor de parâmetros um componente estocástico, dando liberdade para que variem no tempo e sejam estimados recursivamente no processo de filtragem a cada vez que uma nova observação torna-se disponível.

Neste caso, para a estimação dos parâmetros, foi utilizado o Filtro de Kalman Estendido com o vetor de estados ( $f_t$ ) aumentado pela inclusão do vetor de parâmetros ( $\theta$ ).

As figuras 4.2 e 4.3 mostram o processo de estimação, em que é permitido que os parâmetros sejam reestimados a cada passo.

Os gráficos na figura 4.2 evidenciam a evolução dinâmica dos parâmetros do modelo dinâmico Nelson-Siegel sem variáveis exógenas. Na região mais à esquerda dos gráficos, percebe-se uma maior oscilação das linhas, em que o filtro está "aprendendo". Após esse período, é possível verificar um comportamento mais uniforme.

Importante notar que os coeficientes autorregressivos e os ruídos tendem a convergir para um valor mais ou menos fixo, enquanto as médias apresentam uma maior variabilidade.

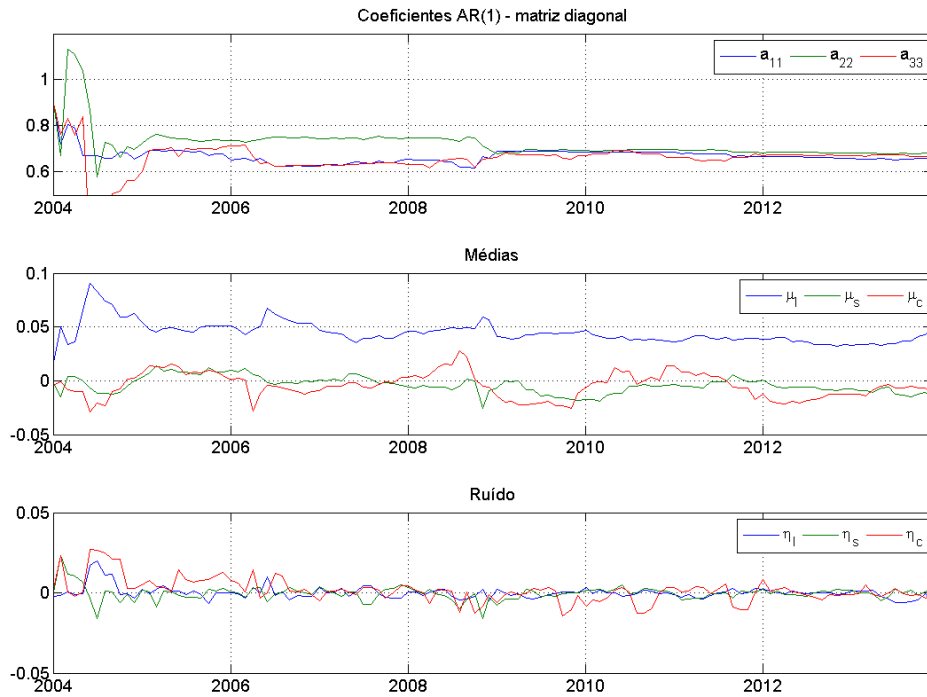


Figura 4.2: Estimação dos parâmetros do modelo DNS para os dados mensais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 usando Filtro de Kalman Estendido com  $\theta_t = \theta_{t-1} + \varepsilon_t$

Na figura 4.2, estão os gráficos da evolução dos parâmetros quando são incluídas as variáveis exógenas ao modelo. O comportamento dos coeficientes autorregressivos diferem um pouco dos anteriores, oscilando em nível mais baixo. No entanto, também convergem para valores mais ou menos fixos. Por outro lado as médias apresentam valores um pouco mais altos do que no modelo anterior. Já os coeficientes das variáveis exógenas (inflação e atividade) mostram um comportamento distintos dos demais, com períodos relativamente longos em tendência, mas aparentemente oscilando em torno de uma média.

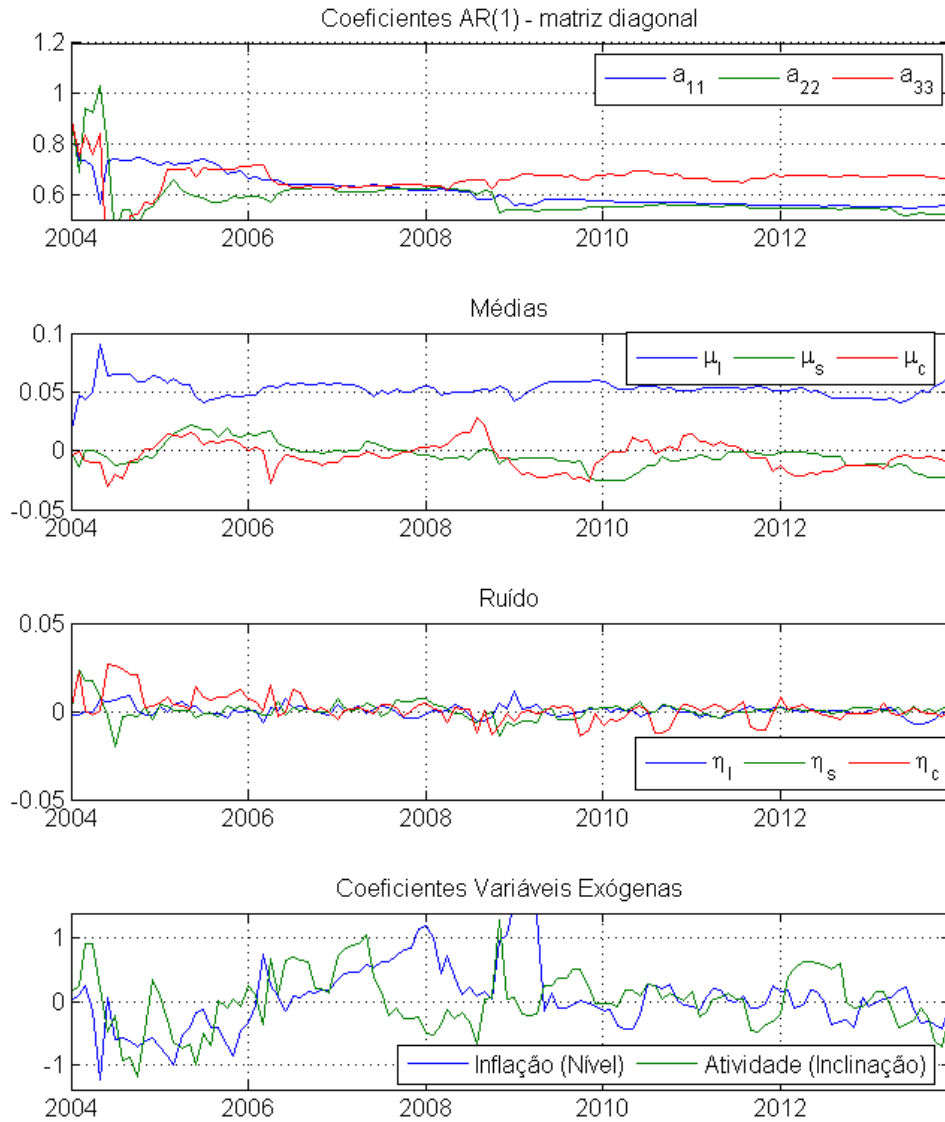


Figura 4.3: Estimação dos coeficientes autorregressivos do modelo DNS com variáveis exógenas para os dados mensais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013 usando Filtro de Kalman Estendido com  $\theta_t = \theta_{t-1} + \varepsilon_t$

## 4.4 Análise dentro da amostra

Para a comparação dos modelos dentro da amostra, foi realizada a análise dos gráficos dos erros das estimativas um passo a frente durante os processos de filtragem.

Para isso, no método de estimação em dois passos, a estimação levou em conta todos

os dados da amostra, totalizando 120 dados para cada vértice e exógena.

Quando o método utilizado foi o Filtro de Kalman Estendido, as estimações dos parâmetros foram condicionadas ao conjunto de informação disponível até o instante da estimação.

Complementarmente à análise gráfica, foram comparados os resultados obtidos com o cálculo da raiz dos erros quadráticos médios (REQM).

O procedimento adotado foi o de análise dos pares (com e sem exógenas) e em seguida foram comparados os resultados obtidos pelos diferentes métodos de estimação de parâmetros.

## 4.5 Análise fora da amostra

Para avaliação fora da amostra, o objeto de análise foi o conjunto das estimativas para as taxas de juros 3, 6, 9 e 12 meses à frente calculadas de forma iterativa e realizadas após o período de estimação dos parâmetros para os modelos com e sem variáveis exógenas e as duas metodologias de estimação de parâmetros.

Conforme Pooter et al.(2010), dado que um modelo que incorpore variáveis macro será utilizado para estimativas das taxas de juros, torna-se também necessária a existência de um modelo para a estimativa dos fatores macroeconômicos. A fim de delimitar o escopo desta pesquisa, foi assumida como hipótese que as trajetórias futuras das variáveis inflação e atividade são conhecidas, desta forma, puderam ser incluídas como entrada nas previsões também de modo iterativo.

Para todos os modelos, o procedimento adotado foi o seguinte:

1. Inicialmente, foram realizadas as estimativas dos parâmetros utilizando 2/3 da amostra total (80 meses de dados);
2. Os parâmetros estimados foram utilizados como entradas em um Filtro de Kalman;
3. Foi realizada a filtragem dos dados a partir do primeiro dado até o último a fim de obter as séries das médias filtradas dos estados latentes (nível, inclinação e curvatura);
4. Fora da amostra (40 meses de dados), foram realizadas sucessivas previsões para os fatores nível, inclinação e curvatura para 3, 6, 9 e 12 meses à frente de maneira iterativa, totalizando respectivamente 37, 34, 31 e 28 previsões, de forma que o cálculo de cada estimativa dependesse apenas da última média filtrada, da estimativa imediatamente anterior e das estimativas para as variáveis exógenas disponíveis também no passo imediatamente anterior (por hipótese, conhecidas);
5. As previsões dos estados 3, 6, 9 e 12 meses a frente foram convertidas em estimativas de taxas para cada vértice;

6. As estimativas das taxas de juros para cada vértice foram confrontadas com os valores observados.

Em seguida, serão descritos os modelos utilizados como *benchmark*, bem como as métricas utilizadas para testar a performance fora da amostra.

#### 4.5.1 *Benchmarks*

Conforme Pooter et al.(2010), estudos empíricos indicam que modelos parcimoniosos tendem a superar modelos mais sofisticados quando avaliados fora da amostra. Para isso, além dos modelos já descritos, serão incluídos como *benchmarks*, um modelo de passeio aleatório, e um modelo autorregressivo AR(1), ambos aplicados sobre os vértices da estrutura a termo. Em particular, o modelo de passeio aleatório será útil para compor algumas das métricas de desempenho.

##### Autorregressivo AR(1)

Segundo Pooter et al.(2010), embora os resultados indiquem que a hipótese nula para a raiz unitária das taxas de juros em nível não possa ser rejeitada, a hipótese da não-estacionariedade para os juros é difícil de interpretar sob um ponto de vista econômico. Ao menos em tese, as taxas não flutuam livremente já que sofrem influência de política monetária e fiscal, e além disso, tendem a reverter à média. Desta forma, o segundo caso consiste em um modelo autorregressivo de primeira ordem:

$$y_t(\tau) = c(\tau) + \phi(y_{t-1}(\tau)) + \varepsilon_t(\tau), \quad \varepsilon_t(\tau) \sim N(0, \sigma^2(\tau)) \quad (4.14)$$

para  $t = 1, \dots, N$ , em que  $\tau$  representa o prazo para a taxa de juros e  $c(\tau)$ ,  $\phi(\tau)$  e  $\sigma(\tau)$  são parâmetros escalares estimados pelo método dos mínimos quadrados ordinários. Na aplicação, as previsões  $h$ -passos a frente são construídas de forma iterativa.

##### Passeio Aleatório

O modelo mais simples consiste em um passeio aleatório sem *drift* aplicado sobre cada vértice da base de dados. Matematicamente, temos:

$$y_t(\tau) = y_{t-1}(\tau) + \varepsilon_t(\tau) \quad \varepsilon_t(\tau) \sim N(0, \sigma^2(\tau)) \quad (4.15)$$

para  $t = 1, \dots, N$ , em que  $\tau$  representa o prazo para a taxa de juros.

Neste modelo, portanto, a previsão  $h$ -passos a frente ( $\hat{y}_{t+h}(\tau)$ ) é simplesmente o último valor observado. Estudos tais como Pooter et al. (2010) e Caldeira et al.(2014) salientam a importância de utilizar tal modelo quando o objeto de estudo é a capacidade preditiva

da ETTJ, já que as taxas de juros são, em geral, não estacionárias. Adicionalmente, a literatura relata frequentemente a dificuldade em superar este *benchmark* consistentemente. (Ang e Piazzesi, 2003; Diebold e Li, 2006; Pooter et al., 2010; Caldeira et al., 2014).

### 4.5.2 Medidas de Desempenho

Para comparar a capacidade dos modelos em prever as taxas de juros foram aplicadas algumas medidas populares para avaliação de performance fora da amostra baseadas no erro de predição. Desta forma, foi utilizado método de análise semelhante aos aplicados em Pooter et al.(2010) e Caldeira et al.(2014). As métricas aplicadas serão descritas abaixo.

#### Raiz do Erro Quadrático Médio

Dado um conjunto de  $P$  previsões fora da amostra para um horizonte de  $h$  períodos à frente, calcula-se a Raiz do Erro Quadrático Médio para o prazo  $\tau$  e o modelo  $m$  da seguinte forma:

$$REQM_m(\tau) = \sqrt{\frac{1}{P} \sum_{t=1}^P (\hat{y}_{t+h|t,m}(\tau) - y_{t+h}(\tau))^2} \quad (4.16)$$

em que  $y_{t+h}(\tau)$  é a taxa para o prazo  $\tau$  observada no tempo  $t + h$  e  $\hat{y}_{t+h|t,m}(\tau)$  é a correspondente estimativa realizada no tempo  $t$ .

#### Traço da Raiz do Erro Quadrático Médio

O Traço da Raiz do Erro Quadrático Médio é uma medida que sintetiza os REQMs de todos os vértices em um só número. Desta forma, é calculado conforme fórmula abaixo.

$$TREQM_m(\tau) = \sqrt{\frac{1}{N} \frac{1}{P} \sum_{i=1}^N \sum_{t=1}^P (\hat{y}_{t+h|t,m}(\tau) - y_{t+h}(\tau))^2} \quad (4.17)$$

Conforme Caldeira et al.(2014), um ponto fraco das avaliações de performance utilizando o REQM e o TERQM é que estes indicadores condensam o desempenho em toda a amostra, impedindo que seja possível perceber em quais situações determinado modelo apresenta maiores ou menores erros de previsão. Desta forma, uma análise completar é a visualização gráfica do Erro Quadrático Acumulado e do Erro Quadrático Relativo Acumulado.

#### Erro Quadrático Acumulado

O Erro Quadrático Acumulado é descrito pela fórmula abaixo.



$$EQA_{m,T}(\tau) = \sum_{t=1}^T (\hat{y}_{t+h|t,m}(\tau) - y_{t+h}(\tau))^2 \quad (4.18)$$

Através da sua representação gráfica é possível analisar o desempenho das previsões ao longo de toda a janela de avaliação.

### Erro Quadrático Relativo Acumulado

Uma técnica bastante utilizada para avaliação de previsões, consiste em verificar o desempenho relativo a algum outro modelo ao longo do tempo.

$$EQRA_{m,T}(\tau) = \sum_{t=1}^T [(\hat{y}_{t+h|t,benchmark}(\tau) - y_{t+h}(\tau))^2 - (\hat{y}_{t+h|t,m}(\tau) - y_{t+h}(\tau))^2] \quad (4.19)$$

Desta forma, o Erro Quadrático Relativo Acumulado produz uma linha ascendente quando o modelo avaliado apresenta um desempenho favorável em relação a um *benchmark* e descendente quando pior.

# Capítulo 5

## Resultados

### 5.1 Dados Utilizados

#### 5.1.1 Taxas de Juros

Como base de dados para o trabalho, foram utilizadas as séries mensais de fechamento das taxas de juros para os prazos fixos de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60 e 72 meses derivadas dos contratos de swaps DI x Pré da BM&F Bovespa. Os dados foram extraídos brutos e então interpolados para as maturidades desejadas utilizando-se o método *flat-forward*.

O período amostral utilizado foi de janeiro de 2004 à dezembro de 2013, totalizando 10 anos, ou 120 dados para cada um dos vértices.

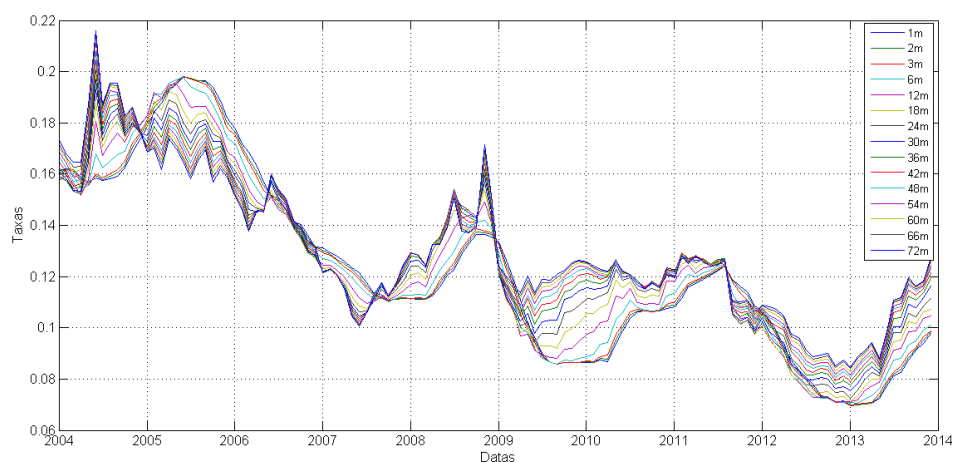


Figura 5.1: Séries temporais mensais dos vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses dos juros no Brasil no período de jan/2004 a dez/2013

Na tabela 5.1 encontram-se as estatísticas descritivas das séries utilizadas. Nela, é possível identificar alguns dos fatos estilizados frequentemente encontrados em estudos

envolvendo a estrutura a termo das taxas de juros. A curva média, é côncava e crescente, e as taxas de longo prazo são mais estáveis quando comparadas à extremidade curta da curva. Em geral, as taxas são altamente autocorrelacionadas e os vértices de curto prazo possuem maior persistência em relação aos mais longos (autocorrelação de primeira ordem). Além disso, a persistência diminui ao longo do tempo, como é possível inferir a partir da última coluna da tabela.

Prazo	Média	DesvPad	Curt.	Assim.	Máx.	Mín.	$\rho_1$	$\rho_{12}$
1	0.1229	0.0355	2.3250	0.5051	0.1978	0.0697	0.9950	0.6524
2	0.1228	0.0354	2.3357	0.5115	0.1979	0.0697	0.9950	0.6499
3	0.1228	0.0353	2.3465	0.5175	0.1981	0.0698	0.9947	0.6463
6	0.1230	0.0350	2.3295	0.5124	0.1979	0.0705	0.9932	0.6387
12	0.1239	0.0340	2.2691	0.4606	0.1952	0.0710	0.9883	0.6209
18	0.1250	0.0328	2.2214	0.4021	0.1923	0.0711	0.9837	0.6095
24	0.1262	0.0316	2.2190	0.3700	0.1928	0.0725	0.9788	0.6044
30	0.1274	0.0305	2.2572	0.3643	0.1967	0.0754	0.9742	0.6067
36	0.1283	0.0297	2.3304	0.3811	0.2003	0.0776	0.9699	0.6104
42	0.1290	0.0290	2.4213	0.4143	0.2042	0.0791	0.9656	0.6154
48	0.1295	0.0286	2.5249	0.4602	0.2075	0.0806	0.9623	0.6241
54	0.1298	0.0283	2.6562	0.5150	0.2112	0.0817	0.9594	0.6337
60	0.1303	0.0280	2.7837	0.5683	0.2136	0.0828	0.9580	0.6404
66	0.1306	0.0277	2.8798	0.6122	0.2148	0.0841	0.9565	0.6428
72	0.1308	0.0276	2.9699	0.6514	0.2159	0.0846	0.9547	0.6414

Tabela 5.1: Estatísticas descritivas das séries mensais das taxas de juros para os vértices estudados

### 5.1.2 Séries das Variáveis Macro

Neste trabalho, as variáveis macro serão incorporadas ao modelo para a ETTJ segundo a interpretação de Diebold e Rudebusch (2013), baseando-se nas evidências apresentadas de que o nível da curva é afetado pela inflação e a inclinação responde aos choques na atividade.

No grupo de inflação, foram escolhidas as séries mensais das variações anualizadas ( $\ln(x_t/x_{t-12})$ ) do Índice de Preços ao Consumidor Amplo (IPCA, IBGE), do Índice Nacional de Preços ao Consumidor (INPC, IBGE) e do Índice Geral de Preços Médio (IGPM, FGV). No grupo de atividade, foram usadas as variações anuais das séries de desocupação da pesquisa mensal de emprego (Desemprego, PME-IBGE), a relação População Economicamente Ativa/População em Idade Ativa da Pesquisa mensal de emprego (PEA/PIA, PME-IBGE), o Índice de Atividade Econômica do Banco Central (IBC-Br, Bacen) e a série da Pesquisa Industrial Mensal de Produção Física (PIM, IBGE).

As estatísticas descritivas das séries estão apresentadas na tabela 5.2.

Variável	Média	DesvPad	Curt.	Assim.	Máx.	Mín.	$\rho_1$	$\rho_{12}$
Desemprego	-0.0000	0.0085	3.0044	-0.2208	0.0193	-0.0226	0.7499	-0.5014
PEA/PIA	-0.0829	0.0867	2.3956	0.2637	0.1295	-0.2610	0.8811	-0.4636
IBC-Br	0.0368	0.0330	2.8972	-0.4882	0.1039	-0.0521	0.7438	-0.3539
PIM	0.0245	0.0672	4.1600	-0.6513	0.1736	-0.1861	0.8110	-0.4493
IPCA	0.0536	0.0115	2.4783	-0.2453	0.0776	0.0291	0.9642	0.0383
INPC	0.0530	0.0121	2.7924	-0.4510	0.0827	0.0256	0.9537	-0.0677
IGPM	0.0595	0.0367	2.3316	-0.0720	0.1407	-0.0173	0.9672	-0.4848

Tabela 5.2: Estatísticas descritivas das séries mensais macroeconômicas utilizadas

Para reduzir a dimensionalidade do sistema, foi empregada uma análise de componentes principais em cada um dos grupos e utilizado o primeiro componente da série de inflação e o primeiro componente da série de atividade. Inicialmente as séries foram normalizadas a fim de apresentarem média zero e variância unitária, em seguida foi aplicada a técnica.

A tabela 5.3 resume a aplicação para o grupo de Atividade. Os valores em cada linha representam os autovetores, ou seja, os pesos que cada variável possui em cada um dos componentes. Na última linha da tabela está o percentual da variância explicada (autovalor normalizado) por cada componente. Para a atividade, aproximadamente 55% do total das variações nas séries são explicados apenas pelo primeiro componente. Os pesos positivos para o IBC-Br, PIM e PEA/PIA e negativo para Desemprego na primeira coluna indicam que é possível interpretar o primeiro componente como um índice de atividade.

Variável	1º Componente	2º Componente	3º Componente	4º Componente
PEA/PIA	0.1366	0.8315	-0.5341	-0.0683
Desemprego	-0.3834	0.5445	0.7452	0.0355
IBC-Br	0.6500	0.0987	0.2281	0.7181
PIM	0.6417	0.0484	0.3278	-0.6917
% Var. Explicada	0.5495	0.2833	0.1534	0.0137

Tabela 5.3: Análise dos Componentes Principais para as variáveis macroeconômicas de atividade.

A análise de componentes principais para o grupo de preços está na tabela 5.4. A interpretação é análoga à anterior. Neste caso, o primeiro componente explica 80% do total da variação dos preços da economia. O peso positivo dados para todos os índices permite que a série resultante seja entendida como inflação.

As séries dos primeiros componentes dos grupos de inflação e atividade estão representados no gráfico abaixo.

Variável	1ª Componente	2ª Componente	3ª Componente
IPCA	0.5865	-0.5086	-0.6303
INPC	0.6110	-0.2331	0.7566
IGPM	0.5317	0.8289	-0.1740
% Var. Explicada	0.8019	0.1532	0.0448

Tabela 5.4: Análise dos Componentes Principais para as variáveis macroeconômicas de inflação.

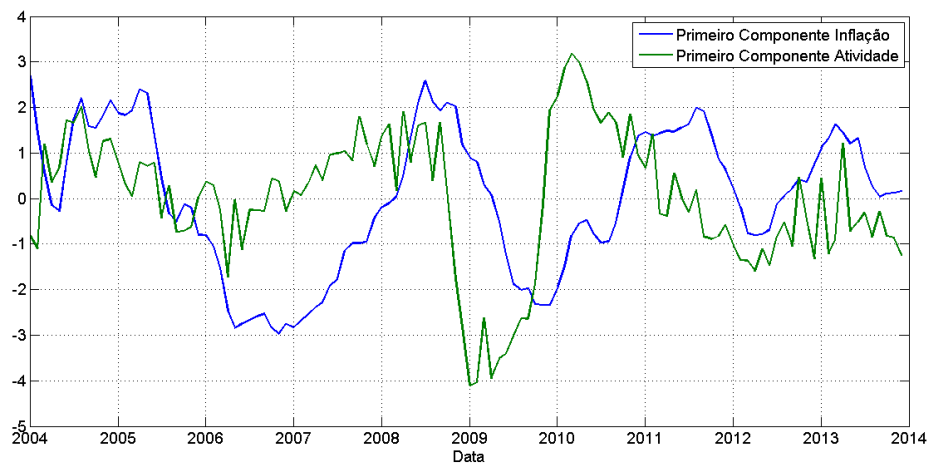


Figura 5.2: Séries temporais mensais dos primeiros componentes de Atividade e Inflação no período de jan/2004 a dez/2013

## 5.2 Resultados Dentro da Amostra

Para avaliar os modelos dentro da amostra, foram comparados os erros de estimação das taxas um passo a frente durante o processo de filtragem, conforme descrito na seção 4.4. Para isso, os estados estimados (nível, inclinação e curvatura) foram convertidos em taxas estimadas, e estas comparadas com as taxas observadas.

A tabela 5.5 contém um resumo dos Erros Quadráticos Médios apresentados pelos quatro modelos analisados (Two-step DNS, Two-step DNSX, EKF DNS e EKF DNSX) para os vértices de 1, 6, 12, 18, 24, 36 e 60 meses.

Percebe-se que para os dois pares de modelos analisados, a inclusão das variáveis de inflação e atividade implicou em uma redução nos Erros Quadráticos Médios ao longo da amostra. Os resultados indicam que a inclusão dessas variáveis foi capaz de adicionar informação ao filtro, gerando estimativas mais precisas um mês a frente nos dois casos.

Em relação às duas primeiras linhas da tabela, em que os parâmetros foram estimados com o método em dois estágios, a redução do erro foi mais expressiva para os vértices de menor maturidade, porém a melhora foi apenas marginal, sendo que nos mais longos

quase não houve melhora.

Para o Filtro de Kalman Estendido, observou-se que as estimativas ficaram relativamente mais precisas para a extremidade longa da curva, no entanto, os resultados foram expressivos para todos os prazos avaliados, com destaque para o último vértice da curva.

Outro fato importante pode ser observado comparando-se os resultados obtidos pelo two-step DNS e o EKF DNS. Apesar dos parâmetros para o método em dois passos terem sido realizadas com toda a mostra e os parâmetros do Filtro de Kalman Estendido serem estimados *on-line*, com a informação disponível até o instante imediatamente anterior, é nítida a superioridade do segundo. Este resultado significa que estimar os parâmetros a cada passo geram melhores estimativas, portanto, uma abordagem não-linear pode ser um tratamento mais adequado para o modelo dinâmico Nelson-Siegel.

As figuras 5.3 e 5.4 permitem uma análise visual dessas conclusões.

MODELO	1m	6m	12m	18m	24m	36m	60m
2-STEP DNS	0.4688	0.5340	0.6112	0.6511	0.6864	0.7304	0.8073
2-STEP DNSX	0.4007	0.4770	0.5563	0.6051	0.6511	0.7106	0.7974
EKF DNS	0.2031	0.2033	0.1976	0.1569	0.1539	0.1721	0.1308
EKF DNSX	0.1466	0.1600	0.1665	0.1104	0.0966	0.1104	0.0675

Tabela 5.5: Raiz do Erro Quadrático Médio para as previsões dentro da amostra para 1 mês a frente para os modelos analisados.

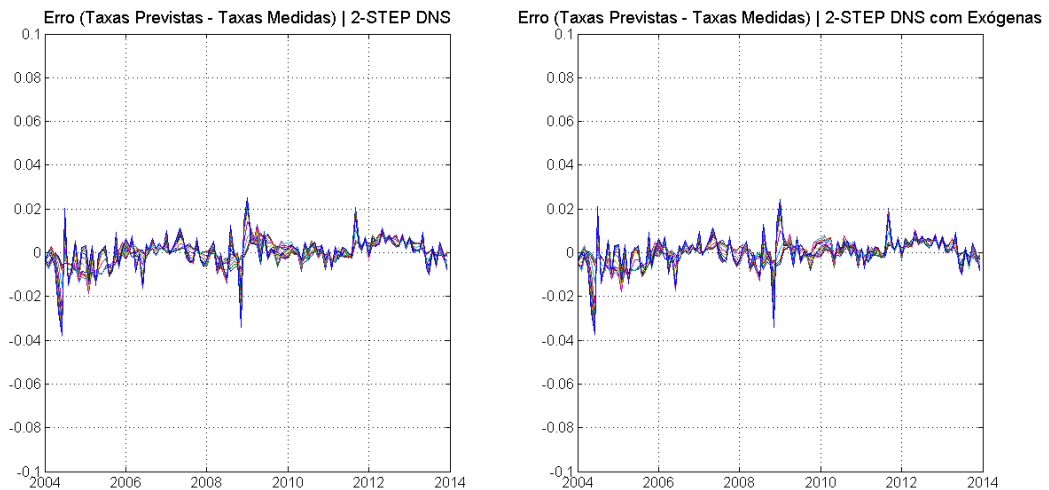


Figura 5.3: Série dos erros de estimação dentro da amostra para os vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013, com parâmetros estimados pelo método 2-STEP DNS

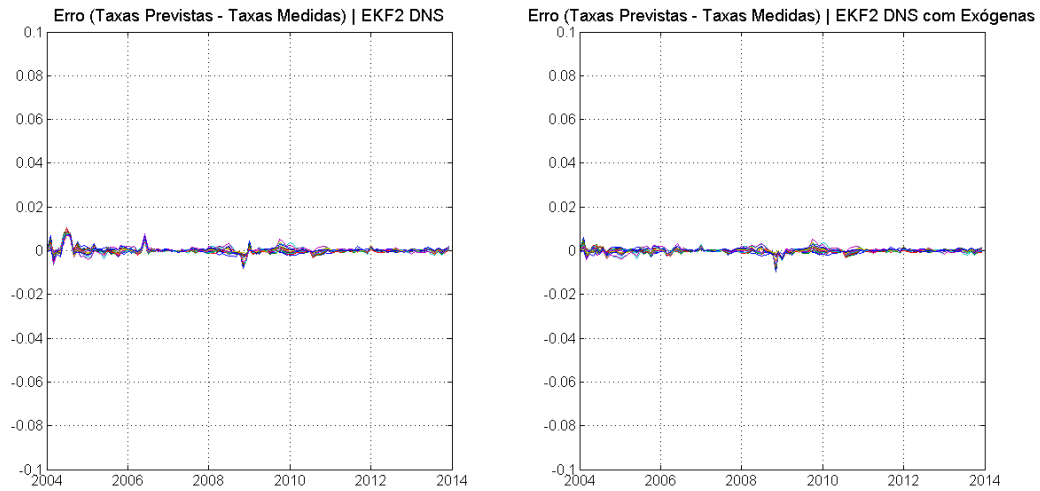


Figura 5.4: Série dos erros de previsão do Filtro de Kalman dentro da amostra para os vértices de 1, 2, 3, 6, 12, 18, 24, 30, 36, 42, 48, 54, 60, 66 e 72 meses, janela de estimação de jan/04 a dez/2013, com parâmetros estimados usando Filtro de Kalman Estendido com  $\theta_t = \theta_{t-1} + \eta$

### 5.3 Resultados Fora da Amostra

As tabela 5.6 resume os erros de estimação para 3 meses à frente para alguns vértices selecionados, para os modelos estudados e os *benchmarks* propostos. As figuras 5.5, 5.6 exibem os erros cumulativos absoluto e em relação aos erros do passeio aleatório.

Para as previsões três meses à frente, percebe-se que, de forma geral, os modelos que incluíam variáveis macro produziram resultados piores do que os modelos sem exógenas, à exceção das taxas com maturidades mais longas no caso em que os parâmetros foram estimados com o Filtro de Kalman Estendido.

Além disso, todas as quatro variantes do modelo Nelson-Siegel proporcionaram estimativas ruins quando comparadas ao modelo autorregressivo aplicado nos vértices e ao passeio aleatório, que simplesmente supõe que as taxas sigam constantes.

MODELO	TREQM	1m	6m	12m	18m	24m	36m	60m
2-STEP DNS	1.4878	1.5587	1.5450	1.5904	1.6127	1.5826	1.4791	1.3649
2-STEP DNSX	1.5745	1.5282	1.5602	1.6455	1.6960	1.6852	1.6037	1.5154
EKF DNS	1.5855	1.3372	1.3998	1.5998	1.7411	1.7806	1.7285	1.5874
EKF DNSX	1.6485	1.8216	1.5599	1.5670	1.6554	1.6955	1.6812	1.5867
AR(1)	0.9859	0.8484	0.8947	0.9594	0.9963	1.0222	1.0441	1.0350
RW	0.9581	0.8581	0.9051	0.9787	1.0162	1.0321	1.0197	0.9526

Tabela 5.6: Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 3 meses à frente fora da amostra para vértices selecionados

Os gráficos de erro absoluto e erro relativo acumulados indicam que os modelos não produziram boas previsões para a janela, perdendo constantemente dos modelos referência.

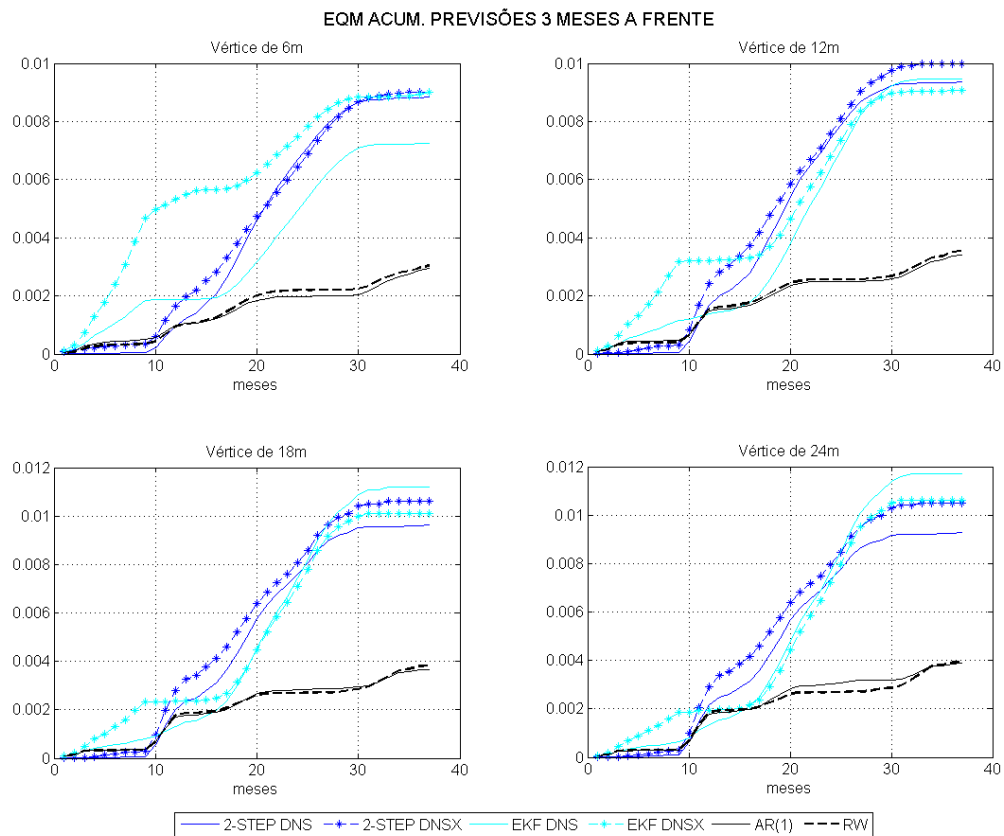


Figura 5.5: Erro quadrático médio acumulado para previsões 3 meses à frente para os vértices de 6, 12, 18 e 24 meses



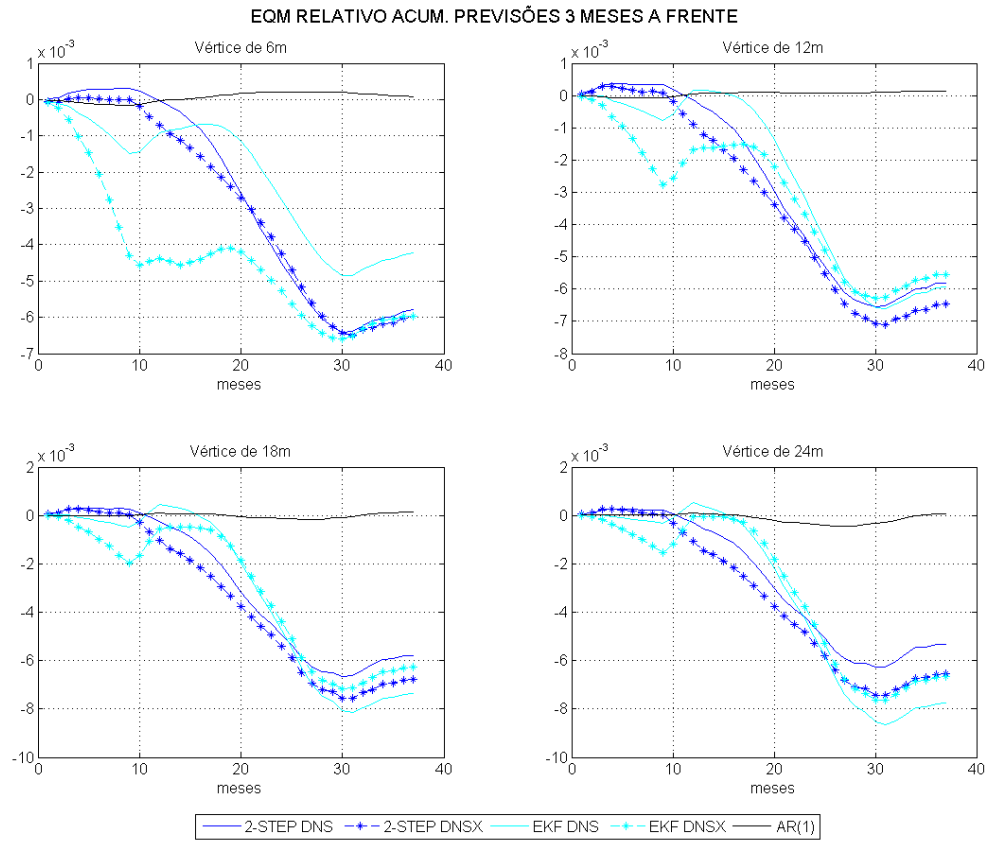


Figura 5.6: Erro quadrático relativo acumulado para previsões 3 meses a frente para os vértices de 6, 12, 18 e 24 meses

Para as medidas de desempenho das estimativas 6 meses à frente, mostradas na tabela 5.7, os resultados são similares aos anteriores. De forma geral, adicionar variáveis macro não produziu bons resultados para o método em dois passos. Para os estimados com EKF, os prazos mais longos foram previstos com maior precisão pelo modelo com variáveis exógenas. Percebe-se, no entanto, que nenhum dos modelos superou as previsões do passeio aleatório. Destaca-se também o melhor desempenho dos modelos estimados com EKF em relação aos estimados em dois passos.

MODELO	TREQM	1m	6m	12m	18m	24m	36m	60m
2-STEP DNS	2.5075	2.6673	2.6728	2.7135	2.7141	2.6477	2.4756	2.2568
2-STEP DNSX	2.6085	2.5316	2.6306	2.7539	2.8111	2.7849	2.6585	2.4879
EKF DNS	2.1119	1.7497	1.8542	2.1265	2.3165	2.3716	2.3124	2.1242
EKF DNSX	2.0186	2.1279	1.8454	1.9144	2.0586	2.1197	2.1053	1.9730
AR(1)	1.6812	1.5686	1.6045	1.6456	1.6826	1.7054	1.7411	1.7209
RW	1.6704	1.6238	1.6633	1.7247	1.7617	1.7617	1.7305	1.6092

Tabela 5.7: Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 6 meses à frente fora da amostra para vértices selecionados

As figuras 5.7 e 5.8 mostram os erros acumulados, onde ficam evidentes as melhores previsões dos métodos cujos parâmetros foram estimados com o EKF. Percebe-se também que não é possível concluir que exista qualquer benefício em incorporar as variáveis de inflação e atividade aos modelos tradicionais.

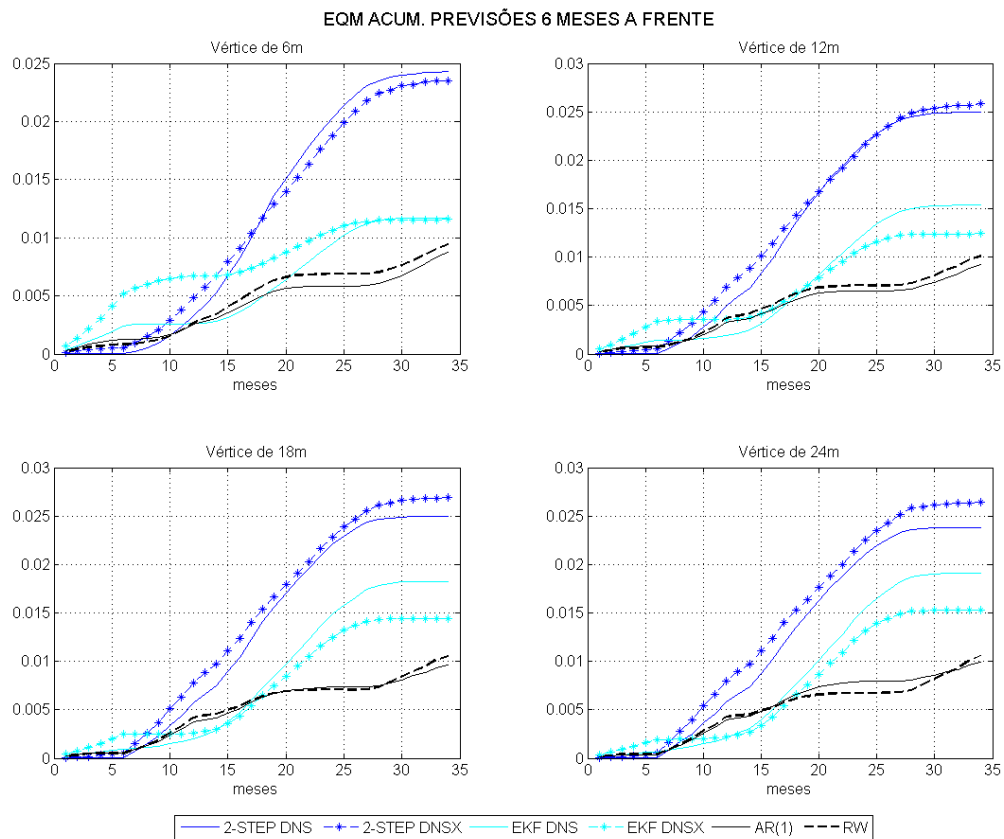


Figura 5.7: Erro quadrático médio acumulado para previsões 6 meses à frente para os vértices de 6, 12, 18 e 24 meses

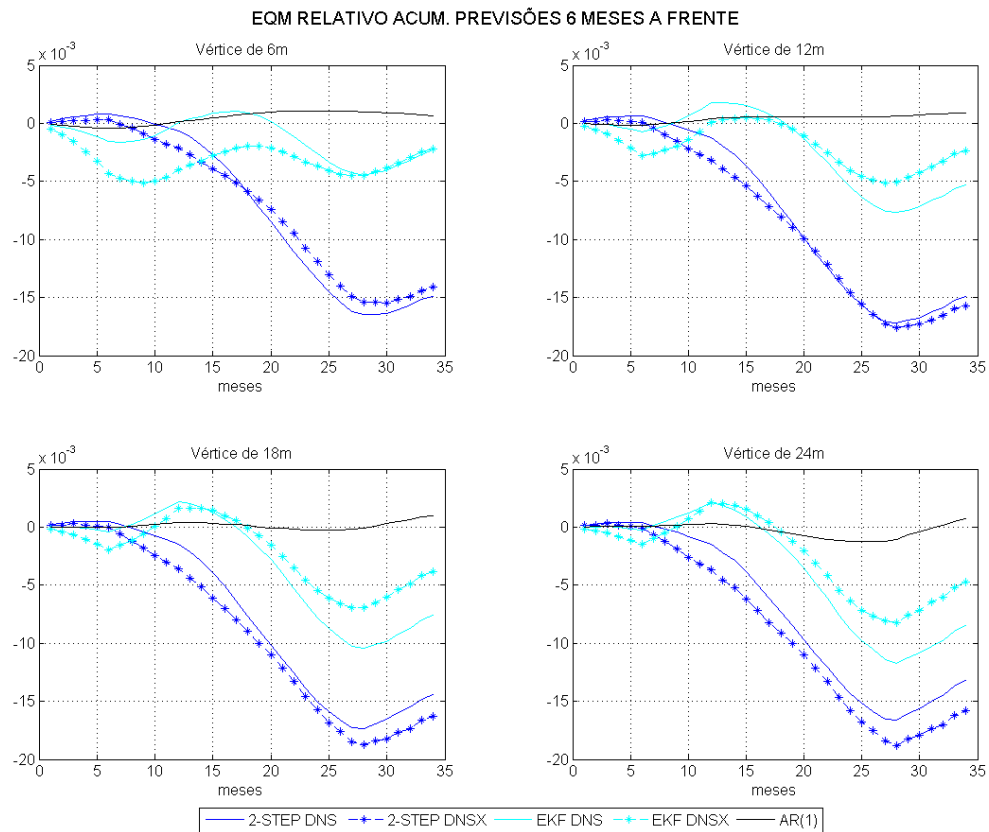


Figura 5.8: Erro quadrático relativo acumulado para previsões 6 meses à frente para os vértices de 6, 12, 18 e 24 meses

Na tabela 5.8 está o resumo dos resultados para as previsões 9 meses à frente. Mais uma vez, fica evidente os melhores resultados dos modelos estimados com o Filtro de Kalman Estendido. Para este horizonte de previsão, no entanto, pode-se verificar que as estimativas do modelo EKF DNSX, que incorpora exógenas, fica marginalmente melhor do que seu par.

MODELO	TREQM	1m	6m	12m	18m	24m	36m	60m
2-STEP DNS	3.2765	3.4909	3.5216	3.5660	3.5548	3.4639	3.2316	2.9260
2-STEP DNSX	3.3559	3.2348	3.4007	3.5609	3.6289	3.5949	3.4283	3.1897
EKF2 DNS	2.3082	1.7871	1.9653	2.3137	2.5440	2.6150	2.5603	2.3572
EKF2 DNSX	2.1026	2.0393	1.8055	1.9726	2.1742	2.2589	2.2586	2.1192
AR(1)	2.2257	2.0748	2.1268	2.1951	2.2419	2.2676	2.3098	2.2644
RW	2.2139	2.2356	2.2822	2.3628	2.3902	2.3599	2.2672	2.0417

Tabela 5.8: Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 9 meses à frente fora da amostra para vértices selecionados

Nas figuras 5.9 e 5.10 plota-se erros acumulados para as previsões de 9 meses. Uma

observação importante está no gráfico de erros relativos, onde percebe-se a clara deterioração das estimativas a partir de 12 meses do início da janela para os vértices de 12, 18 e 24 meses para os modelos com estimação por EKF. Esta deterioração poderia estar relacionada a uma mudança em algum dos parâmetros fixados.

Também nos gráficos, é possível ver que as diferenças entre os desempenhos dos modelos com e sem exógenas são relativamente pequenas para esse horizonte de análise.

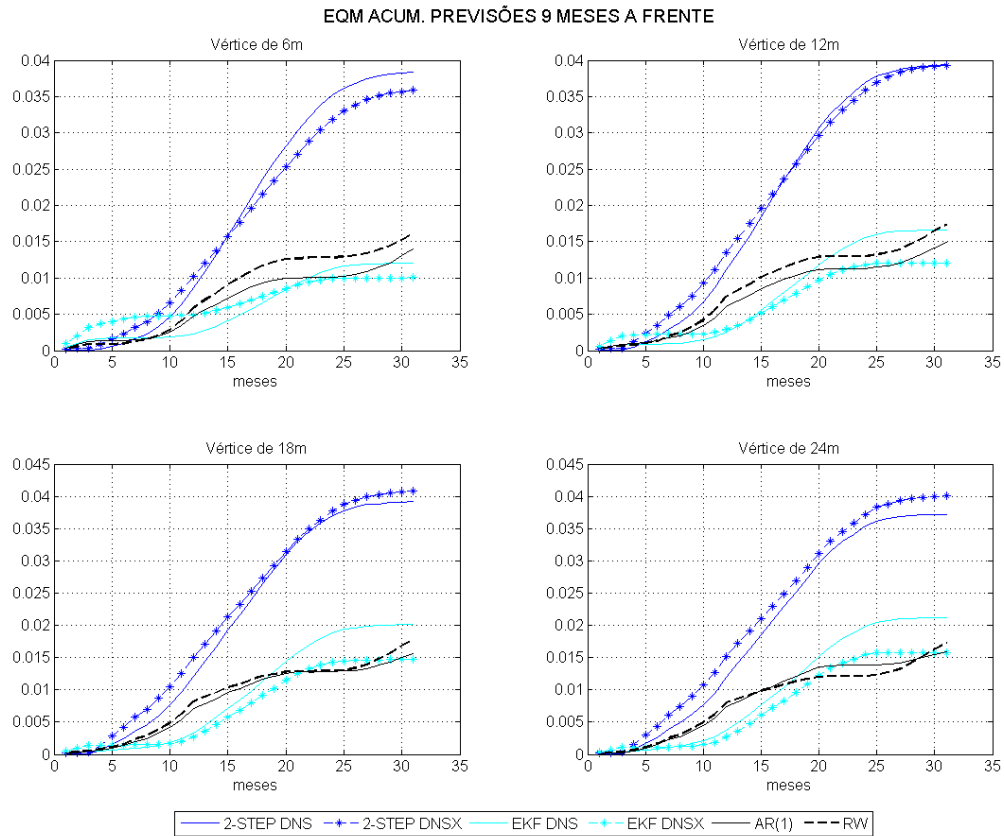


Figura 5.9: Erro quadrático acumulado para previsões 9 meses à frente para os vértices de 6, 12, 18 e 24 meses

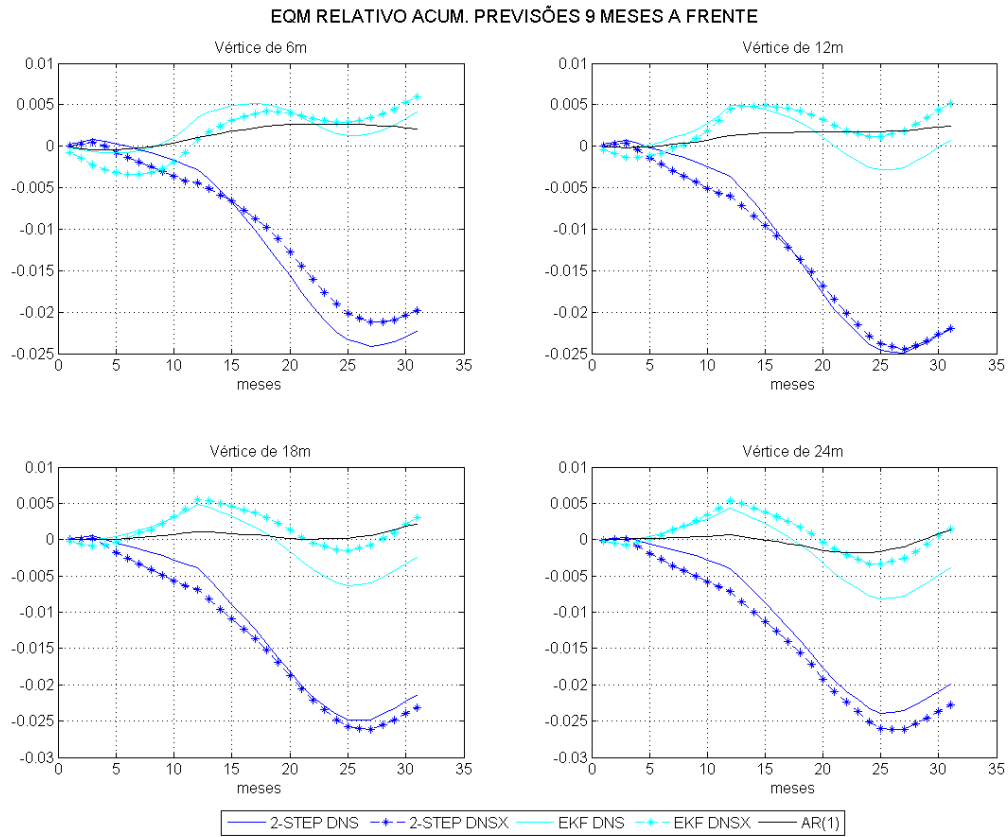


Figura 5.10: Erro quadrático relativo acumulado para previsões 9 meses à frente para os vértices de 6, 12, 18 e 24 meses

Finalmente, a tabela 5.9 sintetiza a análise das estimativas das taxas de juros para 12 meses à frente. Para esse prazo, é possível perceber que o modelo EKF DNSX, apresentou performance superior a do modelo sem exógenas para todas as maturidades com exceção do vértice de 6 meses.

Assim como para o horizonte de 9 meses, os gráficos das figuras 5.11 e 5.12 mostram que as estimativas sofrem deterioração em relação ao passeio aleatório após 12 meses de janela.

Além disso, mais uma vez as estimativas produzidas pelos modelos calibrados com o Filtro de Kalman Estendido mostram-se superiores quando comparados às previsões dos modelos de dois passos.

MODELO	TREQM	1m	6m	12m	18m	24m	36m	60m
2-STEP DNS	3.8671	4.1137	4.1731	4.2241	4.2048	4.0943	3.8157	3.4403
2-STEP DNSX	3.9126	3.7593	3.9810	4.1699	4.2448	4.2014	4.0019	3.7057
EKF DNS	2.4076	1.6380	1.9533	2.4004	2.6767	2.7677	2.7223	2.5139
EKF DNSX	2.0984	1.7120	1.5943	1.9195	2.2029	2.3259	2.3533	2.2241
AR(1)	2.6505	2.3722	2.4476	2.5687	2.6663	2.7300	2.8070	2.7338
RW	2.6352	2.7084	2.7518	2.8421	2.8686	2.8224	2.6967	2.3748

Tabela 5.9: Traço da Raiz do Erro Quadrático Médio e Raiz do Erro Quadrático Médio dos modelos estudados para estimativas 12 meses à frente fora da amostra para vértices selecionados

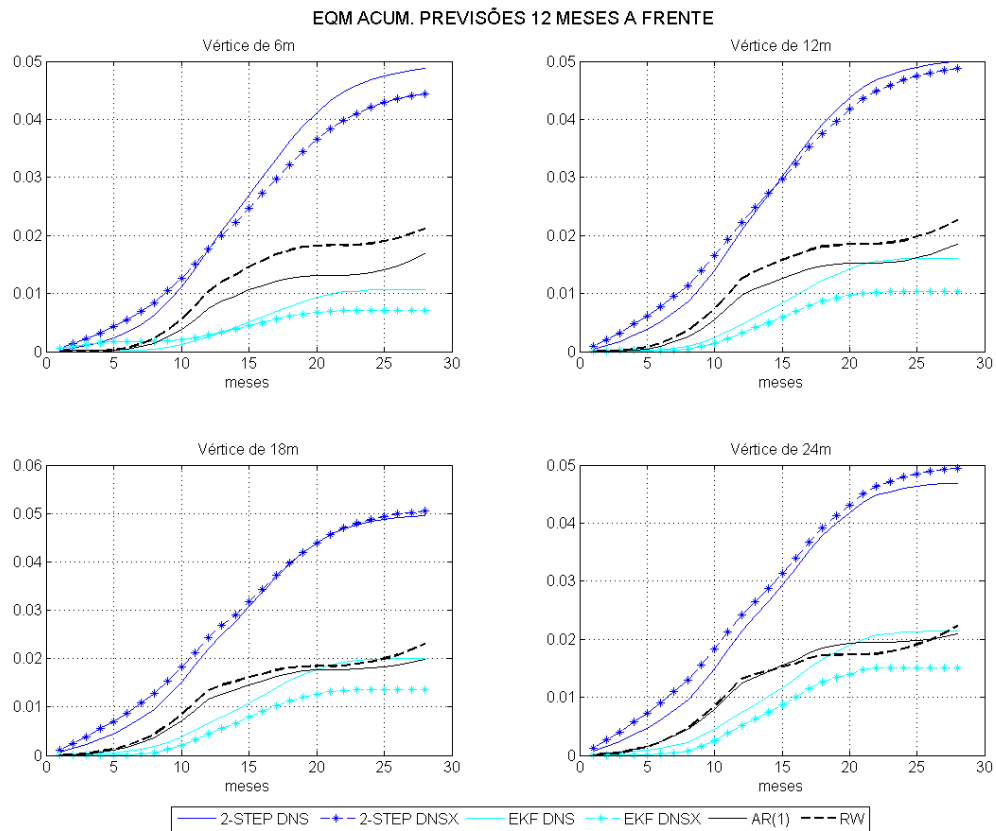


Figura 5.11: Erro quadrático acumulado para previsões 12 meses à frente para os vértices de 6, 12, 18 e 24 meses

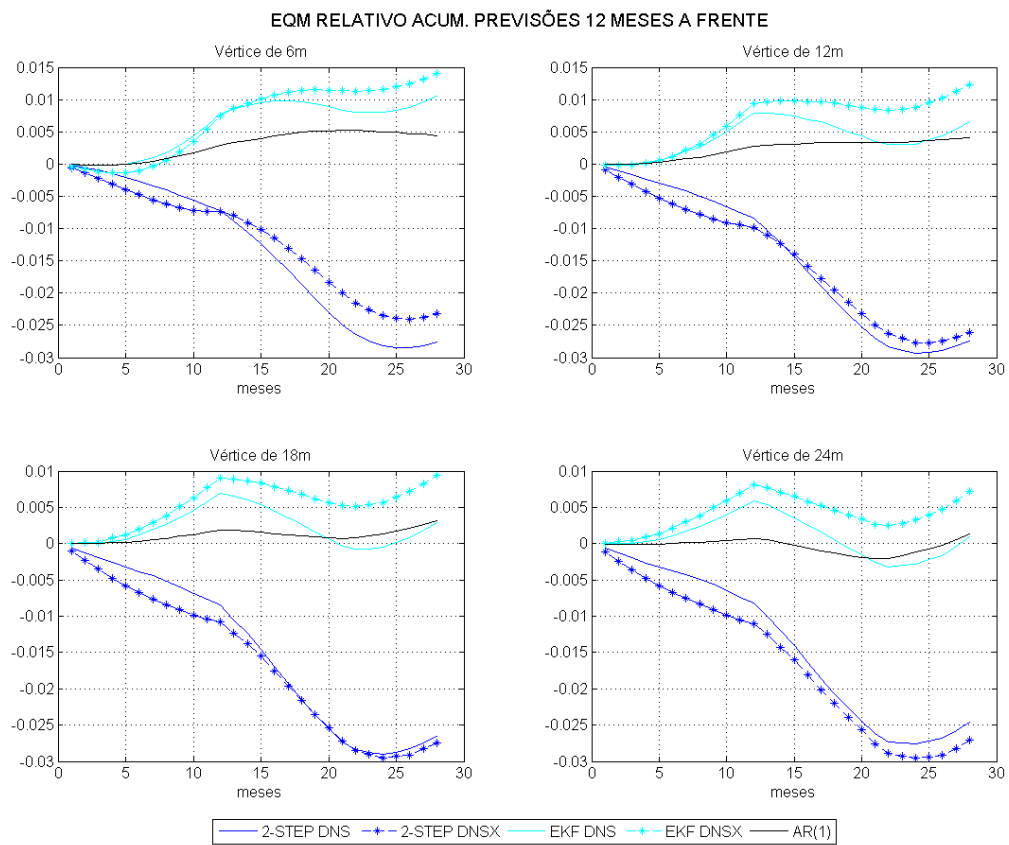


Figura 5.12: Erro quadrático relativo acumulado para previsões 12 meses à frente para os vértices de 6, 12, 18 e 24 meses

# Capítulo 6

## Conclusões

### 6.1 Conclusões

Este trabalho teve como objetivo a comparação da versão dinâmica do modelo Nelson-Siegel para a estrutura a termo das taxas de juros, com uma variante em que são introduzidas variáveis exógenas de inflação e atividade na equação de transição do modelo em espaço de estados.

Em conjunto aos modelos, foram testados dois métodos com diferentes propriedades para a estimação dos parâmetros: o tradicional método em dois passos, conforme introduzido por Diebold e Li (2006) e o Filtro de Kalman Estendido, baseado nos estudos de Ljung e Söderström (1987), no qual estimam-se os parâmetros recursivamente, a cada vez que um novo dado torna-se disponível. Apesar de comum na literatura de engenharia, há poucos trabalhos em finanças que utilizam o FKE para este fim.

Verificou-se, para a janela de dados estudada, que o processo de filtragem quando os parâmetros são estimados iterativamente (FKE) gerou estimativas mais precisas um passo à frente para as taxas de juros do que quando os parâmetros foram estimados pelo método de dois passos e fixados, mesmo dentro da amostra. Também foi encontrado que as estimativas dos modelos com variáveis exógenas foram marginalmente superiores dentro da amostra, sugerindo que a inclusão de variáveis macroeconômicas pode ser capaz de adicionar informação relevante ao filtro, gerando portanto, um maior poder preditivo.

Fora da amostra, contudo, a comparação entre os modelos com e sem a inclusão de fatores exógenos mostrou-se pouco conclusiva, com os pares comportando-se de maneira semelhante. Por outro lado, a utilização do FKE para estimação dos parâmetros mostrou-se consistentemente superior ao método em dois passos para praticamente todos os horizontes de tempo estudados.

Não parece possível, portanto, concluir que os modelos testados com variáveis macro sejam realmente mais precisos do que os modelos na forma tradicional, já que os benefícios da inclusão dos índices de inflação e atividade como variáveis exógenas foram apenas



marginais. No entanto, tratar a dinâmica do modelo Nelson-Siegel como um problema não-linear, em que seja permitido aos parâmetros algum grau de variabilidade, produziu bons resultados dentro e fora da amostra, mesmo utilizando um método relativamente simples como o Filtro de Kalman Estendido.

Uma série de extensões são possíveis a este estudo. Em relação às interações e dinâmicas conjuntas entre as taxas de juros e variáveis macroeconômicas, uma possibilidade seria permitir uma maior complexidade entre as correlações das variáveis e parâmetros, que aqui foram tratadas como independentes. Outro estudo relevante seria trabalhar com a imposição de restrições de não-arbitragem ao modelo. Além disso, sugere-se a utilização de técnicas para aumento da frequência dos dados para que seja possível trabalhar com janelas diferentes de estimação e previsão. Finalmente, a fim de obter melhores estimativas na filtragem de sistemas não-lineares, a utilização de métodos mais precisos, como o Filtro de Kalman *Unscented* ou ainda o PEM (*Prediction Error Method*) também é recomendada.

# Referências Bibliográficas

- ANG, A.; PIAZZESI, M. A no-arbitrage vector autoregression of term structure dynamics with macroeconomics and latent variables. *Journal of Monetary Economics*, v. 50, p. 745–787, 2003.
- CALDEIRA, J. F.; MOURA, G. V.; SANTOS, A. A. P. Bond portfolio optimization: a dynamic heteroskedastic factor model approach. *Working Paper, Department of Economics, Universidade Federal de Santa Catarina*, v. 146, 2012.
- CALDEIRA, J. F.; LAURINI, M. P.; PORTUGAL, M. S. Inferência bayesiana aplicada ao modelo dinâmico de Nelson-Siegel com volatilidade estocástica nos fatores. *Revista Brasileira de Econometria*, v. 30, p. 123–161, 2010.
- CALDEIRA, J.; MOURA, G.; SANTOS, A. A. P. Predicting the yield curve using forecasts combinations. *Computational Statistics and Data Analysis*, 2014.
- CHRISTENSEN, J. H.; DIEBOLD, F. X.; RUDEBUSCH, G. D. The affine arbitrage free class of Nelson-Siegel term structure models. *Journal of Econometrics*, v. 164, p. 4–20, 2011.
- DIEBOLD, F. X.; LI, C. Forecasting the term structure of government bond yields. *Journal of Econometrics*, v. 130, p. 337–364, 2006.
- DIEBOLD, F. X.; LI, C.; JI, L. A three-factor yield curve model: Non-affine structure, systematic risk sources, and generalized duration. *Long-Run Growth and Short-Run Stabilization: Essays in Memory of Albert Ando*, v. Edward Elgar, p. 240–274, 2006.
- DIEBOLD, F. X.; LI, C.; YUE, V. Global yield curve dynamics and interactions: A dynamic nelson-siegel approach. *Journal of Econometrics*, v. 146, p. 351–363, 2008.
- DIEBOLD, F. X.; PIAZZESI, M.; RUDEBUSCH, G. D. Modelling bond yields in finance and macroeconomics. *American Economic Review*, v. 95, p. 415–420, January 2005.
- DIEBOLD, F. X.; RUDEBUSCH, G. D. *Yield curve modeling and forecasting: The dynamic Nelson-Siegel approach*. 1. ed. Princeton University Press, 2013.
- DIEBOLD, F. X.; RUDEBUSCH, G. D.; ARUOBA, S. B. The macroeconomy and the yield curve: A dynamic latent factor approach. *Journal of Econometrics*, v. 131, p. 309–338, 2006.
- DURBIN, J.; KOOPMAN, S. J. *Time series analysis by state space methods*. 2nd. ed. Oxford University Press, 2012.

- ESTRELLA, A.; HARDOUVELIS, G. The term structure as a predictor of real economic activity. *Journal of Finance*, v. 46, p. 555–76, 1991.
- ESTRELLA, A.; MISHKIN, F. S. Predicting U.S. recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, v. 80, p. 45–61, 1998.
- FULOP, A. *Filtering methods*. Springer, 2011. v. 1, p. 439–467.
- GOODWIN, C. G.; SIN, K. S. *Adaptative filtering prediction and control*. 1st. ed. Dover Publications, 2009.
- GREWAL, M.; ANDREWS, A. P. *Kalman filtering: Theory and practice using MATLAB*. 2nd. ed. John Wiley and Sons, 2001.
- KOOPMAN, S. J.; MALLEE, M. I. P.; VAN DER WEL, M. Analyzing the term structure of interest rates using the dynamic Nelson-Siegel model with time-varying parameters. *Journal of Business and Economic Statistics*, v. Vol. 28, p. 329–343, 2010.
- KURMANN, A.; OTROK, C. News shocks and the slope of the term structure of interest. *Working Paper, FEDERAL RESERVE BANK OF ST. LOUIS*, 2012.
- LAURINI, M. P. Macroeconometria e a estrutura a termo de taxas de juros - problemas (ainda) em aberto. *Economia & Tecnologia*, v. 25, Abril/Junho 2011.
- LI, C.; NIU, L.; ZENG, G. A generalized arbitrage-free nelson-siegel term structure model with macroeconomic fundamentals. March 2011.
- LITTERMAN, R.; SCHEINKMAN, J. Common factors affecting bond returns. *Journal of Fixed Income*, v. 1, p. 54–61, 1991.
- LJUNG, L.; SÖDERSTRÖM, T. *Theory and practice of recursive identification*. 1st. ed. MIT Press, 1987.
- LUTKEPOHL, H. New introduction to multiple time series analysis. 2005.
- MARTIN, V.; HURN, S.; HARRIS, D. *Econometric modelling with time series: Specification, estimation and testing*. 1. ed. Cambridge University Press, 2013.
- NELSON, C.; SIEGEL, A. Parsimonious modeling of the yield curve. *The Journal of Business*, v. 60, p. 473–489, October 1987.
- PIAZZESI, M. *Chapter 12: Affine term structure models*. Elsevier, 2010. v. 1, p. 691–766.
- POOTER, M.; RAVAZZOLO, F.; VAN DIJK, D. Term structure forecasting using macro factors and forecast combination. *International Finance Discussion Papers*, v. 993, January 2010.
- RACHEV, S. T.; MITTNIK, S.; FABOZZI, F. J.; FOCARDI, S. M.; JASIC, T. *Financial econometrics: From basics to advanced modeling techniques*. 1. ed. Princeton University Press, 2007.
- RUDEBUSCH, G.; WU, T. A macro-finance model of the term structure, monetary policy, and the economy. *Economic Journal*, v. 118, p. 906–926, 2008.

SÄRKKÄ, S.; , HARTIKAINEN, J.; SOLIN, A. EKF/UKF toolbox for Matlab v1.3. 2011.

SVENSSON, L. E. O. Estimating forward interest rates with the extended Nelson and Siegel method. *Sveriges Riksbank Quarterly Review*, p. 13–26, 1995.