

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
FACULDADE DE CIÊNCIAS ECONÔMICAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

FELIPE NASCIMENTO RUPERTI

UMA ABORDAGEM EM MACROFINANÇAS  
PARA A PREVISÃO DOS MOVIMENTOS  
DA TAXA DE CÂMBIO DO BRASIL

Porto Alegre  
2018

**FELIPE NASCIMENTO RUPERTI**

**UMA ABORDAGEM EM MACROFINANÇAS  
PARA A PREVISÃO DOS MOVIMENTOS  
DA TAXA DE CÂMBIO DO BRASIL**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada.

Orientador: Prof. Dr. João Frois Caldeira

Porto Alegre

2018

### CIP - Catalogação na Publicação

Nascimento Rupert, Felipe  
Uma Abordagem em Macrofinanças Para a Previsão dos  
Movimentos da Taxa de Câmbio do Brasil / Felipe  
Nascimento Rupert. -- 2018.  
62 f.  
Orientador: João Frois Caldeira.

Dissertação (Mestrado) -- Universidade Federal do  
Rio Grande do Sul, Faculdade de Ciências Econômicas,  
Programa de Pós-Graduação em Economia, Porto Alegre,  
BR-RS, 2018.

1. Taxa de Câmbio. 2. Previsão. 3. Nelson-Siegel.  
4. Fator de Esquecimento. 5. TVP-VAR. I. Caldeira,  
João Frois, orient. II. Título.

## **AGRADECIMENTOS**

Sou extremamente grato aos meus pais, Norma e Nerbe, pela minha formação ética e por todo o apoio e incentivo emocional, intelectual e financeiro ao longo da minha trajetória.

Ao professor João Frois Caldeira pela elucidação de dúvidas e, ainda mais, pelas boas ideias para o desenvolvimento deste trabalho.

Aos grandes amigos que fiz em Porto Alegre, José, Leandro, Luiza e Ícaro, cujo apoio, incentivo, troca de ideias e de experiências foram bastante importantes durante essa etapa.

Ao Erick, Allan e Nara que me acolheram durante períodos de mudanças.

Ao CNPq, pela contribuição financeira.

**FELIPE NASCIMENTO RUPERTI**

**UMA ABORDAGEM EM MACROFINANÇAS  
PARA A PREVISÃO DOS MOVIMENTOS  
DA TAXA DE CÂMBIO DO BRASIL**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título de Mestre em Economia, com ênfase em Economia Aplicada.

Aprovada em Porto Alegre, no dia 4 de dezembro de 2018.

BANCA EXAMINADORA:

---

**Profa. Dra. Andreza Aparecida Palma**  
UFSCar

---

**Dr. Leonardo Riegel Sant'Anna**  
UFRGS

---

**Prof. Dr. Marcelo Savino Portugal**  
UFRGS

## RESUMO

Este trabalho de pesquisa, baseado em Chen e Tsang (2011), considera a taxa de câmbio como uma variável macroeconômica, responsável por equilibrar os preços relativos de bens e serviços no mercado internacional e, também, como uma variável financeira, incorporando em seu preço expectativas e riscos. Levando isso em conta, um dos principais objetivos é analisar se a abordagem conjunta em macrofinanças pode melhorar a previsão dos movimentos da taxa de câmbio para o caso do Brasil e, ao mesmo tempo, contribuir para a literatura, ao permitir que os coeficientes dos modelos analisados variem no tempo, por meio da especificação econométrica TVP-VAR não estrutural. Com o objetivo de lidar com a "maldição da dimensionalidade", associada à estimação de modelos TVP-VAR e, paralelamente, de modo a evitar a escolha de somente um modelo ao longo do tempo, fatores de esquecimento, juntamente com os algoritmos de Seleção Dinâmica de Modelos (DMS) e de Combinação Dinâmica de Modelos (DMA) são implementados. Diferentes especificações para a estrutura TVP-VAR são comparadas de acordo com os fundamentos para a taxa de câmbio: a *macro-only*, com as variáveis macro selecionadas a partir do diferencial entre regras de Taylor; a *yields-only*, com os fatores latentes de Nelson-Siegel, que são utilizados como *proxies* de forma a refletir as expectativas e riscos incorporados na curva de juros relativa; e, por fim, a *macro-yields*, com os dois tipos de fundamentos. O período considerado se inicia em janeiro de 2015 e termina em junho de 2017. O resultado principal apresenta evidências de que a abordagem conjunta em macrofinanças em uma estrutura TVP-VAR melhora a previsão da taxa de câmbio se comparada com o passeio aleatório, principal *benchmark*. Em geral, o modelo TVP-VAR-DMA e TVP-VAR-DMS, ambos com a condição de passeio aleatório para o comportamento da matriz de coeficientes, apresentaram as melhores performances dentre os modelos considerados.

**Palavras-chave:** 1. Taxa de Câmbio. 2. Previsão. 3. Nelson-Siegel. 4. Fator de Esquecimento. 5. TVP-VAR. 6. DMA. 7. DMS.

## ABSTRACT

This research, based on Chen and Tsang (2011), considers the exchange rate both as a macroeconomic variable, responsible to equilibrate the relative prices of goods and services in the international markets, and as a financial variable, incorporating in its price expectations and risks. Taking this into account, one of the main objectives is to analyze whether the joint macro-finance approach can improve the forecasting of the movements of the Brazilian real. Besides, we contribute to the previous literature by allowing the coefficients to evolve over time using a non-structural Time-Varying Parameter Vector Autoregressive Model (TVP-VAR). In order to overcome the "curse of dimensionality" and to deal with the uncertainty in choosing the best forecasting model, forgetting factors, along with the Dynamic Model Selection (DMS) and the Dynamic Model Averaging (DMA) algorithms are implemented. Different specifications for the TVP-VAR structure are compared according to exchange rate fundamentals: the macro-only, with macroeconomic variables derived from the Taylor-Rule differential between Brazil and the United States; the yields-only, with Nelson-Siegel latent factors, used as proxies to reflect expectations and risks extracted from the relative yield curve; and, lastly, the macro-yields, with both macroeconomic and finance variables. The period covered goes from January 2005 to June 2017. The main result indicates that the macro-finance approach when combined with time varying parameters in a TVP-VAR structure improved exchange rate forecasting, especially when compared to the random walk. Overall, the TVP-VAR-DMA and TVP-VAR-DMS models, along with the simulated random walk condition for the coefficient matrix, had the best performances among the models considered.

**Keywords:** 1. Exchange Rate. 2. Forecasting. 3. Nelson-Siegel. 4. Forgetting Factor. 5. TVP-VAR. 6. DMA. 7. DMS.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Volume médio diário em dólar transacionado entre 1998 e 2016 no mercado global de câmbio - Fonte: BIS . . . . .	11
Figura 2 – Inter-relação entre a Taxa de Câmbio, a Regra de Taylor, a UIP e a Curva e Juros. . . . .	18
Figura 3 – Fatores de Encolhimento $\gamma$ e de Esquecimento $\lambda_{ff}$ dos modelos TVP-VAR. . . . .	44
Figura 4 – Algoritmo DMA/DMS: probabilidades de dos modelos TVP-VAR a cada período. . . . .	45
Figura 5 – ETTJ do Brasil e dos Estados Unidos - Variação em 12 meses (%). . .	50
Figura 6 – ETTJ Relativa - Variação em 12 meses (%). . . . .	51
Figura 7 – Fatores Latentes de Nelson-Siegel a partir da ETTJ Relativa . . . . .	51



## LISTA DE TABELAS

Tabela 1 – MAFE - Relativo ao modelo TVP-VAR-DMA . . . . .	46
Tabela 2 – Somatório do log das probabilidades preditivas em comparação ao modelo TVP-VAR-DMA . . . . .	47
Tabela 3 – Teste de Dickey-Fuller Aumentado . . . . .	52
Tabela 4 – Testes de Raiz Unitária - KPSS e Phillips-Perron . . . . .	53
Tabela 5 – Estatísticas Descritivas . . . . .	53
Tabela 6 – VAR - <i>Akaike Information Criterion</i> (AIC) . . . . .	54
Tabela 7 – MAFE - Valor Total . . . . .	55
Tabela 8 – Somatório do log das probabilidades preditivas - Valor Total . . . . .	56

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>10</b>
<b>1.1</b>	<b>Objetivos</b>	<b>16</b>
<b>1.2</b>	<b>Divisão do Trabalho</b>	<b>16</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>17</b>
<b>2.1</b>	<b>A Taxa de Câmbio</b>	<b>18</b>
2.1.1	A Regra de Taylor	18
2.1.2	A Paridade Descoberta da Taxa de Juros e o Prêmio de Risco	20
2.1.3	O Modelo para a Taxa de Câmbio Nominal	21
<b>3</b>	<b>REVISÃO DE MODELOS</b>	<b>22</b>
<b>3.1</b>	<b>O Modelo de Nelson-Siegel (NS)</b>	<b>22</b>
<b>3.2</b>	<b>A Abordagem de Diebold e Li - O Modelo de Fatores Latentes Dinâmicos de Nelson-Siegel (DNS)</b>	<b>22</b>
<b>3.3</b>	<b>O Modelo DNS em Formato de Estado-Espaço</b>	<b>23</b>
3.3.1	O Modelo DNS-Only e o Modelo DNS Aumentado por Fatores Macroeconômicos (MA-DNS)	23
<b>3.4</b>	<b>A Abordagem de Chen e Tsang: O Modelo MA-DNS Relativo (R-MA-DNS)</b>	<b>27</b>
3.4.1	O Filtro de Kalman:	30
<b>4</b>	<b>METODOLOGIA</b>	<b>32</b>
<b>4.1</b>	<b>O Modelo TVP-VAR</b>	<b>32</b>
4.1.1	Estimação de Modelos TVP-VAR com Fatores de Esquecimento	33
<b>4.2</b>	<b>Os Modelos DMA e DMS</b>	<b>35</b>
4.2.1	A Combinação Dinâmica de Modelos (DMA)	35
4.2.2	A Seleção Dinâmica de Modelos (DMS)	37
<b>4.3</b>	<b>Estratégia de Estimação - Modelos Considerados</b>	<b>38</b>
<b>4.4</b>	<b>Previsão</b>	<b>40</b>
4.4.1	Medidas de Desempenho e Comparação	40
<b>5</b>	<b>DADOS</b>	<b>42</b>
<b>5.1</b>	<b>Dados Utilizados</b>	<b>42</b>
5.1.1	As Taxas de Juros	42
5.1.2	As Variáveis Macroeconômicas	42
<b>5.2</b>	<b>Transformações</b>	<b>43</b>
<b>6</b>	<b>RESULTADOS</b>	<b>44</b>
<b>6.1</b>	<b>Estimativas</b>	<b>44</b>

6.2	Previsão . . . . .	45
7	CONCLUSÃO . . . . .	48
A	DADOS: . . . . .	50
A.1	A Curva de Juros e os Fatores Latentes de Nelson-Siegel . . . . .	50
A.2	Testes de Raiz Unitária e Estatísticas Descritivas . . . . .	52
B	VAR . . . . .	54
C	PREVISÃO . . . . .	55
	REFERÊNCIAS . . . . .	57

## 1 INTRODUÇÃO

Movimentos nas taxas de câmbio são importantes para o estudo em economia por impactar direta e indiretamente as decisões dos agentes econômicos em relação ao consumo, poupança, investimento, oferta e demanda de trabalho, além de afetar as políticas macroeconômicas formuladas pelos governos. Prever a taxa de câmbio com precisão tem sido um dos maiores desafios dentro da área de macrofinanças devido à sua complexidade inerente. Conforme mencionado em [Chen e Tsang \(2011\)](#) a taxa de câmbio apresenta um caráter dual: configura uma importante variável macroeconômica, equilibrando os preços relativos de bens e serviços no mercado internacional e, ao mesmo tempo, representa um ativo financeiro, incorporando em seu preço expectativas e prêmios de risco.

O encerramento do padrão dólar-ouro em 1971<sup>1</sup>, tornando-se o dólar uma moeda fiduciária, e o fim dos acordos de Bretton Woods em 1973<sup>2</sup>, em que o G-10, grupo de economias desenvolvidas, passou a adotar o regime de câmbio flutuante, marcaram uma nova era de expansão e de reconfiguração do mercado global de câmbio. Até o final do século XX, a maior parte dos países relevantes economicamente tornaram suas moedas flutuantes, incluindo o Brasil, em 1999<sup>3,4</sup>. A aceleração, durante a última década, da integração entre os mercados internacionais e o aperfeiçoamento e a difusão de instrumentos financeiros derivativos, como *swaps*, opções e *forwards*, além de ter elevado a mais de 5 trilhões de dólares o volume nominal de transações diárias entre moedas<sup>5</sup>, montante muito superior ao verificado em outros mercados, como o de ações, também tornou mais relevante o desenvolvimento do câmbio como uma nova classe de ativo financeiro.

Dois grupos de modelos para a determinação da taxa nominal de câmbio surgiram logo após o fim da era Bretton Woods: o monetário e o baseado nos saldos em carteira. A primeira classe de modelos reconhece o papel do câmbio como o preço relativo entre moedas, dependendo do equilíbrio entre a demanda e o estoque monetário. Os principais autores são [Bilson \(1980\)](#) e [Frenkel \(1976\)](#), que assumiram a hipótese de preços flexíveis, e [Dornbusch \(1976\)](#) e [Frankel \(1979\)](#), que incorporaram rigidez de preços. O segundo grupo de modelos, que surgiu a partir de [Frankel \(1984\)](#), estabelece que os ativos denominados em diferentes moedas não são substitutos perfeitos, implicando que os retornos de títulos entre países convertidos para uma mesma moeda podem divergir devido à existência de um prêmio de risco<sup>6</sup>.

<sup>1</sup> <[https://www.federalreservehistory.org/essays/gold\\_convertibility\\_ends](https://www.federalreservehistory.org/essays/gold_convertibility_ends)>

<sup>2</sup> <<https://history.state.gov/milestones/1969-1976/nixon-shock>>

<sup>3</sup> <<http://www.bcb.gov.br/rex/MerCambio/Port/cambio991/1999-1PolíticaCambial.asp>>

<sup>4</sup> O regime adotado foi o câmbio flutuante "sujo", que permite a intervenção do Banco Central no mercado cambial de forma pontual com o objetivo de suavizar os movimentos na taxa de câmbio.

<sup>5</sup> De acordo com o relatório do Banco de Compensações Internacionais (BIC) sobre o volume médio de transações no mercado de câmbio em 2016: <<https://www.bis.org/publ/rpfx16.htm>>

<sup>6</sup> [James, Marsh e Sarno \(2012, p.49\)](#).

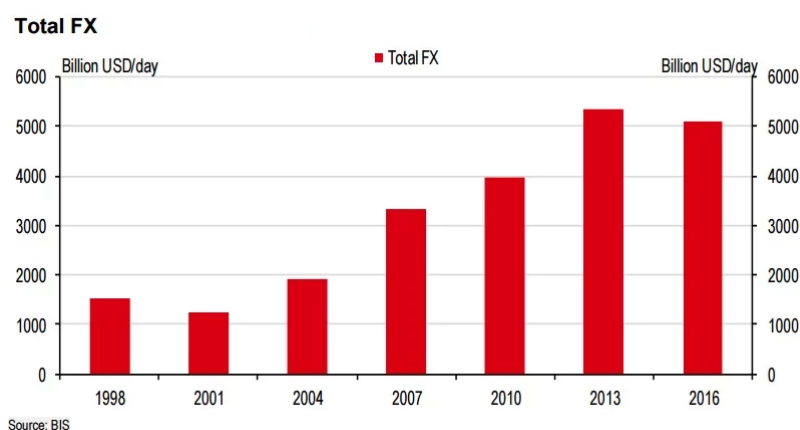


Figura 1 – Volume médio diário em dólar transacionado entre 1998 e 2016 no mercado global de câmbio - Fonte: BIS

A semelhança entre as duas classes de modelos é que ambas passaram a levar em conta o papel das expectativas dos agentes em sua formulação. A resolução dos sistemas de equações para a taxa de câmbio esperada, a partir das hipóteses iniciais dos modelos, resultava na determinação da taxa de câmbio *spot* como o valor dos fundamentos econômicos descontados a valor presente<sup>7</sup>. Entretanto, a abordagem empírica usualmente assumia hipóteses simplificadoras e incompletas para o termo relacionado às expectativas. Outro ponto a ser levado em conta é que, apesar de considerar o papel do câmbio como um ativo financeiro cujo preço também inclui um termo de risco, a literatura em macroeconomia inicialmente concebia o risco como não observável, dando pouca ênfase para a sua especificação<sup>8</sup>, e, assim, deixava de captar movimentos importantes de curto prazo na dinâmica da taxa de câmbio.

A década de 1980 marcou um período de desesperança em relação à capacidade preditiva de modelos cambiais. Meese e Rogoff (1983) avaliaram o desempenho de modelos monetários, com e sem preços flexíveis, um modelo com rigidez de preços que incorporava a conta corrente em sua formulação, além de outros modelos empíricos<sup>9</sup>. Os autores verificaram que nenhum desses modelos conseguia superar o passeio aleatório em previsões fora da amostra. Desde então, tornou-se padrão a comparação da capacidade de previsão de modelos cambiais com a do passeio aleatório, *random walk* (RW). Os estudos, ao longo das duas décadas seguintes, também não encontraram evidências significativas a favor dos modelos dos anos 1970 e 1980.

Paralelamente, desenvolveu-se uma literatura em finanças interessada em captar a influência das expectativas e do risco na determinação das taxas de juros de títulos governamentais, para diferentes maturidades, a partir da modelagem da curva de juros,

<sup>7</sup> Para maiores informações vide James, Marsh e Sarno (2012, p.45-54).

<sup>8</sup> Chen e Tsang (2011, p.01).

<sup>9</sup> Modelos ARIMA e VAR.

também conhecida por estrutura a termo da taxa de juros (ETTTJ). A princípio, os estudos consideravam somente algumas maturidades ou partes da ETTJ, perdendo informação relevante, ou sobreparametrizavam a sua estimação, obtendo modelos com baixo poder preditivo. Uma solução proposta com base em [Nelson e Siegel \(1987\)](#) foi modelar a curva de juros de forma parcimoniosa, por meio da extração de fatores latentes<sup>10</sup>. Desde então, diversos estudos surgiram<sup>11</sup>, sendo um dos mais importantes o proposto por [Diebold e Li \(2006\)](#)<sup>12</sup> que, por meio de modificações na formulação dos fatores latentes de Nelson-Siegel (NS), permitiu melhorar a capacidade de ajuste e de previsão da ETTJ, além de também possibilitar a interpretação dos três fatores como o nível, a inclinação e a curvatura da curva de juros, de modo semelhante à interpretação conferida por [Litterman e Scheinkman \(1991\)](#)<sup>13</sup>.

Nos anos 1990 surgiram estudos focados na influência das expectativas e do risco, captados por meio de movimentos de parte da ETTJ<sup>14</sup>, sobre as variáveis macro<sup>15</sup>. Por sua vez, nos anos 2000 emergiu uma linha de pesquisa interessada em analisar a influência das variáveis macroeconômicas sobre a curva de juros<sup>16</sup>, representada pelos fatores latentes. Havia, portanto, uma lacuna no estudo das relações mútuas entre a ETTJ e as variáveis macro, que começou a ser preenchida a partir de [Diebold, Rudebusch e Aruoba \(2006\)](#), em que o impacto entre a curva de juros, captada pelos fatores latentes de NS, e variáveis macroeconômicas importantes, como a inflação e a atividade do produto, foram analisadas por meio de uma estrutura VAR (*Vector Autoregression*)<sup>17</sup>. Os autores encontraram efeitos significativos nos dois sentidos em termos de previsão, prevalecendo os impactos das variáveis macro sobre os movimentos da ETTJ<sup>18</sup>.

Uma nova classe de modelo cambial despontou com [Engel e West \(2005\)](#) ao incorporar explicitamente a função de reação da autoridade monetária por meio de formulações para a Regra de Taylor. Os autores assumiram o formato internacional, baseado em [Clarida, Gali e Gertler \(1998\)](#), em que a pequena economia aberta, além de estabilizar a

<sup>10</sup> No caso em questão, os autores estimaram três fatores latentes.

<sup>11</sup> Por exemplo: [Knez, Litterman e Scheinkman \(1994\)](#), [Duffie e Kan \(1996\)](#) e [Dai e Singleton \(2000\)](#).

<sup>12</sup> Vide [Caldeira, Moura e Portugal \(2010\)](#) para uma aplicação para o Brasil.

<sup>13</sup> Os autores, ao aplicar o procedimento do "*Principal Component Analysis (PCA)*" para estudar a relação entre os fatores comuns e os retornos para os títulos de renda fixa dos Estados Unidos, constataram que cada um dos três fatores poderia ser interpretado como nível, inclinação e curvatura da ETTJ.

<sup>14</sup> Essa literatura em macro ainda não utilizava os fatores latentes, e, portanto, capturava apenas parte dos movimentos da curva de juros, por meio da escolha pré determinada de somente algumas maturidades a serem analisadas.

<sup>15</sup> Por exemplo: [Mishkin \(1990\)](#), [Barr e Campbell \(1997\)](#) e [Estrella e Mishkin \(1998\)](#).

<sup>16</sup> Por exemplo: [Ang e Piazzesi \(2003\)](#) e [Wu \(2001\)](#).

<sup>17</sup> O modelo dinâmico para os fatores latentes de Nelson-Siegel expandido para incluir variáveis macroeconômicas, *Dynamic Macro-Augmented Nelson-Siegel (MA-DNS)*, utiliza uma estrutura em formato de estado-espaco de modo a estimar de forma dinâmica os fatores NS e, ao mesmo tempo, captar as inter-relações entre as variáveis refletidas nos coeficientes da matriz relacionada à estrutura VAR da equação de transição.

<sup>18</sup> Outros estudos semelhantes são: [Levant e Ma \(2016\)](#), para o Reino Unido, [Lange \(2013\)](#), para o Canadá, e [Bernz \(2014\)](#) para o Brasil.

taxa de inflação e do produto, também tem como objetivo manter a paridade do poder de compra da moeda fixada em determinado nível. Estudos como os de [Alquist e Chinn \(2008\)](#), [Engel \*et al.\* \(2007\)](#) e [Molodtsova e Papell \(2009\)](#) deram uma nova esperança para a modelagem da taxa de câmbio ao verificar que modelos baseados na Regra de Taylor melhoravam a capacidade preditiva tanto dentro como fora da amostra, revertendo quase duas décadas em que se considerava que os modelos cambiais falhavam empiricamente.

Apesar dos avanços na literatura, [Chen e Tsang \(2011\)](#) notaram uma desconexão entre a abordagem empírica em macroeconomia e em finanças no tratamento da taxa de câmbio nominal: enquanto os modelos pós Bretton-Woods passaram a incorporar o papel das expectativas na determinação da taxa de câmbio, pouca ênfase era dada para o prêmio de risco, considerado não observável, e, por isso, costumavam falhar em captar a dinâmica de curto prazo. Na prática, esses modelos assumiam um formato restritivo ou incompleto para as expectativas, enquanto alguns estudos recentes em finanças capturavam o impacto do risco sistêmico embutido na ETTJ sobre o excesso de retorno cambial, porém sem estabelecer conexões entre o risco e a conjuntura macroeconômica e, como consequência, deixavam de captar as inter-relações entre o risco, as variáveis macro e as expectativas<sup>19</sup>.

[Chen e Tsang \(2011\)](#), aproveitando-se dos avanços na literatura e com o objetivo de aproximar as abordagens em finanças e em macroeconomia, desenvolveram um modelo para a taxa de câmbio baseado na combinação da Regra de Taylor e da paridade descoberta da taxa de juros, *uncovered interest rate parity* (UIP)<sup>20</sup>, com os fatores latentes de Nelson-Siegel incluídos como proxies de forma a refletir as expectativas do mercado e o risco cambial. As variáveis são representadas de forma relativa, ou seja, pela diferença entre os valores observados em uma pequena economia aberta<sup>21</sup> em relação aos Estados Unidos. Baseado em [Diebold, Rudebusch e Aruoba \(2006\)](#), o modelo assume uma estrutura VAR em formato de estado-espço em que se torna possível analisar os impactos das variáveis macro e dos fatores latentes entre si, porém, com a inclusão da taxa de câmbio, principal variável de interesse.

[Chen e Tsang \(2011\)](#), por fim, compararam três especificações para a taxa de câmbio com o passeio aleatório<sup>22</sup>: um modelo somente com os fatores latentes NS (*yields-only*); outro unicamente com as variáveis macro (*macro-only*) e, por fim, o modelo com todas as variáveis (*macro-yields*). Os autores concluíram que as três especificações superaram o passeio aleatório e que o modelo *macro-yields* apresentou, em média<sup>23</sup>, o melhor

<sup>19</sup> Como citado, a partir de [Diebold, Rudebusch e Aruoba \(2006\)](#) houve uma evolução no estudo das inter-relações entre as variáveis macro e a curva de juros, por meio dos fatores latentes de NS, entretanto, o modelo original não tinha como foco o estudo da taxa de câmbio.

<sup>20</sup> Uma das hipóteses, conforme [Bekaert, Wei e Xing \(2007\)](#), é que o risco cambial se relaciona ao risco de mercado embutido na curva de juros por meio da UIP.

<sup>21</sup> Os países analisados foram o Canadá, o Reino Unido e o Japão.

<sup>22</sup> principal *benchmark*.

<sup>23</sup> O modelo *macro-yields* apresentou melhor performance, em termo de erros de previsão, para o Reino Unido e para o Japão enquanto que para o Canadá o modelo *macro-only* se saiu melhor.

desempenho, tanto dentro como fora da amostra, indicando a relevância conjunta das expectativas e do risco, captadas por meio dos fatores latentes de Nelson-Siegel, e dos fundamentos macroeconômicos. Para o estudo de caso no Brasil, destaca-se o realizado por [Mariani \(2015\)](#) que acrescentou uma estrutura estocástica para as volatilidades dos fatores latentes de Nelson-Siegel ao modelo de [Chen e Tsang \(2011\)](#), utilizando estimação Bayesiana baseada em processos MCMC (*Markov Chain Monte Carlo*). O autor estimou duas especificações para o modelo e, ao comparar as estatísticas MAFE (*Mean Absolute Forecasting Error*) e MSFE (*Mean Square Forecasting Error*), observou que somente o segundo modelo superou o passeio aleatório. No entanto, ao utilizar o teste de Diebold-Mariano, o resultado encontrado foi que nenhum dos modelos estimados era significativamente superior ao passeio aleatório.

Mais recentemente, [Kavtaradze e Mokhtari \(2017\)](#) realizaram uma revisão da literatura sobre os principais modelos para a previsão da taxa de câmbio<sup>24</sup> e concluíram que aqueles que apresentavam melhor capacidade preditiva fora da amostra eram os modelos FAVAR (*Factor-Augmented Vector Autoregressive*) junto com os TVP-VAR (*Time-Varying Parameter VAR*), apresentando boa performance, no longo prazo, e média, no curto-prazo, enquanto que os modelos baseados na Regra de Taylor apresentavam médio poder de previsão para ambos os horizontes<sup>25</sup>. Os modelos FAVAR foram desenvolvidos de forma a reduzir a dimensão de modelos VAR: em uma primeira etapa, realiza-se a extração de fatores latentes que capturam as relações em comum entre o conjunto maior de variáveis consideradas. Em uma segunda etapa, esse número reduzido de fatores é incluído em uma estrutura VAR.

Por sua vez, o modelo TVP-VAR impõe uma estrutura estocástica à matriz de coeficientes permitindo captar mudanças nas inter-relações entre as variáveis ao longo do tempo, como quebras estruturais, além de choques. Contudo, como apontado por [Koop e Korobilis \(2013\)](#), a estimação de modelos TVP-VAR e, principalmente, a realização de previsão, mesmo para os modelos de reduzida dimensão, costumam impor grande custo computacional, devido à necessidade de se realizar simulações junto à estimação por Filtro de Kalman. Uma solução adotada pela literatura<sup>26</sup> com o objetivo de lidar com a "maldição da dimensionalidade" é impor uma estrutura com decaimento temporal para as matrizes de variâncias/covariâncias do modelo, o que é realizado pela inclusão de fatores de esquecimento. Esse procedimento viabiliza a estimação de modelos TVP-VAR de maior dimensão sem a necessidade de simulações a partir de métodos MCMC.

A utilização de fatores de esquecimento possibilitou o desenvolvimento de novos modelos econométricos para a abordagem TVP-VAR, destacando-se a combinação dinâmica de modelos, *Dynamic Model Averaging* (DMA), desenvolvida por [Raftery, Kárný e](#)

<sup>24</sup> Os autores também avaliam modelos para a previsão da taxa de inflação.

<sup>25</sup> [Kavtaradze e Mokhtari \(2017\)](#), p. 20).

<sup>26</sup> Discutida em profundidade por [Raftery, Kárný e Ettlér \(2010\)](#).



Ettler (2010), e a seleção dinâmica de modelos, *Dynamic Model Selection* (DMS). Ambos os procedimentos lidam o problema da incerteza na escolha do melhor modelo para previsão, por meio de algoritmos que permitem a seleção de um modelo ou de uma combinação de modelos distintos, a cada período de tempo. A diferença é que, enquanto o DMA pondera diferentes modelos, dando maior peso para aqueles com maior função de densidade de probabilidade, o DMS escolhe o modelo com a melhor capacidade preditiva.

Koop e Korobilis (2013), com o intuito de analisar o poder de previsão para algumas variáveis macroeconômicas importantes, estimaram distintas especificações para modelos TVP-VAR, inclusive considerando diferentes dimensões<sup>27</sup>, e compararam esses modelos entre si e, também, com alguns *benchmarks*. Os autores apresentaram as seguintes conclusões: o aumento do número de variáveis melhorou, em geral, a capacidade preditiva em comparação com modelos de menor dimensão<sup>28</sup>; o mesmo ocorreu ao se assumir heteroscedasticidade; a performance de modelos TVP-VAR foi superior a dos *benchmarks*, dentre eles o passeio aleatório e o modelo VAR-OLS (de pequena dimensão)<sup>29</sup>; a especificação DMA apresentou, em geral, resultados levemente superiores aos do modelo DMS e, por fim, os autores observaram que os modelos DMA e DMS foram, em média, superiores<sup>30</sup>, tanto para horizontes curtos, como para horizontes longos.

Dentre as aplicações para o Brasil destacam-se dois estudos: o de Caldeira, Moura e Santos (2015) e o de Ferreira e Palma (2015). O primeiro replica o artigo de Koop e Korobilis utilizando a metodologia do *model confidence set* (MCS) que permite a comparação da performance de um conjunto de modelos sem ser necessário escolher um *benchmark* específico. Os autores observam que modelos TVP-VAR apresentam melhor capacidade preditiva, inclusive, comparados a modelos FAVAR. O segundo estudo encontra bons resultados de previsão para a taxa de inflação a partir de uma curva de Phillips estimada por meio da abordagem DMA. Por fim, até o momento a pesquisa em econometria para o Brasil tem desenvolvido abordagens TVP-VAR (DMA e DMS) com enfoque na estimação e previsão de certas variáveis, como a taxa de juros, a inflação e o crescimento do produto, havendo espaço para a realização de novos estudos com foco em outra variável fundamental: a taxa de câmbio.

<sup>27</sup> Os autores analisaram a capacidade de previsão a partir de três diferentes dimensões para o modelo TVP-VAR: pequena (3), média (7) e grande, com 25 variáveis.

<sup>28</sup> Esse fato foi observado para a taxa de inflação e para a taxa de juros, entretanto, o modelo de menor dimensão se saiu melhor para o crescimento do produto.

<sup>29</sup> No entanto, para a especificação VAR-OLS, com uma defasagem, a capacidade de previsão para longos horizontes foi, em média, equivalente (às vezes superando) a dos modelos TVP-VAR com maior poder preditivo, como o DMA e o DMS.

<sup>30</sup> Performando significativamente melhor do que o passeio aleatório.

## 1.1. Objetivos

Este trabalho tem como foco replicar as três abordagens encontradas em [Chen e Tsang \(2011\)](#) para a previsão dos movimentos da taxa de câmbio do Brasil contribuindo para a literatura ao permitir que os coeficientes relacionados aos fundamentos variem no tempo, impondo uma estrutura TVP-VAR. De forma a lidar com o problema da dimensionalidade e, simultaneamente, de modo a permitir que diferentes modelos sejam escolhidos ao longo do tempo, de acordo com o poder de previsão, fatores de esquecimento e os algoritmos para a Seleção Dinâmica de Modelos (DMS) e para a Combinação Dinâmica de Modelos (DMA) são implementados, seguindo a metodologia de [Koop e Korobilis \(2013\)](#).

As principais perguntas a serem respondidas, em termos de previsão, são:

- Os diferentes fundamentos e abordagens econométricas superam o passeio aleatório (principal *benchmark*)?
- A abordagem conjunta *macro-yields* é superior às abordagens individuais?
- Há ganhos de previsão ao se permitir que os coeficientes variem no tempo, por meio de uma estrutura TVP-VAR?
- Há ganhos de previsão ao se permitir que os modelos variem ao longo do tempo por meio da implementação dos algoritmos DMA e DMS?
- Quais modelos apresentaram, em média, a melhor performance?

## 1.2. Divisão do Trabalho

A seção 2 aborda os fundamentos teóricos do modelo de [Chen e Tsang \(2011\)](#), elucidando a relação entre as variáveis macro, o risco, as expectativas e a curva de juros. A seção 3 realiza uma revisão de cada modelo empírico de forma a ser possível compreender melhor o modelo TVP-VAR, desenvolvido neste trabalho. A seção 4, metodologia, apresenta em detalhes a estrutura TVP-VAR, a sua estimação, os fatores de esquecimento e as abordagens DMA e DMS. A seção 5 aborda os dados utilizados e as transformações realizadas sobre a base de dados. A seção 6 expõe os resultados e, a seção 7, a conclusão.

## 2 FUNDAMENTAÇÃO TEÓRICA

Como declarado por [Chen e Tsang \(2011\)](#)[p. 8, tradução nossa]<sup>1</sup>: ..."a literatura recente em macrofinanças relaciona a condição de curto prazo da taxa de juros, como sendo um instrumento de política monetária, à ideia que os juros, para todas as maturidades, são médias ajustadas pelo risco das taxas esperadas de curto prazo". Conforme [Diebold, Piazzesi e Rudebusch \(2005\)](#), por meio da teoria das expectativas aplicada à ETTJ, as taxas de curto prazo estão interligadas às taxas de prazo mais longo, ao pressupor que os juros de diferentes maturidades são médias ajustadas pelo risco dos valores esperados das taxas de curto prazo.

A conexão entre as variáveis macroeconômicas e a curva de juros pode ser explicada de acordo com [Gürkaynak e Wright \(2012\)](#)[p.332, tradução nossa]: "se as taxas de curto prazo reagem à inflação e ao hiato do produto, como é o caso da Regra de Taylor, então a ETTJ deverá refletir as expectativas dos agentes em relação a essas variáveis no futuro"<sup>2</sup>. Nesse sentido, os fundamentos da política monetária estarão contidos na curva de juros e, dessa forma, métodos que extraíam fatores a partir da ETTJ poderão ser utilizados como *proxies* para captar as expectativas dos agentes e o risco de mercado<sup>3</sup>.

Seguindo essa lógica, a fundamentação teórica de [Chen e Tsang \(2011\)](#) baseada na conexão entre a taxa de câmbio nominal, a curva de juros e as variáveis macroeconômicas, advém de duas considerações: em finanças, a hipótese da Paridade Descoberta da Taxa de Juros (UIP)<sup>4</sup> conecta diretamente os movimentos da taxa de câmbio ao diferencial da curva de juros entre países. Por sua vez, a Regra de Taylor estabelece que a taxa de curto prazo é o principal instrumento das autoridades monetárias para estabilização macroeconômica, cuja função de reação inclui, para uma pequena economia aberta, a estabilização da taxa de câmbio real, e, dessa forma, interliga indiretamente a taxa de câmbio nominal às variáveis macroeconômicas e às expectativas.

A abordagem em macrofinanças proposta pelos autores, portanto, extrai três fatores latentes de Nelson-Siegel (NS) a partir do diferencial entre países da ETTJ com o objetivo de captar as expectativas em relação aos fundamentos macroeconômicos e, também, o risco relativo. Simultaneamente, o diferencial entre as funções de reação das autoridades monetárias e a UIP vinculam a curva de juros relativa, às variáveis macroeconômicas, incluindo a taxa de câmbio nominal. A inter-relação entre a taxa de câmbio,

<sup>1</sup> ..."the recent macro-finance literature relates the fact that the short rate is a monetary policy instrument to the idea that yields of all maturities are risk-adjusted averages of expected short rates".

<sup>2</sup> ..."If short term interest rates are in turn driven by inflation and the output gap, as in the Taylor rule, then the term structure of interest rates ought to reflect expectations of future".

<sup>3</sup> A metodologia a ser seguida, baseada em [Chen e Tsang \(2011\)](#) e [Chen e Tsang\\* \(2013\)](#) utiliza os fatores latentes do modelo de Nelson e Siegel de forma a obter *proxies* para as expectativas e riscos, evitando a necessidade de modelagem de cada termo da curva de juros, o que pode gerar problemas de sobre-especificação.

<sup>4</sup> "Uncovered Interest Rate Parity".

o diferencial da Regra de Taylor, a UIP e a ETTJ relativa pode ser sintetizada pela figura 2, abaixo.

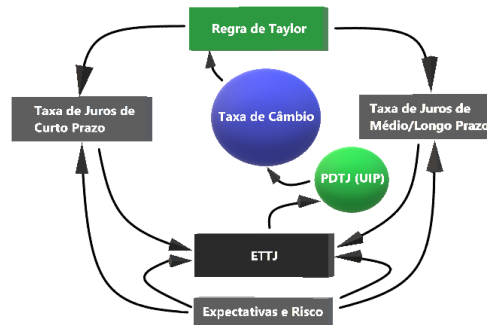


Figura 2 – Inter-relação entre a Taxa de Câmbio, a Regra de Taylor, a UIP e a Curva e Juros.

## 2.1. A Taxa de Câmbio

O modelo de [Chen e Tsang \(2011\)](#) para a taxa de câmbio se estrutura a partir de duas considerações: a primeira, obtida do diferencial entre as regras de Taylor e, a segunda, da hipótese da paridade descoberta da taxa de juros (UIP).

### 2.1.1. A Regra de Taylor

A primeira parte do modelo para a taxa de câmbio nominal é derivada do modelo diferencial para a Regra de Taylor de [Clarida, Galí e Gertler \(1998\)](#), adaptado por [Engel e West \(2005\)](#)[pp. 494-496] com enfoque na taxa de câmbio. Dois países são considerados: o primeiro é uma grande economia aberta, cuja função de reação da autoridade monetária segue uma Regra de Taylor em formato clássico, com estabilização do produto e ancoragem das expectativas de inflação. O segundo país é uma pequena economia aberta, cuja política monetária apresenta um objetivo adicional: a estabilização da taxa real de câmbio<sup>5</sup>. De acordo com a literatura os EUA serão considerados a grande economia aberta ou país estrangeiro, e, o Brasil, uma pequena economia aberta<sup>6</sup>, ou país local.

A regra de Taylor para o país local (pequena economia aberta) é dada por:

$$i_t = \mu_t + \beta_y \tilde{y}_t + \beta_\pi \pi_t^e + \delta q_t \quad (2.1)$$

<sup>5</sup> Conforme [Engel e West \(2005, pp.495\)](#): "we shall assume that monetary authorities target the PPP level of the exchange rate".

<sup>6</sup> O Brasil, apesar de ter uma participação reduzida do comércio sobre o total do produto, pode, ainda assim, ser considerado uma economia aberta, pois o seu mercado financeiro interno é diretamente influenciado pelos movimentos no mercado internacional.

Para o país estrangeiro (grande economia aberta), a regra apresenta a formulação padrão:

$$i_t^* = \mu_t^* + \beta_y \tilde{y}_t^* + \beta_\pi \pi_t^{e*} \quad (2.2)$$

Tomando a diferença entre 2.1 e 2.2, temos a equação que relaciona o diferencial da taxa de juros à regra de Taylor relativa:

$$i_t - i_t^* = \beta_y (\tilde{y}_t - \tilde{y}_t^*) + \beta_\pi (\pi_t^e - \pi_t^{e*}) + \delta q_t + v_t \quad (2.3)$$

Em que:

\*: é o sobrescrito que representa o país estrangeiro.

$i_t$ : representa a taxa de juros de curto prazo, sendo o instrumento sob domínio da autoridade monetária.

$\tilde{y}_t$ : é o hiato do produto.

$\pi_t^e$ : é a taxa de inflação esperada.

$q_t = s_t + p_t^* - p_t$ : é a taxa de câmbio real.

$s_t$ : é a taxa de câmbio nominal.

$p_t^* - p_t$ : é a diferença entre o nível de preços do país estrangeiro e do país local.

$\mu_t$ : é o termo que absorve as metas de inflação e de produto e também o equilíbrio da taxa de juros real.

$v_t$ : é o diferencial entre os termos  $\mu_t$ .

De acordo com a teoria, assume-se que a autoridade monetária aumente a taxa de juros de curto prazo de acordo com desvios positivos do hiato do produto, da inflação esperada e da taxa real de câmbio. Logo, espera-se que  $\beta_y > 0$ ,  $\delta q_t > 0$  e que  $\beta_\pi > 1$ , devido à maior preocupação dos Bancos Centrais em ancorar as expectativas dos agentes em relação à inflação esperada.

*OBS*<sub>1</sub>: todas as variáveis, exceto as taxas de juros, serão representadas em logaritmo.

*OBS*<sub>2</sub>: por simplicidade notacional, assume-se que os bancos centrais do país local e do país estrangeiro apresentem os mesmos coeficientes de reação  $\beta_{\tilde{y}}$  e  $\beta_\pi$ <sup>7</sup>.

<sup>7</sup> De forma semelhante ao realizado em [Chen e Tsang \(2011\)](#)[p.5] e [Chen e Tsang\\* \(2013\)](#)[p.187].

### 2.1.2. A Paridade Descoberta da Taxa de Juros e o Prêmio de Risco

De acordo com Terra (2015) a paridade descoberta da taxa de juros estabelece que os agentes não utilizam o mercado futuro de câmbio, ou seja, os agentes não realizam *hedge*, assimilando o risco cambial, e que eles somente trocam moeda no momento da conversão. Ao incluir o termo de risco cambial dado por  $\hat{\rho}_t^m$ , considera-se que o retorno dos ativos do país local pode divergir do retorno do país estrangeiro após a conversão pela taxa de câmbio, conforme a versão fraca da UIP:

$$i_t^m - i_t^{m,*} - E_t \Delta s_{t+m} = \hat{\rho}_t^m, \forall t \quad (2.4)$$

O termo à esquerda da equação 2.4 representa o diferencial descoberto da taxa de juros.

Reorganizando, temos:

$$i_t^m - i_t^{m,*} = E_t \Delta s_{t+m} + \hat{\rho}_t^m, \forall t \quad (2.5)$$

Ao assumir a teoria das expectativas racionais, de que os agentes não cometem erros de previsão sistemáticos, pode-se utilizar a taxa de câmbio efetiva em  $t + m$  como proxy para  $E s_{t+m}$ :

$$E_t \Delta s_{t+m} = E[s_{t+m}] - s_t \equiv s_{t+m} - s_t \quad (2.6)$$

A existência do prêmio de risco, implica, por definição, excesso de retorno cambial, cuja equação é dada por:

$$XR_{t+m} = i_t^{m,*} - i_t^m + \Delta s_{t+m} (\equiv \hat{\rho}_t^{m,*} + \epsilon_{t+m}) \quad (2.7)$$

Sob expectativas racionais e considerando-se a validade da UIP, assume-se que o termo de erro dado por  $\epsilon_{t+m}$  se comporte como um ruído branco. Nesse caso, em média, o diferencial descoberto da taxa de juros será equivalente ao prêmio de risco cambial. Por sua vez, o termo de risco cambial se relaciona com o diferencial entre o risco latente de se manter títulos do país local e do risco de se manter títulos estrangeiros em carteira, conforme a seguinte equação:

$$\hat{\rho}_t^m = a_0 + a_{m,F} \rho_t^{m,*} - a_{m,H} \rho_t^m + \varsigma_t \quad (2.8)$$

Em que  $\rho_t^m$  representa o prêmio de risco relacionado a se manter títulos e ativos do país local em carteira,  $a_{m,F}$  e  $a_{m,H}$  são os coeficientes, para a maturidade  $m$ , relacionados aos prêmios de risco do país estrangeiro e local, respectivamente. Além disso,  $\varsigma_t$

representa o componente de erro que se comporta como um ruído branco e, assume-se, por simplificação, que  $a_{m,F} = a_{m,H} = a_m$ .

A ideia presente na equação 2.8 é que os mesmos fatores de risco que afetam os títulos de cada país, também afetam o risco de se manter moeda em carteira, estabelecendo uma relação entre o risco cambial e o risco relacionado aos títulos, refletidos na curva de juros.

### 2.1.3. O Modelo para a Taxa de Câmbio Nominal

Iterando a regra de Taylor relativa, dada pela equação 2.3, utilizando a equação 2.5, relacionada à UIP, e reorganizando os termos, obtém-se o modelo de precificação da taxa de câmbio a valor presente, conforme abaixo:

$$s_t = \lambda \sum_{j=0}^{\infty} \psi^j E_t(f_{t+j}^{TR} | I_t) + \varsigma \sum_{j=0}^{\infty} \psi^j E_t(\hat{\rho}_t^j | I_t) + \epsilon_t \quad (2.9)$$

Em que  $f_t^{TR} = [p_t - p_t^*, \tilde{y}_t - \tilde{y}_t^*, \pi_t^e - \pi_t^{e*}]'$  representa o vetor com os fundamentos do modelo. Os indivíduos formam as suas expectativas em relação ao futuro com base em toda a informação disponível até o presente, representada por  $I_t$ . Considera-se que o termo de erro  $\epsilon_t$  incorpore choques, como aqueles que incidem sobre o termo de risco, e assume-se que seja não correlacionado com as variáveis macro e com o risco relativo de se manter ativos do país local e em relação ao país estrangeiro. Deve-se ressaltar que, como discutido, o termo de risco relativo  $\hat{\rho}_t$  e as expectativas se refletem na ETTJ.

Empiricamente a taxa de câmbio apresenta raiz unitária e, da mesma forma como realizado em [Chen e Tsang\\* \(2013\)](#), adota-se a sua diferenciação, dada por  $ds_t$ , que será o principal objeto de estudo. Outras variáveis foram diferenciadas, de acordo com testes de raiz unitária apresentados na tabela 4. Para maiores informações, acessar a seção 5. Dados, subseção 5.1. Transformações.

### 3 REVISÃO DE MODELOS

#### 3.1. O Modelo de Nelson-Siegel (NS)

Nelson e Siegel (1987), com o objetivo de modelar de forma parcimoniosa a curva de juros dos Estados Unidos, captando diferentes formatos (monotônico, curva inclinada ou em formato de S), formularam uma representação paramétrica a partir das taxas *forward*, para cada maturidade  $m$ , por meio da estimação de três parâmetros  $\beta$ , mais um parâmetro de decaimento exponencial,  $\tau$ .

$$r(m) = \beta_0 + \beta_1 e^{-m/\tau} + \beta_2 [(-m/\tau) e^{-m/\tau}] \quad (3.1)$$

A equação 3.1 é representada por uma constante somada a uma função de Laguerre, permitindo a generalização para modelos de ordem superior. Os autores integraram a equação para todas as maturidades e, ao dividir por  $m$ , obtiveram o valor médio para cada termo da curva de juros:

$$R(m) = \beta_0 + (\beta_1 + \beta_2)(1 - e^{-m/\tau})/(m/\tau) - \beta_2 e^{-m/\tau} \quad (3.2)$$

#### 3.2. A Abordagem de Diebold e Li - O Modelo de Fatores Latentes Dinâmicos de Nelson-Siegel (DNS)

Diebold e Li (2006) apontaram dois obstáculos na formulação original do modelo de NS:

First, conceptually, it would be hard to provide intuitive interpretations of the factors in the original Nelson-Siegel framework. Second, operationally, it would be difficult to estimate the factors precisely, because the high coherence in the factors produces multicollinearity (DIEBOLD; LI, 2006, p. 343).

De forma a lidar com o problema, os autores adotaram uma abordagem dinâmica e propuseram uma reformulação para a equação 3.2 tornando possível interpretar os coeficientes de NS como fatores latentes relacionados ao nível  $L_t$ , à inclinação  $S_t$  e à curvatura  $C_t$  da ETTJ<sup>1</sup>.

$$i_t(m) = L_t + S_t \left( \frac{1 - e^{-\lambda m}}{\lambda m} \right) + C_t \left( \frac{1 - e^{-\lambda m}}{\lambda m} - e^{-\lambda m} \right) + v_t \quad (3.3)$$

<sup>1</sup> A interpretação é semelhante à realizada em Litterman e Scheinkman (1991) ao aplicar o procedimento do "Principal Component Analysis (PCA)" para modelar a ETTJ dos Estados Unidos extraindo três componentes, que foram interpretados como nível, inclinação e curvatura da curva de juros.



O subscrito  $t$  denota a unidade de tempo,  $i_t$  é a taxa de juros zero cupom,  $v_t$  é o termo de erro e o parâmetro  $\lambda$  é o fator de decaimento exponencial do modelo, sendo que valores pequenos produzem decaimentos lentos enquanto que o inverso produz decaimentos rápidos. É possível notar semelhanças entre a formulação NS original da equação 3.2 e a de Diebold e Li (2006), equação 3.3, pois  $\beta_0 = L$ ,  $\beta_1 + \beta_2 = S$  e  $\beta_2 = C$ .

Os fatores de carregamento, junto com o fator de decaimento, são responsáveis pelo grau de persistência temporal dos fatores latentes sobre a curva de juros. De acordo com os autores, o termo  $L_t$ , por ter carregamento igual a um, apresenta persistência temporal para qualquer maturidade  $m$  considerada, podendo ser interpretado como um fator de longo prazo. O carregamento de  $S_t$ , por sua vez, apresenta decaimento rápido conforme aumenta a maturidade, sendo  $S_t$  relacionado ao curto prazo. Por sua vez, o carregamento de  $C_t$  apresenta valor zero para maturidades de curto prazo, tornando-se positivo para maturidades de médio prazo e decaindo novamente para zero para maturidades de longo prazo<sup>2</sup>, o que implica a relação do fator  $C_t$  com o médio prazo.

Diebold e Li (2006), com o intuito de obter os fatores latentes variantes no tempo, em um primeiro momento, calcularam o valor de  $\lambda$  de forma a maximizar o carregamento do fator da curvatura  $C_t$  na maturidade de 30 meses, e, no caso, o valor encontrado foi de 0.0609. Em um segundo momento, o modelo dado pela equação 3.3 foi estimado por mínimos quadrados ordinários (OLS), para cada período, gerando os coeficientes latentes dinâmicos. Por fim, previsões foram realizadas a partir dos fatores latentes em uma estrutura VAR(1).

### 3.3. O Modelo DNS em Formato de Estado-Espaço

#### 3.3.1. O Modelo DNS-Only e o Modelo DNS Aumentado por Fatores Macroeconômicos (MA-DNS)

Diebold, Rudebusch e Aruoba (2006) formularam o modelo DNS (*Dynamic Nelson-Siegel*) com duas configurações em formato de estado-espço. O primeiro, considerando somente os fatores latentes da curva de juros (DNS-Only) e, o segundo, adicionando variáveis macroeconômicas ao modelo (MA-DNS)<sup>3</sup>.

O objetivo dos autores foi verificar se a inclusão dos fatores macro melhorava de forma significativa a capacidade de previsão da curva de juros, além de estudar a inter-relação entre os fatores NS e algumas variáveis macroeconômicas importantes, como a inflação, a capacidade ociosa da indústria e a taxa de juros básica (*fed funds rate*) dos Estados Unidos, por meio de uma estrutura VAR(1) não estrutural para a equação de transição. A vantagem da modelagem em estado-espço, via aplicação do Filtro de Kal-

<sup>2</sup> Pode-se verificar que  $\lim_{m \rightarrow 0} \frac{1 - e^{-\lambda m}}{\lambda m} = \lim_{m \rightarrow \infty} \frac{1 - e^{-\lambda m}}{\lambda m} = 0$ .

<sup>3</sup> *Macro-Augmented Dynamic Nelson-Siegel*.

man, é que o procedimento retorna estimativas baseadas na maximização da função de verossimilhança e, também, valores ótimos filtrados e suavizados dos parâmetros do modelo. Como apontado em (DIEBOLD; RUDEBUSCH; ARUOBA, 2006, p. 313) "...the one-step Kalman filter approach of this paper is preferable to the two-step Diebold-Li approach, because the simultaneous estimation of all parameters produces correct inference via standard theory".

### 3.3.1.1. O Modelo DNS em Formato de Estado-Espaço: DNS-Only

A formulação em estado-espço do modelo para a curva de juros como função dos fatores DNS é dada por:

$$\text{Fundamentos: } f_t = (L_t, S_t, C_t)'$$

$$\text{Equação de Mensuração: } \bar{y} \equiv y - \Lambda\mu = \Lambda(f_t - \mu) + \epsilon_t$$

$$\text{Equação de Transição: } f_t - \mu = A(f_{t-1} - \mu) + \eta_t$$

Matriz dos Fatores de Carregamento da Equação de Mensuração:

$$\Lambda = \begin{pmatrix} 1 & \frac{1-e^{-\lambda m_1}}{\lambda m_1} & \frac{1-e^{-\lambda m_1}}{\lambda m_1} - e^{-\lambda m_1} \\ 1 & \frac{1-e^{-\lambda m_2}}{\lambda m_2} & \frac{1-e^{-\lambda m_2}}{\lambda m_2} - e^{-\lambda m_2} \\ \vdots & \vdots & \vdots \\ 1 & \frac{1-e^{-\lambda m_n}}{\lambda m_n} & \frac{1-e^{-\lambda m_n}}{\lambda m_n} - e^{-\lambda m_n} \end{pmatrix}$$

**Equação de Mensuração** em formato matricial:

$$\begin{pmatrix} i_t(m_1) \\ i_t(m_2) \\ \vdots \\ i_t(m_n) \end{pmatrix} - \Lambda \begin{pmatrix} \mu_L \\ \mu_S \\ \mu_C \end{pmatrix} = \Lambda \begin{pmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{pmatrix} + \begin{pmatrix} \epsilon_t(m_1) \\ \epsilon_t(m_2) \\ \vdots \\ \epsilon_t(m_n) \end{pmatrix}$$

**Equação de Transição** em formato matricial:

$$\begin{pmatrix} L_t - \mu_L \\ S_t - \mu_S \\ C_t - \mu_C \end{pmatrix} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{pmatrix} \begin{pmatrix} L_{t-1} - \mu_L \\ S_{t-1} - \mu_S \\ C_{t-1} - \mu_C \end{pmatrix} + \begin{pmatrix} \eta_t(L) \\ \eta_t(S) \\ \eta_t(C) \end{pmatrix}$$

Condições para Estimação por filtro de Kalman:

$$\begin{bmatrix} \eta_t \\ \epsilon_t \end{bmatrix} \sim WR \left[ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{pmatrix} Q & 0 \\ 0 & H \end{pmatrix} \right]$$

$$E[f_0\epsilon_t] = 0 \quad \text{e} \quad E[f_0\eta_t] = 0$$

Em que:

- $y_t$ : é o vetor com as taxas de juros dos títulos zero cupom para cada uma das maturidades  $m \in [1, n]$ . Nada mais é do que a representação vetorial do lado esquerdo da equação 3.3 para todas as maturidades em determinado ponto no tempo  $t$ .
- $\bar{y}_t$ : é um ajuste ao vetor  $y$ , contendo variáveis dependentes da equação de mensuração, de forma a compatibilizá-las como desvios da média.
- $\mu$ : representa o vetor com as médias das séries de tempo.
- $\Lambda$ : é a matriz com os fatores de carregamento do modelo NS.
- $A$ : é a matriz de coeficientes relacionada ao sistema VAR(1) não estrutural da equação de transição.
- $\epsilon_t$  e  $\eta_t$ : são os vetores das perturbações associados às equações de mensuração e de transição, respectivamente. De forma a ser possível estimar o modelo por Filtro de Kalman, assume-se que ambos os termos sejam ruídos brancos e que sejam ortogonais entre si para todos os períodos de tempo  $t, s$  e, também, para o estado inicial  $f_0$ .
- $Q$ : é a matriz de variância-covariância para a equação de transição. Assume-se que seja não diagonal, de forma a possibilitar que os choques entre as variáveis do vetor de fundamentos sejam correlacionados.
- $H$ : é a matriz de variância-covariância para a equação de mensuração, especificada como uma matriz diagonal, de forma que os desvios nas taxas de juros de diferentes maturidades em relação à curva de juros sejam não correlacionados. A especificação para os termos de erro e para as matrizes de variância-covariância permanecem inalteradas.

### 3.3.1.2. O Modelo DNS Aumentado por Fatores Macroeconômicos - MA-DNS

O modelo MA-DNS (*Macro-Augmented Dynamic Nelson-Siegel*), proposto por [Diebold, Rudebusch e Aruoba \(2006\)](#) com a finalidade de analisar a inter-relação entre as variáveis macroeconômicas e os fatores DNS para os Estados Unidos, é formulado por

meio da inclusão de variáveis macro (taxa de juros de referência, capacidade utilizada da indústria e taxa de inflação) no vetor de fundamentos  $f_t$  do modelo DNS-Only e do incremento correspondente nas dimensões das matrizes e vetores  $A$ ,  $\mu$ ,  $\eta_t$ ,  $Q$  e de adaptações da matriz  $\Lambda$ . Devido às semelhanças entre o modelo MA-DNS e o modelo R-MA-DNS, os detalhes sobre a sua formulação em formato de estado-espço será apresentada na próxima subseção. Abaixo descreve-se, de acordo com o autores, o impacto dos choques dos fatores NS sobre as variáveis macroeconômicas<sup>4</sup>:

1. Curvatura: quase nenhum impacto sobre as variáveis macro.
2. Inclinação: aumento proporcional sobre a taxa de juros de referência (demonstrando a relação entre inclinação e política monetária). Segundo os autores, existem duas possíveis explicações:
  - A autoridade monetária reage às mudanças de curto prazo na ETTJ, mudando a taxa de referência.
  - O mercado antecipa as decisões da autoridade monetária devido às fricções institucionais, que criam uma defasagem entre a avaliação de conjuntura econômica e a tomada de decisão pelos *policemakers*.
3. Nível: eleva a capacidade utilizada, a taxa de juros de referência e a taxa de inflação. O motivo é que o nível está associado a uma percepção do mercado por um aumento da taxa de inflação no futuro. A reação das autoridades monetárias de acomodar somente uma parte da inflação esperada tem um impacto negativo sobre a taxa de juros real ex-ante, o que contribui para acelerar a atividade econômica.

Por outro lado, o efeito dos choques das variáveis macroeconômicas sobre os fatores da curva de juros verificado pelos autores são:

1. Curvatura: quase nenhum impacto.
2. Inclinação: é impactada de forma significativa pelas variáveis macroeconômicas. Uma elevação da taxa de juros de referência torna a ETTJ mais achatada, afetando negativamente o fator de inclinação. Choques positivos sobre o produto e sobre o nível de preços também têm efeito negativo, estando de acordo com a função de estabilização da inflação e do produto por parte da autoridade monetária, que, nesse caso, eleva a taxa de juros de referência.
3. Nível: se eleva de acordo com choques positivos nas variáveis macroeconômicas. Surpresas inflacionárias têm efeito de longo prazo, enquanto que o impacto de um aumento da taxa de juros de referência apresenta dois possíveis efeitos:

<sup>4</sup> Para mais detalhes, vide [Diebold, Rudebusch e Aruoba \(2006, pp.325-326\)](#).

- Efeito positivo, caso os agentes acreditem que a reação da autoridade monetária foi inesperada, indicando maior preocupação com a inflação futura.
- Efeito nulo ou negativo, caso haja confiança na política monetária, podendo os agentes econômicos anteciparem, inclusive, uma redução da meta para a inflação futura.

### 3.4. A Abordagem de Chen e Tsang: O Modelo MA-DNS Relativo (R-MA-DNS)

Chen e Tsang (2011) reformularam o modelo MA-DNS em formato de estado-espço para o contexto internacional com o objetivo de estudar as inter-relações entre as variações cambiais, as expectativas, o risco e a conjuntura macroeconômica. Os componentes do modelo R-MA-DNS (*Relative Macro Augmented Dynamic Nelson-Siegel*), sendo eles os macro e os fatores latentes de NS, são computados pela diferença entre o valor observado em uma pequena economia aberta, no caso do estudo em questão, no Canadá, Reino Unido e no Japão, em relação ao verificado nos Estados Unidos, considerados uma grande economia aberta.

Os autores compararam o poder de previsão da taxa de câmbio a partir de três especificações para os seus fundamentos: a primeira, com as variáveis macro (*Macro-Only*); a segunda, com os fatores latentes NS (*Yields-Only*); e, a terceira, com todas as variáveis (Macro-Yields). Os resultados apontaram que o modelo Macro-Yields apresentou melhor performance para a libra esterlina e também para o iene enquanto que o modelo Macro se saiu melhor em prever movimentos do dólar canadense. Todos os modelos superaram o passeio aleatório, em termos dos MSFEs, nos horizontes de previsão considerados: 3, 6, 12 e 24 meses.

O modelo R-MA-DNS é baseado na reestruturação da equação 3.3 que agora passa a considerar a curva de juros em formato relativo, conforme a seguinte equação:

$$i_t(m)^R = L_t^R + S_t^R \left( \frac{1 - e^{-\lambda m}}{\lambda m} \right) + C_t^R \left( \frac{1 - e^{-\lambda m}}{\lambda m} - e^{-\lambda m} \right) + v_t \quad (3.4)$$

Em que o sobrescrito  $R$  indica que o fator está representado em sua forma relativa,  $i_t(m)^R = i_t(m) - i_t(m)^*$ , ou seja, da pequena economia aberta em relação à grande economia aberta, no caso, em relação aos Estados Unidos.

Há três tipos de abordagem para os fundamentos: a primeira, Yields-Only, trata-se do modelo DNS-Only, entretanto, com os fatores sendo representados em sua forma relativa. A segunda, Macro-Only, considera somente as variáveis macroeconômicas em formato relativo em uma estrutura VAR de ordem 1, estimada pela equação de transição. Já a abordagem conjunta Macro-Yields representa o modelo MA-DNS-Relativo completo, que considera ambos os fundamentos em formato relativo, em que na equação de men-

suração são estimados os fatores NS dinâmicos e, na equação de mensuração em formato VAR(1), a inter-relação entre os fundamentos, captada pela matriz de coeficientes.

Abaixo apresenta-se o modelo R-MA-DNS (*Relative Macro Augmented Dynamic Nelson-Siegel Model*) em formato de estado espaço:

Fundamentos (3 especificações):

- *Macro-Only*:  $f_t^M = (ds_t, \pi_t^R, \tilde{y}_t^R)'$
- *Yields-Only*:  $f_t^Y = (ds_t, L_t^R, S_t^R, C_t^R)'$
- **Macro-Yields**:  $f_t^{MY} = (ds_t, \pi_t^R, \tilde{y}_t^R, L_t^R, S_t^R, C_t^R)'$

**Equação de Mensuração:**  $\bar{y} \equiv y - \Lambda\mu = \Lambda(f_t^{MY} - \mu) + \epsilon_t$

Em que  $\Lambda$  é a matriz com os fatores de carregamento, que deve ser ajustada de acordo com o vetor de fundamentos. No caso em que  $f_t = f_t^{MY}$  temos a inclusão de três colunas iguais a zero:

$$\Lambda = \begin{pmatrix} 0 & 0 & 0 & 1 & \frac{1-e^{-\lambda m_1}}{\lambda m_1} & \frac{1-e^{-\lambda m_1}}{\lambda m_1} - e^{-\lambda m_1} \\ 0 & 0 & 0 & 1 & \frac{1-e^{-\lambda m_2}}{\lambda m_2} & \frac{1-e^{-\lambda m_2}}{\lambda m_2} - e^{-\lambda m_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & \frac{1-e^{-\lambda m_n}}{\lambda m_n} & \frac{1-e^{-\lambda m_n}}{\lambda m_n} - e^{-\lambda m_n} \end{pmatrix}_{14,6}$$

**Equação de Transição:**  $f_t^{MY} - \mu = A(f_{t-1}^{MY} - \mu) + u_t ; u_t \sim iidN(0, Q)$

**Equação de Mensuração em formato matricial:**

$$\begin{pmatrix} i_t(m_1) - i_t(m_1)^* \\ i_t(m_2) - i_t(m_2)^* \\ \vdots \\ i_t(m_n) - i_t(m_n)^* \end{pmatrix}_{n,1} - \Lambda \begin{pmatrix} \mu_{ds} \\ \mu_{\pi^R} \\ \mu_{\tilde{y}^R} \\ \mu_{L^R} \\ \mu_{S^R} \\ \mu_{C^R} \end{pmatrix}_{6,1} = \Lambda \begin{pmatrix} ds_{t-1} - \mu_{ds} \\ \pi_{t-1}^R - \mu_{\pi^R} \\ \tilde{y}_{t-1}^R - \mu_{\tilde{y}^R} \\ L_{t-1}^R - \mu_{L^R} \\ S_{t-1}^R - \mu_{S^R} \\ C_{t-1}^R - \mu_{C^R} \end{pmatrix}_{6,1} - \begin{pmatrix} \epsilon_t(m_1) \\ \epsilon_t(m_2) \\ \vdots \\ \epsilon_t(m_n) \end{pmatrix}_{n,1}$$

**Equação de Transição em formato matricial:**

$$\begin{pmatrix} ds_t - \mu_{ds} \\ \pi_t^R - \mu_{\pi^R} \\ \tilde{y}_t^R - \mu_{\tilde{y}^R} \\ L_t^R - \mu_{L^R} \\ S_t^R - \mu_{S^R} \\ C_t^R - \mu_{C^R} \end{pmatrix}_{6,1} = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} & a_{1,5} & a_{1,6} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} & a_{2,5} & a_{2,6} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} & a_{3,5} & a_{3,6} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} & a_{4,5} & a_{4,6} \\ a_{5,1} & a_{5,2} & a_{5,3} & a_{5,4} & a_{5,5} & a_{5,6} \\ a_{6,1} & a_{6,2} & a_{6,3} & a_{6,4} & a_{6,5} & a_{6,6} \end{pmatrix}_{6,6} \begin{pmatrix} ds_{t-1} - \mu_{ds} \\ \pi_{t-1}^R - \mu_{\pi^R} \\ \tilde{y}_{t-1}^R - \mu_{\tilde{y}^R} \\ L_{t-1}^R - \mu_{L^R} \\ S_{t-1}^R - \mu_{S^R} \\ C_{t-1}^R - \mu_{C^R} \end{pmatrix}_{6,1} + \begin{pmatrix} u_t(\tilde{y}^R) \\ u_t(\pi^R) \\ u_t(ds^R) \\ u_t(L^R) \\ u_t(S^R) \\ u_t(C^R) \end{pmatrix}_{6,1}$$

Condições para Estimação por filtro de Kalman:

$$\begin{bmatrix} \epsilon_t \\ u_t \end{bmatrix} \sim WR \left[ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{pmatrix} \Sigma & 0 \\ 0 & Q \end{pmatrix} \right]$$

$$E[f_0^{MY} \cdot \epsilon_t] = 0 \quad \text{e} \quad E[f_0^{MY} \cdot u_t] = 0$$

Em que:

- $R$ : o sobrescrito indica que o fundamento está expresso pela diferença entre o valor do país local e o do país estrangeiro.
- $y_t^R$ : é o vetor contendo o diferencial entre países das taxas de juros dos títulos zero cupom para cada uma das maturidades  $m \in [1, n]$ . Nada mais é do que a representação vetorial do lado esquerdo da equação 3.3 para todas as maturidades em determinado ponto no tempo  $t$ .
- $\bar{y}_t$ : é um ajuste ao vetor  $y$ , contendo variáveis dependentes da equação de mensuração, de forma a compatibilizá-las como desvios da média.
- $\mu$ : representa o vetor das médias das séries de tempo.
- $ds_t$ : representa a variação da taxa de câmbio.
- $\pi_t^R$ : é o diferencial das taxas de inflação.
- $\tilde{y}_t^R$ : é o diferencial da atividade econômica, representada pelo hiato do produto.
- $A$ : é a matriz de coeficientes relacionada ao sistema VAR(1) não estrutural da equação de transição, permitindo captar a inter-relação entre as variáveis macro e os fatores latentes da curva de juros.

- $Q$ : é a matriz de variância-covariância para a equação de transição. Assume-se que seja não diagonal.
- $\Sigma$ : é a matriz de variância-covariância para a equação de mensuração, especificada como uma matriz diagonal, de forma a evitar que os desvios das taxas de juros de diferentes maturidades em relação à curva de juros sejam não correlacionados, como discutido por [Lange \(2014\)](#).

As condições impostas sobre o comportamento dos resíduos ( $\epsilon_t$  e  $u_t$ ), de transversalidade em relação ao estado inicial e sobre as matrizes de variância-covariância das equações de mensuração e de transição ( $\Sigma$  e  $Q$ ) para estimação por Filtro de Kalman são semelhantes às do modelo *DNS-Only*.

### 3.4.1. O Filtro de Kalman:

De acordo com [Aiube \(2013\)](#):

O filtro de Kalman é um conjunto de equações matemáticas que constitui um processo recursivo eficiente de estimação, uma vez que o erro quadrático é minimizado. Através da observação da variável denominada "variável de observação" outra variável, não observável, denominada "variável de estado" pode ser estimada eficientemente. Podem ser estimados os estados passados, o estado presente e mesmo previstos os estados futuros. ([AIUBE, 2013](#), p.77–78)

Conforme [Kim, Nelson \*et al.\* \(1999\)](#) o Filtro de Kalman apresenta duas etapas recursivas, uma de predição e outra de atualização (correção), totalizando seis equações, mais o ganho de Kalman.

#### 1. Equações de Predição:

- Predição *a priori* das variáveis de estado:

$$f_{t|t-1} - \mu = A(f_{t-1|t-1} - \mu)$$

- Predição *a priori* da matriz de covariâncias das variáveis de estado:

$$V_{t|t-1} = AV_{t-1|t-1}A' + Q$$

- Resíduos da estimativa *a priori*:

$$\eta_{t|t-1} = y_t^* - y_{t|t-1}^* = y_t^* - \Lambda(f_{t|t-1} - \mu)$$

- Matriz de variância do erro de predição *a priori*:

$$S_{t|t-1} = \Lambda V_{t|t-1} \Lambda' + \Sigma \equiv E[\eta_{t|t-1}^2]$$

#### 2. Equações de Atualização (Correção):



- o Estado atualizado (estimativa *a posteriori*):

$$f_{t|t} = f_{t|t-1} + K_t \eta_{t|t-1}$$

- o Covariância atualizada: (estimativa *a posteriori*)

$$V_{t|t} = V_{t|t-1} - K_t \Lambda V_{t|t-1}$$

### 3. O Ganho de Kalman:

- o  $K_t = V_{t|t-1} \Lambda' (S_{t|t-1})^{-1} = V_{t|t-1} \Lambda' (\Lambda V_{t|t-1} \Lambda' + \Sigma)^{-1}$

De acordo com [Aiube \(2013, pp. 77–78\)](#) as equações de predição "...são responsáveis pelo avanço das variáveis de estado e das covariâncias no tempo para se obter, desta forma, as estimativas anteriores (a priori) para o próximo instante" e, sobre as equações de atualização: "...são responsáveis pela retroalimentação, ou seja, incorporam uma nova informação da variável observável nas estimativas anteriores para obter um ganho (ou melhoria) na estimação posterior".

Em um primeiro momento obtêm-se os valores das equações de predição com as estimativas *a priori*. Em seguida, utiliza-se essa informação para determinar o ganho de Kalman. Por fim, obtêm-se os valores atualizados para as variáveis de estado e para a sua matriz de covariâncias, melhorando a estimação. O procedimento ocorre de forma iterativa para cada novo período de tempo.

#### 3.4.1.1. A Maximização da Função de Verossimilhança

Segundo [Aiube \(2013\)](#), se os termos de erro das equações de mensuração, de transição e do vetor de estado inicial forem gaussianos, então a distribuição condicional de  $y_t$  em relação à informação até  $t - 1$ , dada por  $Y_{t-1}$ , também será gaussiana, sendo possível utilizar o Filtro de Kalman para a estimação dos parâmetros por meio da maximização da função de verossimilhança.

Os parâmetros são obtidos, de acordo com [Commandeur e Koopman \(2007\)](#), por meio da maximização da função de verossimilhança, em relação ao vetor de hiperparâmetros  $\Theta$  (que depende do vetor de estados não observados), de forma iterativa, para cada período de tempo.

$$\ln L(y|\Theta) = -\frac{NT}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=1}^T \ln[S_t] - \frac{1}{2} \sum_{t=1}^T u'_{t|t-1} S_t^{-1} u_{t|t-1}$$

$$\hat{\Theta} = \arg.\max_{\Theta} \ln L(y|\Theta)$$

## 4 METODOLOGIA

### 4.1. O Modelo TVP-VAR

Este trabalho propõe uma alteração no modelo R-MA-DNS de [Chen e Tsang \(2011\)](#) de forma a permitir que os coeficientes da matriz de transição do modelo R-MA-DNS variem no tempo, o que se espera que venha a melhorar o poder de previsão para a taxa de câmbio em comparação aos modelos com coeficientes fixos.

Três especificações para os fundamentos, representados pelo vetor  $f_t$ , são mostradas abaixo<sup>1</sup>:

- *Macro-Only*:  $f_t^M = (ds_t, \pi_t^R, \tilde{y}_t^R)'$
- *Yields-Only*:  $f_t^Y = (ds_t, L_t^R, S_t^R, C_t^R)'$
- *Macro-Yields*:  $f_t^{MY} = (ds_t, \pi_t^R, \tilde{y}_t^R, L_t^R, S_t^R, C_t^R)'$

A estimação dinâmica dos fatores latentes de Nelson-Siegel consiste em aplicar Mínimos Quadrados Ordinários (OLS) para todas as maturidades em um ponto específico no tempo e, depois, iterar o procedimento para todo o período considerado.

A representação em formato de estado-espço do modelo TVP-VAR:

**Equação de Mensuração:**

$$f_t - \mu = Z_t A_t + \epsilon_t; \quad \epsilon_t \sim iidN(0, \Sigma_t) \quad (4.1)$$

Em que:

$$Z_t = \begin{pmatrix} z_t' & 0 & \dots & 0 \\ 0 & z_t' & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & z_t' \end{pmatrix}_{M,k}$$

**Equação de Transição:**

$$A_t = A_{t-1} + u_t; \quad u_t \sim iidN(0, Q_t) \quad (4.2)$$

As condições para as equações de transição e de mensuração são:

$$\begin{bmatrix} u_t \\ \epsilon_t \end{bmatrix} \sim N \left[ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{pmatrix} Q_t & 0 \\ 0 & \Sigma_t \end{pmatrix} \right]$$

<sup>1</sup> Para maiores detalhes sobre cada uma das variáveis incluídas nos fundamentos, vide o modelo R-MA-DNS: seção 3.4.

$$E[A_0\epsilon_t] = 0 \quad \text{e} \quad E[A_0u_t] = 0$$

Em que:

- $t \in [1, T]$  representa o período de tempo.
- $f_t$  é o vetor de fundamentos com dimensão  $M \times 1$ , em que  $M$  é o número de variáveis consideradas (séries temporais).
- $z_t$  é o vetor que inclui um intercepto mais  $t - 1$  até  $p$  defasagens do vetor de fundamentos  $f_t$  em relação à média,  $\mu$ .
- $Z_t$  é uma matriz composta por  $M$  repetições do vetor  $z_t$ , apresentando dimensão  $M \times k$ , em que  $k = M(1 + M \cdot p)$ .  $Z_t$  permite a inclusão de defasagens acima de 1 período na equação de mensuração.
- $A_t$  é o vetor de estado, com dimensão  $k \times 1$ .
- $Q_t$  e  $\Sigma_t$  são as matrizes de variância-covariância das equações de transição e de mensuração, respectivamente. Existem duas condições para os erros das equações 4.1 e 4.2: ambos são conjuntamente independentes e ambos são não serialmente correlacionados com o estado inicial  $A_0$ .

#### 4.1.1. Estimação de Modelos TVP-VAR com Fatores de Esquecimento

Com a finalidade de contornar o problema da dimensionalidade, [Koop e Korobilis \(2013\)](#) recorreram à inclusão de fatores de esquecimento como forma de aproximar as estimativas para as matrizes  $Q_t$  e  $\Sigma_t$ . De acordo com [Raftery, Kárný e Ettlér \(2010\)](#) o fator de esquecimento assume a fórmula do decaimento exponencial em que a  $i$ -ésima informação mais recente apresenta peso na estimação dada por  $\lambda_{ff}^i$  e o total de informação recente efetivamente utilizado é dado por  $horizon = \frac{1}{1-\lambda_{ff}}$ . No limite, quando  $\lambda_{ff} = 1$ , considera-se toda a amostra, e os coeficientes passam a ser fixos. O subscrito  $ff$  (*forgetting factor*) serve para diferenciar o fator de esquecimento do fator de decaimento do modelo NS.

Um dos principais passos envolve a distribuição de  $A_{t-1}|I_{t-1} \sim N(A_{t-1|t-1}, V_{t-1|t-1})$ , ou seja, a distribuição condicional das variáveis de estado até  $t - 1$ , dada toda a informação passada observada até o período anterior  $I_{t-1}$ . O filtro então é utilizado para obter o estado filtrado  $A_t|I_{t-1} \sim N(A_{t|t-1}, V_{t|t-1})$  e, ao incluir o fator de esquecimento  $\lambda_{ff} \in (0, 1]$ , é possível substituir a expressão  $V_{t|t-1} = V_{t-1|t-1} + Q_t$  por  $V_{t|t-1} = \frac{1}{\lambda_{ff}} V_{t-1|t-1}$ , não sendo

mais necessário a direta e custosa estimação de  $Q_t$ . Além disso, [Koop e Korobilis \(2013\)](#) se baseiam em [Park, Jun e Kim \(1991\)](#) para estimar o parâmetro  $\lambda_{ff,t}$  de forma variável no tempo<sup>2</sup>.

De maneira semelhante [Koop e Korobilis \(2013\)](#) reduzem a matriz de variância-covariância da equação de mensuração  $\Sigma_t$  por meio de uma média móvel exponencial ponderada (EWMA) utilizando o fator de decaimento<sup>3</sup>  $\kappa$ :  $\hat{\Sigma}_t = \kappa\hat{\Sigma}_{t-1} + (1 - \kappa)\tilde{\epsilon}_t\tilde{\epsilon}_t'$  em que  $\tilde{\epsilon}_t$  é o erro de predição da primeira etapa do filtro de Kalman.

#### 4.1.1.1. O Filtro de Kalman com Fatores de Esquecimento

Conforme [Koop e Korobilis \(2013\)](#) o Filtro de Kalman, com a inclusão de fatores de esquecimento, segue as seguintes etapas:

##### 1. Equações de Predição:

- Predição das variáveis de estado:

$$A_{t|t-1} = A_{t-1|t-1}$$

- Estimação do fator de esquecimento:

$$\lambda_{ff,t} = \lambda_{ff,min} + (1 - \lambda_{ff,min})L^{ft};$$

$$\text{Em que: } f_t = -NINT(\tilde{\epsilon}_t'\tilde{\epsilon}_t)$$

- Predição da variância condicional de  $A_t$ :

$$V_{t|t-1} = \frac{1}{\lambda_{ff,t}}V_{t-1|t-1}$$

Agora não há a necessidade de se calcular a matriz de covariâncias  $Q_t$  devido à inclusão do fator de esquecimento  $\lambda_{ff,t}$ .

- Resíduos da equação de mensuração:

$$\tilde{\epsilon}_t = (f_t - \mu) - Z_t A_{t|t-1}$$

- Estimação da variância da equação de mensuração:

$$\hat{\Sigma}_t = \kappa\hat{\Sigma}_{t-1} + (1 - \kappa)\tilde{\epsilon}_t\tilde{\epsilon}_t'; \text{ para } t = 1, \hat{\Sigma}_1 = \kappa\hat{\Sigma}_0$$

Em que  $\kappa$  é o fator de decaimento.

##### 2. Equações de Atualização:

- Estado atualizado:  $A_{t|t} = A_{t|t-1} + K_t\tilde{\epsilon}_t$

- Covariância atualizada:  $V_{t|t} = V_{t|t-1} - K_t Z_t V_{t|t-1}$

##### 3. O Ganho de Kalman:

- $K_t = V_{t|t-1}Z_t'(Z_t V_{t|t-1}Z_t' + \hat{\Sigma}_t)^{-1}$

<sup>2</sup> Conforme mostrado na seção do Filtro de Kalman com Fatores de Esquecimento.

<sup>3</sup> O padrão da literatura é chamar o fator de esquecimento relacionado à volatilidade da equação de mensuração de fator de decaimento.

O Filtro de Kalman, quando o estado inicial é desconhecido, depende de *priors*, ou seja, de uma formulação inicial para os parâmetros, de forma a ser inicializado. Entretanto, como apontado por Kilian e Lütkepohl (2017, cap. 5) a especificação completa da distribuição *a priori* costuma ser impraticável, sendo necessário reduzir o número de parâmetros para um conjunto menor de hiperparâmetros. O *prior* de Minnesota, dentro da abordagem Bayesiana, adaptado em Koop e Korobilis (2013), é um dos principais métodos para enfrentar o problema da dimensionalidade ao impor uma estrutura para os parâmetros *a priori*.

## 4.2. Os Modelos DMA e DMS

Raftery, Kárný e Ettlér (2010) ao lidar, dentro da área de engenharia, com a incerteza na escolha do melhor modelo, a partir de diferentes especificações, para a previsão da largura de uma tira<sup>4</sup>, desenvolveram o método bayesiano da combinação dinâmica de modelos (DMA), que ponderava as novas previsões geradas por cada um dos modelos conferindo maior peso de acordo com as performances preditivas observadas até o momento. O DMA aproveita a informação gerada por diferentes modelos, reduzindo a perda de dados relevantes, e, simultaneamente, também permite que a ponderação de cada modelo varie, levando em conta que um mesmo modelo pode apresentar diferentes performances em períodos distintos.

Por sua vez, Koop e Korobilis (2013), com o intuito de contornar o problema da dimensionalidade dentro da área de macroeconometria, realizaram algumas alterações na formulação DMA original possibilitando diferentes dimensões (número de variáveis) nas estruturas TVP-VAR (3, 7 e 25 variáveis foram analisadas). Os autores desenvolveram o procedimento de seleção dinâmica de modelos (DMS), a partir da abordagem DMA, que, em vez de ponderar os modelos de acordo com sua capacidade preditiva, escolhe somente a especificação que apresenta o maior valor de  $\pi_{t|t-1,j}$ <sup>5</sup>. Além disso, os autores modificaram o modelo DMA de forma a proporcionar diferentes escolhas para os *priors* e para os fatores de esquecimento, a cada período, o que também é implementado pelo algoritmo DMS.

### 4.2.1. A Combinação Dinâmica de Modelos (DMA)

Segundo Caldeira, Moura e Santos (2015) a ideia principal do modelo DMA (*Dynamic Model Averaging*) consiste em:

<sup>4</sup> Trata-se de um procedimento de laminação a frio.

<sup>5</sup> probabilidade preditiva dada toda informação disponível até o período anterior para o modelo  $j$  pertencente a  $J$  (conjunto de modelos).

A partir de previsões oriundas de todos os modelos, a abordagem DMA calcula a probabilidade de que o modelo  $j$  é o melhor modelo para prever observações do período  $t$ , dadas as informações disponíveis em  $t - 1$ , ou seja,  $Prob(L_t = j|A_{t-1}) = \pi_{t|t-1,j}$ . Uma vez calculado  $\pi_{t|t-1,j}$ , essas probabilidades podem ser usadas para fazer combinação das previsões dos modelos. (CALDEIRA; MOURA; SANTOS, 2015, p. 412)

O procedimento desenvolvido por Raftery, Kárný e Ettlér (2010) expande o Filtro de Kalman ao incluir uma etapa específica para a correção e para a atualização das  $J$  especificações do TVP-VAR. Assume-se que o processo, por trás do sistema, apresenta mudanças pouco frequentes e que a sua evolução é determinada pela matriz de transição  $Q = q_{j,l}$ , com dimensão  $J * J$ , equivalente ao número de especificações distintas do TVP VAR, e a probabilidade do modelo  $l$  ocorrer, dado que no período anterior a estrutura  $j$  havia ocorrido, é dada por:  $q_{j,l} = P[L_t = l|L_{t-1} = j]$ .

As etapas adicionais de predição e de atualização do Filtro de Kalman para os modelos, a partir de uma probabilidade inicial *a priori*  $\pi_{0|0,j}$ , são:

- o Etapa de predição para os modelos:

$$\pi_{t|t-1,j} \equiv P[L_t = j|y_{t-1}] = \sum_{l=1}^J \pi_{t-1|t-1,l} * q_{j,l}$$

Uma forma de evitar especificar explicitamente a matriz de transição relacionada a  $q_{j,l}$ , e, assim, de reduzir custos computacionais, é por meio da inclusão do fator de esquecimento  $a$ , um pouco menor do que 1, na equação acima:

$$\pi_{t|t-1,j} = \frac{\pi_{t-1|t-1,j}^a}{\sum_{l=1}^J \pi_{t-1|t-1,l}^a}$$

- o Etapa de atualização para os modelos:

$$\pi_{t|t,j} = \frac{w_{t,j}}{\sum_{l=1}^J w_{t,l}} = \frac{\pi_{t|t-1,j} p_j(f_t|f_{t-1})}{\sum_{l=1}^J \pi_{t|t-1,l} p_l(f_t|f_{t-1})}$$

Em que  $w_t$  é a ponderação de cada especificação do TVP-VAR com base na contribuição preditiva esperada de cada modelo no período anterior. Conforme Caldeira, Moura e Santos (2015, pg. 412): "...  $p_j(f_t|f_{t-1})$  é a densidade preditiva (verossimilhança preditiva) do modelo  $j$  avaliada em  $f_t$ . Esta densidade preditiva é obtida através do filtro de Kalman e possui fórmula fechada".

Segundo Koop e Korobilis (2013) a probabilidade preditiva para dado modelo  $j$  pode ser representada da seguinte forma aproximada  $\pi_{t|t-1,j} \propto \prod_{i=1}^{t-1} [p_j(y_{t-i}|y_{t-i-1})]^{a^i}$  e,

portanto, o modelo  $j$  receberá um peso maior em  $t$  se até o período anterior  $t - 1$  a sua performance preditiva, medida pela função de densidade preditiva  $p_j(y_{t-i}|y_{t-i-1})$ , tenha sido elevada em comparação aos outros modelos.

Por fim, de acordo com [Raftery, Kárný e Ettlér \(2010\)](#), a etapa de previsão um passo a frente, do modelo DMA, para a variável de observação, é dada por:

$$\hat{f}_t^{DMA} = \sum_{j=1}^J \pi_{t|t-1,j} \hat{f}_t^{(j)}$$

#### 4.2.2. A Seleção Dinâmica de Modelos (DMS)

O algoritmo DMS (*Dynamic Model Selection*) procede de maneira semelhante ao DMA, entretanto, selecionando a especificação  $j$  que apresenta maior poder preditivo, sendo que agora  $\hat{f}_t^{DMA} = \hat{f}_t^{(j=best)}$ . Os autores utilizam o DMS de forma a realizar seleção dinâmica: do fator de esquecimento  $\lambda_{ff}$  (*Dynamic Forgetting Selection* - DFS), do hiperparâmetro de encolhimento  $\gamma$  (*Dynamic Shrinkage Selection* - DSS) e, por fim, para escolher a melhor especificação do TVP-VAR de acordo com as suas diferentes dimensões (*Dynamic Dimension Selection* - DDS).

##### 4.2.2.1. O Modelo DSS (*Dynamic Shrinkage Selection*)

[Koop e Korobilis \(2013\)](#) impuseram uma estrutura para a distribuição *a priori* dos parâmetros  $A_{0|0}$  a partir de algumas alterações do *prior* de Minnesota. Os autores selecionaram o parâmetro de encolhimento para o TVP-VAR de forma dinâmica, escolhendo, para cada período de tempo  $t$ , o valor de  $\gamma$  com maior poder preditivo. Dado um modelo TVP-VAR de dimensão fixa, utiliza-se a distribuição normal para o comportamento dos *priors*  $A_0$ . Além disso, a estacionariedade dos dados permite fixar  $E[A_0] = 0$ .

Define-se a matriz de Minnesota da covariância dos *priors* da seguinte forma:

$$\underline{V}_i = \begin{cases} \frac{\gamma}{r^2} & \text{para coeficientes com defasagem } r=1\dots p, \\ \underline{a} & \text{para os interceptos.} \end{cases}$$

Em que a covariância *a priori* é diagonal, de forma que  $var(A_0) = \underline{V}$  e  $\underline{V}_i$  são os elementos da diagonal. Além disso,  $p$  é o total de defasagens e  $\gamma$  apresenta 7 valores possíveis para o hiperparâmetro de encolhimento, cuja especificação encontra-se na seção [4.3: "Estratégia de Estimação"](#).

De acordo com os autores a estimação dos *priors* e dos valores ótimos para os hiperparâmetros de encolhimento se baseia nas seguintes etapas:

1. Inicialização para todos os valores de  $\gamma$ ,  $\Sigma_0$  e  $A_0^j \sim N(0, \underline{V}^j)$  e, além disso, assume-se que a probabilidade do modelo inicial é não-informativa, apresentando o mesmo peso para os 7 diferentes valores de  $\gamma$  ( $\#J = 7$ ):  $\pi_{0|0,j} = \frac{1}{J}$ <sup>6</sup>.
2. Obter a probabilidade de predição:  $\pi_{t|t-1,j} = \frac{\pi_{t-1|t-1,j}^a}{\sum_{l=1}^J \pi_{t-1|t-1,l}^a}$
3. Rodar a etapa de predição do Filtro de Kalman para cada um dos modelos  $j$ .
4. Obter a função de densidade de probabilidade preditiva um passo à frente para  $f_{t|t-1} - \mu \sim N(Z_t^{(j)} A_{t|t-1}^{(j)} - \mu, \hat{\Sigma}_t^{(j)} + Z_t^{(j)} V_{t|t-1}^{(j)} Z_t'^{(j)})$
5. Estimar a probabilidade atualizada:  $\pi_{t|t,j} = \frac{\pi_{t|t-1,j} p_j(y_t|y_{t-1})}{\sum_{l=1}^J \pi_{t|t-1,l} p_l(y_t|y_{t-1})}$

Por fim, o parâmetro de encolhimento que confere maior probabilidade para o modelo em cada período é escolhido.

#### 4.2.2.2. O Modelo DFS (Dynamic Forgetting Selection)

A aplicação do modelo DMS para a seleção dos valores para os fatores de esquecimento  $\lambda_{ff}$  segue o mesmo procedimento do modelo DSS. A estimação de  $\lambda_{ff}$  encontra-se na segunda equação de predição do Filtro de Kalman com Fatores de Esquecimento, subseção 4.1.1.1.

#### 4.2.2.3. O Modelo DDS (Dynamic Dimension Selection)

Após a realização da escolha dos *priors* e dos parâmetros de encolhimento, os autores utilizam o DMS para escolher, para cada período de tempo  $t$ , qual modelo TVP-VAR (pequeno, médio ou grande) apresenta melhor capacidade de previsão. O procedimento é semelhante ao DSS, exceto que a dimensão entre os modelos é variável e, por isso, as densidades preditivas  $N(Z_t^{(j)} A_{t|t-1}^{(j)} - \mu, \hat{\Sigma}_t^{(j)} + Z_t^{(j)} V_{t|t-1}^{(j)} Z_t'^{(j)})$  são calculadas a partir do VAR de menor dimensão.

### 4.3. Estratégia de Estimação - Modelos Considerados

Os modelos considerados para previsão são quase todos subespecificações do modelo de referência TVP-VAR-DMS ( $A_{T+h} = A_T$ ). Este modelo considera  $\alpha = 0.99$  (o que será comum a todas as outras subespecificações), realiza o DSS escolhendo, para cada período,  $\gamma \in \{10^{-10}, 10^{-5}, 0.001, 0.005, 0.01, 0.05, 0.1\}$ , além de também executar o DFS, obtendo o valor ótimo de  $\lambda_{ff} \in [0.96, 1.0]$ . Seguindo Morgan *et al.* (1996),  $\kappa$  é fixado em 0.96, sendo o modelo principal heteroscedástico. Para a realização de previsões assume-se

<sup>6</sup> Na prática, os autores utilizaram o menor valor de  $\gamma_0 = \min[\gamma_j] = 0.1$  para a amostra de treinamento.



que  $A_{T+h} = A_T$ , ou seja, que não há mudança nos coeficientes da estrutura VAR entre os períodos  $T$  (última observação) e  $T + h$  (período a ser previsto). Dessa forma, os coeficientes da equação de mensuração são estimados de forma iterativa, considerando não haver mudanças nos coeficientes da equação de transição. Duas defasagens foram escolhidas<sup>7</sup>, ajudando a captar a persistência temporal na evolução das variáveis.

Por sua vez, os modelos DMA e DMS consideram quatro modelos com dimensões (fundamentos) diferentes, sendo eles: *Macro-Only*, *Yields-Only* e o modelo completo, *Macro-Yields*, além de um quarto, o *Exchange-Only*. Cada modelo, exceto o último, é gerado a partir de especificações do vetor de fundamentos  $f_t$ , de acordo com [Chen e Tsang \(2011\)](#). Deve-se ressaltar que o TVP-VAR-DMS e DMA apresentarão dimensões diferentes e, por isso, as densidades preditivas não serão comparáveis. [Koop e Korobilis \(2013\)](#), com o objetivo de lidar com o problema, consideraram somente a densidade preditiva do modelo de menor dimensão, cujas variáveis também estavam incluídas nos modelos de dimensão maior. Entretanto, este trabalho segue de forma diferente, pois o modelo de menor dimensão não está incluído no modelo de dimensão intermediária. A solução foi adicionar uma quarta especificação para  $f_t$  com somente uma dimensão, o *Exchange-Only*<sup>8</sup>, dada pelas defasagens da variável de interesse, a taxa de câmbio, permitindo comparar as densidades preditivas.

### Modelos Considerados:

- TVP-VAR-DMA
- TVP-VAR-DMS
- TVP-VAR sem DDS:
  - *Macro-Only*:
 
$$f_t = (ds_t, \pi_t^R, \tilde{y}_t^R)'$$
  - *Yields-Only*:
 
$$f_t = (ds_t, L_t^R, S_t^R, C_t^R)'$$
  - *Macro-Yields*:
 
$$f_t = (ds_t, \pi_t^R, \tilde{y}_t^R, L_t^R, S_t^R, C_t^R)'$$
- TVP-VAR sem DFS:  $\lambda_{ff} = 0.99$ .

<sup>7</sup> A partir da comparação do critério AIC para diferentes estruturas temporais do VAR para o modelo maior dimensão, o macro-yields. Os resultados podem ser encontrados na tabela 6 do apêndice Dados.

<sup>8</sup> O modelo *Exchange-Only* não será representado individualmente por não ser o objetivo deste trabalho, que é avaliar a performance de previsão para a taxa de câmbio, a partir das especificações para os fundamentos de [Chen e Tsang \(2011\)](#).

- TVP-VAR ( $A_{T+h} \sim RW$ ): esta especificação permite que os coeficientes da estrutura VAR variem para previsões fora da amostra realizando simulações, a partir da estrutura em formato de passeio aleatório da equação de estado do TVP-VAR.
- Modelo VAR (com DSS):
  - VAR Heteroscedástico:  $\lambda_{ff} = 1$  e  $\kappa = 0.96$ .
  - VAR Homoscedástico:  $\lambda_{ff} = 1$  e  $\kappa = 1$ .
- Outros modelos para comparação (*Benchmarks*):
  - Passeio Aleatório (RW):  $ds_t = ds_{t-1} + e_t$ ; em que  $e_t \sim N(0, \sigma)$ .  
O valor da previsão h passos a frente será igual ao último valor observado:  
 $\hat{ds}_{T+h} = ds_T$ .
  - VAR-OLS (*Macro-Yields, Macro-Only e Yields-Only*):  
Especificações para os três modelos em formato VAR utilizando a estimação por Mínimos Quadrados Ordinários (OLS) em janela expansiva.

#### 4.4. Previsão

O horizonte de previsão considerado vai de um mês até dois anos ( $h = 24$  meses). Conforme ressaltado por [Koop e Korobilis \(2013\)](#), quando  $h > 1$ , não existe uma fórmula analítica fechada para as densidades preditivas. Logo, é necessário recorrer a simulações, o que é feito de duas formas: definindo  $A_{T+h} = A_T$  ou  $A_{T+h} \sim RW$ . O primeiro caso assume não haver mudança nos coeficientes VAR entre o período  $T$  e  $T+h$ . Já o segundo caso realiza simulações baseadas na estrutura em formato de passeio aleatório da equação de transição para que a matriz de coeficientes varie fora da amostra. Ambos os casos fornecem valores para  $A_{T+h}$  que serão utilizados nas estimativas para o valor condicional  $f_{t+h}|A_{T+h}$  utilizando simulações de forma a aproximar o valor condicional à densidade de probabilidade preditiva.

##### 4.4.1. Medidas de Desempenho e Comparação

Os Erros Absolutos Médios de Previsão (*Mean Absolute Forecasting Error - MAFE*) e o somatório do log das densidades de probabilidade preditiva são utilizados como medidas de desempenho.

De acordo com ([VALI, 2014](#), p. 374-375) o MAFE apresenta a seguinte estrutura:

$$MAFE = \frac{\sum_{i=1}^n |f_i - \hat{f}_i|}{n},$$

Em que  $f_i$  é o valor observado e  $\hat{f}_i$  é o valor estimado de previsão para o vetor de fundamentos. O valor absoluto evita o viés causado quando erros positivos e negativos se cancelam. Quanto menor o valor do MAFE melhor será o poder relativo do modelo em questão<sup>9</sup>.

A segunda métrica de comparação é o log da densidade de probabilidade preditiva  $s$  passos a frente:

$$\log[p(f_t|f_{t-s})] = \log[\int \int p(f_t|A_t)p(f_t|A_{t-s})p(A_{t-s}|f_{1:t-s})dA_t dA_{t-s}]$$

Quanto maior (menos negativo) for o valor em log da densidade preditiva, melhor será o poder de previsão<sup>10,11</sup>.

---

<sup>9</sup> A comparação é realizada pela divisão do MAFE de cada modelo pelo MAFE do modelo TVP-VAR-DMA:  $MAFE_{modelo(j)}/MAFE_{TVP-VAR-DMA}$ .

<sup>10</sup> Dado que a probabilidade encontra-se entre  $z \in (0, 1]$  e que o  $\lim_{z \rightarrow 0} \log(z) = -\infty$  e  $\log(z = 1) = 0$ .

<sup>11</sup> A comparação é realizada pela subtração da métrica de cada modelo em relação à métrica do modelo TVP-VAR-DMA.

## 5 DADOS

O período considerado: janeiro de 2004 a junho de 2017, totalizando 162 observações (frequência mensal) para os dados brutos. Para os dados transformados, o período vai de janeiro de 2005 até junho de 2017, contabilizando 150 meses.

### 5.1. Dados Utilizados

#### 5.1.1. As Taxas de Juros

As taxas de juros para ambos os países incluem maturidades de 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 36, 42, 48 e 60 meses, totalizando 14 maturidades. Para o caso dos Estados Unidos a principal base de dados utilizada será a de [Gürkaynak, Sack e Wright \(2007\)](#)<sup>1,2</sup>, que emprega o método de Svensson por ser flexível na modelagem da curva de juros e, ao mesmo tempo, por evitar o problema de sobreajuste dos dados. Para o caso do Brasil serão consideradas as taxas, obtidas da Anbima, associadas ao mercado interfinanceiro, DI Futuro da BM&F, por apresentar maior liquidez do que o mercado para títulos do governo.

#### 5.1.2. As Variáveis Macroeconômicas

Os dados da produção industrial e da inflação para os Estados Unidos (CPI) foram obtidos por meio da base de dados do FRED (Federal Reserve Economic Data)<sup>3</sup>. Para o Brasil, o IPCA e o índice de atividade da indústria foram obtidos por meio da base de dados do IBGE (SIDRA)<sup>4</sup>. A taxa de câmbio nominal à vista (PTAX) e as expectativas sobre a inflação e a atividade da indústria foram obtidas do Bacen (SGS)<sup>5</sup>.

$ds_t$ : representa a variação log anualizada da série mensal para a taxa spot de câmbio (dólar contra o real).

$\pi_t^R = \pi_t - \pi_t^*$ : representa a diferença da variação da taxa de inflação para o consumidor no país local em relação à variação no país estrangeiro, dada pelo asterisco. As taxas são computadas pela variação log-anualizada, de acordo com a seguinte aproximação:

$$\pi_t = \ln(\text{índice}(\pi_t) - \ln(\text{índice}(\pi_{t-12})).$$

<sup>1</sup> A base de dados atualizada encontra-se no seguinte link:

<https://www.federalreserve.gov/Pubs/feds/2006/200628/200628abs.html>.

<sup>2</sup> As taxas referentes às maturidades de 3,6, 9, 15, 18, 21, 27, 30 e 42 meses foram computadas a partir dos fatores NS disponibilizados na base de dados em [Gürkaynak, Sack e Wright \(2007\)](#).

<sup>3</sup> <https://fred.stlouisfed.org/>

<sup>4</sup> <https://sidra.ibge.gov.br/home/ipca/brasil>

<sup>5</sup> <https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>

$\tilde{y}_t^R = \tilde{y}_t - \tilde{y}_t^*$ : corresponde ao diferencial do hiato do produto entre os dois países utilizando dados mensais log linearizados da série de produção industrial. O hiato é obtido em duas etapas: extração da tendência, que será uma estimativa para o produto potencial, por meio do filtro HP. Em uma segunda etapa, o hiato é calculado de acordo com os desvios do produto observado em relação ao potencial<sup>6</sup>. Por fim, de forma a suavizar as variações, o hiato é computado em relação à média dos 12 últimos meses.

Os dados para a taxa de inflação e para a atividade industrial, divulgados no mês seguinte ao de referência, foram ajustados utilizando-se uma defasagem em relação ao mês de referência. Dessa forma, evita-se melhorar artificialmente as previsões provenientes de informações ainda não disponíveis.

A seção A.1 apresenta os gráficos para as curvas de juros: do Brasil, 5a; dos EUA, gráfico 5b; a ETTJ relativa, gráfico 6; e, por fim, os fatores relativos do modelo NS, gráfico 7.

## 5.2. Transformações

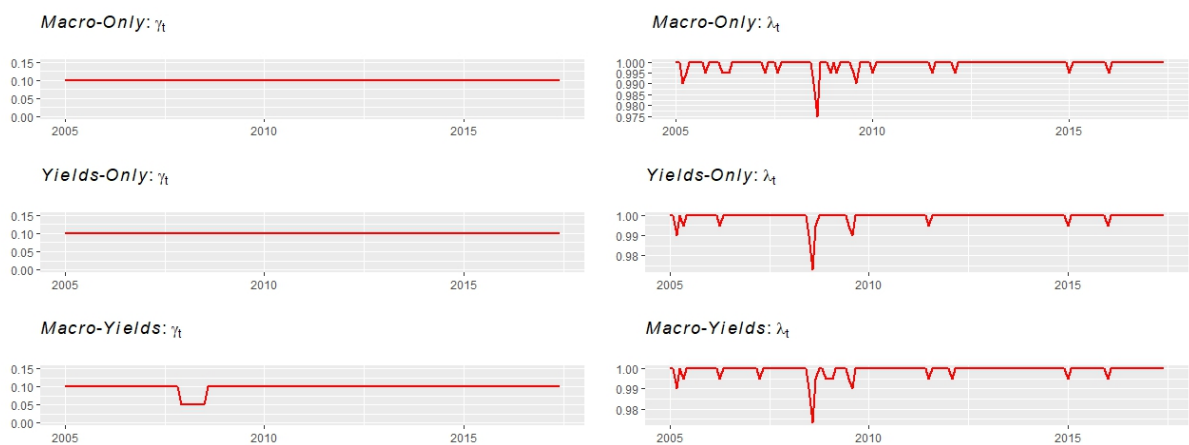
Com o intuito de compatibilizar o procedimento deste trabalho com a metodologia de Koop e Korobilis (2013), as variáveis foram normalizadas, ou seja, computadas pela diferença em relação à média dividida pelo desvio padrão. De forma a evitar dados sobrepostos, tanto as médias como os desvios padrão foram obtidos por meio de uma janela expansiva. Após a padronização, os testes de raiz unitária ADF, da tabela 3, e o KPSS e Phillips-Perron, da tabela 4 do Anexo sobre os Dados, foram utilizados. O critério para a diferenciação se baseou no caso em que 2 dos 3 testes avaliados apresentassem indícios para a não estacionariedade ou presença de raiz unitária, para o nível de significância de 5%. Nesse caso, a diferenciação foi realizada a partir dos dados brutos e a normalização foi implementada posteriormente. As variáveis diferenciadas foram a taxa de inflação relativa e os três fatores da curva de juros. As estatísticas descritivas das variáveis transformadas estão na tabela 5.

<sup>6</sup>  $\tilde{y}_t = (y_t - y_t^{potencial}) / y_t^{potencial}$ . Para mais detalhes, vide: <[http://repositorio.ipea.gov.br/bitstream/11058/2011/1/TD\\_1881.pdf](http://repositorio.ipea.gov.br/bitstream/11058/2011/1/TD_1881.pdf)>

## 6 RESULTADOS

### 6.1. Estimativas

O valor do coeficiente de encolhimento  $\gamma$  costuma se reduzir de acordo com o aumento do número de defasagens consideradas, evitando assim, sobreparametrização. Na figura 3a, devido à inclusão de somente duas defasagens ( $p = 2$ ), os valores se mantiveram no limite superior em quase todo o período considerado, exceto entre 2008 e 2009 para o modelo *Macro-Yields*, podendo refletir o aumento na volatilidade devido à crise financeira.



- (a) Fatores de encolhimento de Minnesota,  $\gamma$ , e- (b) Fatores de esquecimento,  $\lambda_{ff}$ , estimados, a cada período, para as diferentes especificações TVP-VAR.

Figura 3 – Fatores de Encolhimento  $\gamma$  e de Esquecimento  $\lambda_{ff}$  dos modelos TVP-VAR.

Os fatores de esquecimento  $\lambda_{ff}$ , que controlam o grau de variação dos coeficientes, em geral apresentaram valores iguais a 1, o que significa dizer que, na maior parte da amostra, os coeficientes não variaram de um mês para o outro. Entretanto, houve mudança nos coeficientes, sendo o intervalo mais significativo, e com maior redução no valor de  $\lambda_{ff}$ , o observado entre 2008 e 2010, indício de que os modelos conseguiram captar a volatilidade gerada pela crise financeira, além de outras possíveis quebras estruturais. Deve-se ressaltar que até mesmo valores de  $\lambda_{ff}$  ligeiramente inferiores a 1 conferem variabilidade à matriz de coeficientes  $A$ .

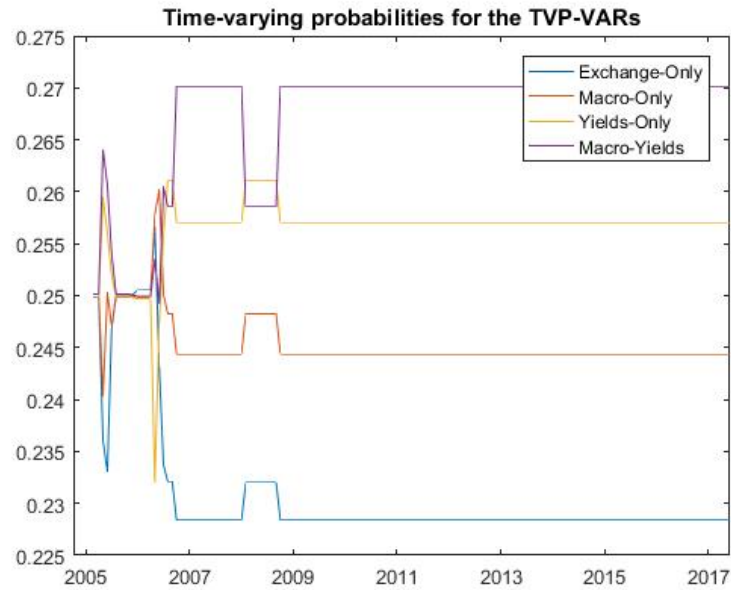


Figura 4 – Algoritmo DMA/DMS: probabilidades de dos modelos TVP-VAR a cada período.

Houve, em geral, estabilidade nas probabilidades de cada modelo, exceto pelo período inicial, entre 2005 até o final de 2006. De 2007 em diante os modelos que apresentaram maior probabilidade foram o *Macro-Yields*, o *Yields-Only*, o *Macro-Only* e o *Exchange-Only*, respectivamente. A escolha mais frequente do modelo *Macro-Yields* pelo algoritmo DMS é um indício de que combinar os fundamentos macro da Regra de Taylor para uma economia aberta com as expectativas embutidas na ETTJ, e captadas pelos fatores latentes Nelson-Siegel, pode melhorar a capacidade preditiva para a taxa de câmbio.

## 6.2. Previsão

A tabela 1 apresenta os erros absolutos médios de previsão (MAFE)<sup>1</sup> de cada modelo, para os horizontes de 1 a 24 meses, divididos pelos valores do modelo TVP-VAR-DMA. A primeira constatação é a de que quase todos os modelos superaram o passeio aleatório, em particular, para  $h > 3$ . As exceções são: o modelo VAR OLS-*Macro-Only* e o VAR OLS-*Macro-Yields*, que tiveram desempenho inferior para quase todo o período de previsão analisado. Por outro lado, os modelos TVP-VAR-DMS (com DFS) sobrepujaram o RW em todos os horizontes.

Outros resultados importantes são: o TVP-VAR-*Macro-Yields* apresentou melhor desempenho, em média, em comparação com os modelos TVP-VAR-*Yields-Only* e TVP-VAR-*Macro-Only*; simular o valor de  $A_{T+h} \sim RW$  não melhorou significativamente a performance dos modelos TVP-VAR-DMA/DMS, mas aumentou consideravelmente a

<sup>1</sup> Os resultados com os valores totais se encontram na tabela 7.

capacidade de previsão dos modelos TVP-VAR-*Yields-Only* e TVP-VAR-*Macro-Yields* para ( $h > 6$ ), piorando para ( $h \leq 6$ ), enquanto que para o modelo TVP-VAR-*Macro-Only* os resultados foram mistos; assumir erros heteroscedásticos não melhorou substancialmente as previsões; os resultados dos modelos TVP-VAR-DMA e TVP-VAR-DMS foram, em geral, similares; os modelos VAR OLS-*Yields-Only*, TVP-VAR-*Yields-Only* e o TVP-VAR-*Macro-Yields* performaram particularmente bem para horizontes acima de 12 meses, enquanto o TVP-VAR-*Macro-Only* se saiu melhor entre 1 e 3 meses. Deve-se ressaltar que nenhum modelo apresentou hegemonia em termos de previsão para os horizontes considerados. Na média, os modelos TVP-VAR-DMA e TVP-VAR-DMS (ambos com DFS), além do VAR OLS-*Yields-Only*, produziram as melhores previsões.

Tabela 1 – MAFE - Relativo ao modelo TVP-VAR-DMA

MAFE									
Modelos	h = 1	h = 3	h = 6	h = 9	h = 12	h = 15	h = 18	h = 21	h = 24
<i>Macro-Yields</i>									
TVP-VAR	1,13	1,00	1,01	1,09	1,12	1,10	1,12	1,06	0,97
TVP-VAR $A_{T+h} \sim RW$	1,20	1,15	1,10	0,94	0,92	0,94	0,96	0,95	0,92
TVP-VAR ( $\lambda_{ff} = 0.99$ )	1,11	1,00	1,02	1,11	1,15	1,13	1,16	1,08	0,99
VAR, heteroskedastic	1,14	1,00	1,00	1,09	1,11	1,09	1,11	1,05	0,96
VAR, homoscedastic	1,14	0,99	1,01	1,08	1,10	1,08	1,09	1,04	0,95
<i>Yields-Only</i>									
TVP-VAR	1,08	1,00	1,06	1,13	1,16	1,15	1,18	1,10	1,02
TVP-VAR $A_{T+h} \sim RW$	1,12	1,09	1,09	0,98	0,93	0,92	0,94	0,93	0,91
TVP-VAR ( $\lambda_{ff} = 0.99$ )	0,99	1,00	1,05	1,14	1,18	1,18	1,18	1,11	1,03
VAR, heteroskedastic	1,08	1,00	1,06	1,13	1,16	1,14	1,17	1,09	1,01
VAR, homoscedastic	1,10	0,98	1,06	1,12	1,14	1,13	1,15	1,08	1,00
<i>Macro-Only</i>									
TVP-VAR	1,01	0,99	1,04	1,12	1,15	1,15	1,15	1,09	1,01
TVP-VAR $A_{T+h} \sim RW$	1,01	1,05	1,04	0,96	0,99	1,04	1,07	1,09	1,15
TVP-VAR ( $\lambda_{ff} = 0.99$ )	0,99	1,00	1,05	1,14	1,18	1,18	1,18	1,11	1,03
VAR, heteroskedastic	1,03	0,98	1,03	1,19	1,07	1,12	1,15	1,08	1,00
VAR, homoscedastic	1,04	1,00	1,03	1,10	1,13	1,12	1,12	1,06	0,98
<i>Modelos DMA/DMS</i>									
TVP-VAR-DMA	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00
TVP-VAR-DMS	0,92	0,96	1,02	1,04	1,04	1,02	1,02	1,05	1,08
TVP-VAR-DMA $A_{T+h} \sim RW$	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,00	1,01
TVP-VAR-DMS $A_{T+h} \sim RW$	0,92	0,96	1,02	1,04	1,04	1,02	1,02	1,06	1,10
TVP-VAR-DMA ( $\lambda_{ff} = 0.99$ )	0,98	0,99	1,01	1,11	1,04	1,17	1,20	1,35	1,64
TVP-VAR-DMS ( $\lambda_{ff} = 0.99$ )	0,92	0,97	1,06	1,10	1,15	1,21	1,36	1,68	2,32
<i>Benchmarks</i>									
VAR OLS - Macro-Yields	1,18	1,23	1,20	1,11	1,17	1,31	1,46	1,61	1,91
VAR OLS - Yields-Only	1,04	1,06	1,05	1,02	0,95	0,93	0,90	0,82	0,99
VAR OLS - Macro-Only	1,02	1,18	1,18	1,08	1,16	1,32	1,46	1,67	1,91
Random Walk	0,97	1,00	1,08	1,17	1,28	1,38	1,38	1,26	1,09



A tabela 2 mostra a diferença entre o somatório do logaritmo das probabilidades preditivas de cada modelo TVP-VAR em relação ao TVP-VAR-DMA<sup>2</sup>, sendo que valores positivo (negativos) indicam desempenho superior (inferior) do modelo considerado (em comparação com o *benchmark*). Os resultados foram: o TVP-VAR-DMA foi, em média, superado em horizontes de 1 a 9 meses, saindo-se melhor para  $h \geq 12$ ; simulações com  $A_{T+h} \sim RW$  aumentaram substancialmente a performance de quase todos os modelos, em particular, do modelo DMA (para  $h > 3$ ); não houve ganhos aparentes a partir da especificação heteroscedástica; os modelos TVP-VAR-*Yields-Only* e TVP-VAR-*Macro-Yields* previram melhor os movimentos da taxa de câmbio do que o TVP-VAR-*Macro-Only*. Por fim, ressalta-se que nenhum modelo apresentou preponderância em termos de previsão, porém, em geral, os algoritmos DMA e DMS performaram relativamente bem, sendo que os modelos TVP-VAR-DMA ( $A_{T+h} \sim RW$ ) e TVP-VAR-DMS se destacaram para horizontes acima de 15 meses.

Tabela 2 – Somatório do log das probabilidades preditivas em comparação ao modelo TVP-VAR-DMA

<i>Sum of Log Predictive Likelihoods</i>									
Modelos	h = 1	h = 3	h = 6	h = 9	h = 12	h = 15	h = 18	h = 21	h = 24
Macro-Yields									
TVP-VAR	18,96	89,47	114,91	64,99	-28,61	-75,28	-80,16	-52,24	-19,97
TVP-VAR $A_{T+h} \sim RW$	14,43	88,05	155,90	179,43	118,61	58,60	21,45	3,22	-7,69
TVP-VAR ( $\lambda_{ff} = 0.99$ )	20,57	89,97	116,83	64,80	-28,94	-76,73	-85,91	-57,31	-24,65
VAR, heteroskedastic	18,61	89,24	114,79	66,04	-26,27	-71,73	-79,89	-50,02	-18,44
VAR, homoscedastic	-9,571	87,8	162,89	169,18	104,82	60,33	37,66	37,29	41,97
Yields-Only									
TVP-VAR	23,33	88,07	98,93	25,70	-72,31	-122,79	-120,85	-94,07	-53,66
TVP-VAR $A_{T+h} \sim RW$	20,14	90,26	161,47	183,16	122,62	62,78	25,02	6,19	-4,77
TVP-VAR ( $\lambda_{ff} = 0.99$ )	30,39	84,29	79,91	12,16	-79,23	-126,82	-123,24	-95,64	-55,59
VAR, heteroskedastic	23,62	87,00	98,89	26,90	-70,73	-121,50	-117,80	-90,65	-52,06
VAR, homoscedastic	-6,941	89,46	160,82	160,52	96,59	54,19	33,64	32,43	34,44
Macro-Only									
TVP-VAR	27,874	84,4	81,08	14,08	-78,48	-123,31	-119,35	-88,26	-47,56
TVP-VAR $A_{T+h} \sim RW$	27,782	90,65	157,89	179,63	116,26	54,22	14,55	-6,44	-19,95
TVP-VAR ( $\lambda_{ff} = 0.99$ )	30,247	84,65	80,62	13,02	-79,61	-124,16	-124,65	-96,55	-55,94
VAR, heteroskedastic	27,473	84,76	82,09	17,84	-77,68	-123,29	-115,21	-84,37	-44,41
VAR, homoscedastic	-4,844	91,02	156,51	159,03	94,9	51,37	32,94	34,34	37,9
Modelos Completos									
TVP-VAR-DMA	0	0	0	0	0	0	0	0	0
TVP-VAR-DMS	32,667	83,48	139,5	152,65	110,98	69,09	43,62	30,48	23,85
TVP-VAR-DMA $A_{T+h} \sim RW$	-10,481	-25,22	44,67	97,11	88,61	64,2	43,33	31,35	22,42
TVP-VAR-DMS $A_{T+h} \sim RW$	32,021	85,57	152,82	172,01	118,28	61,38	24,01	3,99	-8,02
TVP-VAR-DMA ( $\lambda_{ff} = 0.99$ )	7,162	19,56	41,54	67,16	65,18	51,28	35,39	13,91	-8,95
TVP-VAR-DMS ( $\lambda_{ff} = 0.99$ )	35,404	83,38	144,66	160,18	114,29	63,71	29,5	4,72	-12,73

<sup>2</sup> Vide a tabela 8 para os resultados em valores absolutos.

## 7 CONCLUSÃO

O objetivo deste trabalho de pesquisa foi avaliar a capacidade de previsão para os movimentos da taxa de câmbio do Brasil a partir da interação entre a política monetária com as expectativas e riscos refletidos na estrutura a termo da taxa de juros. A escolha dos fundamentos seguiu a abordagem em macrofinanças de [Chen e Tsang \(2011\)](#): as variáveis macroeconômicas são obtidas a partir das funções de reação das autoridades monetárias, representadas pelas Regras de Taylor, enquanto os fatores latentes de Nelson-Siegel são utilizados como *proxies* de forma a captar as expectativas e os riscos contidos na curva de juros relativa. A hipótese da paridade descoberta da taxa de juros (UIP) conecta, por sua vez, a taxa de câmbio com a ETTJ relativa.

Diferentes fundamentos para a taxa de câmbio foram considerados: o *macro-only*, com somente as variáveis macroeconômicas, o *yields-only*, com os fatores latentes NS e o *macro-yields*, com ambos. Os modelos foram estimados por meio de uma estrutura TVP-VAR não estrutural, de modo a captar mudanças nas inter-relações entre as variáveis. A metodologia de [Koop e Korobilis \(2013\)](#) foi implementada com o objetivo de lidar com a "maldição da dimensionalidade" na estimação de modelos TVP-VAR, por meio da utilização de fatores de esquecimento. Além disso, os algoritmos para a seleção dinâmica de modelos (DMS) e para a combinação dinâmica de modelos (DMA) foram empregados de forma a atenuar a perda de informação decorrente da escolha de somente um modelo para todos os períodos.

Os horizontes de previsão analisados foram de 1 a 24 meses. Os modelos cambiais TVP-VAR foram comparados, por meio dos erros médios absolutos de previsão (MAFE), com o passeio aleatório e com especificações VAR estimadas por mínimos quadrados ordinários (OLS) em janela expansiva. Os resultados apontaram que os modelos TVP-VAR-DMA e TVP-VAR-DMS preveem comparativamente bem tanto no curto prazo como em horizontes mais longos. O TVP-VAR-DMS, por sua vez, foi o único modelo que superou o passeio aleatório em todos os horizontes de previsão. As especificações TVP-VAR, em média, performaram melhor que o RW, em especial, para horizontes mais longos.

A comparação entre as diferentes especificações TVP-VAR foi realizada por meio das somas dos logaritmos das densidades preditivas. Nenhum modelo superou os outros em todos os horizontes, entretanto, os modelos que realizaram simulações para os coeficientes a partir do passeio aleatório, em particular, o modelo TVP-VAR-DMA, tiveram melhor desempenho. Em relação aos fundamentos, os modelos *macro-yields* e *yields-only* tenderam a se sair melhor que a especificação *macro-only*. Houve ganhos significativos de performance, em ambas as métricas, ao se permitir que os coeficientes variassem no tempo por meio de uma estrutura VAR, porém, somente para os modelos que simularam o passeio aleatório (TVP-VAR  $A_{T+h} \sim RW$ ), e, deve-se ressaltar, os resultados tenderam

a se inverter para horizontes mais longos:  $h > 15$ .

O desenvolvimento deste trabalho levou à seguinte conclusão: as variáveis macroeconômicas, obtidas a partir de uma regra de Taylor, assim como os fatores latentes de Nelson-Siegel, conjuntamente contêm informação relevante para previsão da taxa de câmbio quando se permite que os coeficientes variem no tempo. Este resultado está de acordo com a literatura recente apontando que modelos fundamentados na Regra de Taylor em combinação com uma estrutura TVP-VAR tendem a superar o passeio aleatório<sup>1</sup>, *benchmark* pós [Meese e Rogoff \(1983\)](#). Por fim, permitir a variação na escolha dos modelos por meio dos algoritmos DMA e DMS melhorou, em geral, a performance preditiva.

---

<sup>1</sup> Por exemplo, vide [Kavtaradze e Mokhtari \(2017, p.25\)](#).

## A DADOS:

### A.1. A Curva de Juros e os Fatores Latentes de Nelson-Siegel

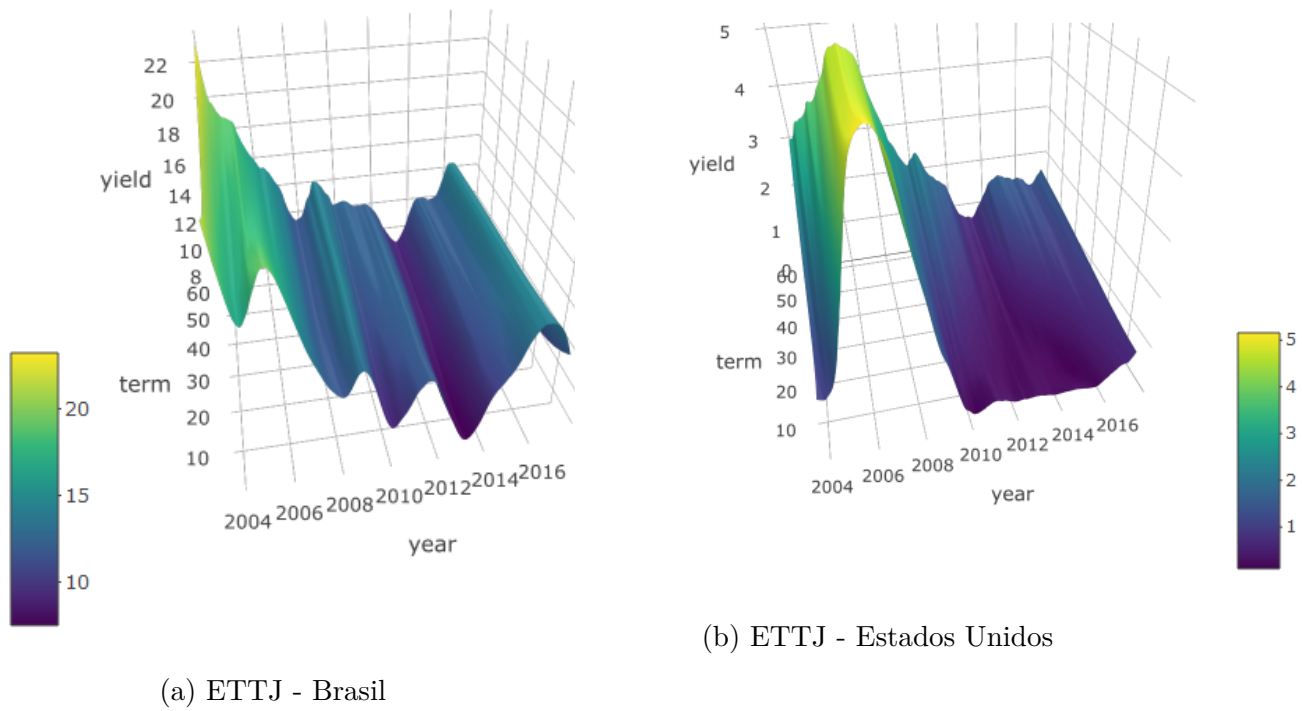


Figura 5 – ETTJ do Brasil e dos Estados Unidos - Variação em 12 meses (%).

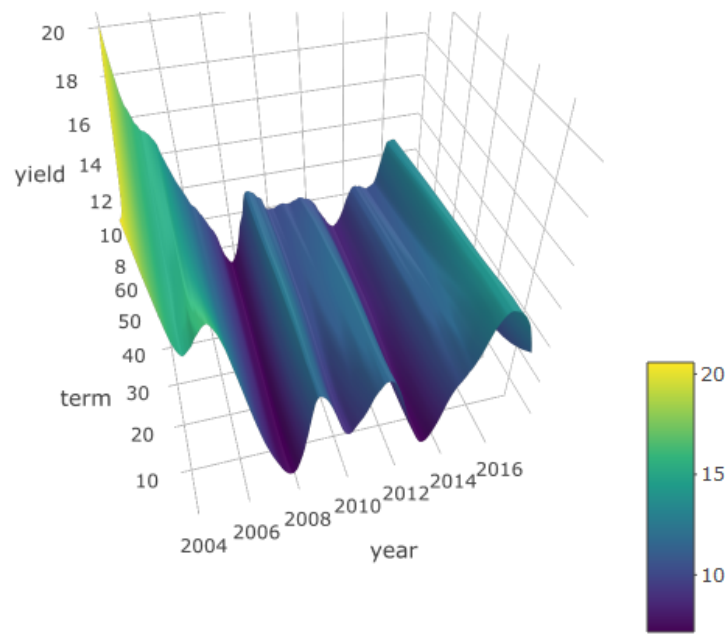


Figura 6 – ETTJ Relativa - Variação em 12 meses (%).

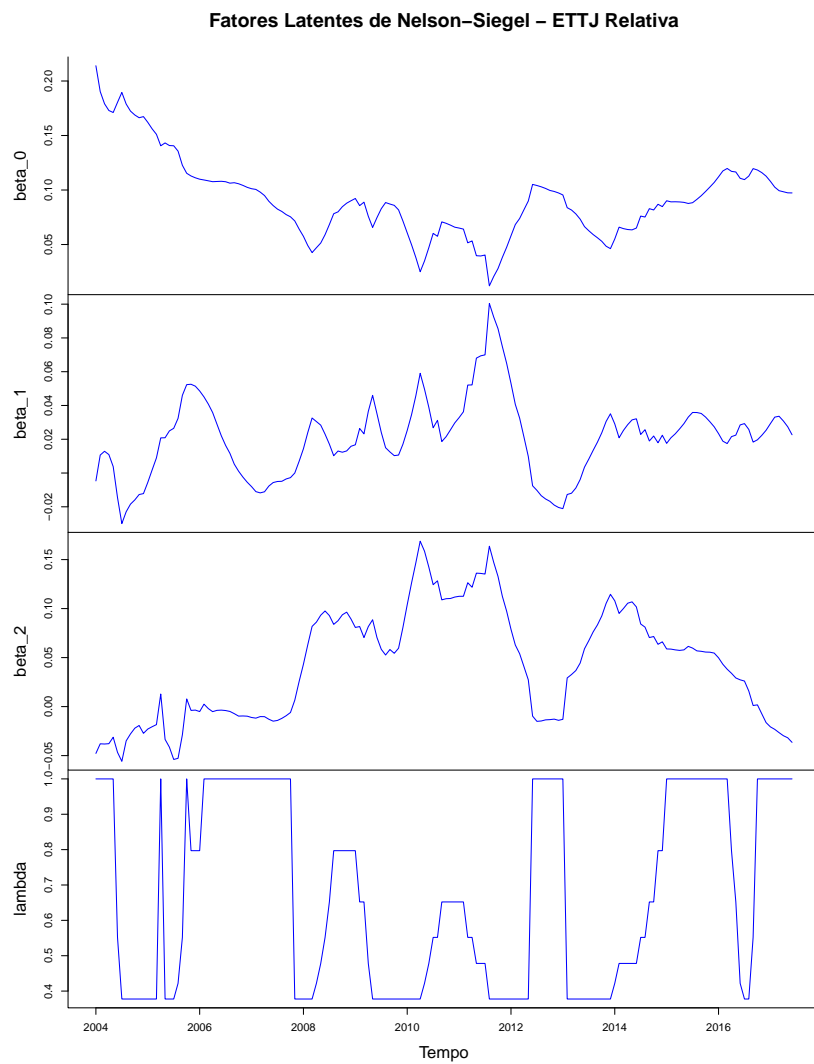


Figura 7 – Fatores Latentes de Nelson-Siegel a partir da ETTJ Relativa

## A.2. Testes de Raiz Unitária e Estatísticas Descritivas

Tabela 3 – Teste de Dickey-Fuller Aumentado

	<i>Augmented Dickey-Fuller (ADF)</i>					
	No Intercept		Intercept		Intercept + Trend	
	Stats	p.value	Stats	p.value	Stats	p.value
$ds$	-3,80	0,01	-3,54	0,01	-3,85	0,02
$\pi^R$	-1,48	0,14	-1,85	0,37	-2,99	0,16
$\pi^{Br}$	-0,95	0,31	-1,72	0,42	-2,24	0,48
$\pi^{US}$	-2,69	0,01	-2,84	0,06	-3,56	0,04
$\tilde{y}_t^R$	-4,46	0,01	-4,46	0,01	-4,55	0,01
$\tilde{y}^{Br}$	-3,04	0,01	-3,20	0,02	-3,71	0,03
$\tilde{y}^{US}$	-2,39	0,02	-2,71	0,08	-2,64	0,31
$L^R$	-1,24	0,22	-1,78	0,40	-4,22	0,01
$S^R$	-1,78	0,08	-2,88	0,05	-2,91	0,20
$C^R$	-2,11	0,04	-2,27	0,21	-2,97	0,17
$\Delta L^R$	-6,24	0,01	-5,84	0,01	-5,86	0,01
$\Delta S^R$	-8,58	0,01	-8,55	0,01	-8,98	0,01
$\Delta C^R$	-8,66	0,01	-8,64	0,01	-8,62	0,01
$L^{Br}$	-1,24	0,22	-1,42	0,53	-3,28	0,08
$S^{Br}$	-1,21	0,23	-1,32	0,57	-2,40	0,41
$C^{Br}$	-2,05	0,04	-2,60	0,10	-3,13	0,11
$L^{US}$	-1,74	0,08	-2,00	0,32	-2,21	0,49
$S^{US}$	-1,44	0,16	-1,44	0,52	-2,48	0,38
$C^{US}$	-2,19	0,03	-2,41	0,16	-2,84	0,23
$\Delta\pi^R$	-12,12	0,01	-12,20	0,01	-9,49	0,01
$\Delta\pi^{Br}$	-5,75	0,01	-5,73	0,01	-4,78	0,01
$\Delta\pi^{Us}$	-12,00	0,01	-12,05	0,01	-9,59	0,01
	H0	H1				
ADF	Not stationary	Otherwise				

Tabela 4 – Testes de Raiz Unitária - KPSS e Phillips-Perron

	<i>KPSS</i>		<i>Phillips-Perron (PP)</i>	
	Intercept	I + Trend	Stats	p-value
	Stats	Stats		
$ds$	0,27	0,08	-16,63	0,16
$\pi^R$	0,63	0,06	-11,27	0,47
$\pi^{Br}$	0,62	0,05	-8,19	0,64
$\pi^{US}$	0,28	0,06	-22,74	0,03
$\tilde{y}^R$	0,11	0,06	-38,73	0,01
$\tilde{y}^{Br}$	0,04	0,04	-23,10	0,03
$\tilde{y}^{US}$	0,11	0,05	-13,51	0,34
$L^R$	0,96	0,08	-21,08	0,05
$S^R$	0,11	0,05	-15,87	0,20
$C^R$	0,56	0,14	-16,78	0,15
$\Delta L^R$	0,05	0,05	-99,72	0,01
$\Delta S^R$	0,45	0,10	-156,52	0,01
$\Delta C^R$	0,11	0,10	-160,79	0,01
$L^{Br}$	0,84	0,09	-18,81	0,08
$S^{Br}$	0,22	0,20	-8,69	0,62
$C^{Br}$	0,53	0,05	-20,51	0,06
$L^{US}$	0,50	0,18	-10,89	0,49
$S^{US}$	0,28	0,25	-5,53	0,80
$C^{US}$	0,26	0,25	-4,27	0,87
$\Delta\pi^R$	0,21	0,06	-134,74	0,01
$\Delta\pi^{Br}$	0,20	0,10	-58,86	0,01
$\Delta\pi^{Us}$	0,14	0,05	-132,32	0,01

	H0	H1
KPSS	Stationarity	Otherwise
PP	Not stationary	Otherwise

KPSS (I) - C.I.:	KPSS (I + trend) - C.I.:
C.10pct: 0,347	C.10pct: 0,119
C.5pct: 0,463	C.5pct: 0,146
C.2.5pct: 0,574	C.2.5pct: 0,176
C.1pct: 0,739	C.1pct: 0,216

Tabela 5 – Estatísticas Descritivas

Statistic	N	Mean	St. Dev.	Min	Max
$ds$	150	-0.450	1.198	-3.458	1.770
$\Delta\pi^R$	150	0.134	1.240	-7.504	6.828
$\tilde{y}^R$	150	-0.085	1.110	-3.415	4.326
$\Delta L^R$	150	0.106	1.352	-5.258	4.834
$\Delta S^R$	150	-0.003	0.640	-2.076	6.784
$\Delta C^R$	150	0.016	0.958	-6.803	3.302

**B VAR**Tabela 6 – VAR - *Akaike Information Criterion* (AIC)

	AIC
p1	-2,25
p2	<b>-2,53</b>
p3	-2,35
p4	-2,17
p5	-2,15
p6	-2,17
p7	-2,09
p8	-1,84
p9	-2,04
p10	-2,23
p11	-2,10
p12	-2,02



## C PREVISÃO

Tabela 7 – MAFE - Valor Total

MAFE									
Modelos	h = 1	h = 3	h = 6	h = 9	h = 12	h = 15	h = 18	h = 21	h = 24
<i>Macro-Yields</i>									
TVP-VAR	0,34	0,64	0,95	1,23	1,31	1,20	1,20	1,16	1,13
TVP-VAR $A_{T+h} \sim RW$	0,36	0,74	1,04	1,06	1,08	1,02	1,03	1,05	1,08
TVP-VAR ( $\lambda_{ff} = 0.99$ )	0,33	0,64	0,97	1,25	1,34	1,23	1,24	1,19	1,16
VAR, heteroskedastic	0,34	0,64	0,95	1,23	1,30	1,20	1,19	1,15	1,13
VAR, homoscedastic	0,34	0,63	0,96	1,22	1,29	1,18	1,17	1,14	1,12
<i>Yields-Only</i>									
TVP-VAR	0,32	0,64	1,00	1,28	1,36	1,26	1,26	1,21	1,19
TVP-VAR $A_{T+h} \sim RW$	0,33	0,70	1,03	1,10	1,10	1,01	1,01	1,02	1,06
TVP-VAR ( $\lambda_{ff} = 0.99$ )	0,30	0,64	1,00	1,29	1,38	1,29	1,26	1,22	1,21
VAR, heteroskedastic	0,32	0,63	1,00	1,27	1,35	1,25	1,25	1,20	1,18
VAR, homoscedastic	0,33	0,62	1,00	1,27	1,34	1,23	1,23	1,19	1,17
<i>Macro-Only</i>									
TVP-VAR	0,30	0,63	0,98	1,26	1,35	1,26	1,23	1,19	1,18
TVP-VAR $A_{T+h} \sim RW$	0,30	0,67	0,98	1,08	1,16	1,13	1,14	1,19	1,34
TVP-VAR ( $\lambda_{ff} = 0.99$ )	0,30	0,63	1,00	1,29	1,38	1,29	1,27	1,22	1,20
VAR, heteroskedastic	0,31	0,63	0,98	1,35	1,26	1,23	1,23	1,19	1,17
VAR, homoscedastic	0,31	0,63	0,98	1,24	1,32	1,23	1,19	1,16	1,14
<i>Modelos DMA/DMS</i>									
TVP-VAR-DMA	0,30	0,64	0,95	1,13	1,17	1,09	1,07	1,10	1,17
TVP-VAR-DMS	0,28	0,61	0,96	1,18	1,22	1,11	1,09	1,15	1,26
TVP-VAR-DMA $A_{T+h} \sim RW$	0,30	0,64	0,95	1,13	1,17	1,09	1,07	1,10	1,18
TVP-VAR-DMS $A_{T+h} \sim RW$	0,28	0,61	0,96	1,17	1,22	1,11	1,09	1,16	1,29
TVP-VAR-DMA ( $\lambda_{ff} = 0.99$ )	0,29	0,63	0,95	1,25	1,22	1,28	1,28	1,48	1,92
TVP-VAR-DMS ( $\lambda_{ff} = 0.99$ )	0,27	0,62	1,01	1,24	1,34	1,32	1,46	1,84	2,72
<i>Benchmarks</i>									
VAR OLS - Macro-Yields	0,35	0,78	1,14	1,25	1,38	1,43	1,56	1,77	2,24
VAR OLS - Yields-Only	0,31	0,67	0,99	1,15	1,12	1,02	0,97	0,90	0,92
VAR OLS - Macro-Only	0,30	0,75	1,12	1,21	1,36	1,45	1,57	1,83	2,23
Random Walk	0,29	0,64	1,02	1,32	1,50	1,51	1,48	1,38	1,28

Tabela 8 – Somatório do log das probabilidades preditivas - Valor Total

*Sum of Log Predictive Likelihoods*

Modelos	h = 1	h = 3	h = 6	h = 9	h = 12	h = 15	h = 18	h = 21	h = 24
<b>Macro-Yields</b>									
TVP-VAR	-75,35	-154,64	-229,25	-302,97	-335,12	-322,72	-293,36	-251,11	-212,76
TVP-VAR $A_{T+h} \sim RW$	-79,88	-156,06	-188,26	-188,53	-187,90	-188,84	-191,75	-195,65	-200,48
TVP-VAR ( $\lambda_{ff} = 0.99$ )	-73,74	-154,14	-227,33	-303,16	-335,45	-324,17	-299,11	-256,18	-217,44
VAR, heteroskedastic	-75,70	-154,87	-229,37	-301,92	-332,78	-319,17	-293,09	-248,89	-211,23
VAR, homoscedastic	-103,88	-156,31	-181,27	-198,78	-201,69	-187,11	-175,54	-161,58	-150,82
<b>Yields-Only</b>									
TVP-VAR	-70,98	-156,04	-245,23	-342,26	-378,82	-370,23	-334,05	-292,94	-246,45
TVP-VAR $A_{T+h} \sim RW$	-74,17	-153,85	-182,69	-184,80	-183,89	-184,66	-188,18	-192,68	-197,56
TVP-VAR ( $\lambda_{ff} = 0.99$ )	-63,92	-159,82	-264,25	-355,80	-385,74	-374,26	-336,44	-294,51	-248,38
VAR, heteroskedastic	-70,69	-157,11	-245,27	-341,06	-377,24	-368,94	-331,00	-289,52	-244,85
VAR, homoscedastic	-101,25	-154,65	-183,34	-207,44	-209,92	-193,25	-179,56	-166,44	-158,35
<b>Macro-Only</b>									
TVP-VAR	-66,44	-159,71	-263,08	-353,88	-384,99	-370,75	-332,55	-287,13	-240,35
TVP-VAR $A_{T+h} \sim RW$	-66,53	-153,46	-186,27	-188,33	-190,25	-193,22	-198,65	-205,31	-212,74
TVP-VAR ( $\lambda_{ff} = 0.99$ )	-64,06	-159,46	-263,54	-354,94	-386,12	-371,60	-337,85	-295,42	-248,73
VAR, heteroskedastic	-66,84	-159,35	-262,07	-350,12	-384,19	-370,73	-328,41	-283,24	-237,20
VAR, homoscedastic	-99,153	-153,09	-187,65	-208,93	-211,61	-196,07	-180,26	-164,53	-154,89
<b>Modelos Completos</b>									
TVP-VAR-DMA	-94,309	-244,11	-344,16	-367,96	-306,51	-247,44	-213,2	-198,87	-192,79
TVP-VAR-DMS	-61,642	-160,63	-204,66	-215,31	-195,53	-178,35	-169,58	-168,39	-168,94
TVP-VAR-DMA $A_{T+h} \sim RW$	-104,79	-269,33	-299,49	-270,85	-217,9	-183,24	-169,87	-167,52	-170,37
TVP-VAR-DMS $A_{T+h} \sim RW$	-62,29	-158,54	-191,34	-195,95	-188,23	-186,06	-189,19	-194,88	-200,81
TVP-VAR-DMA ( $\lambda_{ff} = 0.99$ )	-87,15	-224,55	-302,62	-300,80	-241,33	-196,16	-177,81	-184,96	-201,74
TVP-VAR-DMS ( $\lambda_{ff} = 0.99$ )	-58,91	-160,73	-199,50	-207,78	-192,22	-183,73	-183,70	-194,15	-205,52

## REFERÊNCIAS

- AIUBE, F. A. L. *Modelos quantitativos em Finanças com enfoque em commodities*. [S.l.]: Bookman Editora, 2013. 30, 31
- ALQUIST, R.; CHINN, M. D. Conventional and unconventional approaches to exchange rate modelling and assessment. *International Journal of Finance & Economics*, Wiley Online Library, v. 13, n. 1, p. 2–13, 2008. 13
- ANG, A.; PIAZZESI, M. A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables. *Journal of Monetary economics*, Elsevier, v. 50, n. 4, p. 745–787, 2003. 12
- BARR, D. G.; CAMPBELL, J. Y. Inflation, real interest rates, and the bond market: A study of uk nominal and index-linked government bond prices. *Journal of Monetary Economics*, Elsevier, v. 39, n. 3, p. 361–383, 1997. 12
- BEKAERT, G.; WEI, M.; XING, Y. Uncovered interest rate parity and the term structure. *Journal of International Money and Finance*, Elsevier, v. 26, n. 6, p. 1038–1069, 2007. 13
- BERNZ, B. M. *Modelo Nelson-Siegel dinâmico da estrutura a termo da taxa de juros com fatores exógenos macroeconômicos: uma aplicação ao mercado brasileiro*. Tese (Doutorado), 2014. 12
- BILSON, J. F. *The "speculative efficiency" hypothesis*. [S.l.]: National Bureau of Economic Research Cambridge, Mass., USA, 1980. 10
- CALDEIRA, J.; MOURA, G. V.; PORTUGAL, M. S. Efficient yield curve estimation and forecasting in brazil. 2010. 12
- CALDEIRA, J. F.; MOURA, G. V.; SANTOS, A. A. Previsões macroeconômicas baseadas em modelos tvp-var: evidências para o brasil. *Revista Brasileira de Economia*, SciELO Brasil, v. 69, n. 4, p. 407–428, 2015. 15, 35, 36
- CHEN, Y.-c.; TSANG, K. P. A macro-finance approach to exchange rate determination. 2011. 10, 11, 13, 14, 16, 17, 18, 19, 27, 32, 39, 48
- CHEN, Y.-c.; TSANG\*, K. P. What does the yield curve tell us about exchange rate predictability? *Review of Economics and Statistics*, MIT Press, v. 95, n. 1, p. 185–205, 2013. 17, 19, 21
- CLARIDA, R.; GALI, J.; GERTLER, M. Monetary policy rules in practice: some international evidence. *European economic review*, Elsevier, v. 42, n. 6, p. 1033–1067, 1998. 12, 18
- COMMANDEUR, J. J.; KOOPMAN, S. J. *An introduction to state space time series analysis*. [S.l.]: Oxford University Press, 2007. 31
- DAI, Q.; SINGLETON, K. J. Specification analysis of affine term structure models. *The Journal of Finance*, Wiley Online Library, v. 55, n. 5, p. 1943–1978, 2000. 12
- DIEBOLD, F. X.; LI, C. Forecasting the term structure of government bond yields. *Journal of econometrics*, Elsevier, v. 130, n. 2, p. 337–364, 2006. 12, 22, 23

- DIEBOLD, F. X.; PIAZZESI, M.; RUDEBUSCH, G. D. Modeling bond yields in finance and macroeconomics. *The American Economic Review*, American Economic Association, v. 95, n. 2, p. 415, 2005. [17](#)
- DIEBOLD, F. X.; RUDEBUSCH, G. D.; ARUOBA, S. B. The macroeconomy and the yield curve: a dynamic latent factor approach. *Journal of econometrics*, Elsevier, v. 131, n. 1, p. 309–338, 2006. [12](#), [13](#), [23](#), [24](#), [25](#), [26](#)
- DORNBUSCH, R. Expectations and exchange rate dynamics. *Journal of political Economy*, The University of Chicago Press, v. 84, n. 6, p. 1161–1176, 1976. [10](#)
- DUFFIE, D.; KAN, R. A yield-factor model of interest rates. *Mathematical finance*, Wiley Online Library, v. 6, n. 4, p. 379–406, 1996. [12](#)
- ENGEL, C. *et al.* Exchange rate models are not as bad as you think [with comments and discussion]. *NBER Macroeconomics annual*, University of Chicago Press, v. 22, p. 381–473, 2007. [13](#)
- ENGEL, C.; WEST, K. D. Exchange rates and fundamentals. *Journal of political Economy*, The University of Chicago Press, v. 113, n. 3, p. 485–517, 2005. [12](#), [18](#)
- ESTRELLA, A.; MISHKIN, F. S. Predicting us recessions: Financial variables as leading indicators. *The Review of Economics and Statistics*, MIT Press, v. 80, n. 1, p. 45–61, 1998. [12](#)
- FERREIRA, D.; PALMA, A. A. Forecasting inflation with the phillips curve: A dynamic model averaging approach for brazil. *Revista Brasileira de Economia*, SciELO Brasil, v. 69, n. 4, p. 451–465, 2015. [15](#)
- FRANKEL, J. A. On the mark: A theory of floating exchange rates based on real interest differentials. *The American Economic Review*, JSTOR, v. 69, n. 4, p. 610–622, 1979. [10](#)
- FRANKEL, J. A. Tests of monetary and portfolio balance models of exchange rate determination. In: *Exchange Rate Theory and Practice*. [S.l.]: University of Chicago Press, 1984. p. 239–260. [10](#)
- FRENKEL, J. A. A monetary approach to the exchange rate: doctrinal aspects and empirical evidence. *the scandinavian Journal of economics*, JSTOR, p. 200–224, 1976. [10](#)
- GÜRKAYNAK, R. S.; SACK, B.; WRIGHT, J. H. The us treasury yield curve: 1961 to the present. *Journal of monetary Economics*, Elsevier, v. 54, n. 8, p. 2291–2304, 2007. [42](#)
- GÜRKAYNAK, R. S.; WRIGHT, J. H. Macroeconomics and the term structure. *Journal of Economic Literature*, American Economic Association, v. 50, n. 2, p. 331–367, 2012. [17](#)
- JAMES, J.; MARSH, I.; SARNO, L. *Handbook of exchange rates*. [S.l.]: John Wiley & Sons, 2012. v. 2. [10](#), [11](#)
- KAVTARADZE, L.; MOKHTARI, M. Factor models and time-varying parameter framework for forecasting exchange rates and inflation: A survey. *Journal of Economic Surveys*, Wiley Online Library, 2017. [14](#), [49](#)

- KILIAN, L.; LÜTKEPOHL, H. *Structural vector autoregressive analysis*. [S.l.]: Cambridge University Press, 2017. [35](#)
- KIM, C.-J.; NELSON, C. R. *et al.* State-space models with regime switching: classical and gibbs-sampling approaches with applications. *MIT Press Books*, The MIT press, v. 1, 1999. [30](#)
- KNEZ, P. J.; LITTERMAN, R.; SCHEINKMAN, J. Explorations into factors explaining money market returns. *The Journal of Finance*, Wiley Online Library, v. 49, n. 5, p. 1861–1882, 1994. [12](#)
- KOOP, G.; KOROBILIS, D. Large time-varying parameter vars. *Journal of Econometrics*, Elsevier, v. 177, n. 2, p. 185–198, 2013. [14](#), [15](#), [16](#), [33](#), [34](#), [35](#), [36](#), [37](#), [39](#), [40](#), [43](#), [48](#)
- LANGE, R. H. The canadian macroeconomy and the yield curve: A dynamic latent factor approach. *International Review of Economics & Finance*, Elsevier, v. 27, p. 261–274, 2013. [12](#)
- LANGE, R. H. The small open macroeconomy and the yield curve: A state-space representation. *The North American Journal of Economics and Finance*, Elsevier, v. 29, p. 1–21, 2014. [30](#)
- LEVANT, J.; MA, J. Investigating united kingdom’s monetary policy with macro-factor augmented dynamic nelson–siegel models. *Journal of Empirical Finance*, Elsevier, v. 37, p. 117–127, 2016. [12](#)
- LITTERMAN, R. B.; SCHEINKMAN, J. Common factors affecting bond returns. *The Journal of Fixed Income*, Institutional Investor Journals, v. 1, n. 1, p. 54–61, 1991. [12](#), [22](#)
- MARIANI, L. A. *Modelos macro-financeiros com o uso de fatores latentes do tipo Nelson-Siegel*. Tese (Doutorado) — Universidade de São Paulo, 2015. [14](#)
- MEESE, R. A.; ROGOFF, K. Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of international economics*, Elsevier, v. 14, n. 1-2, p. 3–24, 1983. [11](#), [49](#)
- MISHKIN, F. S. What does the term structure tell us about future inflation? *Journal of monetary economics*, Elsevier, v. 25, n. 1, p. 77–95, 1990. [12](#)
- MOLODTSOVA, T.; PAPELL, D. H. Out-of-sample exchange rate predictability with taylor rule fundamentals. *Journal of international economics*, Elsevier, v. 77, n. 2, p. 167–180, 2009. [13](#)
- MORGAN, J. *et al.* Riskmetrics technical document. New York, 1996. [38](#)
- NELSON, C. R.; SIEGEL, A. F. *Parsimonious Modeling of Yield Curves*. 1987. 473–489 p. [12](#), [22](#)
- PARK, D.; JUN, B.; KIM, J. Fast tracking rls algorithm using novel variable forgetting factor with unity zone. *Electronics Letters*, IET, v. 27, n. 23, p. 2150–2151, 1991. [34](#)
- RAFTERY, A. E.; KÁRNÝ, M.; ETTLER, P. Online prediction under model uncertainty via dynamic model averaging: Application to a cold rolling mill. *Technometrics*, Taylor & Francis, v. 52, n. 1, p. 52–66, 2010. [14](#), [15](#), [33](#), [35](#), [36](#), [37](#)

---

TERRA, C. *Principles of International Finance and Open Economy Macroeconomics: Theories, Applications, and Policies*. [S.l.]: Academic Press, 2015. 20

VALI, S. *Principles of Mathematical Economics*. [S.l.]: Springer, 2014. 40

WU, T. Monetary policy and the slope factor in empirical term structure estimations. 2001. 12