



Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós-graduação em Economia
Área: Economia Aplicada

Previsibilidade da taxa de câmbio com sentimento das notícias

Lucas-Matheus Souza de Araújo Lima

João Pessoa, PB
2020

Lucas-Matheus Souza de Araújo Lima

Previsibilidade da taxa de câmbio com sentimento das notícias

Dissertação apresentada ao Programa de Pós-graduação em Economia da Universidade Federal da Paraíba - UFPB, na área de Economia Aplicada, para obtenção do título de Mestre em Ciências.

Universidade Federal da Paraíba - UFPB
Centro de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Economia

Orientador: Prof. Dr. Luiz Renato Régis de Oliveira Lima
Coorientador: Prof. Dr. Edilean Kleber da Silva Bejarano Aragón

João Pessoa, PB
2020

Catálogo na publicação
Seção de Catalogação e Classificação

L732p Lima, Lucas-Matheus Souza de Araújo.

Previsibilidade da taxa de câmbio com sentimento das notícias / Lucas-Matheus Souza de Araújo Lima. - João Pessoa, 2020.

60 f. : il.

Orientação: Luiz Renato Régis de Oliveira Lima.

Coorientação: Edilean Kleber da Silva Bejarano Aragón.
Dissertação (Mestrado) - UFPB/CCSA.

1. Taxa de câmbio. 2. Predição. 3. Sentimento das notícias. 4. Regra de Taylor. I. Lima, Luiz Renato Régis de Oliveira. II. Aragón, Edilean Kleber da Silva Bejarano. III. Título.

UFPB/BC

CDU 339.743



Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Economia

Campus Universitário I – Cidade Universitária – CEP 58.059-900 – João Pessoa – Paraíba
Telefone: (83) 3216-7482 – <http://www.ccsa.ufpb.br/ppge> – E-mail: ppge.ccsa@gmail.com

Ata da reunião da Banca Examinadora designada para avaliar o trabalho de dissertação do mestrando **LUCAS-MATHEUS SOUZA DE ARAÚJO LIMA** submetida para obtenção do grau de mestre em Economia área de concentração em **Economia Aplicada**.

Aos vinte e oito dias do mês de fevereiro do ano dois mil e vinte, às quatorze horas, na Sala 01 do Bloco da Pós-Graduação, do Centro de Ciências Sociais Aplicadas, da Universidade Federal da Paraíba, reuniu-se em cerimônia pública os membros da Banca Examinadora, constituída pelos professores doutores **LUIZ RENATO REGIS DE OLIVEIRA LIMA** (Orientador) da Universidade Federal da Paraíba, **CASSIO DA NOBREGA BESARRIA**, (Examinador Interno) da Universidade Federal da Paraíba e **MARIA DANIELLA DE OLIVEIRA PEREIRA DA SILVA** (Examinadora Externa), da Universidade Federal da Paraíba, a fim de examinarem o candidato ao grau de mestre em Economia, área de concentração em **Economia Aplicada**, **LUCAS-MATHEUS SOUZA DE ARAÚJO LIMA**. Além dos examinadores e do examinando, compareceram também, representantes do Corpo Docente e do Corpo Discente. Iniciando a sessão, o professor **LUIZ RENATO REGIS DE OLIVEIRA LIMA**, na qualidade de presidente da Banca Examinadora, comunicou aos presentes qual o fim da reunião e os procedimentos de encaminhamento da mesma. A seguir concedeu à palavra ao candidato, para que fizesse oralmente a exposição do trabalho, apresentado sob o título: **“Previsibilidade da taxa de câmbio com sentimento das notícias.”**. Concluída a exposição, o senhor presidente, solicitou que fosse feita a arguição por cada um dos examinadores. Concedida a palavra ao candidato, para que respondesse e esclarecesse às questões levantadas. Terminadas as arguições, a Banca Examinadora, passou a proceder a avaliação e ao julgamento do candidato. Em seguida, o senhor presidente comunicou aos presentes que a Banca Examinadora, por unanimidade, **aprovou** a dissertação apresentada e defendida com o conceito **APROVADO**, concedendo assim, o grau de **Mestre em Economia**, área de concentração em **Economia Aplicada**, ao mestrando **LUCAS-MATHEUS SOUZA DE ARAÚJO LIMA**. E, para constar, eu, Ricardo Madeira Cataldi, secretário *ad hoc* do Programa de Pós-Graduação em Economia, lavrei a presente ata, que assino junto com os membros da Banca Examinadora. João Pessoa, 28 de fevereiro de 2020.

P/Cássio da Nobrega Besarria
Prof. Dr. LUIZ RENATO REGIS DE OLIVEIRA LIMA
Orientador – UFPB

P/Cássio da Nobrega Besarria
Prof. Dr. CASSIO DA NOBREGA BESARRIA
Examinador Interno – UFPB

P/Cássio da Nobrega Besarria
Prof.^a Dr.^a MARIA DANIELLA DE OLIVEIRA PEREIRA DA SILVA
Examinadora Externa – UFPB

Ricardo MC
Ricardo Madeira Cataldi
Secretário *ad hoc* – PPGE/CCSA/ UFPB

Agradecimentos

Agradeço a minha família, principalmente aos meus pais, Flávio Delson e Maria Helena, e minha avó, Maria Lúcia, por todo apoio e suporte dado durante essa etapa acadêmica.

Ao meu orientador, Prof. Dr. Luiz Renato Lima, e coorientador, Prof. Dr. Edilean Aragón, por toda paciência, tempo disponível e orientação para a realização deste trabalho. Agradeço também ao Prof. Dr. Lucas Godeiro, que ajudou de forma fundamental para finalização da dissertação.

Aos professores Dr. Cássio Bessaria e Dr^a Maria Daniella, agradeço o aceite para compor a banca avaliadora e as contribuições sobre tópicos relativos ao meu interesse de pesquisa.

Agradeço aos discentes e docentes Programa de Pós-Graduação em Economia da UFPB, pelo espaço para discussões e contribuições acadêmicas. A todos que de alguma forma contribuíram para a realização desta dissertação, meu sincero obrigado.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Resumo

Este trabalho tem como objetivo estudar a previsibilidade fora da amostra e um passo à frente das variações da taxa de câmbio nominal com sentimento das notícias, utilizando notícias financeiras do The New York Times e do The Wall Street Journal na construção desses sentimentos para o período de janeiro de 1980 a maio de 2017. Os resultados sugerem forte poder preditivo do sentimento das notícias gerado com o dicionário variante no tempo, podendo ser utilizado como único preditor ou como preditor adicional ao modelo com base na regra de Taylor convencional. Além de demonstrar que o método do dicionário variante no tempo é mais adequado para grandes amostras textuais em relação a dicionários fixos quando o objetivo do trabalho passa pelo poder preditivo. Quando utilizado como único preditor, foi observado superioridade de performance na predição sobre o passeio aleatório para todos os países analisados e sobre a regra de Taylor convencional para 7 dos 9 países. Para os modelos ampliados da regra de Taylor com sentimento das notícias, foi observado superioridade de performance para todos os países em relação aos modelos convencionais da regra de Taylor. Todos esses resultados a 1% de significância.

Palavras-chave: Taxa de câmbio; predição; sentimento das notícias; regra de Taylor.

Abstract

This paper aims to study the out-of-sample predictability of nominal exchange rate variations with news sentiment, we use financial news from The New York Times and The Wall Street Journal to construct these sentiments for the period from January 1980 to May 2017. The results suggest a strong predictive power of the news sentiment generated with the time variant dictionary, which can be used as the only predictor or as an additional predictor to the model based on the conventional Taylor rule. Besides demonstrating that the time-varying dictionary method is more suitable for large textual samples compared to fixed dictionaries when the objective of the work passes through predictive power. When used as the only predictor, superiority of performance was observed in the prediction on random walk for all analyzed countries and on the conventional Taylor rule for 7 of 9 countries. For extended Taylor rule models with news sentiment, superior performance was observed for all countries in relation to the conventional Taylor rule models. All of these results at 1% significance.

Keywords: Exchange rate; forecast; news sentiment; Taylor rule.

Lista de ilustrações

Figura 1 – Coeficientes da inflação dos Estados Unidos com base no modelo ampliado da regra de Taylor com sentimentos do dicionário variante no tempo.	31
Figura 2 – Coeficientes da inflação do país estrangeiro com base no modelo ampliado da regra de Taylor com sentimentos do dicionário variante no tempo.	32
Figura 3 – Coeficientes da inflação dos Estados Unidos com base na especificação da regra de Taylor selecionada.	43
Figura 4 – Coeficientes da inflação do país estrangeiro com base na especificação da regra de Taylor selecionada.	44
Figura 5 – Coeficientes da inflação dos Estados Unidos com base na especificação da regra de Taylor indicada pela literatura.	45
Figura 6 – Coeficientes da inflação do país estrangeiro com base na especificação da regra de Taylor indicada pela literatura.	46
Figura 7 – Dicionários fixos: sentimentos gerais.	47
Figura 8 – Dicionários fixos: sentimentos negativos e de incerteza.	48
Figura 9 – Dicionário variante no tempo: sentimento geral por país.	49
Figura 10 – Evolução do RMSPE por país para o intervalo de termos testados (1000 a 10000).	50
Figura 11 – Coeficientes da inflação dos Estados Unidos e Reino Unido por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	53
Figura 12 – Coeficientes da inflação dos Estados Unidos e Japão por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	54
Figura 13 – Coeficientes da inflação dos Estados Unidos e Suécia por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	55
Figura 14 – Coeficientes da inflação dos Estados Unidos e Coreia do Sul por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	56
Figura 15 – Coeficientes da inflação dos Estados Unidos e Canadá por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	57
Figura 16 – Coeficientes da inflação dos Estados Unidos e Dinamarca por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	58
Figura 17 – Coeficientes da inflação dos Estados Unidos e Noruega por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	59
Figura 18 – Coeficientes da inflação dos Estados Unidos e México por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	60
Figura 19 – Coeficientes da inflação dos Estados Unidos e Brasil por especificação da regra de Taylor convencional e ampliada com dicionários fixos.	61

Lista de tabelas

Tabela 1 – Palavras por classificação e dicionário	17
Tabela 2 – Resultados dos modelos só com sentimento das notícias - Parte 1	24
Tabela 3 – Resultados dos modelos só com sentimento das notícias - Parte 2	25
Tabela 4 – Tabela de correlação	26
Tabela 5 – Resultados dos modelos ampliados da regra de Taylor com sentimento das notícias - Parte 1	27
Tabela 6 – Resultados dos modelos ampliados da regra de Taylor com sentimento das notícias - Parte 2	28
Tabela 7 – Fontes dos dados	38
Tabela 8 – Tamanho das amostras por país estrangeiro	39
Tabela 9 – Especificações da regra de Taylor estimadas	40
Tabela 10 – Resultados da regra de Taylor convencional - Parte 1	41
Tabela 11 – Resultados da regra de Taylor convencional - Parte 2	42
Tabela 12 – Amostra de termos positivos e negativos por data de referência - Parte 1	51
Tabela 13 – Amostra de termos positivos e negativos por data de referência - Parte 2	52

Sumário

1	INTRODUÇÃO	10
2	METODOLOGIA	13
2.1	Dados	13
2.2	Regra de Taylor	13
2.3	Análise textual	15
2.3.1	Dicionário fixo	15
2.3.2	Dicionário variante no tempo	17
2.4	Regra de Taylor ampliada	19
2.5	Método de estimação	19
2.6	Testes de performance	20
3	RESULTADOS E DISCUSSÕES	21
3.1	Consistência dos modelos ampliados	29
4	CONCLUSÕES	33
	REFERÊNCIAS	34
	APÊNDICES	37
	APÊNDICE A – FONTE E INTERVALO DOS DADOS	38
	APÊNDICE B – REGRA DE TAYLOR	40
	APÊNDICE C – SENTIMENTO DAS NOTÍCIAS	47
	APÊNDICE D – NÚMERO DE TERMOS	50
	APÊNDICE E – AMOSTRA DE TERMOS	51
	APÊNDICE F – COEFICIENTES DOS MODELOS AMPLIADOS	53

1 Introdução

Explicar as variações da taxa de câmbio foi um desafio que permaneceu por anos em aberto, principalmente utilizando os fundamentos macroeconômicos. Desde os trabalhos seminais de [Meese e Rogoff \(1983a\)](#) e [Meese e Rogoff \(1983b\)](#), dentre as tentativas de obter resultados superiores ao passeio aleatório temos o trabalho de [Molodtsova e Papell \(2009\)](#), onde foi obtido êxito na predição da variação da taxa de câmbio com os fundamentos macroeconômicos e a partir dele a literatura tomou um novo caminho no desenvolvimento desta fronteira.

Como exposto por [Carriero, Kapetanios e Marcellino \(2009\)](#) até o momento do seu trabalho, obter performance superior ao do passeio aleatório sem constante no multi-horizonte, partindo dos fundamentos macroeconômicos, foi uma tarefa sem êxito na literatura até então, sendo que quando o interesse é em horizontes curtos essa dificuldade é intensificada.

[Molodtsova e Papell \(2009\)](#), utilizando os fundamentos da regra de Taylor, obtiveram performances estatisticamente superiores ao passeio aleatório sem constante na predição um passo à frente da taxa de câmbio, sendo observada previsibilidade no curto prazo para 11 das 12 taxas de câmbio testadas. Outros fundamentos macroeconômicos foram utilizados, gerando os modelos convencionais da política monetária, da paridade do poder de compra e da taxa de juros, mas a regra de Taylor demonstrou superioridade em relação aos demais modelos. Na sequência, [Nikolsko-Rzhevskyy e Prodan \(2012\)](#) conseguiram êxito em obter performances superiores no multi-horizonte em relação ao passeio aleatório utilizando os fundamentos macroeconômicos.

A especificação do modelo de predição é um ponto de suma importância na literatura. No trabalho de [Molodtsova e Papell \(2009\)](#) são testadas 16 especificações para a regra de Taylor, onde a taxa de câmbio real foi desconsiderada na especificação que gerou melhores performances. Já [Morales-Arias e Moura \(2013\)](#) incluíram preditores oriundos do retorno/volatilidade dos ativos, de índices cíclicos e de confiança, foi observado que combinações de previsões a partir de informações mais amplas melhoram a precisão da previsão em relação aos preditores isolados e acabam por minimizar a incerteza *ex-ante*. Ainda nessa linha do estudo das especificações, [Wang, Morley e Ordóñez \(2016\)](#) observaram que os preços dos ativos são importante na modelagem com base nos fundamentos macroeconômicos, sendo observado superioridade nas performances em relação aos modelos convencionais e do passeio aleatório.

Dada a abordagem linear, a literatura vem indicando a regra de Taylor como uma das melhores dentre os fundamentos macroeconômicos para predição da variação da taxa de câmbio. Dessa forma, utilizando essa esfera da literatura que investiga especificações alternativas para a regra de Taylor na modelagem, um dos objetivos desse trabalho visa a estudar novas especificações para o mesmo na predição da variação da taxa de câmbio, mas considerando informações contidas nas notícias financeiras.

Paralelo a essas abordagens ditas convencionais, uma outra passou a estudar as informações ou os efeitos informacionais de textos sobre o comportamento da taxa de câmbio, isso com base em uma literatura mais abrangente de análise textual que considera diversas fontes de dados, como notícias, atas ou relatórios. Dentro dessa linha na literatura de previsão da taxa de câmbio, um dos primeiros trabalhos a utilizar tal fonte de dados foi o [Galati e Ho \(2003\)](#), que investigaram a influência de notícias macroeconômicas na zona do euro e nos Estados Unidos sobre taxa câmbio entre o dólar e o euro, isso para os primeiros anos da União Monetária Europeia, considerando a origem e a classificação positiva ou negativa das notícias, assim foi observado

que as notícias macroeconômicas influenciavam as variações da taxa de câmbio estudada, mas de forma assimétrica.

Para a relação entre os fundamentos macroeconômicos e a taxa de câmbio, [Ehrmann e Fratzscher \(2005\)](#) estudaram e identificaram que notícias macroeconômicas (especificamente de anúncios) têm forte efeito sobre a variação da taxa de câmbio diário entre o dólar e o euro, isso para o período de 1993 a 2003. Na mesma linha, [Feuerriegel, Wolff e Neumann \(2016\)](#) utilizaram um modelo de *overshooting* de Dornbush para analisar as variações da taxa de câmbio, sendo uma versão convencional e outra ampliada com sentimento das notícias com o intuito de captar novas informações, os autores observaram que a variável de sentimento das notícias tem poder explicativo sobre a predição da taxa de câmbio.

Com o objetivo de investigar os efeitos das notícias macroeconômicas de manchetes dos jornais sobre as taxas de câmbio, [Caporale, Spagnolo e Spagnolo \(2017\)](#) identificaram um efeito transbordamento e assimétrico das notícias para os países do BRICS. Já [Caporale, Spagnolo e Spagnolo \(2018\)](#) observaram um efeito transbordamento de volatilidade entre as notícias e as taxas de câmbio para um grupo específico de países. Estes últimos sendo trabalhos que trazem mais evidências sobre a relação entre as notícias e as variações da taxa de câmbio.

Tendo como base a literatura de previsão com fundamentos macroeconômicos, [Bulut \(2018\)](#) buscou investigar como dados oriundos do Google Trends podem auxiliar nas predições da taxa de câmbio. Com dados de 2004 a 2014, cobrindo 11 países da OECD e da zona do euro, o autor observou que as predições sobre a direção das variações da taxa de câmbio com base no Google Trends tem boa performance, já para os retornos mensais foi identificada para uma das especificações testadas performance superior ao passeio aleatório em 5 das 11 taxas de câmbio, resultado esse não sendo pior do que o observado do modelo com base na paridade do poder de compra, que foi o modelo que gerou melhores resultados dentre os baseados nos fundamentos macroeconômicos. Vale destacar que não foi utilizado a regra de Taylor neste trabalho.

Sobre o uso de sentimento das notícias, que é uma das abordagens na análise de texto e que foi utilizada neste trabalho, podemos citar [Garcia \(2013\)](#) que com o foco nos períodos de recessão identificou que, para estes, as notícias têm poder sobre a previsibilidade do retorno das ações. Temos também o trabalho do [Shapiro, Sudhof e Wilson \(2018\)](#), que utilizando sentimentos das notícias observaram que as variáveis construídas melhoraram as previsões, além de que um choque positivo de sentimento se comporta de forma consistente com um choque de demanda agregada, sendo estes exemplos consistentes do uso de sentimento das notícias como preditores de variáveis econômicas. Se expandirmos o conceito para sentimento textual, não restringindo apenas as notícias, também podemos citar [Hájek \(2018\)](#), que usou para a previsão de retornos de ativos.

No trabalho de [Lima, Godeiro e Mohsin \(2019\)](#) foi desenvolvido um método para gerar sentimento textual por meio de um dicionário variante no tempo com base em métodos de regularização, tendo como objetivo de prever crescimento do produto com base nas minutas do FED. A abordagem de fronteira desenvolvida pelos autores é altamente compatível com o problema de previsão da taxa de câmbio, dado que o método é robusto a mudanças do padrão textuais das fontes, que é um problema para o uso de séries textuais com um período muito longo.

Unindo o ramo da literatura que estuda novas especificações para a modelagem na predição da taxa de câmbio com a que estuda as informações ou os efeitos informacionais das notícias, sendo a parte de análise textual, este trabalho busca investigar se as notícias, com base em dicionários fixos e variante no tempo, têm poder preditivo sobre as variações da taxa de câmbio, e se sua integração à regra de Taylor na forma de uma especificação alternativa melhora

os resultados em relação ao passeio aleatório e a uma regra de Taylor convencional, sendo estes os objetivos do trabalho.

A estrutura do trabalho está dividida em 4 capítulos: o primeiro é esta introdução que capta a motivação e seu referencial teórico; seguido do [Capítulo 2](#) que detalha tanto a construção das variáveis como a abordagem metodológica utilizada; e por fim, temos os capítulos [3](#) e [4](#) onde são demonstrados os resultados observados e as conclusões do trabalho.

2 Metodologia

2.1 Dados

As variáveis macroeconômicas utilizadas foram construídas com base nos dados obtidos no *International Financial Statistic* (IFS) do FMI e na [OECD](#)¹. Seguindo [Molodtsova e Papell \(2009\)](#), a taxa de inflação é uma taxa anual com base nas diferenças dos últimos 12 meses do índice de preços ao consumidor, já o hiato do produto na forma percentual foi construído por meio do filtro HP² utilizando a produção industrial com ajuste sazonal como *proxy* para o produto.

Devido à natureza das variáveis de sentimento geradas foi utilizada a taxa de câmbio média do período em detrimento do fim do período, considerando que o sentimento construído é com base nas notícias financeiras do respectivo mês, ou seja, o sentimento é relativo ao período em sua totalidade.

A base de dados textual tem como fonte o Down Jones Factiva, construído com as notícias financeiras do The New York Times e The Wall Street Journal, sendo que para o primeiro destes seu uso é recorrente na literatura, podendo ser observado nos trabalhos de [Engelberg, Reed e Ringgenberg \(2012\)](#), [Garcia \(2013\)](#) ou [Shapiro, Sudhof e Wilson \(2018\)](#). Já em [Liu e McConnell \(2013\)](#), temos o uso de ambos os jornais.

O banco de dados é mensal e cobre o período de janeiro de 1980 a maio de 2017, abrangendo os Estados Unidos, Reino Unido, Japão, Suécia, Coreia do Sul, Canadá, Dinamarca, Noruega, México e Brasil. Para selecionar os países analisados, foi inicialmente tomado como base os mais utilizados na literatura, após a retirada dos integrantes da zona do euro e os que não haviam disponibilidade dos dados, foi acrescentado o Brasil e México devido, respectivamente, ao seu uso anterior no trabalho de [Gaglianone e Marins \(2017\)](#) e a sua estreita relação com os Estados Unidos, mesmo tais países não tendo um longo histórico com taxas de câmbio flutuantes. Devido a dados faltantes para alguns países em relação à especificação da regra de Taylor indicada por [Molodtsova e Papell \(2009\)](#), o tamanho máximo da amostra varia entre pares de países analisados, o tamanho da amostra pode ser observado nos seus respectivos resultados ou na [Tabela 8 no Apêndice A](#).

2.2 Regra de Taylor

De acordo com [Taylor \(1993\)](#), o banco central segue uma regra de política monetária que pode ser expressa como:

$$i_t^* = \pi_t + \phi(\pi_t - \pi^*) + \gamma y_t + r^* \quad (2.1)$$

onde i_t^* é a taxa de juros de curto prazo, π_t a inflação corrente, π^* a meta de inflação, y_t o hiato do produto e r^* é a taxa de juros de equilíbrio, respectivamente.

¹ A informação mais detalhada sobre as fontes dessas variáveis pode ser observada na [Tabela 7 no Apêndice A](#).

² [Hodrick e Prescott \(1997\)](#), foi utilizado lambda de 14400, como feito por [Molodtsova e Papell \(2009\)](#).

Seguindo os passos indicados por [Molodtsova e Papell \(2009\)](#) para construir a equação da diferença da taxa de juros, a partir da equação (1), sendo $\lambda = (1 + \phi)$ e $\mu = r^* - \phi\pi^*$, a regra de política monetária pode ser expressa por:

$$i_t^* = \mu + \lambda\pi_t + \gamma y_t \quad (2.2)$$

para μ é suposto que r^* e π^* são constantes no tempo, que é uma forte suposição, dado que já foram trazidas evidências de que a meta de inflação para os Estados Unidos varia no tempo, segundo [Cogley, Primiceri e Sargent \(2010\)](#), e para o $\lambda > 1$ indica que a taxa de juros sobe quando a inflação sobe, satisfazendo as condições da regra de Taylor.

Considerando que a taxa de juros observável se ajusta parcialmente em relação à meta³, temos:

$$i_t = (1 - \rho)i_t^* + \rho i_{t-1} + v_t \quad (2.3)$$

com a equação da diferença da taxa de juros⁴ sendo expressa como:

$$i_t - \tilde{i}_t = \alpha_0 + \alpha_1\pi_t + \alpha_2y_t + \alpha_3i_{t-1} - (\alpha_4 + \alpha_5\tilde{\pi}_t + \alpha_6\tilde{y}_t + \alpha_7\tilde{i}_{t-1}) + \zeta_t \quad (2.4)$$

Partindo da paridade de juros a descoberta e supondo que seu coeficiente linear e angular são, respectivamente, 0 e -1 , podemos expressar a equação de predição da variação da taxa de câmbio sendo:

$$E_t S_{t+1} - S_t = \omega_0 - \omega_1\pi_t + \omega_2\tilde{\pi}_t - \omega_3y_t + \omega_4\tilde{y}_t - \omega_5i_{t-1} + \omega_6\tilde{i}_{t-1} + \psi_{t+1} \quad (2.5)$$

sendo S_t o logaritmo da taxa de câmbio nominal do dólar americano expresso em moeda estrangeira.

No trabalho de [Molodtsova e Papell \(2009\)](#) foram considerados 4 parâmetros de especificação para os modelos, podendo ser heterogêneo/homogêneo, simétrico/assimétrico, com/sem suavização e com/sem constante. Sobre estas especificações, os modelos são heterogêneos quando os coeficientes das mesmas variáveis (como inflação ou hiato do produto) para os pares de países são diferentes. Já o homogêneo, indica um coeficiente para o par de variáveis dos países, dessa forma implicando trabalhar com a diferença dos termos. O modelo é simétrico quando o número de variáveis para cada país é igual, sendo assimétrico quando o número é diferente, basicamente indicando a ausência ou presença da taxa real de câmbio na equação de predição. O termo de suavização do modelo corresponde a presença da taxa de juros de curto com uma defasagem do par de países.

Devido a discussão recente levantada por [Engel et al. \(2019\)](#) sobre a validade e consistência dos resultados do modelo com base na regra de Taylor para predição da variação da taxa de câmbio, foi comparado os resultados da regra de Taylor convencional, considerando variação em 3 dos 4 parâmetros entre as especificações utilizadas por [Molodtsova e Papell \(2009\)](#)⁵, em relação ao passeio aleatório para o grupo de países e amostra deste trabalho.

³ [Molodtsova e Papell \(2009\)](#) utilizaram como base [Clarida, Gali e Gertler \(1998\)](#).

⁴ O tilde indica país estrangeiro.

⁵ Foi fixado a constante aos modelos e testado as demais especificações, que seriam, respectivamente, heterogêneos ou homogêneos, simétricos ou assimétricos e com ou sem suavização.

O melhor resultado foi observado para a especificação da regra de Taylor heterogênea, assimétrica e com suavização⁶. Dessa forma, a equação de predição da variação da taxa de câmbio com base na regra de Taylor convencional pode ser expressa como:

$$\Delta S_{t+1} = \omega - \omega_{\pi us} \pi_t + \omega_{\pi f} \tilde{\pi}_t - \omega_{y us} y_t + \omega_{y f} \tilde{y}_t + q_t - \omega_i i_{t-1} + \omega_{\tilde{i} f} \tilde{i}_{t-1} + \eta_t \quad (2.6)$$

Para a Coreia do Sul, usamos a especificação da regra de Taylor heterogênea, simétrica e com suavização por cauda da ausência de dados da variável da taxa real de câmbio nas fontes utilizadas para o período. Assim, seguimos a especificação indicada por [Molodtsova e Papell \(2009\)](#).

2.3 Análise textual

Na análise textual temos o uso recorrente de dicionários fixos no processo de transformação de dados textuais em numéricos, que convencionalmente utiliza o sentido das palavras para obter um sentimento textual, sendo uma abordagem consolidada na literatura. Seu uso traz consigo imposições fortes, como o sentido e o conjunto de palavras analisadas são fixos no tempo, sendo essas imposições que tornam a abordagem rígida em relação aos dados textuais ao longo do tempo, principalmente para longos períodos.

Uma abordagem que contorna a rigidez imposta pelos dicionários fixos é a do dicionário variante no tempo, proposta por [Lima, Godeiro e Mohsin \(2019\)](#), que com base no poder preditivo da palavra e em métodos de aprendizado de máquina tornou possível que o dicionário utilizado varie ao longo do tempo.

Este trabalho utiliza dois tipos de dicionários, os fixos e o variante no tempo, seja como forma de validação e/ou comparação do resultado. A abordagem com dicionários fixo é descrita a seguir na [subseção 2.3.1](#), já na [subseção 2.3.2](#) temos a descrição da abordagem do dicionário variante no tempo.

2.3.1 Dicionário fixo

Seguindo [Feuerriegel, Wolff e Neumann \(2016\)](#), para construção dos sentimentos das notícias financeiras são utilizados 3 dicionários diferentes e que são amplamente utilizados na literatura, o *General Inquirer* (GI), *Henry's Financial dictionary* (HE) e *Loughran-McDonald dictionary* (LM), que iremos nos referir, respectivamente, como os dicionários “GI”, “HE” e “LM”, sendo os dois últimos com base em suas versões dos trabalhos [Henry \(2008\)](#) e [Loughran e McDonald \(2011\)](#)⁷, respectivamente. No trabalho de referência, o dicionário base é o HE, utilizando o GI e o LM como robustez, além da questão de possível robustez, utilizar os 3 pode gerar a possibilidade de obter diferentes *insights* em relação aos sentimentos gerados por cada dicionário, considerando que seus propósitos são diferentes.

O *General Inquirer* (GI) é um dicionário desenvolvido por [Stone, Dunphy e Smith \(1966\)](#) que é de propósito geral, sendo amplamente utilizado na literatura que comporta o uso de

⁶ Os resultados de todas as especificações testadas previamente podem ser observadas nas tabelas 10 e 11 no [Apêndice B](#). No mesmo apêndice, pode ser observado os coeficientes da inflação ao longo do tempo dos Estados Unidos e do país estrangeiro da especificação da regra de Taylor utilizada neste trabalho como base e a indica como a melhor por [Molodtsova e Papell \(2009\)](#).

⁷ Existe versões mais recentes deste dicionário, porém optamos por esta versão devido ao forte uso na literatura em detrimento das mais recente que não são testadas tanto quanto a de 2011

dicionário e lista de palavras, como nos trabalhos de Tetlock (2007), Tetlock, Saar-Tsechansky e Macskassy (2008), Engelberg, Reed e Ringgenberg (2012) e Feuerriegel, Wolff e Neumann (2016), mas segundo Kearney e Liu (2014) existem algumas ressalvas sobre este dicionário, se for considerado o contexto financeiro, por exemplo, algumas palavras possuem erros de classificação sobre seu sentido, como "*liability*" e "*tax*" que são consideradas negativas no GI mas dentro do contexto financeiro não são.

O *Henry's Financial dictionary* (HE) é um dicionário estritamente financeiro, feito para este contexto específico, exemplos de sua utilização são nos trabalhos de Henry (2006), Henry (2008), Henry e Leone (2009), Price et al. (2012) e Feuerriegel, Wolff e Neumann (2016). Já o *Loughran-McDonald dictionary* (LM) foi desenvolvido para suprir os erros de classificação no contexto financeiro observados no GI. Além de incorporar termos financeiros, expandiu as listas de palavras e possibilitou uma gama maior de classificação/dicionário de palavras, como seu dicionário de incerteza, alguns dos trabalhos que o utilizaram foram Ferris, Hao e Liao (2012), Garcia (2013), Loughran e McDonald (2013), Ferguson et al. (2015) e Feuerriegel, Wolff e Neumann (2016).

Sobre o uso de dicionários com diferentes propósitos, Henry e Leone (2015) argumentam que para dados textuais com linguagem especializada é indicado o uso de dicionário de domínio específico em detrimento dos de propósito geral, devido a alguns fatores como erros de classificação, já citados anteriormente. Uma resposta a esse uso cruzado seria captar informações com baixo poder preditivo. Os dados textuais utilizados são de notícias financeiras, mas não necessariamente a linguagem utilizada é estritamente especializada, dado que o público receptor dessas notícias é amplo e diversificado, o dicionário mais indicado seria o LM, que mesmo sendo de propósito específico ele também capta informação de propósito geral. A natureza das informações textuais utilizadas, porém, não veta o uso dos demais dicionários, como o GI que é de propósito geral.

O pré-processamento aplicado aos textos foi basicamente o fornecido pelo pacote utilizado⁸, que é uma limpeza simples, lidando com a remoção de *stop words* e filtrando para o radical da palavra, que é o processo conhecido como *stemming*. Após esse pré-processamento, é aplicada uma regra utilizando as listas de palavras do dicionário por mês para gerar os sentimentos. As métricas utilizadas são simples: para sentimentos positivos, negativos e de incerteza, a regra é dada pela frequência relativa da palavra, métrica conhecida como *term-frequency* (tf); para o sentimento geral, a regra é dada por:

$$Sentimento_t = \frac{Positivas_t - Negativas_t}{Total_t}$$

Dadas as regras utilizadas neste trabalho para gerar sentimentos, uma limpeza simples acaba por inflar o total de palavras que é usado como denominador em todas as regras, acabando por reduzir em nível todos os sentimentos gerados, mas o comportamento da série é conservado.

Uma métrica bastante utilizada na literatura de análise textual é a frequência relativa da palavra ponderada pelo inverso da frequência do documento, conhecido como *term frequency - inverse document frequency* (TF-IDF), sugerida por Loughran e McDonald (2011), essa métrica penaliza tanto palavras raras como palavras comuns, além de ter outras funcionalidades, como por exemplo, identificar documentos de interesse. Como discutido por Henry e Leone (2015), quando o exercício requer apenas análise de conteúdo, os dados textuais de interesse já foram selecionados e a análise não passa pelo poder preditivo da palavra. Assim, a ponderação igual

⁸ SentimentAnalysis.

pode ser usada em detrimento da ponderação pelo IDF, já que nesse contexto os resultados são parelhos, além de exigir menor poder computacional para execução e facilitar a replicação, justificando o uso da métrica empregada.

A [Tabela 1](#) a seguir indica o número de palavras por classificação e dicionário, como pode ser observado na mesma, para o dicionário LM o número de palavras negativas é 6.1 vezes maior que o número de palavras positivas, característica essa que não é observada nos outros dicionários e que facilita o sentimento geral ser negativo.

Tabela 1 – Palavras por classificação e dicionário

	Positiva	Negativa	Incerteza
<i>General Inquirer</i>	1316	1746	-
<i>Henry's Financial dictionary</i>	53	44	-
<i>Loughran-McDonald dictionary</i>	145	885	129

Fonte: Elaboração própria.

Na predição um passo à frente com o sentimento das notícias financeiras com dicionário fixo⁹, podemos expressar a equação de predição da variação da taxa de câmbio com sentimento das notícias sendo:

$$\Delta S_{t+1} = Z_t' \phi + \epsilon_{t+1} \quad (2.7)$$

onde S_t é o logaritmo da taxa de câmbio, Z_t é o vetor de variáveis de sentimento utilizadas, ϕ é o vetor de coeficientes e ϵ_t é o resíduo.

Para cada equação estimada é considerado apenas um dicionário, ou seja, sem cruzamento de sentimentos de diferentes dicionários, e como o sentimento geral é uma combinação do positivo com o negativo, são utilizados até dois sentimentos por modelo, o sentimento geral com o negativo ou positivo.

2.3.2 Dicionário variante no tempo

Inicialmente, antes de falar do método em si, temos que falar do pré-processamento dos dados textuais, considerando que na abordagem anterior é feita sobre a palavra (unigramas) e nessa é sobre termos (unigramas, bigramas ou trigramas), podendo ser uma ou mais palavras em conjunto.

A importação é feita em formato de *Corpus* e o pré-processamento passa por remoção de *stop words*, pontuação, números, espaços em branco e manipulação do texto para ter apenas letras minúsculas¹⁰ e o radical das palavras, trabalhando com o *stemming* da palavra, além de considerar até trigramas, ou seja, leva em consideração palavras que fazem sentimento conjuntamente, como

⁹ As séries de sentimento das notícias financeiras geradas podem ser observadas no [Apêndice C](#).

¹⁰ Esse processo é necessário, pois, a linguagem R que é utilizada neste trabalho é sensível a palavras maiúsculas e minúsculas, sendo *case sensitive*.

"stock market", que é um bigrama e tem sentido econômico único. Dessa forma, a referência aqui é por termos, considerando que o termo pode conter mais de uma palavra. Com o tratamento dos dados textuais realizado, o formato final trabalhado é o *bag of words*.

Na abordagem do dicionário variante no tempo (TVD) é utilizado o poder preditivo dos termos, seguindo a metodologia utilizada por Lima, Godeiro e Mohsin (2019). O dicionário variante no tempo é uma matriz $X_t^* \subset X_t$ contendo os termos com maior poder preditivo em relação a nossa variável dependente, que é a variação da taxa de câmbio ΔS_{t+1} . Dessa forma, aplicamos *Elastic Net* na equação de predição, dado que o nosso horizonte é de um passo à frente ($h = 1$), podemos representar tal equação como sendo:

$$\Delta S_{t+1} = W_t' \beta + X_t' \phi + \epsilon_t \quad (2.8)$$

onde ΔS_{t+1} é a variação do logaritmo da taxa de câmbio, W_t é o vetor de preditores predeterminados¹¹, X_t é o vetor que contém os termos, já $\hat{\beta}$ e $\hat{\phi}$ são os coeficientes estimados pela função objetivo que pode ser apresentada como:

$$\min_{\beta, \phi} \sum_t (\Delta S_{t+1} - W_t' \beta - X_t' \phi)^2 + \lambda_1 \|\phi\|_{l_1} + \lambda_2 \|\phi\|_{l_2} \quad (2.9)$$

sendo $\|\cdot\|_{l_1}$ e $\|\cdot\|_{l_2}$, respectivamente, as normas l_1 e l_2 , onde l_1 penaliza o valor absoluto dos coeficientes, enquanto o l_2 penaliza o valor ao quadrado dos coeficientes. A estimação dessa equação é feita de forma recursiva¹². Quando $\lambda_1 = 0$ ficamos com a penalização em l_2 , que é o estimador Ridge, para $\lambda_2 = 0$ temos a penalização em l_1 , sendo o estimador LASSO. Já quando λ_1 e λ_2 são maiores que zero, detemos o *Elastic Net*.

A utilização do *Elastic Net*, proposto por Zou e Hastie (2005), é justificado pela necessidade de lidar com duas características importantes na abordagem, as quais são: a alta dimensionalidade de X_t , ou seja, $p \gg T$, dado que esse vetor contém todas as palavras utilizadas como preditores, isso faz necessário um método de seleção que é uma característica do LASSO; e, a correlação entre palavras importantes que são preditores, dado que o Ridge não penaliza para zero os coeficientes, é um método com boas propriedades para lidar com essa característica, diferentemente do LASSO. Esse par de características demandam propriedades de ambos os métodos, tanto LASSO como Ridge. Dessa forma, o *Elastic Net* é o método mais compatível com a abordagem, garantindo a seleção das palavras mais preditivas sem o problema de descarte de outras palavras devido a correção que são tão preditivas quanto as selecionadas.

Uma questão importante nesse método é a escolha dos lambdas, já que eles influenciam diretamente na seleção dos termos mais preditivos. Neste trabalho foi feito o mesmo processo de escolha dos lambdas utilizado por Lima, Godeiro e Mohsin (2019), que é um procedimento de validação cruzada desenvolvido por Trevor Hastie e Junyang Qian, onde é feita uma seleção ex ante a fim de evitar *overfitting* do modelo¹³.

A estimação de forma recursiva com o *Elastic Net* acaba por selecionar ao longo do tempo as palavras com maior poder preditivo, dado que passa por re-estimação, fazendo com que o dicionário varie com o tempo, onde um $X_t^* \subset X_t$ é utilizado com os termos que são os melhores preditores.

¹¹ Não foi utilizado o vetor de preditores predeterminados nesse processo de construção dos sentimentos, só os termos em função da variação do logaritmo da taxa de câmbio, ΔS_{t+1} .

¹² Foi utilizado 120 observações para amostra inicial.

¹³ λ_1 e λ_2 são combinações diferentes de um α e λ internos do *Elastic Net*, com base nos resultados e nas características desejadas de seleção do LASSO, foi usado $\alpha = 0.15$ de base.

Dado que o interesse é trabalhar com o sentimento textual, utilizamos essas estimações para identificar se os termos têm efeito positivo ou negativo sobre nossa variável dependente com base no sinal dos seus coeficientes. Com essa informação podemos identificar os termos positivos e negativos dentro do X_t^* e extrair o sentimento geral, obedecendo regra semelhante do dicionário fixo, a diferença é que o total passa a ser a soma dos termos positivos e negativos. Por fim, seguindo [Jiang et al. \(2019\)](#), é feita uma suavização por meio de uma média móvel.

Como uma forma de reduzir o número de palavras em X_t antes de aplicarmos o método do dicionário variante no tempo, foi aplicado a métrica já mencionada do *term frequency - inverse document frequency* (TF-IDF), que pode ser usada para descartar termos muito frequentes e raros. Como o número de termos escolhidos pelo ranqueamento dos melhores TF-IDF também é um hiperparâmetro, a escolha foi feita com base na melhor raiz do erro quadrático médio (RMSE) dos resultados obtidos utilizando um intervalo 1000 a 10000 termos, separados por incremento de 100 termos¹⁴.

A partir do sentimento geral construído com os termos com maior poder preditivo^{15, 16}, podemos representar a equação de predição da variação da taxa de câmbio com sentimento geral para previsão um passo à frente, como sendo:

$$\Delta S_{t+1} = D_t' \phi + \epsilon_{t+1} \quad (2.10)$$

onde D_t é o vetor de sentimento textual baseado nas notícias financeiras do The New York Times e do The Wall Street Journal, sendo $D_t = g(X_t^*)$, com seu respectivo vetor de coeficientes ϕ .

2.4 Regra de Taylor ampliada

Com os sentimentos gerais produzidos pelos dicionários fixos e variante no tempo, podemos ampliar a equação (2.6) no molde ao feito por [Morales-Arias e Moura \(2013\)](#) e [Wang, Morley e Ordóñez \(2016\)](#), que testaram novos preditores e ampliaram a especificação da equação de predição com base nos fundamentos macroeconômicos. Dessa forma, é gerado uma nova equação de predição da variação da taxa de câmbio com base na regra de Taylor, que pode ser expressa como sendo:

$$\Delta S_{t+1} = W_t' \beta + N_t' \phi + \epsilon_{t+1} \quad (2.11)$$

onde W_t é o vetor de preditores da regra de Taylor, N_t pode ser o vetor de preditores Z_t ou D_t , que são, respectivamente, oriundos dos dicionários fixos e variante no tempo, com β e ϕ sendo o vetor dos coeficientes. Sendo esta à equação referida da especificação da regra de Taylor ampliada.

2.5 Método de estimação

Seguindo [Molodtsova e Papell \(2009\)](#), para obter a predição fora da amostra um passo à frente dos modelos é empregado o método de Mínimos Quadrados Ordinários (MQO) com

¹⁴ O número de termos utilizados foram 3300, 1400, 3600, 1000, 1000, 2600, 2000, 5300 e 5600, respectivamente, para o Reino Unido, Japão, Suécia, Coreia do Sul, Canadá, Dinamarca, Noruega, México e Brasil. Os resultados do RMSE por país para o número de termos testado podem ser observados no [Apêndice D](#).

¹⁵ As séries de sentimento das notícias financeiras geradas podem ser observadas no [Apêndice C](#).

¹⁶ No [Apêndice E](#) podemos observar amostras dos termos com efeitos positivos e negativos do dicionário variante no tempo por país e datas de referência selecionadas, com o intuito de observar diferenças do país e do tempo nos termos usados.

janela móvel, e com base em [Clark e West \(2006\)](#) é utilizado uma janela de 120 observações para treinar o modelo.

2.6 Testes de performance

Para comparar as performances fora da amostra em relação ao do passeio aleatório sem constante e aos *nested models*, foi utilizado o teste de Clark e West (CW) com correção, com base nos trabalhos [Clark e West \(2006\)](#) e [Clark e West \(2007\)](#). Já para comparações entre *non-nested models*, o teste de Diebold e Mariano (DM)¹⁷, baseado em [Diebold e Mariano \(1995\)](#), foi empregado.

Para ambos os testes de comparação de performance utilizados, CW e DM, a hipótese nula é que o modelo base e o modelo alternativo têm performances estaticamente iguais, com a hipótese alternativa é de que o modelo alternativo tem performance superior ao modelo base.

¹⁷ Foi especificado o erro ao quadrado como o método da função de perda (função de custo).

3 Resultados e discussões

Para facilitar a visualização e demonstração dos resultados também é utilizado notações para os modelos bases (*benchmarks*), na identificação dos dicionários e para o sentimento de incerteza, tanto nas tabelas como no texto. Dessa forma, temos o passeio aleatório sem constante referido por “RW”, regra de Taylor por “TR”, para os dicionários fixos temos “GI”, “HE” e “LM”, que representam, respectivamente, o *General Inquirer*, *Henry’s Financial* e *Loughran-McDonald*, sendo o “TVD” para identificar o dicionário variante no tempo e para o sentimento de incerteza é “Incert”. Como a predição é um passo à frente, todos os resultados são referentes a previsibilidade no curto prazo.

As tabelas 2 e 3 contêm os resultados dos modelos só com sentimentos das notícias, indicando sua precisão por meio do RMSPE e os resultados dos testes de comparação de performance da predição, onde são comparados com o passeio aleatório sem constante e a regra de Taylor convencional. Sobre a precisão, o modelo com sentimento geral do dicionário variante no tempo obteve a melhor precisão em todos os país analisados. Já os resultados entre os dicionários fixos, para 7 dos 9 países, temos o sentimento geral do GI ou do LM sendo o que gerou a melhor precisão.

Para os resultados com os dicionários fixos sobre a comparação de performances em relação ao RW, observando os resultados a 5% de significância, apenas o Japão, Coreia do Sul e México tiveram pelo menos 1 modelo com sentimento obtendo superioridade na comparação, observando a 10% temos pelo menos 1 modelo o Canadá, Dinamarca e os demais países citados a 5% de significância. Já seus resultados em relação ao TR, podemos observar que para 5 dos 9 países pelo menos 1 modelo com sentimento que obteve superioridade de performance a 5% de significância, reduzindo o rigor para 10% continuamos com o mesmo grupo de países, sendo o Reino Unido, Japão, Coreia do Sul, Dinamarca e México. Os resultados expressivos foram do dicionário variante no tempo, que obteve superioridade a 1% de significância sobre o RW em todos os países analisados, para o mesmo nível de significância obteve superioridade sobre a regra de Taylor convencional em 7 dos 9 países. Se for considerado 5% de significância para os resultados dos testes do TVD, temos sua superioridade para todos os países na comparação com o RW e TR.

Os primeiros resultados indicam fortes evidências que modelos só com sentimentos das notícias, construídos com dicionário variante no tempo, possuem alta previsibilidade na predição um passo à frente da variação da taxa de câmbio. Isso quando comparados aos respectivos modelos base, que seriam o passeio aleatório sem constante e a regra de Taylor convencional. Os resultados observados, considerando as diferenças do trabalhos desde grupo de países a tamanho da amostra, foram mais expressivos dos que os identificados em [Bulut \(2018\)](#). Trabalho esse que é um dos recentes na literatura que usam métodos não convencionais na previsão da variação da taxa de câmbio, em que os modelos tiveram como fonte os dados do Google Trends sendo direcionados por termos dentro dos fundamentos macroeconômicos. Para os resultados com dicionários fixos é apontado baixa previsibilidade, principalmente em relação ao passeio aleatório.

As evidências e os resultados observados nas tabelas 2 e 3, até mesmo aqueles mais fracos, podem ser questionados sobre a correlação, entre as variáveis de sentimento em relação a variação da taxa de câmbio, ser a origem desses resultados. Como uma forma de observar

tal questionamento, a [Tabela 4](#) apresenta os coeficientes de correlação entre estas variáveis. Os valores desses coeficientes em relação aos sentimentos gerados por dicionários fixos indicam uma correlação baixa ou muito baixa, tornando fraco esse questionamento para esse grupo de sentimentos. Para o sentimento com base no dicionário variante no tempo, foi observado correlações moderadas para 5 dos 9 países, com os demais tendo uma correção fraca. Os coeficientes de correlação do TVD sendo maiores do que os produzidos dos dicionários fixos é um resultado esperado, dado que o sentimento do TVD é construído em função da variável dependente, que no caso é variação da taxa de câmbio, sendo correlacionadas por natureza. Assim, apesar dessa característica, as correlações fracas e moderadas não podem ser usadas como argumento para refutar os resultados gerados pelo dicionário variante no tempo.

A diferença entre os resultados dos modelos com dicionários fixos e variante no tempo podem ser explicados em parte pela variação ao longo do tempo do vocabulário das notícias do The New York Times e do The Wall Street Journal, que é a base textual utilizada neste trabalho. [Lima, Godeiro e Mohsin \(2019\)](#) também observaram a mesma diferença entre dicionário fixo e variante no tempo na predição do crescimento do produto utilizando as minutas do FED com fonte textual. Dessa forma, a um indicativo do uso do método do dicionário variante no tempo em detrimento dos dicionários fixos para análise de sentimento textual com grandes amostras, considerando fonte textual especializada ou não.

Nas tabelas [5](#) e [6](#) podem ser observados os resultados dos modelos ampliados da regra de Taylor com sentimento das notícias por dicionário, fixos e variante no tempo, indicando sua precisão e o resultado da comparação com o RW e o TR com base no teste CW.

Para os resultados com dicionário fixo, os testes de performance indicam superioridade de pelo menos 1 modelo ampliado para 7 dos 9 países, considerando 5% de significância, em relação ao RW. Para 10% de significância é observado 3 dos 9 países tendo todos os modelos ampliados superiores ao passeio aleatório sem constante, sendo o Canadá, Noruega e México. Na comparação com a regra de Taylor convencional, a ampliação do modelo com sentimentos das notícias por dicionário fixo resultaram em melhora de performance na predição. Para 7 dos 9 países, foi observado melhora em todas as especificações dos modelos com dicionário fixo, resultado esse considerando 5% de significância.

Se for observado os resultados por dicionário individualmente, tanto o HE e LM, conseguiram performance superior para 4 dos 9 países considerando 5% de significância em relação ao passeio aleatório. Quando comparados seus modelos ampliados com a regra de Taylor convencional, o modelo ampliado com o sentimento geral do GI obteve resultados fracos, superioridade para apenas 1 país a 5% e 3 países a 10% de significância, enquanto os ampliados pelo sentimento geral do HE e LM obtiveram superioridade, respectivamente, para 5 e 4 dos 9 países a 5% de significância. Para o sentimento de incerteza do dicionário LM os resultados foram fracos em relação aos modelos base, o passeio aleatório e a regra de Taylor convencional.

O modelo da regra de Taylor ampliado com sentimento geral do dicionário variante no tempo obteve, novamente, a melhor precisão para todos os países analisados, além da superioridade de performance na predição um passo à frente da variação da taxa de câmbio em relação ao passeio aleatório sem constante e a regra de Taylor convencional, também para todos os 9 países. Todos os resultados de comparações de performance entre especificações foram observados a 1% de significância.

Como visto anteriormente, o sentimento geral construído com o dicionário variante no tempo têm alto poder preditivo sobre a previsão da variação da taxa de câmbio no curto prazo. Além disso, também foi observado o indicativo de que essa variável possui a capacidade de

melhorar a performance da predição dos modelos com base nos fundamentos da regra Taylor, onde o modelo ampliado foi superior ao convencional para todos os países a 1% de significância. Vale ressaltar os resultados do Brasil, que só modelos com sentimento do TVD conseguiram superar o passeio aleatório, podendo ser um indicio de que o método do TVD também tem boa performance para séries de tempo mais curtas.

Tabela 2 – Resultados dos modelos só com sentimento das notícias - Parte 1

		Modelos:											
Benchmark		Geral - GI		Geral - HE		Geral - LM		Incert - LM		(Geral+Incert) - LM		Geral - TVD	
Reino Unido	RW TR	0.02375	**	0.02398	*	0.02377	**	0.02389	**	0.02398	*	0.02147	*** ***
Japão	RW TR	0.02616	**	0.02575	** ***	0.02564	** **	0.02628	**	0.02575	** **	0.02263	*** ***
Suécia	RW TR	0.02717		0.02741		0.02727		0.02729		0.02754		0.02271	*** **
Coreia do Sul	RW TR	0.03151	*	0.03161	*	0.03138	** **	0.03186	*	0.03147	** *	0.02982	** **
Canadá	RW TR	0.01696		0.01708	*	0.01694	*	0.01713		0.01702	*	0.01528	*** **

Fonte: Elaboração própria.

Nota: A tabela reporta os RMSPE e os resultados dos testes de comparação de performance, indicados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para as comparações com o RW foi utilizado o teste CW, enquanto para a comparação com o TR foi utilizado o teste DM, com suas hipóteses nulas indicando igualdade e a alternativa indicando superioridade de performance da predição alternativa. Sobre o tamanho da amostra, Japão e Suécia possuem, respectivamente, 382 e 246 observações, os demais países possuem 449 observações.

Tabela 3 – Resultados dos modelos só com sentimento das notícias - Parte 2

		Modelos:													
		Benchmark		Geral - GI		Geral - HE		Geral - LM		Incert - LM		(Geral+Incert) - LM		Geral - TVD	
Dinamarca	RW TR	0.02393	**	0.02377	* **	0.02383	**	0.02387	**	0.02387	**			0.02095	*** ***
Noruega	RW TR	0.02572		0.02577		0.02571		0.02564		0.02583				0.02177	*** ***
México	RW TR	0.03400	*** **	0.03527	*** **	0.03383	***	0.03424	*** **	0.03402	*** **			0.03029	*** ***
Brasil	RW TR	0.03521		0.03612		0.03626		0.03547		0.03660				0.02968	*** ***

Fonte: Elaboração própria.

Nota: A tabela reporta os RMSPE e os resultados dos testes de comparação de performance, indicados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para as comparações com o RW foi utilizado o teste CW, enquanto para a comparação com o TR foi utilizado o teste DM, com suas hipóteses nulas indicando igualdade e a alternativa indicando superioridade de performance da predição alternativa. Sobre o tamanho da amostra, Dinamarca e Brasil possuem, respectivamente, 364 e 221 observações, os demais países possuem 449 observações.

Tabela 4 – Tabela de correlação

ΔS_{t+1}	Sentimentos:				
	Geral - GI	Geral - HE	Geral - LM	Incerteza - LM	Geral - TVD
Reino Unido	-0.0317	0.0235	-0.0776	-0.0212	0.4495
Japão	0.1125	0.2008	0.2108	-0.0755	0.5060
Suécia	-0.0631	-0.0069	-0.0568	0.0658	0.5709
Coreia do Sul	-0.1182	-0.0624	-0.1863	-0.0504	0.5307
Canadá	-0.1070	-0.0543	-0.1456	0.0180	0.4202
Dinamarca	0.0130	0.1454	0.0877	0.0684	0.4647
Noruega	-0.0104	0.0926	-0.0094	0.0173	0.5170
México	0.0818	-0.1331	-0.1615	0.0004	0.4126
Brasil	-0.1694	-0.1570	-0.2525	0.0666	0.6017

Fonte: Elaboração própria.

Nota: A tabela reporta os coeficientes de correlação de Pearson entre cada variável de sentimento com a da variação da taxa de câmbio por país. Sobre o tamanho da amostra, Japão, Suécia, Dinamarca e Brasil possuem, respectivamente, 382, 246, 364 e 221 observações, os demais países possuem 449 observações.

Tabela 5 – Resultados dos modelos ampliados da regra de Taylor com sentimento das notícias - Parte 1

		Modelos ampliados com:											
	Benchmark	Geral - GI		Geral - HE		Geral - LM		Incert - LM		(Geral+Incert) - LM		Geral - TVD	
Reino Unido	RW	0.02418	*	0.02440	**	0.02411	**	0.02436	***	0.02430	*	0.02161	***
	TR		***		***		***				***		***
Japão	RW	0.02731	***	0.02660	**	0.02689	*	0.02719	***	0.02697	*	0.02312	***
	TR				***		***				***		***
Suécia	RW	0.02817		0.02787	*	0.02779		0.02814		0.02796		0.02285	***
	TR												***
Coreia do Sul	RW	0.03294	***	0.03287	***	0.03290	***	0.03339	**	0.03309	***	0.02978	***
	TR												***
Canadá	RW	0.01695	***	0.01699	***	0.01689	***	0.01717	**	0.01698	***	0.01560	***
	TR		**		**		***		**		**		***

Fonte: Elaboração própria.

Nota: A tabela reporta os RMSPE e os resultados dos testes de comparação de performance, indicados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para as comparações com o RW foi utilizado o teste CW, enquanto para a comparação com o TR foi utilizado o teste DM, com suas hipóteses nulas indicando igualdade e a alternativa indicando superioridade de performance da predição alternativa. Sobre o tamanho da amostra, Japão e Suécia possuem, respectivamente, 382 e 246 observações, os demais países possuem 449 observações.

Tabela 6 – Resultados dos modelos ampliados da regra de Taylor com sentimento das notícias - Parte 2

		Modelos ampliados com:											
Benchmark		Geral - GI		Geral - HE		Geral - LM		Incert - LM		(Geral+Incert) - LM		Geral - TVD	
Dinamarca	RW TR	0.02509	***	0.02474	**	0.02489	***	0.02500	***	0.02490	***	0.02139	***
Noruega	RW TR	0.02591	*	0.02594	***	0.02582	**	0.02582	*	0.02596	**	0.02209	***
México	RW TR	0.03510	*	0.03464	***	0.03486	***	0.03497	*	0.03472	**	0.03238	***
Brasil	RW TR	0.04092		0.03973		0.04142		0.04018		0.04183		0.03079	***

Fonte: Elaboração própria.

Nota: A tabela reporta os RMSPE e os resultados dos testes de comparação de performance, indicados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para as comparações com o RW foi utilizado o teste CW, enquanto para a comparação com o TR foi utilizado o teste DM, com suas hipóteses nulas indicando igualdade e a alternativa indicando superioridade de performance da predição alternativa. Sobre o tamanho da amostra, Dinamarca e Brasil possuem, respectivamente, 364 e 221 observações, os demais países possuem 449 observações.

3.1 Consistência dos modelos ampliados

Os modelos de previsão da variação da taxa de câmbio com base nos fundamentos da regra de Taylor estão sujeitos à consistência na sua aplicação, que seria responder de forma esperada na variação da taxa de câmbio em forma de apreciação ou depreciação, dada uma variação da inflação. A verificação dessa consistência passa por observar os coeficientes da inflação ao longo do tempo, dos Estados Unidos e do país estrangeiro, nas estimações feitas com janela móvel.

Na literatura a consistência é esperada por parte dos coeficientes da inflação dos Estados Unidos, como pode ser observado nos trabalhos de [Molodtsova e Papell \(2009\)](#) e [Engel et al. \(2019\)](#), dado que sua política monetária pode ser caracterizada com base nesse fundamento macroeconômico a partir de meados da década de 1980. Dessa forma, como a série começa em 1980 para a maioria dos países, não existe a expectativa de observar os sinais esperados desses coeficientes para os primeiros anos da série, dado que os primeiros resultados dos coeficientes possuem na sua amostra para estimação anos incompatíveis com essa caracterização da política monetária. Para os coeficientes da inflação dos demais países, não é esperado que os resultados indiquem uma forte consistência, dado que a caracterização da política monetária com uma regra de Taylor não é tão evidente quanto é nos Estados Unidos.

Assim, dada a estrutura dos fundamentos macroeconômicos da regra de Taylor, a equação de predição utilizada e a taxa de câmbio sendo expressa em termos da moeda estrangeira, é esperado coeficientes positivos para a inflação os Estados Unidos, dado que um aumento (redução) da sua inflação levaria a uma apreciação (depreciação) da sua moeda. Para o país estrangeiro na equação são esperados coeficientes negativos.

Os resultados dos modelos da regra de Taylor ampliados com sentimento das notícias, tanto dos dicionários fixos como variante no tempo, em relação a precisão e superioridade de performance, é observado apenas a consistência dos modelos ampliados com o sentimento gerado pelo dicionário variante no tempo, que foi o modelo que gerou os melhores resultados, superando fortemente as demais especificações¹⁸.

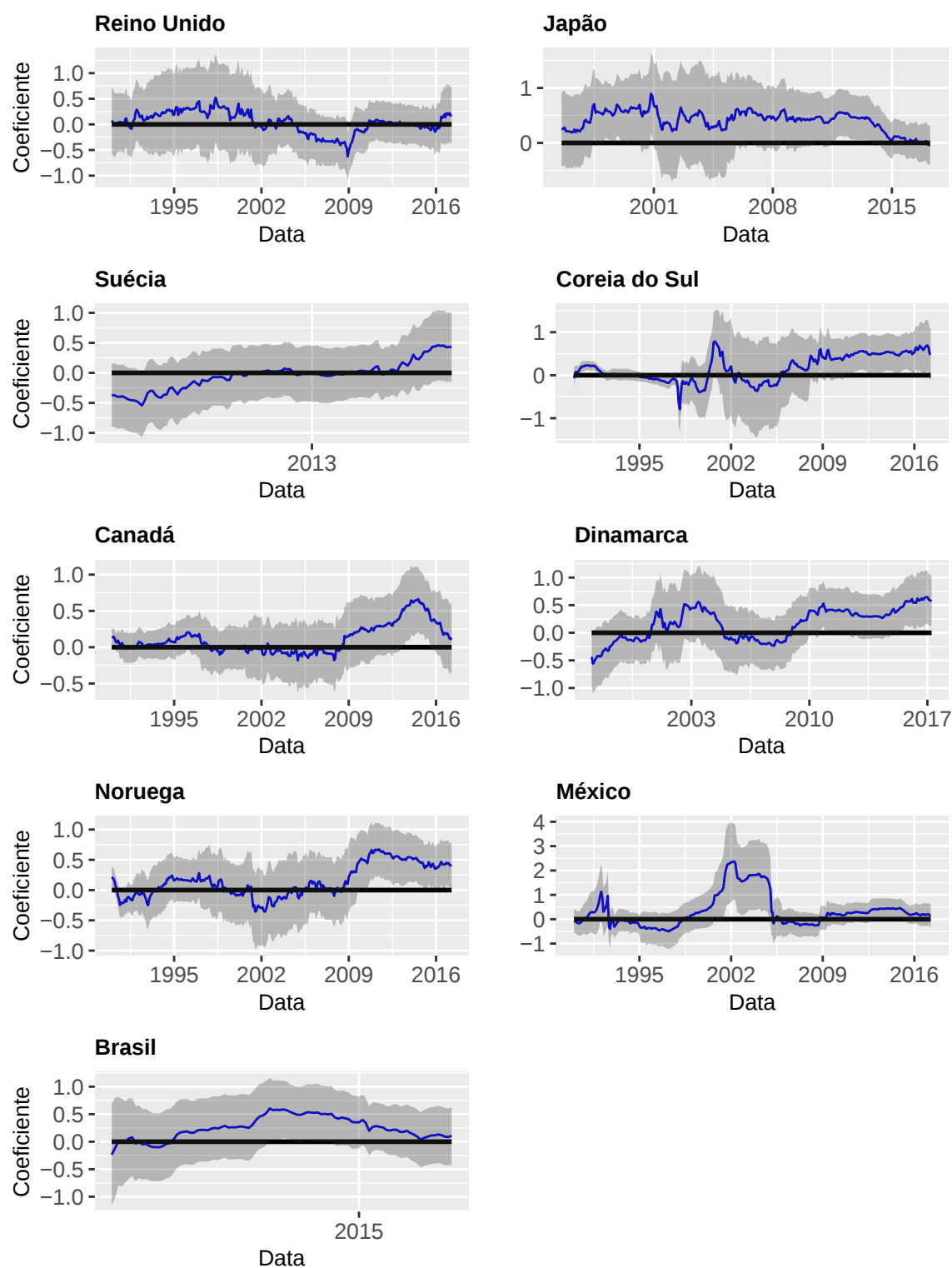
As figuras 1 e 2 demonstram, respectivamente, os coeficientes inflação dos Estados Unidos e do país estrangeiro ao longo das estimações para os 9 países analisados, juntamente com seus respectivos intervalos de confiança a 95%, oriundos e construídos com base nas estimações com janela móvel da equação (2.11).

Na [Figura 1](#), dos coeficientes dos Estados Unidos, temos o coeficiente esperado de forma predominante em 2 dos 9 países, o intervalo de confiança não podemos descartar o zero como resultado para a maior parte da série. Se for observado os resultados dos coeficientes a partir de 2009, período pós crise de 2008, o padrão do sinal esperado esperado é identificado na maioria dos países, como Coreia do Sul, Canadá, Dinamarca, Noruega e México, além do Japão e Brasil que tem seus padrões sendo predominantemente esperado. Assim, podemos ter um indício de que a consistência dos modelos com esse fundamento macroeconômico foi estabilizada a partir do período em questão. Assim, os coeficientes da inflação dos Estados Unidos são consistentes para 7 dos 9 países a partir de 2009, mas para a série completa só é observada consistência para 2 dos 9 países.

¹⁸ As figuras contendo os coeficientes da inflação dos Estados Unidos e do país estrangeiro dos modelos da regra de Taylor ampliados com sentimentos para os países analisados, gerados por dicionários fixos, podem ser observados no [Apêndice F](#). Resultados estes semelhantes ao observado no modelo da regra de Taylor tradicional utilizado como base.

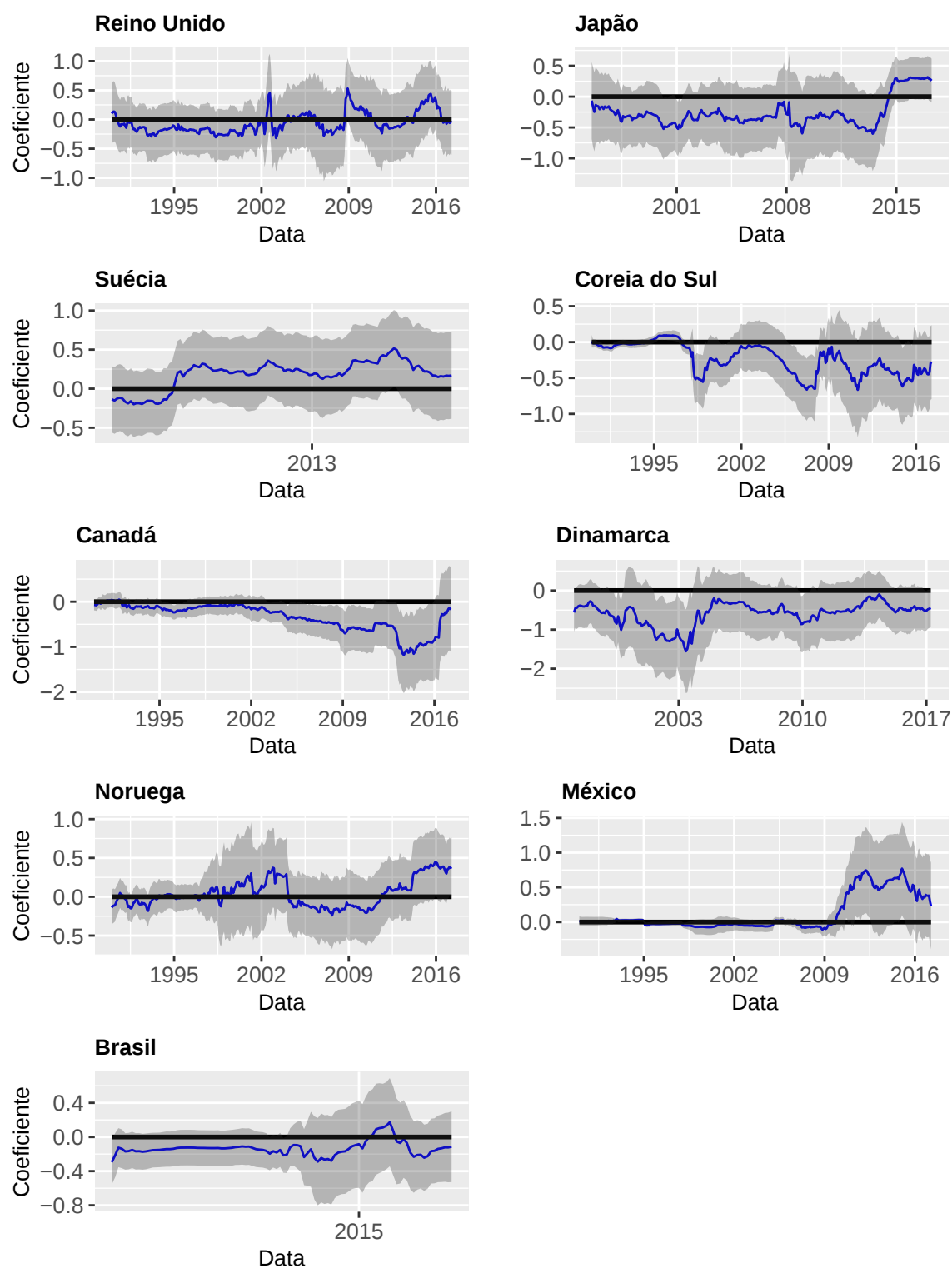
Para os resultados dos coeficientes do país estrangeiro, observados na [Figura 2](#), temos a predominância do coeficiente esperado para 5 dos países analisados, embora que observando seus intervalos de confiança, não pode ser observados alguma predominância para todos os países analisados, dado que os intervalos de confiança em sua maioria incluem o zero.

De forma geral, não é observado fortes evidências do indicativo de consistência do uso da regra de Taylor, tanto para os coeficientes da inflação dos Estados Unidos e dos países estrangeiros, nos modelos de predição da variação da taxa de câmbio considerando seus respectivos intervalos a 95% de confiança. Quando observado os resultados para um período mais recente, no pós crise de 2008, temos padrão predominantemente consistente da inflação dos Estados Unidos para 7 países, respectivamente, das equações da Japão, Coreia do Sul, Canadá, Dinamarca, Noruega, México e Brasil.



Fonte: Elaboração própria.

Figura 1 – Coeficientes da inflação dos Estados Unidos com base no modelo ampliado da regra de Taylor com sentimentos do dicionário variante no tempo.



Fonte: Elaboração própria.

Figura 2 – Coeficientes da inflação do país estrangeiro com base no modelo ampliado da regra de Taylor com sentimentos do dicionário variante no tempo.

4 Conclusões

A maior contribuição desse trabalho é em demonstrar que o sentimento das notícias, construído a partir do método do dicionário variante no tempo e utilizando como fonte textual as notícias financeiras do The New York Times e do The Wall Street Journal, tem alto poder preditivo para a predição da variação da taxa de câmbio, tanto como único preditor ou utilizado de forma conjunta com os preditores da regra de Taylor na equação de predição.

A capacidade de predição desse preditor foi avaliada nos modelos só com sentimento geral desse dicionário, onde foi observado de forma predominante a superioridade estatística de performance na predição em relação a dos modelos base, respectivamente, superioridade ao passeio aleatório sem constante em todos os países e em relação a regra de Taylor convencional para 7 dos 9 países, como observado no capítulo anterior. Sobre a precisão deste modelo, foi observado os melhores RMSPE para todos os países analisados, isso considerando também a precisão de todos os modelos ampliados, indicando que só com esse sentimento geral podemos ter uma precisão melhor em relação aos demais modelos. Sobre os modelos ampliados com sentimento geral do dicionário variante no tempo, foi observado padrão predominantemente consistente dos coeficientes dos Estados Unidos apenas para 7 dos 9 países quando considerado período recente após a crise de 2008, não sendo assegurado sua consistência em relação ao intervalo total das estimações.

Uma contribuição derivada desses resultados é sobre o próprio método do dicionário variante no tempo, além do indicativo de preferência de seu uso para grandes amostras textuais, o método obteve êxito em selecionar bons preditores para a nossa variável dependente, corroborando sua estrutura metodológica. Além do suporte ao uso do método do dicionário variante no tempo, os resultados, tanto da precisão quanto da superioridade da performance, também sugerem a preferência de seu uso em detrimento dos métodos com dicionários fixos, quando o objetivo do trabalho passa pelo poder preditivo de fonte textual.

Referências

- BULUT, L. Google trends and the forecasting performance of exchange rate models. *Journal of Forecasting*, Wiley Online Library, v. 37, n. 3, p. 303–315, 2018. Citado 2 vezes nas páginas 11 e 21.
- CAPORALE, G. M.; SPAGNOLO, F.; SPAGNOLO, N. Macro news and exchange rates in the brics. *Finance Research Letters*, Elsevier, v. 21, p. 140–143, 2017. Citado na página 11.
- CAPORALE, G. M.; SPAGNOLO, F.; SPAGNOLO, N. Exchange rates and macro news in emerging markets. *Research in International Business and Finance*, Elsevier, v. 46, p. 516–527, 2018. Citado na página 11.
- CARRIERO, A.; KAPETANIOS, G.; MARCELLINO, M. Forecasting exchange rates with a large bayesian var. *International Journal of Forecasting*, Elsevier, v. 25, n. 2, p. 400–417, 2009. Citado na página 10.
- 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. Citado na página 14.
- CLARK, T. E.; WEST, K. D. Using out-of-sample mean squared prediction errors to test the martingale difference hypothesis. *Journal of econometrics*, Elsevier, v. 135, n. 1-2, p. 155–186, 2006. Citado na página 20.
- CLARK, T. E.; WEST, K. D. Approximately normal tests for equal predictive accuracy in nested models. *Journal of econometrics*, Elsevier, v. 138, n. 1, p. 291–311, 2007. Citado na página 20.
- COGLEY, T.; PRIMICERI, G. E.; SARGENT, T. J. Inflation-gap persistence in the us. *American Economic Journal: Macroeconomics*, v. 2, n. 1, p. 43–69, 2010. Citado na página 14.
- DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business & economic statistics*, Taylor & Francis, v. 13, n. 3, p. 253–263, 1995. Citado na página 20.
- EHRMANN, M.; FRATZSCHER, M. Exchange rates and fundamentals: new evidence from real-time data. *Journal of International Money and Finance*, Elsevier, v. 24, n. 2, p. 317–341, 2005. Citado na página 11.
- ENGEL, C. et al. The uncovered interest parity puzzle, exchange rate forecasting, and taylor rules. *Journal of International Money and Finance*, Elsevier, v. 95, p. 317–331, 2019. Citado 2 vezes nas páginas 14 e 29.
- ENGELBERG, J. E.; REED, A. V.; RINGGENBERG, M. C. How are shorts informed?: Short sellers, news, and information processing. *Journal of Financial Economics*, Elsevier, v. 105, n. 2, p. 260–278, 2012. Citado 2 vezes nas páginas 13 e 16.
- FERGUSON, N. J. et al. Media content and stock returns: The predictive power of press. *Multinational Finance Journal*, v. 19, n. 1, p. 1–31, 2015. Citado na página 16.
- FERRIS, S. P.; HAO, Q.; LIAO, M.-Y. The effect of issuer conservatism on ipo pricing and performance. *Review of Finance*, Oxford University Press, v. 17, n. 3, p. 993–1027, 2012. Citado na página 16.

- FEUERRIEGEL, S.; WOLFF, G.; NEUMANN, D. News sentiment and overshooting of exchange rates. *Applied Economics*, Taylor & Francis, v. 48, n. 44, p. 4238–4250, 2016. Citado 3 vezes nas páginas 11, 15 e 16.
- GAGLIANONE, W. P.; MARINS, J. T. M. Evaluation of exchange rate point and density forecasts: an application to brazil. *International Journal of Forecasting*, Elsevier, v. 33, n. 3, p. 707–728, 2017. Citado na página 13.
- GALATI, G.; HO, C. Macroeconomic news and the euro/dollar exchange rate. *Economic notes*, Wiley Online Library, v. 32, n. 3, p. 371–398, 2003. Citado na página 10.
- GARCIA, D. Sentiment during recessions. *The Journal of Finance*, Wiley Online Library, v. 68, n. 3, p. 1267–1300, 2013. Citado 3 vezes nas páginas 11, 13 e 16.
- HÁJEK, P. Combining bag-of-words and sentiment features of annual reports to predict abnormal stock returns. *Neural Computing and Applications*, Springer, v. 29, n. 7, p. 343–358, 2018. Citado na página 11.
- HENRY, E. Market reaction to verbal components of earnings press releases: Event study using a predictive algorithm. *Journal of Emerging Technologies in Accounting*, v. 3, n. 1, p. 1–19, 2006. Citado na página 16.
- HENRY, E. Are investors influenced by how earnings press releases are written? *The Journal of Business Communication (1973)*, SAGE Publications Sage CA: Los Angeles, CA, v. 45, n. 4, p. 363–407, 2008. Citado 2 vezes nas páginas 15 e 16.
- HENRY, E.; LEONE, A. J. Measuring qualitative information in capital markets research. *Available at SSRN 1470807*, 2009. Citado na página 16.
- HENRY, E.; LEONE, A. J. Measuring qualitative information in capital markets research: Comparison of alternative methodologies to measure disclosure tone. *The Accounting Review*, American Accounting Association, v. 91, n. 1, p. 153–178, 2015. Citado na página 16.
- HODRICK, R. J.; PRESCOTT, E. C. Postwar us business cycles: an empirical investigation. *Journal of Money, credit, and Banking*, JSTOR, p. 1–16, 1997. Citado na página 13.
- JIANG, F. et al. Manager sentiment and stock returns. *Journal of Financial Economics*, Elsevier, v. 132, n. 1, p. 126–149, 2019. Citado na página 19.
- KEARNEY, C.; LIU, S. Textual sentiment in finance: A survey of methods and models. *International Review of Financial Analysis*, Elsevier, v. 33, p. 171–185, 2014. Citado na página 16.
- LIMA, L. R.; GODEIRO, L.; MOHSIN, M. Time-varying dictionary and the predictive power of fed minutes. *Available at SSRN 3312483*, 2019. Citado 4 vezes nas páginas 11, 15, 18 e 22.
- LIU, B.; MCCONNELL, J. J. The role of the media in corporate governance: Do the media influence managers' capital allocation decisions? *Journal of Financial Economics*, Elsevier, v. 110, n. 1, p. 1–17, 2013. Citado na página 13.
- LOUGHRAN, T.; MCDONALD, B. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, Wiley Online Library, v. 66, n. 1, p. 35–65, 2011. Citado 2 vezes nas páginas 15 e 16.

LOUGHRAN, T.; MCDONALD, B. Ipo first-day returns, offer price revisions, volatility, and form s-1 language. *Journal of Financial Economics*, Elsevier, v. 109, n. 2, p. 307–326, 2013. Citado na página 16.

MEESE, R.; ROGOFF, K. The out-of-sample failure of empirical exchange rate models: sampling error or misspecification? In: *Exchange rates and international macroeconomics*. [S.l.]: University of Chicago Press, 1983. p. 67–112. Citado na página 10.

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. Citado na página 10.

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. Citado 7 vezes nas páginas 10, 13, 14, 15, 19, 29 e 39.

MORALES-ARIAS, L.; MOURA, G. V. Adaptive forecasting of exchange rates with panel data. *International Journal of Forecasting*, Elsevier, v. 29, n. 3, p. 493–509, 2013. Citado 2 vezes nas páginas 10 e 19.

NIKOLSKO-RZHEVSKYY, A.; PRODAN, R. Markov switching and exchange rate predictability. *International Journal of Forecasting*, Elsevier, v. 28, n. 2, p. 353–365, 2012. Citado na página 10.

PRICE, S. M. et al. Earnings conference calls and stock returns: The incremental informativeness of textual tone. *Journal of Banking & Finance*, Elsevier, v. 36, n. 4, p. 992–1011, 2012. Citado na página 16.

SHAPIRO, A. H.; SUDHOF, M.; WILSON, D. Measuring news sentiment. In: FEDERAL RESERVE BANK OF SAN FRANCISCO. [S.l.], 2018. Citado 2 vezes nas páginas 11 e 13.

STONE, P. J.; DUNPHY, D. C.; SMITH, M. S. The general inquirer: A computer approach to content analysis. MIT press, 1966. Citado na página 15.

TAYLOR, J. B. Discretion versus policy rules in practice. In: ELSEVIER. *Carnegie-Rochester conference series on public policy*. [S.l.], 1993. v. 39, p. 195–214. Citado na página 13.

TETLOCK, P. C. Giving content to investor sentiment: The role of media in the stock market. *The Journal of finance*, Wiley Online Library, v. 62, n. 3, p. 1139–1168, 2007. Citado na página 16.

TETLOCK, P. C.; SAAR-TSECHANSKY, M.; MACSKASSY, S. More than words: Quantifying language to measure firms' fundamentals. *The Journal of Finance*, Wiley Online Library, v. 63, n. 3, p. 1437–1467, 2008. Citado na página 16.

WANG, R.; MORLEY, B.; ORDÓÑEZ, J. The taylor rule, wealth effects and the exchange rate. *Review of International Economics*, Wiley Online Library, v. 24, n. 2, p. 282–301, 2016. Citado 2 vezes nas páginas 10 e 19.

ZOU, H.; HASTIE, T. Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)*, Wiley Online Library, v. 67, n. 2, p. 301–320, 2005. Citado na página 18.

Apêndices

APÊNDICE A – Fonte e intervalo dos dados

Tabela 7 – Fontes dos dados

	Fonte	Código
Taxa de câmbio nominal:		
- Todos os países	IFS	ENDA_XDC_USD_RATE
Taxa de juros de curto prazo:		
- Estados Unidos	IFS	FIMM_PA
- Reino Unido	IFS	FIMM_PA
- Japão	IFS	FIMM_PA
- Suécia	IFS	FIMM_PA
- Coreia do Sul	IFS	FIMM_PA
- Canadá	OCDE	
- Dinamarca	OCDE	
- Noruega	OCDE	
- Brasil	IFS	FIMM_PA
- México	IFS	FIMM_PA
Produção industrial com ajuste sazonal:		
- Estados Unidos	IFS	AIP_SA_IX
- Reino Unido	IFS	AIP_SA_IX
- Japão	IFS	AIP_SA_IX
- Suécia	IFS	AIP_SA_IX
- Coreia do Sul	IFS	AIP_SA_IX
- Canadá	OCDE	
- Dinamarca	OCDE	
- Noruega	OCDE	
- Brasil	IFS	AIP_SA_IX
- México	IFS	AIP_SA_IX
Índice de preços ao consumidor:		
- Todos os países	IFS	PCPI_IX
Taxa real de câmbio:		
- Todos os países, exceto a Coreia do Sul (*)	IFS	EREER_IX
Notícias financeiras (<i>The New York Times</i> e <i>The wall Street Journal</i>)	Factiva	

Fonte: Elaboração própria.

Nota: (*) Devido a ausência de dados.

Tabela 8 – Tamanho das amostras por país estrangeiro

	Observações		Total
	Início	Fim	
Reino Unido	Janeiro/1980	Maio/2017	449
Japão	Agosto/1985	Maio/2017	382
Suécia	Dezembro/1996	Maio/2017	246
Coreia do Sul	Janeiro/1980	Maio/2017	449
Canadá	Janeiro/1980	Maio/2017	449
Dinamarca	Fevereiro/1987	Maio/2017	364
Noruega	Janeiro/1980	Maio/2017	449
México	Janeiro/1980	Maio/2017	449
Brasil	Janeiro/1999	Maio/2017	221

Fonte: Elaboração própria.

Nota: Intervalos com base nos dados disponíveis para as especificações da regra de Taylor utilizadas por [Molodtsova e Papell \(2009\)](#).

APÊNDICE B – Regra de Taylor

Tabela 9 – Especificações da regra de Taylor estimadas

Especificações heterogêneas:

simétrico e sem suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi us} \pi_t + \omega_{\pi f} \tilde{\pi}_t - \omega_{y us} y_t + \omega_{y f} \tilde{y}_t + \eta_t$$

simétrico e com suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi us} \pi_t + \omega_{\pi f} \tilde{\pi}_t - \omega_{y us} y_t + \omega_{y f} \tilde{y}_t - \omega_i i_{t-1} + \omega_{if} \tilde{i}_{t-1} + \eta_t$$

assimétrico e sem suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi us} \pi_t + \omega_{\pi f} \tilde{\pi}_t - \omega_{y us} y_t + \omega_{y f} \tilde{y}_t + q_t + \eta_t$$

assimétrico e com suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi us} \pi_t + \omega_{\pi f} \tilde{\pi}_t - \omega_{y us} y_t + \omega_{y f} \tilde{y}_t + q_t - \omega_i i_{t-1} + \omega_{if} \tilde{i}_{t-1} + \eta_t$$

Especificações homogêneas:

simétrico e sem suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi} (\pi_t - \tilde{\pi}_t) - \omega_y (y_t - \tilde{y}_t) + \eta_t$$

simétrico e com suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi} (\pi_t - \tilde{\pi}_t) - \omega_y (y_t - \tilde{y}_t) - \omega_i (i_{t-1} - \tilde{i}_{t-1}) + \eta_t$$

assimétrico e sem suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi} (\pi_t - \tilde{\pi}_t) - \omega_y (y_t - \tilde{y}_t) + q_t + \eta_t$$

assimétrico e com suavização

$$\Delta S_{t+1} = \omega - \omega_{\pi} (\pi_t - \tilde{\pi}_t) - \omega_y (y_t - \tilde{y}_t) + q_t - \omega_i (i_{t-1} - \tilde{i}_{t-1}) + \eta_t$$

Fonte: Elaboração própria.

Nota: Sendo considerado a presença da constante em todas às especificações da tabela.

Tabela 10 – Resultados da regra de Taylor convencional - Parte 1

	Modelos heterogêneos:			
	Simétrico		Assimétrico	
	s/ suavização	c/ suavização	s/ suavização	c/ suavização
Reino Unido	0.077*	0.307	0.041**	0.027**
Japão	0.738	0.431	0.280	0.001***
Suécia	0.134	0.110	0.005***	0.000***
Coreia do Sul	0.555	0.419	-	-
Canadá	0.003**	0.442	0.000***	0.091*
Dinamarca	0.896	0.828	0.194	0.087*
Noruega	0.054*	0.294	0.000***	0.000***
México	0.099*	0.115	0.002***	0.001***
Brasil	0.449	0.403	0.104	0.024**

Fonte: Elaboração própria.

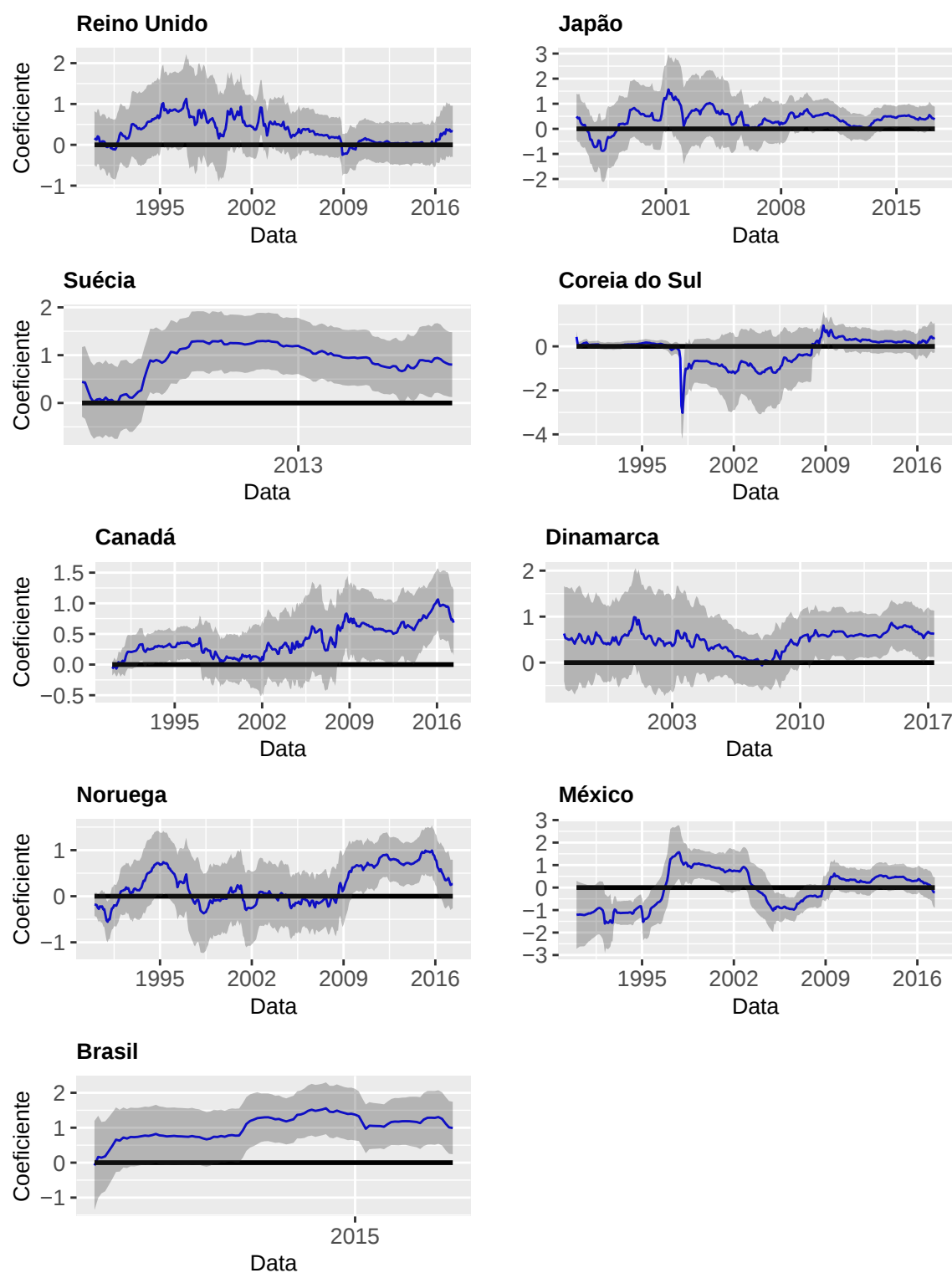
Nota: A tabela reporta os p-valores do teste CW da comparação da performance um passo à frente das previsões em relação ao passeio aleatório sem constante, acompanhados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para os países analisados na tabela, Japão, Suécia, Dinamarca e Brasil possuem, respectivamente, 382, 246, 364 e 221 observações, os demais países possuem 449 observações.

Tabela 11 – Resultados da regra de Taylor convencional - Parte 2

	Modelos homogêneos:			
	Simétrico		Assimétrico	
	s/ suavização	c/ suavização	s/ suavização	c/ suavização
Reino Unido	0.331	0.617	0.040**	0.106
Japão	0.547	0.266	0.160	0.035**
Suécia	0.509	0.783	0.076*	0.127
Coreia do Sul	0.724	0.701	-	-
Canadá	0.028**	0.132	0.000***	0.003***
Dinamarca	0.882	0.773	0.073*	0.025**
Noruega	0.573	0.848	0.003***	0.010***
México	0.076*	0.164	0.001***	0.000***
Brasil	0.317	0.356	0.034**	0.073*

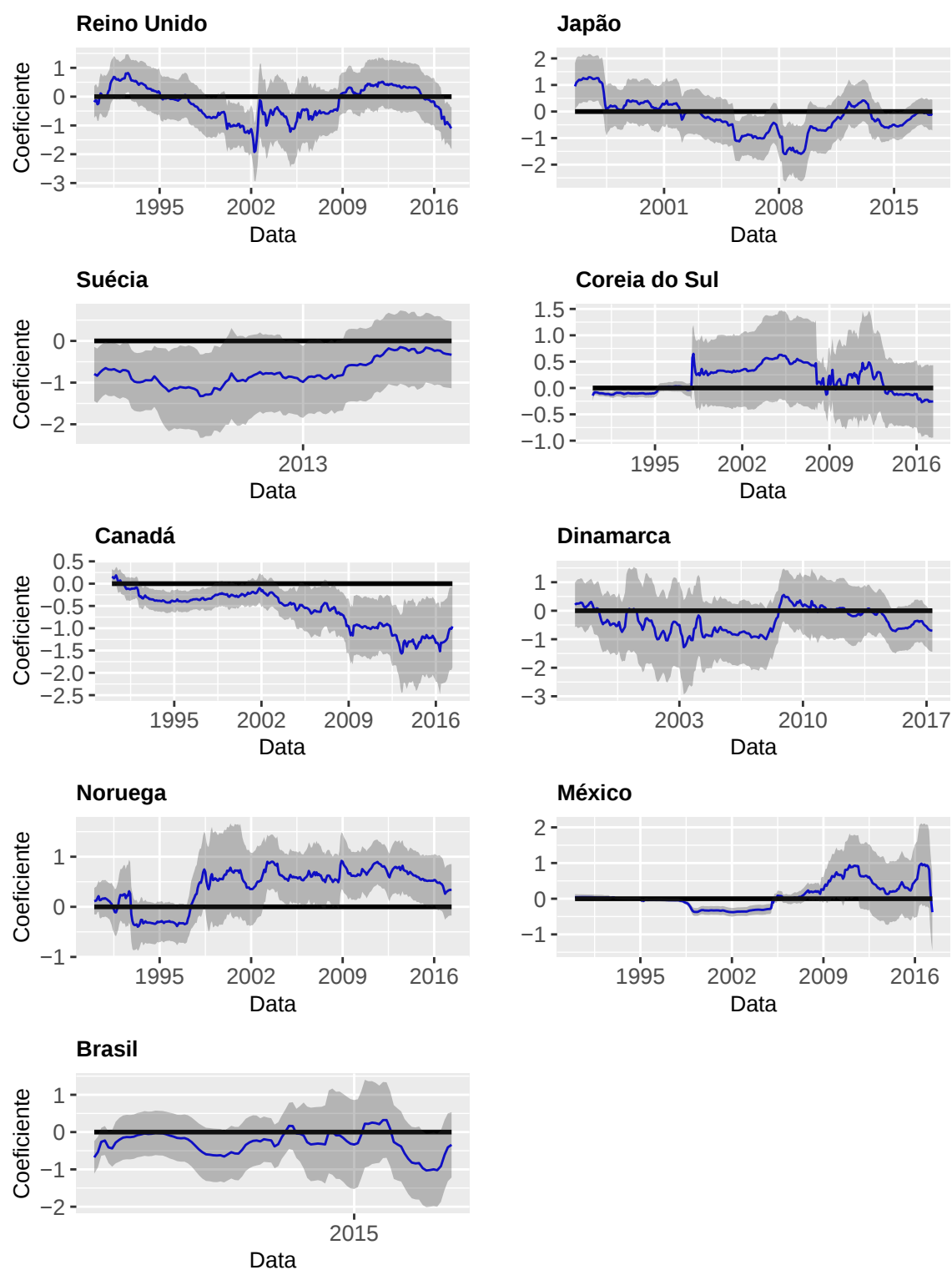
Fonte: Elaboração própria.

Nota: A tabela reporta os p-valores do teste CW da comparação da performance um passo à frente das previsões em relação ao passeio aleatório sem constante, acompanhados por *, ** e ***, que sinalizam sua significância a 10%, 5% e 1%, respectivamente. Para os países analisados na tabela, Japão, Suécia, Dinamarca e Brasil possuem, respectivamente, 382, 246, 364 e 221 observações, os demais países possuem 449 observações.



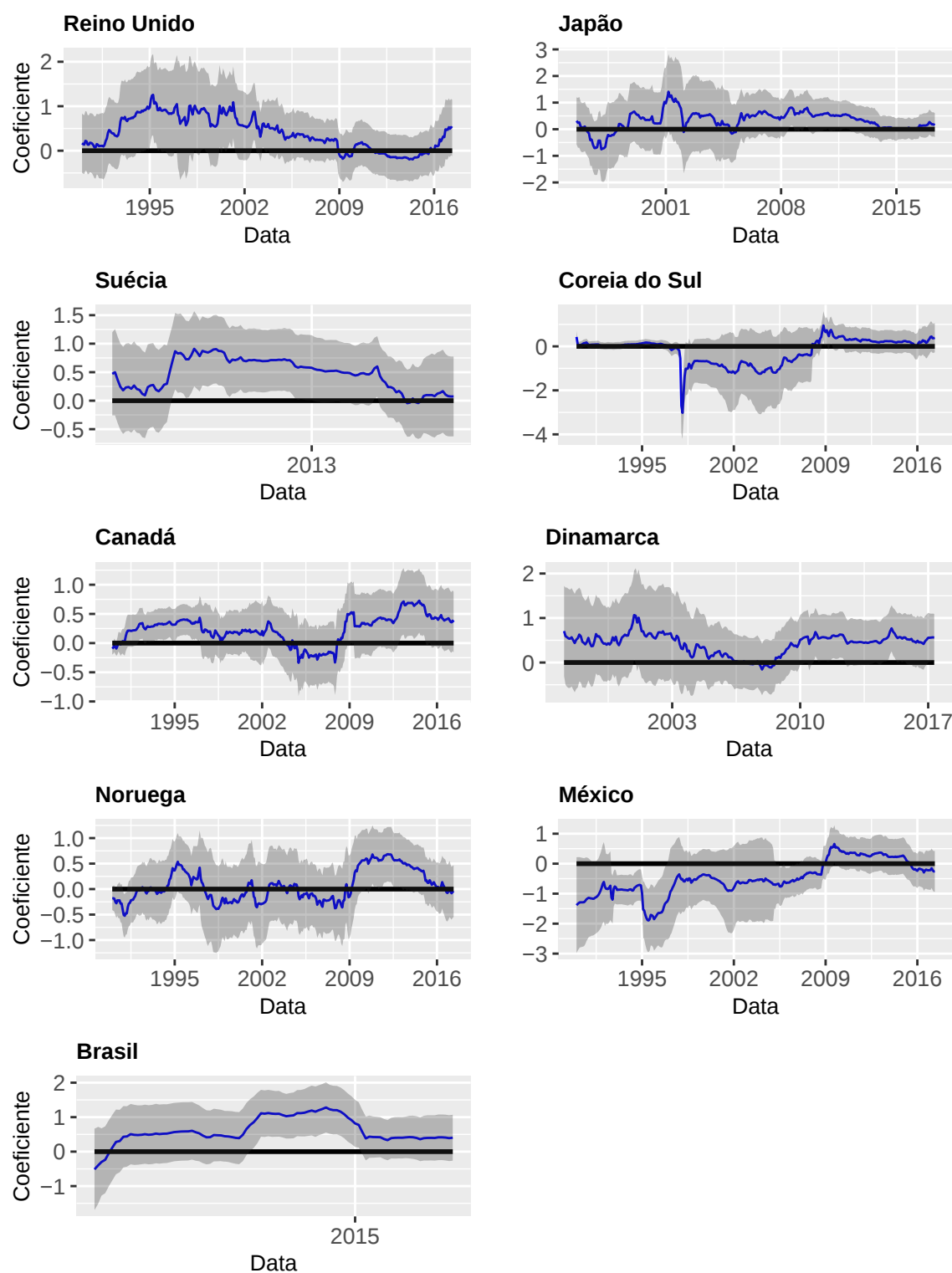
Fonte: Elaboração própria.

Figura 3 – Coeficientes da inflação dos Estados Unidos com base na especificação da regra de Taylor selecionada.



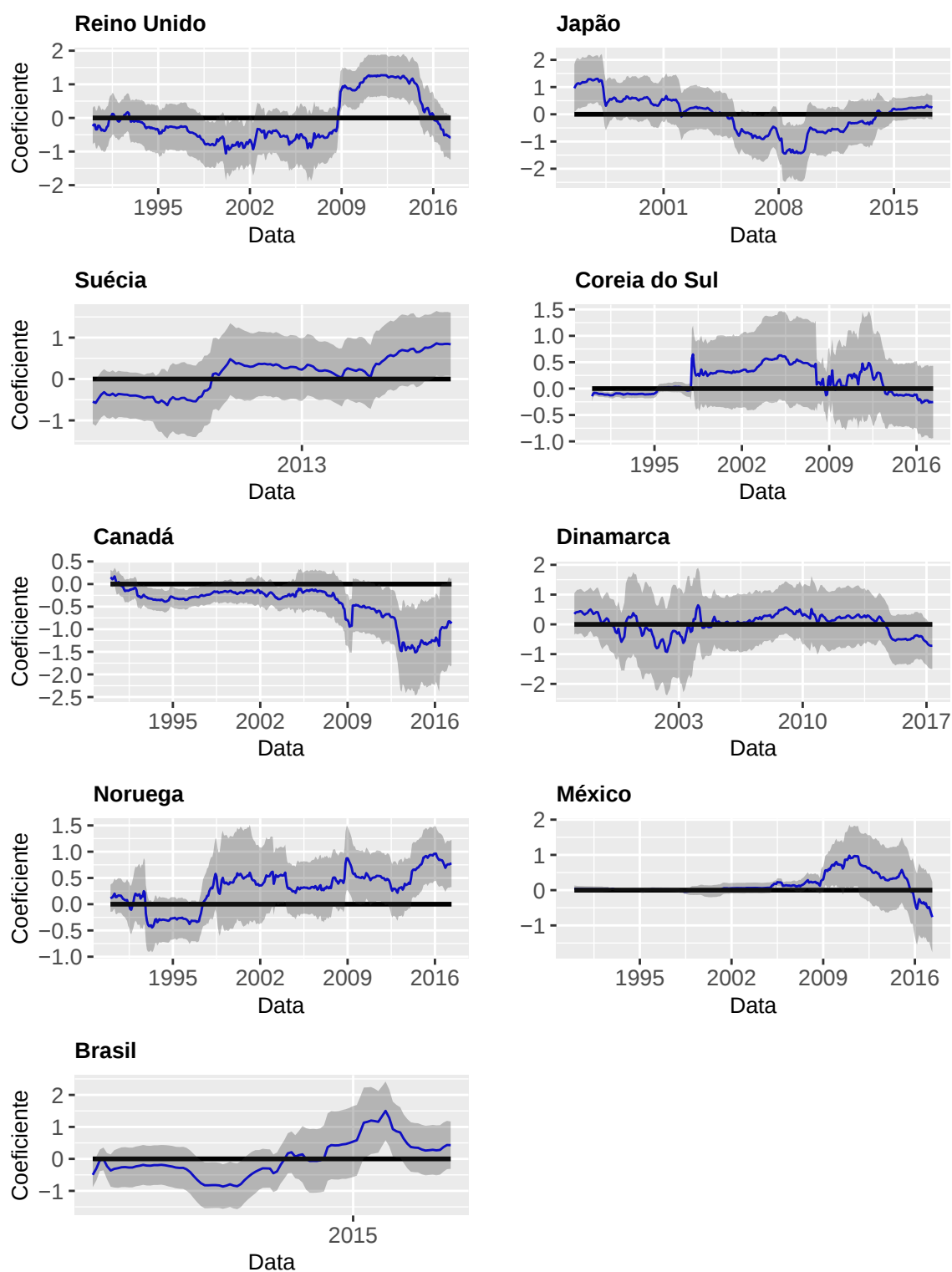
Fonte: Elaboração própria.

Figura 4 – Coeficientes da inflação do país estrangeiro com base na especificação da regra de Taylor selecionada.



Fonte: Elaboração própria.

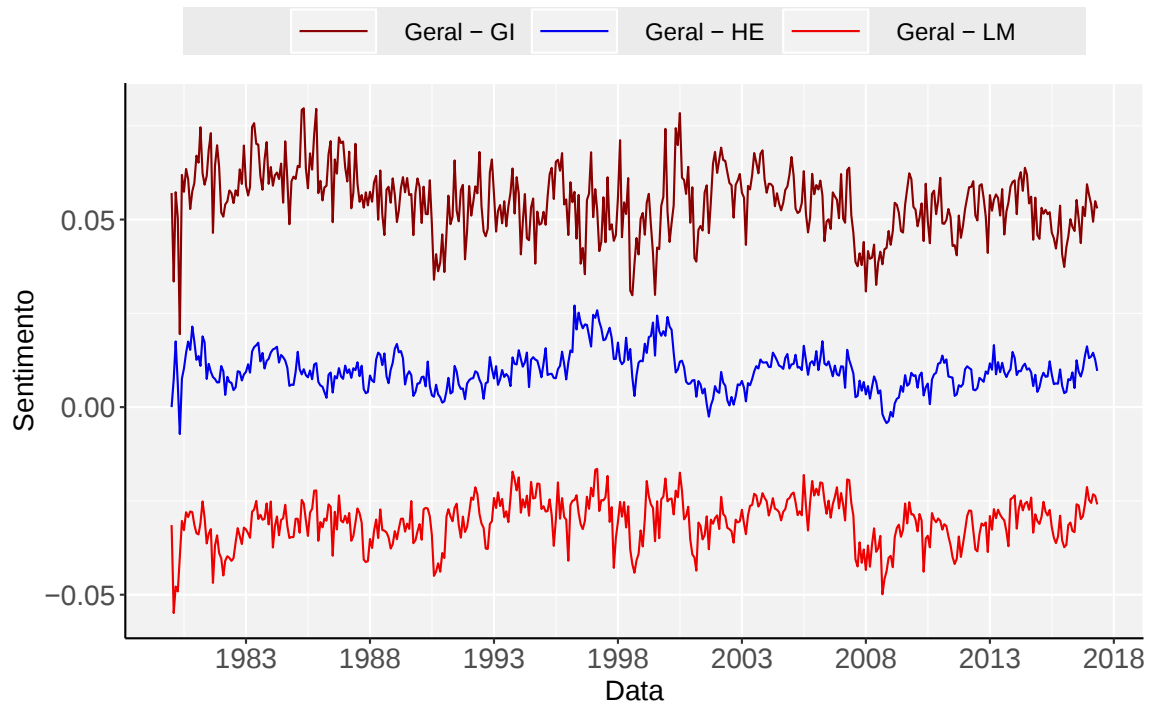
Figura 5 – Coeficientes da inflação dos Estados Unidos com base na especificação da regra de Taylor indicada pela literatura.



Fonte: Elaboração própria.

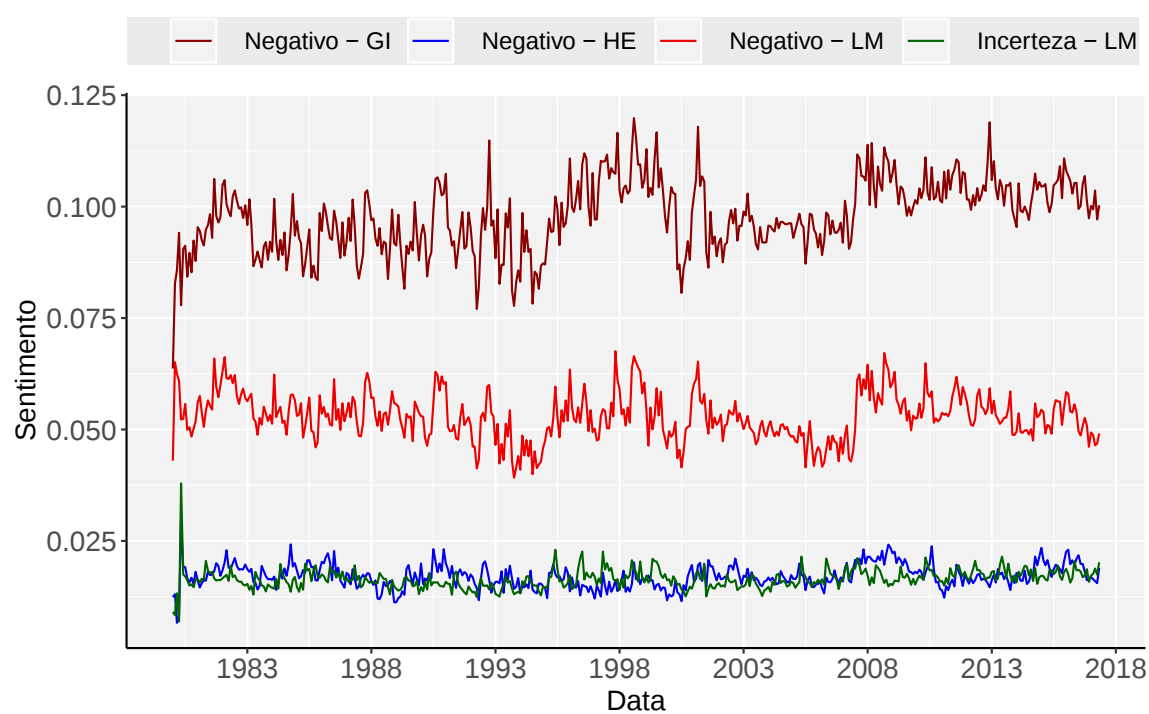
Figura 6 – Coeficientes da inflação do país estrangeiro com base na especificação da regra de Taylor indicada pela literatura.

APÊNDICE C – Sentimento das notícias



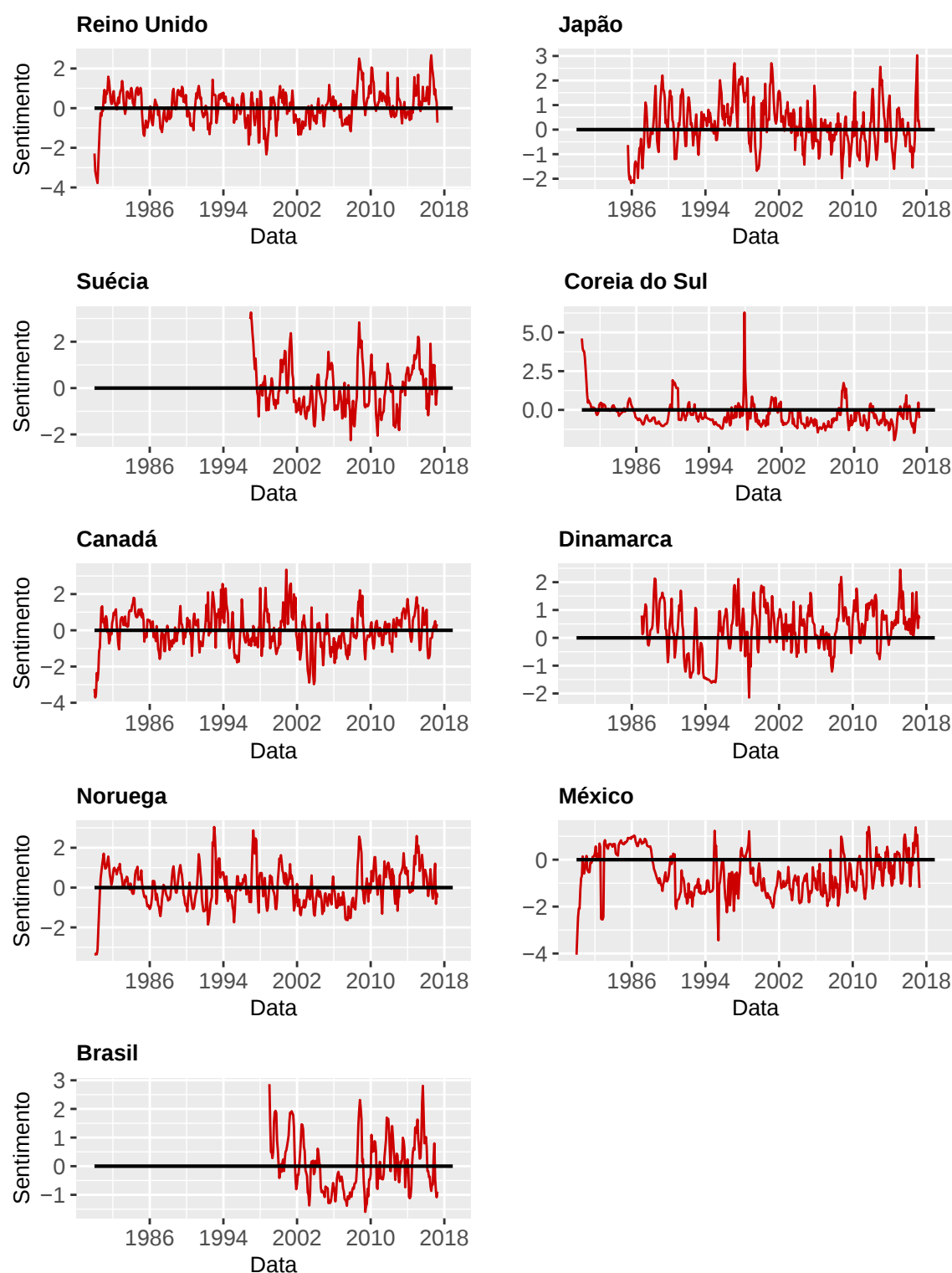
Fonte: Elaboração própria.

Figura 7 – Dicionários fixos: sentimentos gerais.



Fonte: Elaboração própria.

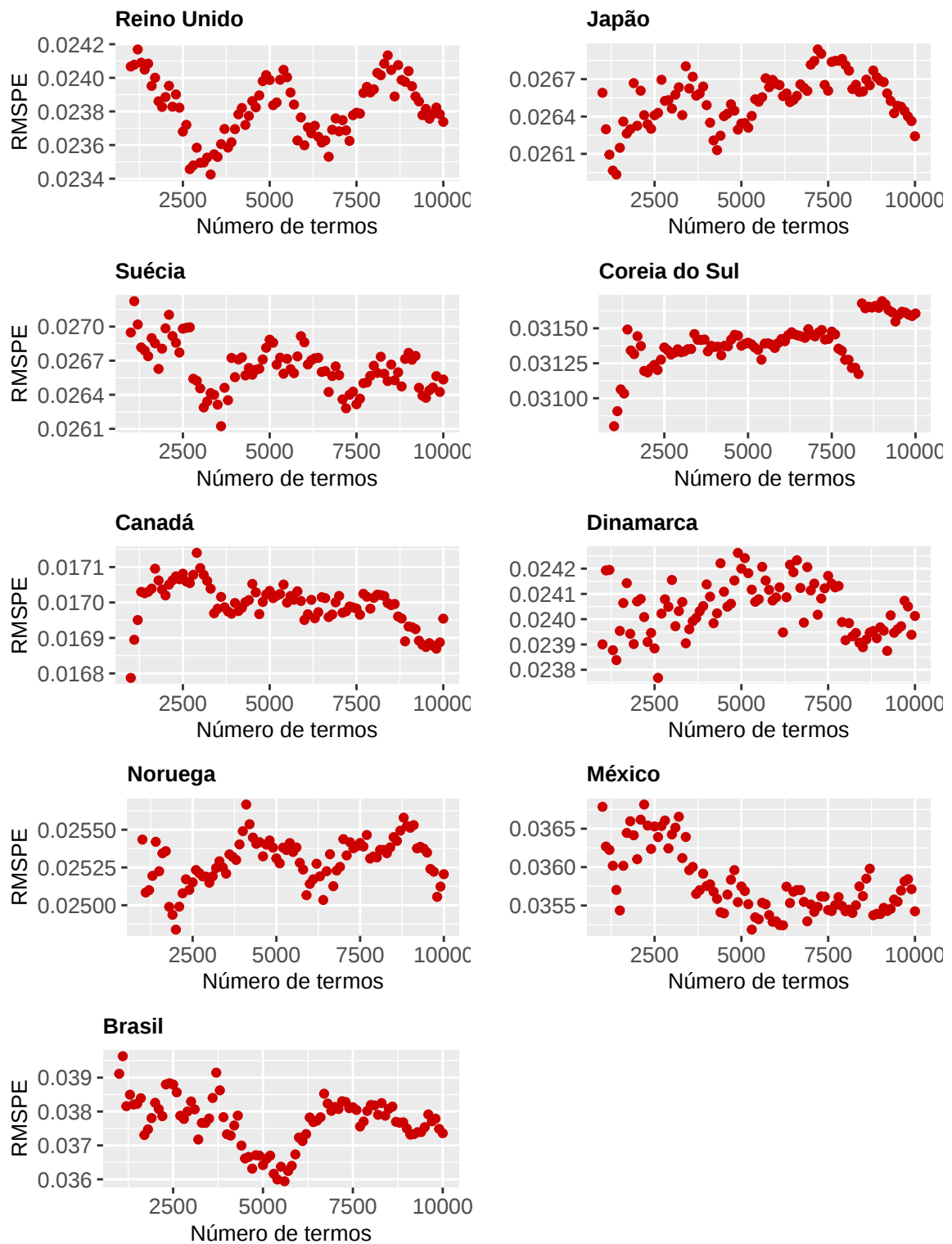
Figura 8 – Dicionários fixos: sentimentos negativos e de incerteza.



Fonte: Elaboração própria.

Figura 9 – Dicionário variante no tempo: sentimento geral por país.

APÊNDICE D – Número de termos



Fonte: Elaboração própria.

Figura 10 – Evolução do RMSPE por país para o intervalo de termos testados (1000 a 10000).

APÊNDICE E – Amostra de termos

Tabela 12 – Amostra de termos positivos e negativos por data de referência - Parte 1

	data referência:			
	2000-01-01		2015-01-01	
	Positivo	Negativo	Positivo	Negativo
Reino Unido	voter retail builder rent ginni	fell inflow intervent overvalu yeltsin	ecommerc soybean stockpric bundesbank coordin	iranian inflow intervent chevron carbon
Japão	social.security wto retire oecd capitalgain	oil.price bloomberg gold rumor broker	economic.growth brent eurozon highyield davo	venezuela fee etf fda auster
Coreia do sul	internet dealer gold euro onlin	bloomberg refinanc deriv buyout venezuela	eurozon asia est bailout money.supply	healthcar emiss transpar kraft iran
Canadá	social.security bloomberg rubin quarter asia	report federal.reserve intervent forb penni	economi valuat gdp ecb bubbl	ethanol diesel shutdown greec etf

Fonte: Elaboração própria.

Nota: A tabela reporta uma amostra dos termos que tiveram efeito e destaque na data de referência, com base no dicionário variante no tempo. Brasil e Suécia não foram expostos por causa da limitação do período dos dados, não possibilitando uma comparação entre dicionários com um intervalo de pelo menos 10 anos.

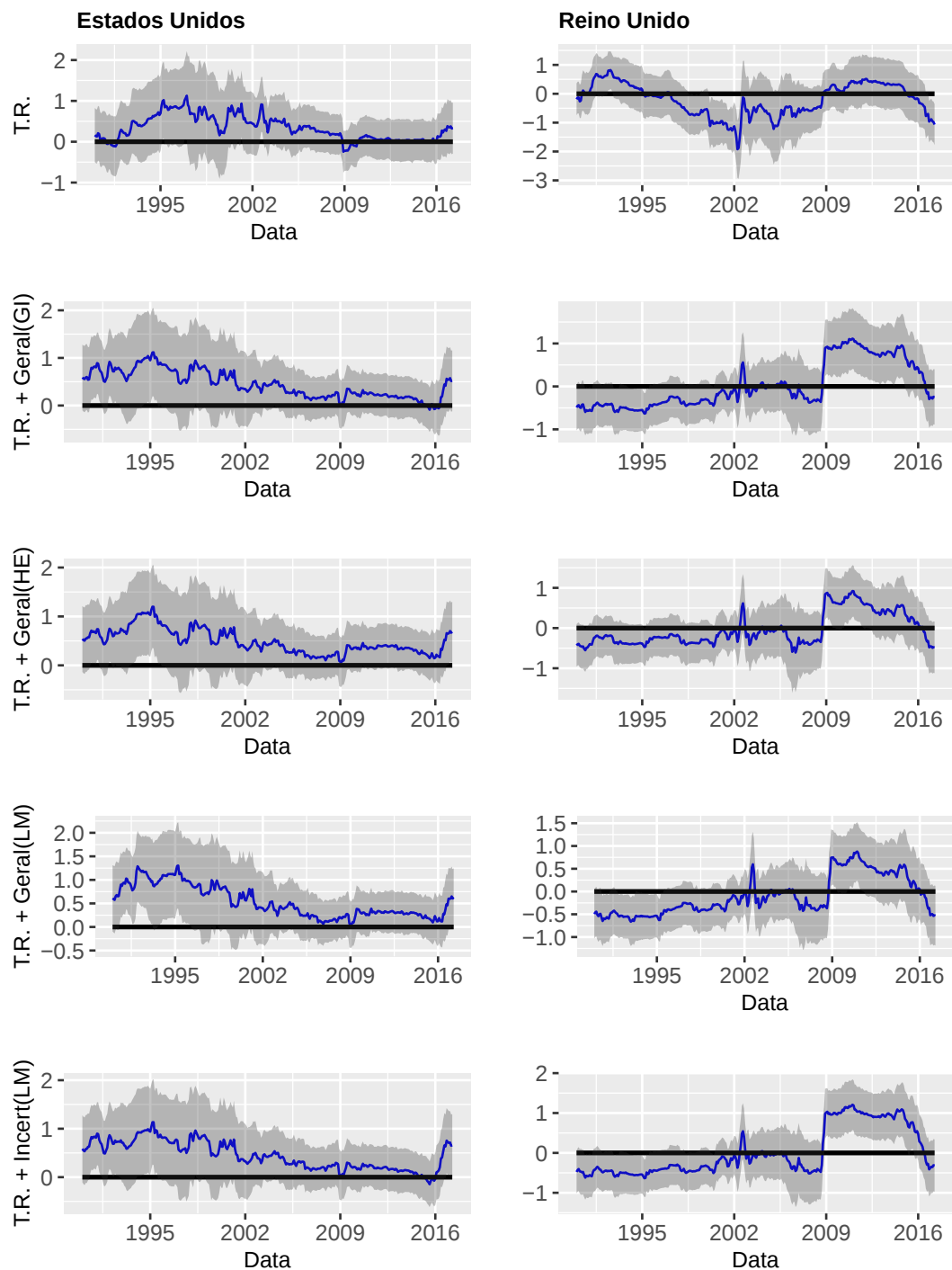
Tabela 13 – Amostra de termos positivos e negativos por data de referência - Parte 2

	data referência:			
	2000-01-01		2015-01-01	
	Positivo	Negativo	Positivo	Negativo
Dinamarca	social.security capitalgain agenda gov fanni	brokerag swiss afraid tourism outflow	eurozon gpd legislatur ecommerc swap	carbon imf dept securit trump
Noruega	voter gov gdp stockhold telecom	fell brokerag ministri intervent tuition	oil.price eurozon economic.growth dividend gdp	ethanol carbon dept intervent tuition
México	supplysid accommod keynesian cutter zedillo	bush payrol marietta	accord ted chrysler tourism singer	underfund adapt abbott greenspan contest

Fonte: Elaboração própria.

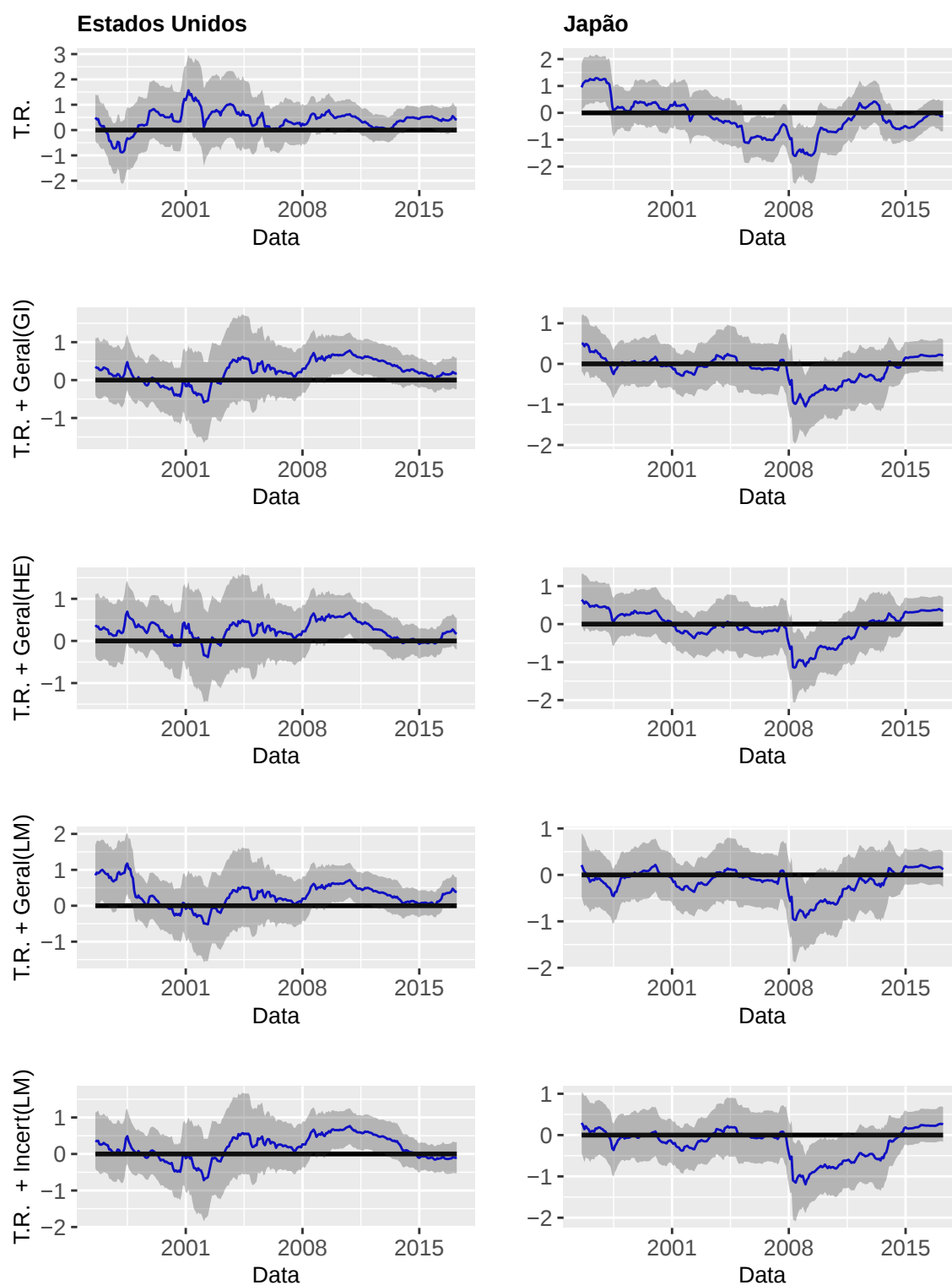
Nota: A tabela reporta uma amostra dos termos que tiveram efeito e destaque na data de referência, com base no dicionário variante no tempo. Brasil e Suécia não foram expostos por causa da limitação do período dos dados, não possibilitando uma comparação entre dicionários com um intervalo de pelo menos 10 anos.

APÊNDICE F – Coeficientes dos modelos ampliados



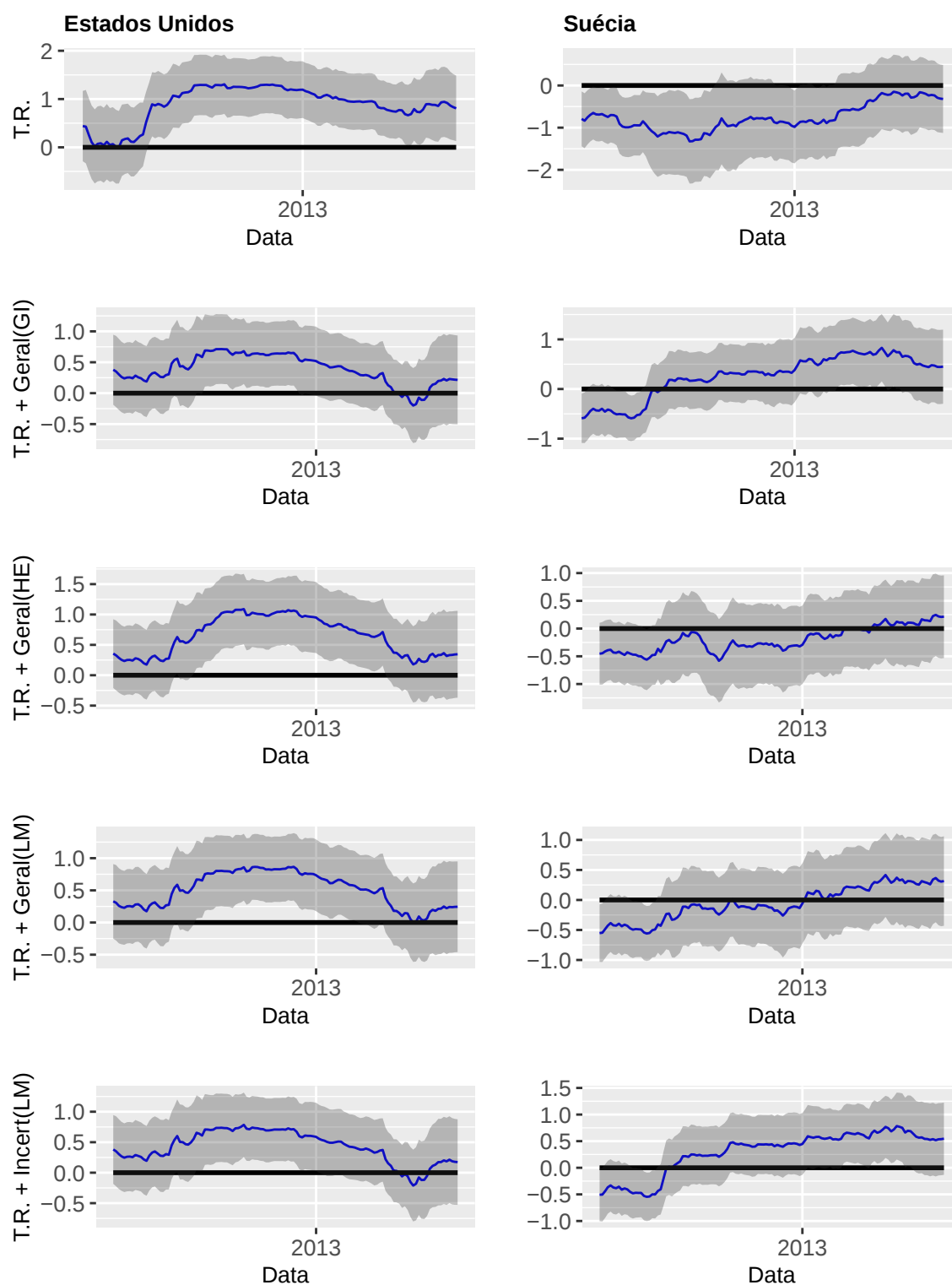
Fonte: Elaboração própria.

Figura 11 – Coeficientes da inflação dos Estados Unidos e Reino Unido por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



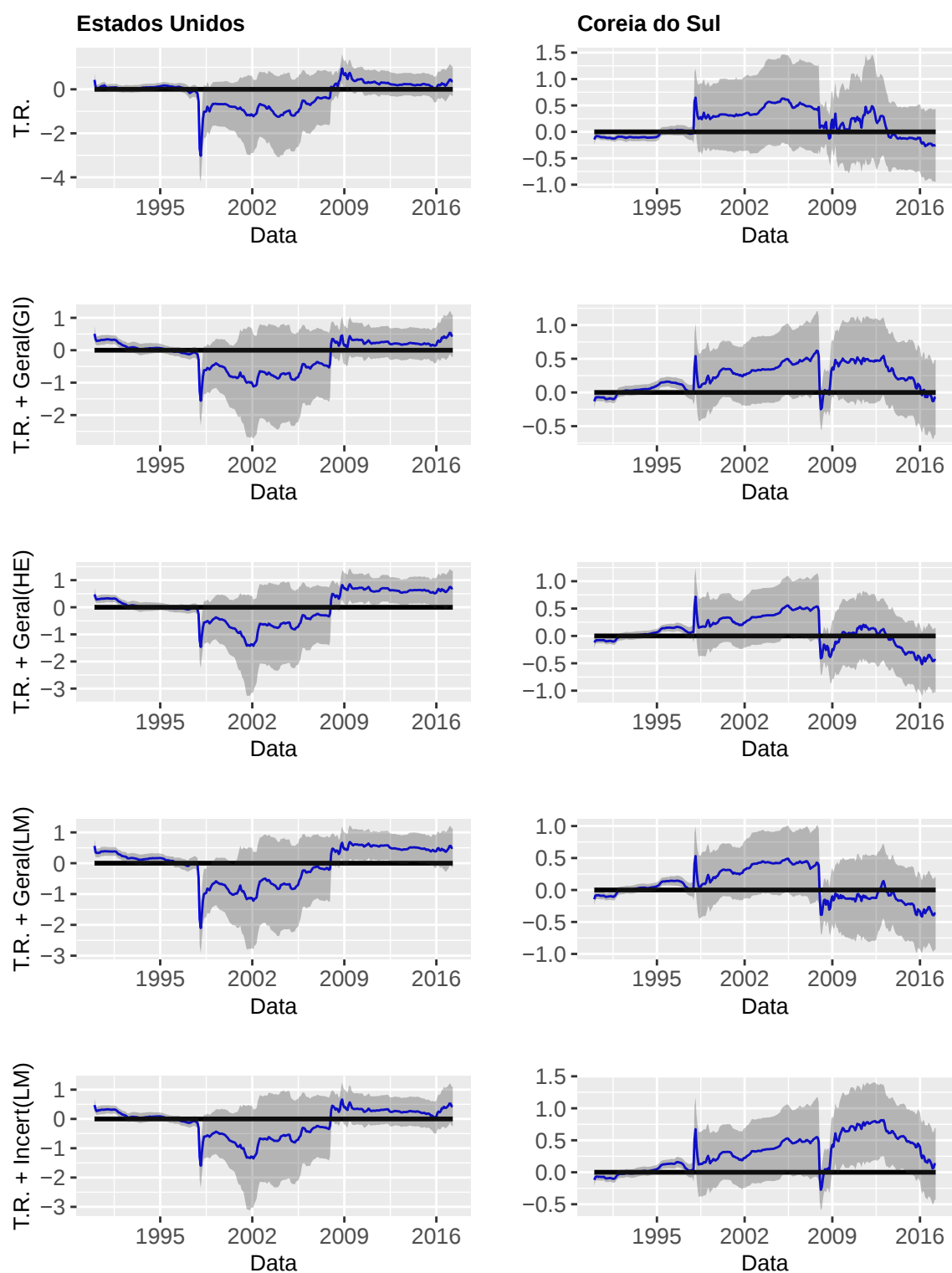
Fonte: Elaboração própria.

Figura 12 – Coeficientes da inflação dos Estados Unidos e Japão por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



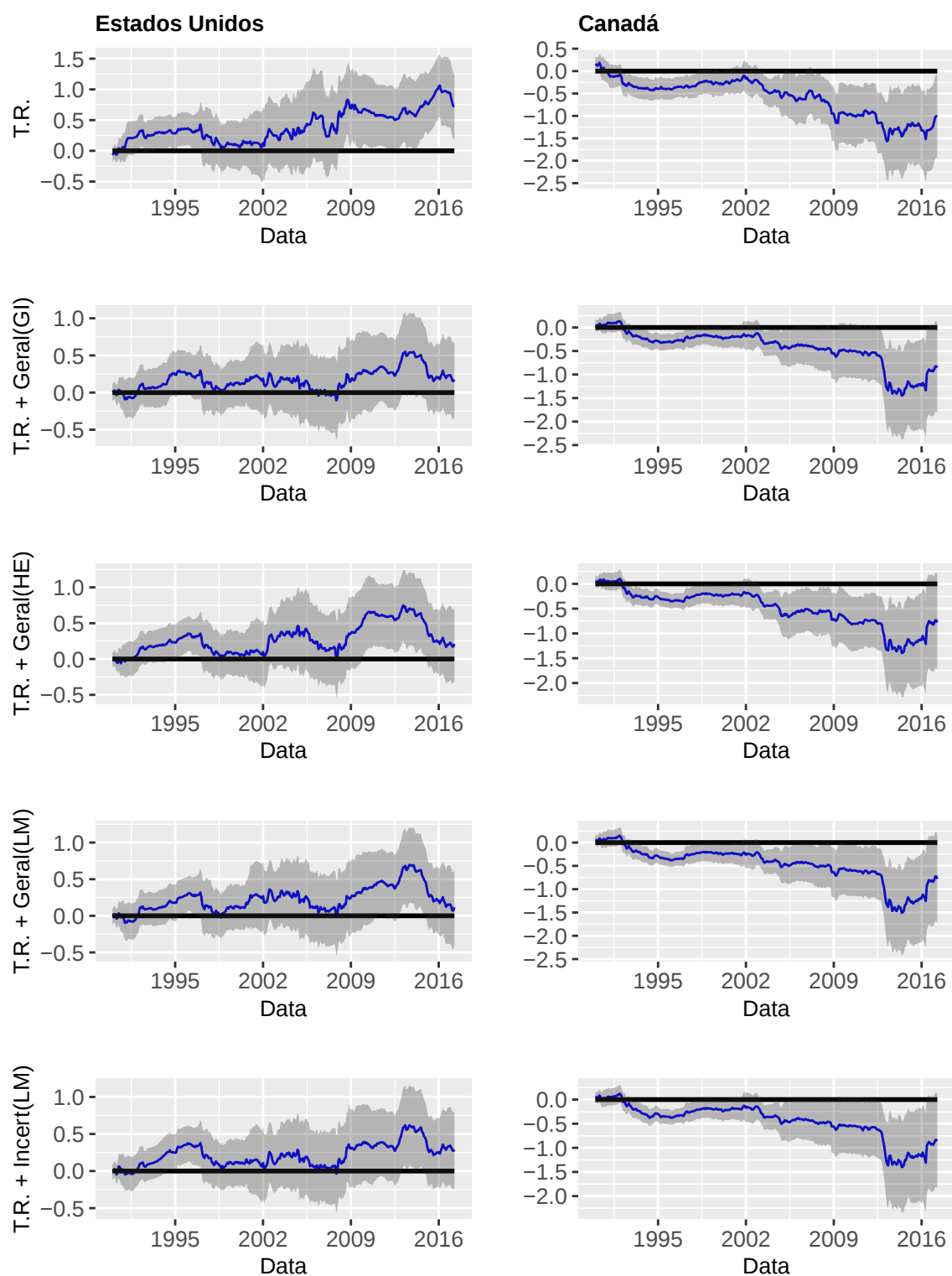
Fonte: Elaboração própria.

Figura 13 – Coeficientes da inflação dos Estados Unidos e Suécia por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



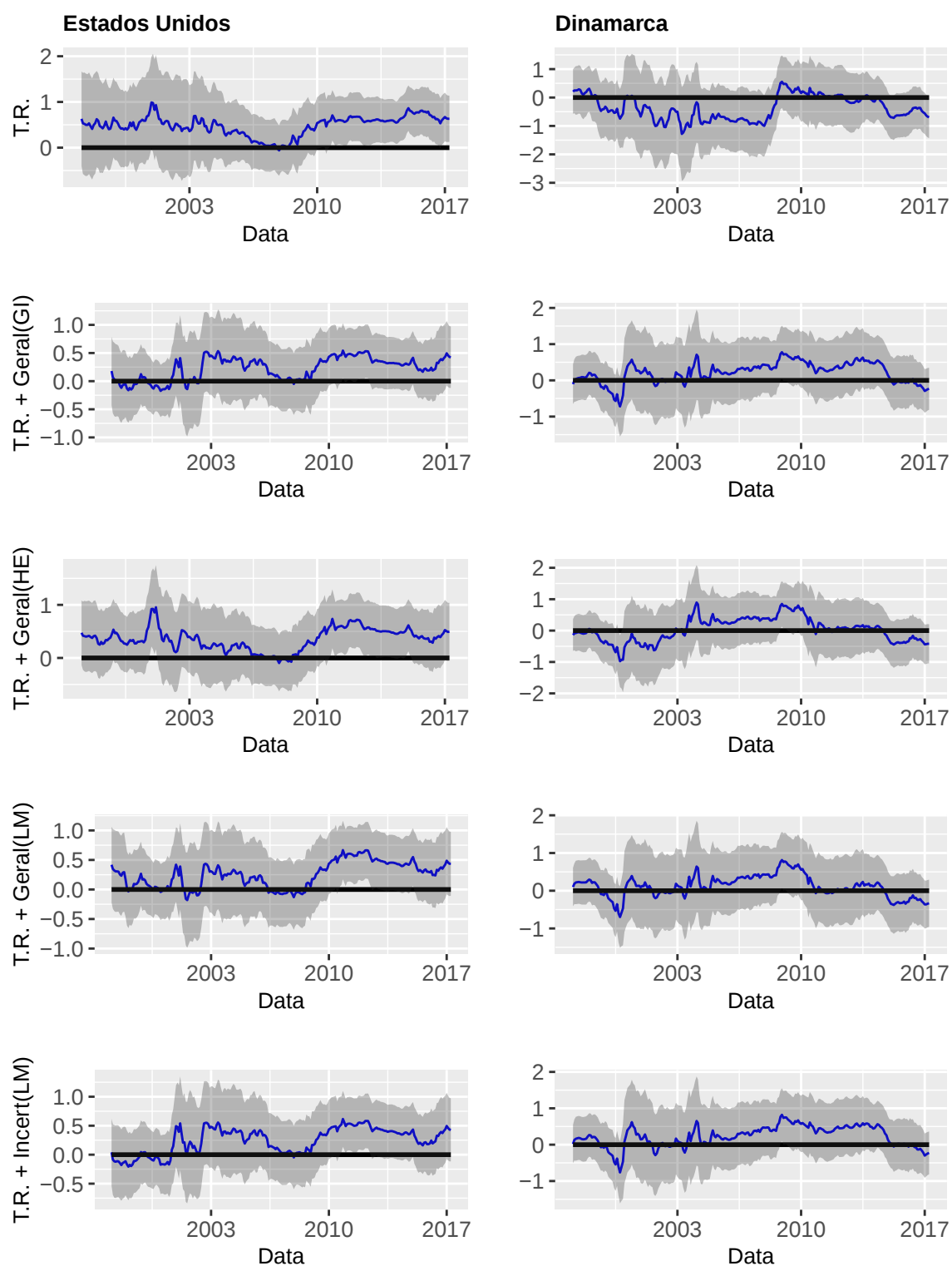
Fonte: Elaboração própria.

Figura 14 – Coeficientes da inflação dos Estados Unidos e Coreia do Sul por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



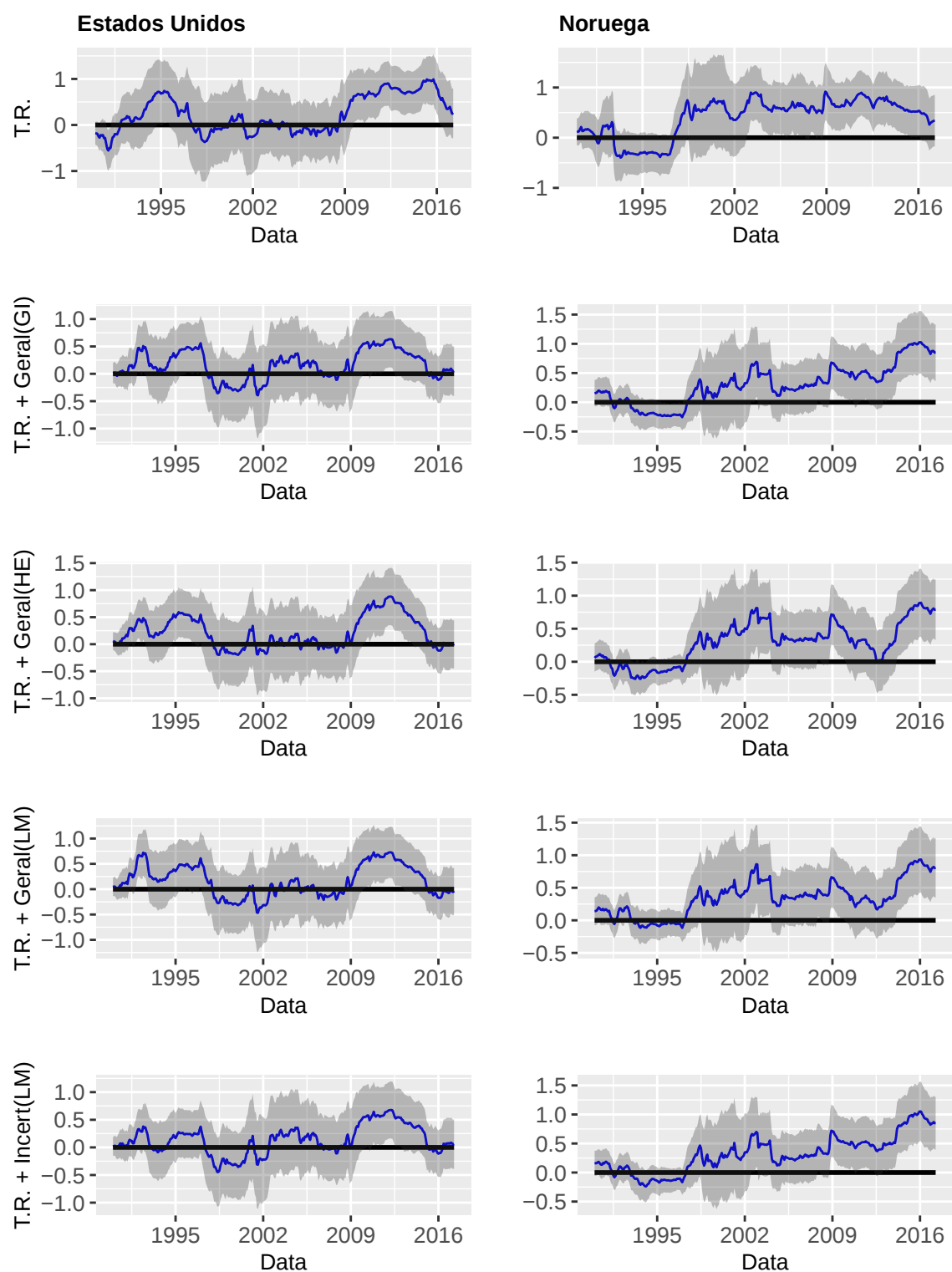
Fonte: Elaboração própria.

Figura 15 – Coeficientes da inflação dos Estados Unidos e Canadá por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



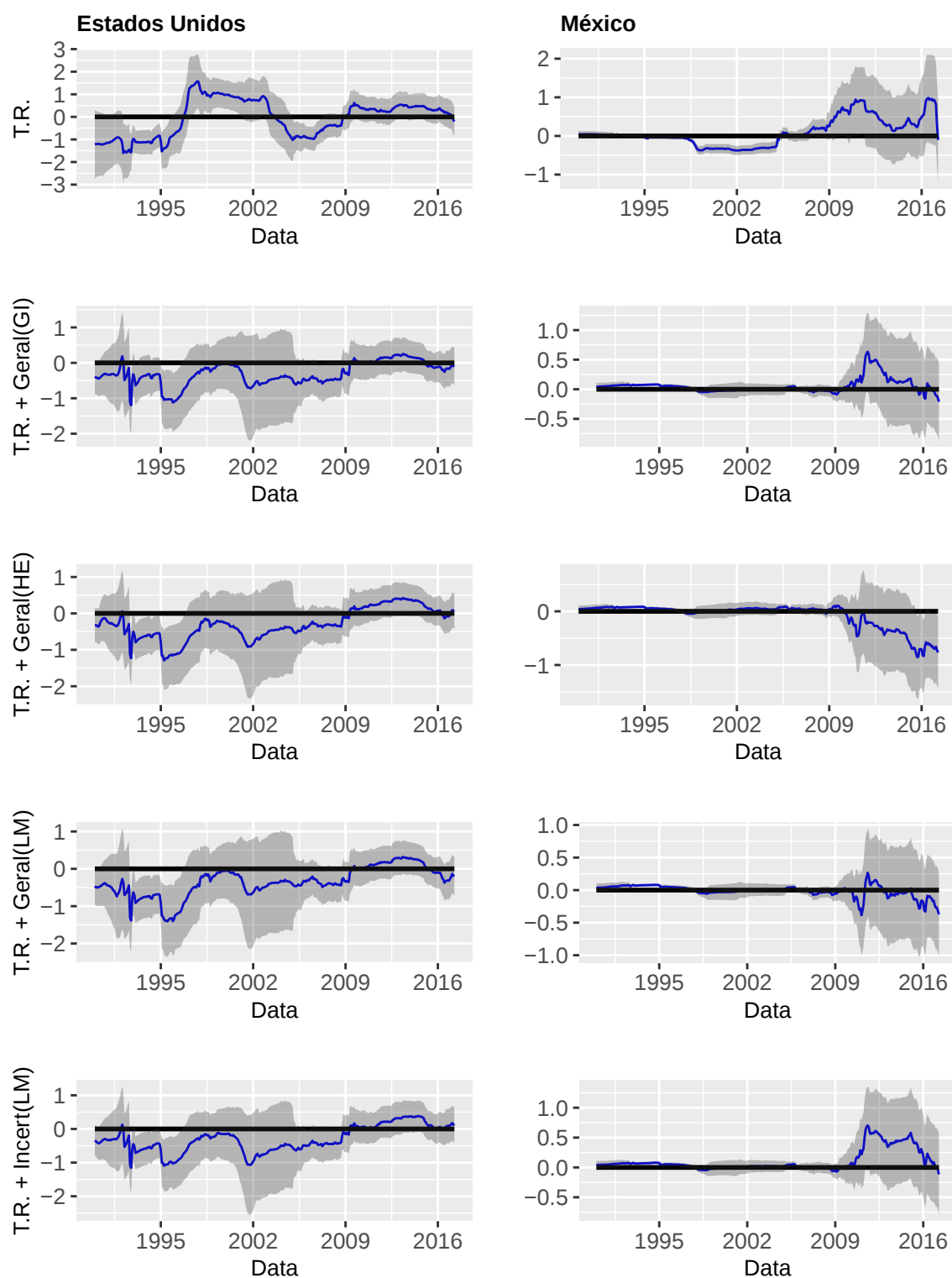
Fonte: Elaboração própria.

Figura 16 – Coeficientes da inflação dos Estados Unidos e Dinamarca por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



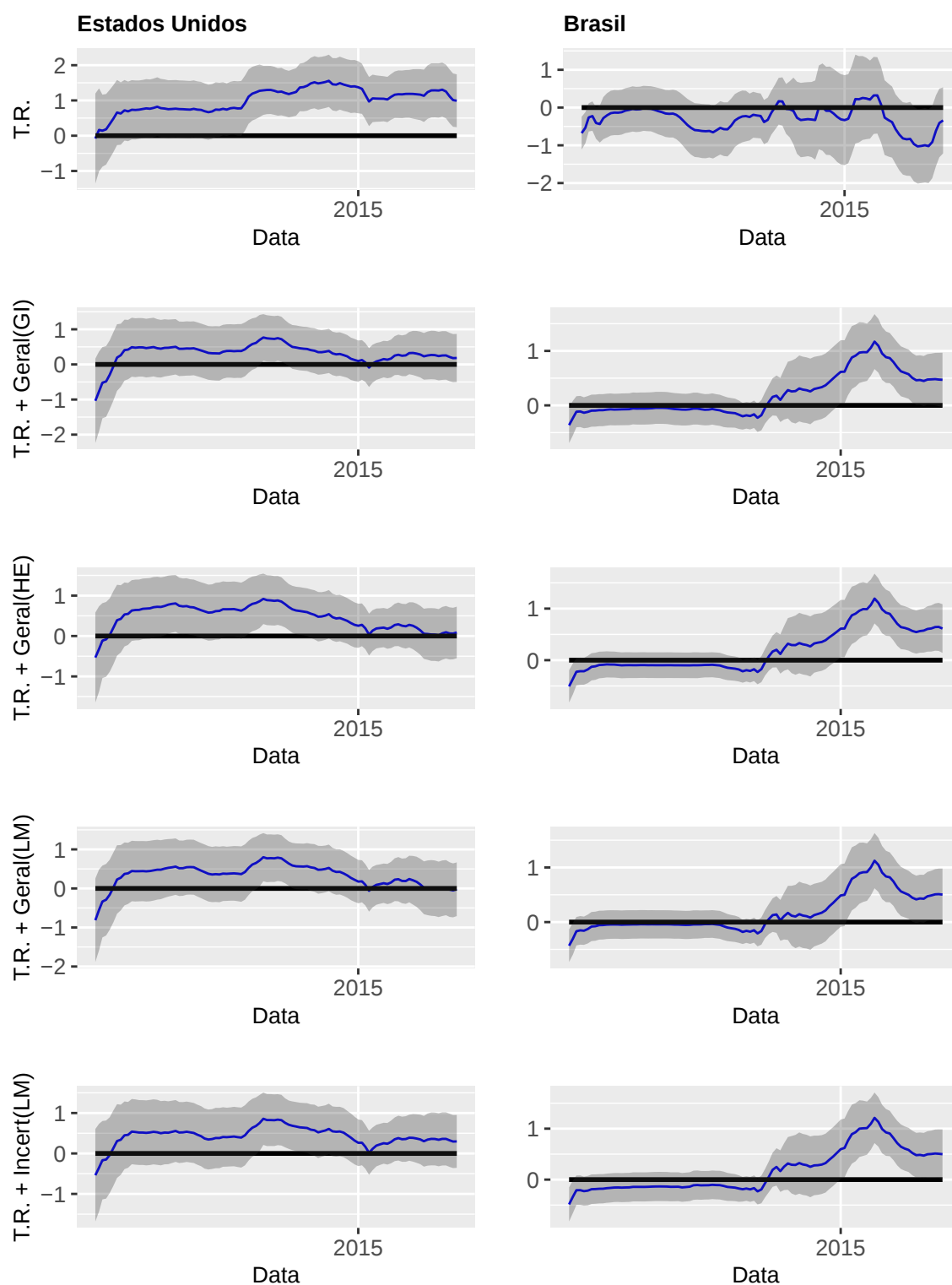
Fonte: Elaboração própria.

Figura 17 – Coeficientes da inflação dos Estados Unidos e Noruega por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



Fonte: Elaboração própria.

Figura 18 – Coeficientes da inflação dos Estados Unidos e México por especificação da regra de Taylor convencional e ampliada com dicionários fixos.



Fonte: Elaboração própria.

Figura 19 – Coeficientes da inflação dos Estados Unidos e Brasil por especificação da regra de Taylor convencional e ampliada com dicionários fixos.