Universidade Federal da Paraíba – UFPB

Centro de Ciências Sociais Aplicadas – CCSA

Programa de Pós-Graduação em Economia – PPGE

Diego Pitta de Jesus

Ensaios em Macroeconomia: Previsão Macroeconômica, Risco Bancário e Preferências do Banco Central.

João Pessoa – PB

2022

Diego Pitta de Jesus

Ensaios em Macroeconomia: Previsão Macroeconômica, Risco Bancário e Preferências do Banco Central.

Tese de doutorado apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba – PPGE/UFPB, em cumprimento às exigências de conclusão do Curso de Doutorado em Economia Aplicada.

Orientador: Prof. Dr. Cássio da Nóbrega Besarria

João Pessoa – PB

2022

Catalogação na publicação
Seção de Catalogação e Classificação

J58e Jesus, Diego Pitta de.
Ensaios em macroeconomia : previsão macroeconômica, risco bancário e preferências do Banco Central / Diego Pitta de Jesus. - João Pessoa, 2022.
108 f. : il.

Orientação: Cássio da Nóbrega Besarria.
Tese (Doutorado) - UFPB/CCSA.

1. Macroeconômica - Previsão de variáveis. 2. Machine Learning. 3. Índice de sentimento. 4. Banco Central do Brasil - Risco bancário. I. Besarria, Cássio da Nóbrega. II. Título.

UFPB/BC CDU 330.101.541(043)

Elaborado por ANNA REGINA DA SILVA RIBEIRO - CRB-15/024

DIEGO PITTA DE JESUS

Ensaios em Macroeconomia: Previsão Macroeconômica, Risco Bancário e Preferências do Banco Central.

Tese apresentada ao Programa de Pós-Graduação em Economia da Universidade Federal da Paraíba - UFPB, em cumprimento às exigências de conclusão do Curso de Doutorado em Economia.

Submetido à apreciação da banca examinadora, sendo aprovado em: 25/02/22

Prof. Dr. Cássio da Nóbrega Besarria
Orientador

Prof. Dr. Edilean Kléber da Silva Bejarano Aragón
Avaliador Interno

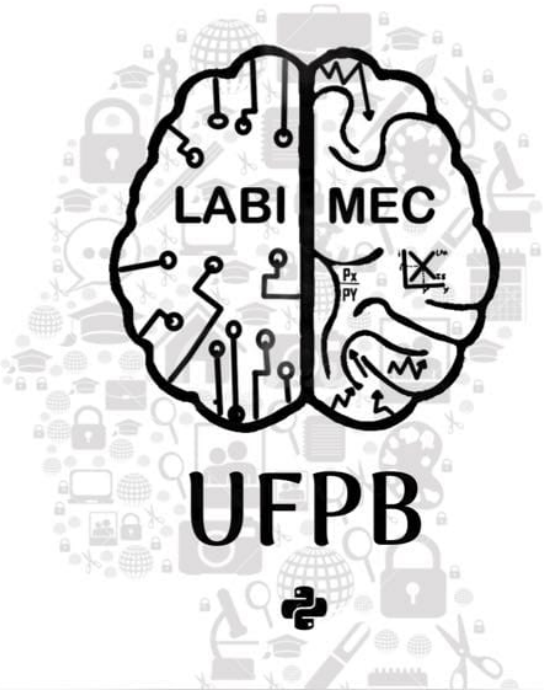
Prof. Dr. Luiz Renato Régis de Oliveira Lima
Avaliador Interno

Prof. Dr. Marcelo Cunha Medeiros
Avaliador Externo

Prof. Dr. Celso José Costa Junior
Avaliador Externo

JOÃO PESSOA – PB

2022


LABI
MEC
UFPB

Lista de ilustrações

Lista de tabelas

Lista de abreviaturas e siglas

IPCA	Índice de Preço ao Consumidor Amplo
IPC	Índice de Preço ao Consumidor
SELIC	Sistema Especial de Liquidação e de Custódia
BCB	Banco Central do Brasil
DSGE	Modelos Dinâmicos Estocásticos de Equilíbrio Geral
FOMC	*Federal Open Market Committee*
PIB	Produto Interno Bruto
CME	*Chicago Mercantile Exchange*
RI	Relatório de Inflação
RF	*Random Forest*
FM	Modelo de Fatores
SVM	*Support Vector Machines*
LASSO	*Least Absolute Shrinkage and Selection Operator*
LARS	*Least Angle Regression*
ML	*Machine Learning)*
COPOM	Comitê de Política Monetária
IBGE	Instituto Brasileiro de Geografia e Estatística
FGV	Fundação Getúlio Vargas
Fipe	Fundação Instituto de Pesquisas Econômica
Anbima	Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais
Funcex	Fundação Centro de Estudos do Comércio Exterior
SECEX	Secretaria de Comércio Exterior
MTE	Ministério do Trabalho
Fiesp	Federação das Indústrias do Estado de São Paulo

BOW	*bag-of-words*
VAR	Vetores Autorregressivos
PDF	*Portable Document Format*
LM	Loughran e Mcdonald
ARMA	Modelo Autoregressivo de Média Móvel
PCA	Análise de Componentes Principais
RMSE	*Root Mean Square Deviation*
MSE	*Mean Square Deviation*
DM	Diebold-Mariano
B3	Brasil, Bolsa, Balcão
LOGIT	Regressão Logística
NB	*Naive Bayes*
AB	*Adaptive Boosting*
DT	*Decision Trees*
RSS	Soma dos Quadrados dos Resíduos
ITR	Formulário de Informação Trimestral
DFP	Demonstrações Financeiras Padronizadas
CVM	Comissão de Valores Mobiliários
SBF	Sentimento Textual Bancário de Dicionário Fixo
SBV	Sentimento Textual Bancário de Dicionário Variante
ROA	Retorno Sobre os Ativos
IBOV	Índice Ibovespa
HP	Hodrick-Prescott
FP	Falso Positivo
FN	Falso Negativo
TFNP	Teoria Fiscal do Nível de Preços

DBGG	Dívida Bruta do Governo Geral
MQO	Mínimos Quadrados Ordinários
GMM	Método Generalizado de Momento
DSGE	Modelo Dinâmico Estocástico de Equilíbrio Geral
MCMC	*Monte Carlo Markov Chain*
IIM	Índice de Incerteza Macroeconômica

Sumário

Resumo

Capítulo 1 – Narrativas do Banco Central e Previsões Macroeconômicas: Usando Análise Textual de *Machine Learning*

O objetivo do artigo é verificar se o tom dos relatórios produzidos pelo Banco Central do Brasil contêm informações que podem ser utilizadas para melhorar a precisão das projeções dos indicadores macroeconômicos para um trimestre à frente. Assim, construímos preditores para a inflação e o crescimento do PIB obtidos a partir de análises textuais da Ata do Copom e do Relatório de Inflação. Para a criação dos índices de sentimento, usamos uma abordagem tradicional de dicionário de léxico fixo e uma nova abordagem que usa o aprendizado de máquina para gerar um dicionário variante no tempo. Em seguida, testamos o poder preditivo das novas variáveis para indicadores macroeconômicos para um período à frente. Também testamos se esses novos preditores são capazes de melhorar o desempenho dos modelos de previsão. Os resultados mostram que as melhores previsões foram obtidas com os modelos que utilizaram a série de pontuação textual de dicionário variante no tempo. O fato aconteceu porque esse tipo de dicionário é capaz de incorporar novos termos que aparecem nos relatórios. Também descobrimos que as previsões de crescimento médio do PIB do mercado podem ser melhoradas com índices de sentimento. Mas, isso não foi verificado para a inflação

Palavras-chave: Previsão Macroeconômica. *Machine Learning*. Índice de Sentimento. Ata do Copom. Relatório da Inflação.

Capítulo 2 – *Machine learning* e Análise de Sentimento: Projetando o Risco de Insolvência Bancária

A principal motivação deste artigo é utilizar técnicas de *machine learning* para construir uma nova métrica de classificação de risco de insolvência para os bancos negociados na B3. Em seguida, será utilizado um conjunto de modelos de predição para projetar a classificação de risco destas instituições. Convencionalmente, a literatura analisa o risco de insolvência bancária a partir dos dados contábeis e variáveis macroeconômicas. Além dessas variáveis, esse artigo irá construir uma série de sentimento do gestor da instituição bancária, via relatórios trimestrais (ITR), e essa será utilizada para melhorar a acurácia das previsões do risco bancário. Os resultados indicam que a classificação de risco bancário, via algoritmo *k-means*, foi capaz de classificar 17% da amostra no grupo de maior risco (1), enquanto 83% da amostra ficou no grupo de menor risco de falência (0). Utilizando a métrica do Z-score verificamos que 65% da amostra faz parte do grupo de baixo risco e 35% da amostra no grupo de risco elevado. Desse modo, o algoritmo *k-means* é mais rigoroso em classificar um banco na categoria de maior risco. Na sequência utilizamos os dados já descritos para projetar o risco de insolvência bancária. Os resultados desta etapa mostraram que o modelo de árvore de decisão apresentou o melhor desempenho para a amostra

de teste. Além disso, constatou-se que a inclusão da variável de sentimento bancário foi capaz de melhorar o desempenho dos modelos de previsão, principalmente, quando o sentimento bancário é construído a partir de um dicionário variante no tempo.

Palavras-chave: Insolvência Bancária. *Machine learning*. *Cluster*. Sentimento Bancário.

Capítulo 3 – O Banco Central do Brazil reage ao Sentimento da Política Fiscal? Uma análise a partir de processamento de linguagem natural

O objetivo do presente trabalho é investigar se o Banco Central vem reagindo ao sentimento da política fiscal. A variável que mede a polaridade da política fiscal foi construída por meio de processamento de linguagem natural e análise de sentimento dos relatórios mensais da dívida pública emitidos pelo Tesouro Nacional. O índice de sentimento foi inserido como variável dependente em duas abordagens para atingir o objetivo do artigo. A primeira é estimação de uma tradicional função de reação do banco central. A segunda é a estimação de um modelo DSGE para se estimar funções de reação e com isto produzir inferências sobre o efeito do sentimento da política fiscal no comportamento da política monetária. Os principais resultados sugerem que o sentimento da política fiscal tem entrado explicitamente no processo decisório da política monetária no Brasil, indicando um possível cenário de dominância fiscal.

Palavras-chave: Análise de Sentimentos. Função de Reação. Sentimento da Política Fiscal. DSGE.

Abstract

Chapter 1 – Central Bank Narratives and Macroeconomic Forecasts: Using Machine Learning Textual Analysis

The aim of the paper is to verify if the tone of the reports produced by the Central Bank of Brazil contain information that can be used to improve the precision of the projections of the macroeconomic indicators for a quarter ahead. Thus, we built predictors for inflation and GDP growth obtained from textual analyzes of the Copom Minutes and Inflation Report. For the creation of sentiment scores, we used a traditional fixed-lexicon dictionary approach and a new approach that uses machine learning to generate a time-varying dictionary. Next, we test the predictive power of the new variables for macroeconomic indicators for a period ahead. We also tested whether these new predictors are able to improve the performance of predictive models. The results show that the best predictions were obtained with the models that used the time-varying dictionary textual score series. The fact happened because this type of dictionary is capable of incorporating new terms that appear in the reports. We also found that market forecasts of average GDP growth can be improved with sentiment scores. But this was not verified for inflation.

Keywords*: Macroeconomic Forecast. Machine Learning. Sentiment Index. Copom minutes. Inflation Report*

Chapter 2 – Machine learning and Sentiment Analysis: Projecting Bank Insolvency Risk

The main motivation of this paper is to use machine learning techniques to build a new insolvency risk rating metric for banks traded on B3. Then, a set of prediction models will be used to project the risk classification of these institutions. Conventionally, the literature analyzes the risk of bank insolvency based on accounting data and macroeconomic variables. In addition to these variables, this work will build a series of sentiment of the bank's manager, via quarterly reports (ITR), and this will be used to improve the accuracy of bank risk forecasts. The results indicate that the banking risk classification, by the k-means algorithm, was able to classify 17% of the sample in the highest risk group (1), while 83% of the sample was in the lowest bankruptcy risk group (0). Using the Z-score metric, we found that 65% of the sample is part of the low-risk group and 35% of the sample is part of the high-risk group. Thus, the k-means algorithm is more rigorous in classifying a bank in the highest risk category. Next, we use the data already described to project the risk of bank insolvency. The results of this step showed that the decision tree model presented the best performance for the sample of test. In addition, it was found that the inclusion of the banking sentiment variable was able to improve the performance of forecast models, especially when banking sentiment is constructed from a time-varying dictionary.

***Keywords**: Bank Insolvency.Machine learning.Cluster.Banking Sentiment*

Chapter 3 – Does the Central Bank of Brazil react to Fiscal Policy Sentiment? An analysis from natural language processing

The objective of the paper is to investigate whether the Central Bank has been reacting to fiscal policy sentiment. The variable that measures the polarity of fiscal policy was constructed through natural language processing and sentiment analysis of monthly public debt reports issued by the National Treasury. The sentiment index was inserted as a dependent variable in two approaches to achieve the objective of the work. The first is the estimation of a traditional central bank reaction function. The second is the estimation of a DSGE model to estimate reaction functions and thus produce inferences about the effect of fiscal policy sentiment on monetary policy behavior. The main results suggest that fiscal policy sentiment has explicitly entered the monetary policy decision-making process in Brazil, indicating a possible scenario of fiscal dominance.

Keywords*: Sentiment Analysis. Reaction Function. Sentiment of Fiscal Policy.DSGE*

Parte I

Previsão Macroeconômica

1 Narrativas do Banco Central e Previsões Macroeconômicas: Usando Análise Textual de *Machine Learning*

1.1 Introdução

A previsão de variáveis macroeconômicas, em particular indicadores-chave como crescimento do PIB, inflação e taxas de juros, são insumos fundamentais para o planejamento orçamentário do governo, a formulação de políticas do banco central e as decisões dos empresários. O uso de abordagens de séries temporais para previsão macroeconômica ganhou impulso nas décadas de 1970 e 1980, pois as previsões dos modelos univariados ARIMA (BOX et al., 2015) e Vetores Autoregressivos (VAR) (SIMS, 1980) mostraram desempenho superior aos modelos macroeconômicos estruturais. Durante essa época, os conjuntos de informações usados para formar previsões geralmente continham apenas um pequeno número de variáveis. Entretanto, tais métodos possuem uma limitação importante, em que esses modelos suportam apenas um pequeno número de preditores. Essa situação mudou no início dos anos 2000, quando os pesquisadores começaram a usar dados macroeconômicos de alta dimensão. Nesse cenário, modelos capazes de lidar com um grande numéro de preditores começaram a ganhar destaque. Então, a literatura de previsão macroeconômica passou a utilizar com maior frequência os modelos de fatores e os modelos de encolhimento de *machine learning*.

Dois exemplos que podem ser encontrados na literatura são o conjunto de dados dos EUA que contém 149 variáveis medidas com uma frequência mensal apresentada em Stock e Watson (2002) e o conjunto de dados da área do euro contendo 447 variáveis mensuradas com frequência mensal apresentadas Forni et al. (2003). Nos dois estudos, a utilização de um grande número de preditores em uma estrutura de modelagem de fatores dinâmico apresentou um melhor desempenho nas previsões da produção industrial em relação aos modelos tradicionais de referência. Um fator importante na popularidade dessa abordagem é sua simplicidade, em que os componentes principais fornecem estimativas consistentes dos fatores dinâmicos e podem subsequentemente ser usados em regressões preditivas auxiliares. De acordo com Eickmeier e Ziegler (2008) existe uma extensa literatura que mostra que quando o modelo de fatores é usado com um grande número de preditores, produz boas previsões para variáveis macroeconômicas, como PIB e inflação, para várias economias diferentes.

Apesar de seu sucesso, o modelo de fatores dinâmicos não é a única estrutura para previsão com um grande número de preditores. Os avanços da estatística e da literatura de *machine learning* também foram explorados no contexto macroeconômico. Por exemplo, Mol et

al. (2008) consideram a regressão de ridge e de *least absolute shrinkage and selection operator* LASSO (TIBSHIRANI, 1996) para os dados de Stock e Watson (2002) e obtiveram previsões com desempenho semelhante ao obtido em com o modelo de fatores dinâmicos. Bai e Ng (2008) usam regressão LARS (EFRON et al., 2004) para selecionar um conjunto de preditores. As previsões foram produzidas usando esses preditores selecionados pelo modelo. Bai e Ng (2008) mostram que, pelo menos em alguns períodos dos dados, os métodos baseados no *Least Angle* (LARS) produzem melhores previsões de inflação, renda, vendas no varejo, produção industrial e emprego total em comparação com o modelo de fatores principais.

Também surgiram na literatura trabalhos que usam métodos que respondem pela incerteza do modelo, como agregação de *bootstrap* ou "*bagging*"[1] foram bem-sucedidos na previsão da inflação por Inoue e Kilian (2008). Finalmente, na classe de previsão multivariada, houve um foco em modelos VAR de grande dimensão estimados usando técnicas bayesianas. Exemplos incluem Kadiyala e Karlsson (1997) e, mais recentemente, Bańbura et al. (2010), Carriero et al. (2011) e Koop (2013) que utilizam priores de encolhimento.

No caso do Brasil, nos últimos anos, surgiu um crescente corpo de literatura sobre previsão macroeconômica com destaque para os modelos de *machine learning*. Medeiros e Mendes (2016) consideraram diferentes modelos de alta dimensão para prever a inflação brasileira. Os autores mostraram que as técnicas baseadas no LASSO apresentam os menores erros de previsão para previsões de horizonte curto. Para horizontes mais longos, o *benchmark* de AR é o melhor modelo para previsão de pontos, mesmo que de acordo com os autores, não haja diferenças significativas entre eles. Modelos fatoriais também produzem boas previsões de longo horizonte em alguns casos. Mais recentemente, Garcia et al. (2017) e Medeiros et al. (2019) usaram modelos de alta dimensão para prever a inflação em tempo real do Brasil e mostraram que o desempenho dos modelos de encolhimento é superior em relação às técnicas mais tradicionais.

Barbosa et al. (2020) analisaram o desempenho de modelos fatoriais de alta dimensão para prever quatro variáveis macroeconômicas brasileiras: duas variáveis reais, taxa de desemprego e o índice de produção industrial, e duas variáveis nominais, Índice de Preço ao Consumidor Amplo (IPCA) e Índice de Preços ao Consumidor (IPC). Os autores usaram três tipos de técnicas de aprendizado estatístico: métodos de *shrinkage*, combinações de previsões e seleção de previsores. Os fatores foram extraídos de forma supervisionada e não supervisionada. Os resultados indicaram que métodos de aprendizado estatístico melhoram o desempenho preditivo das variáveis econômicas brasileiras.

Araujo e Gaglianone (2020) realizaram um exercício de previsão fora da amostra, através de uma variedade de técnicas de *machine learning* e modelos econométricos tradicionais. Os resultados encontrados pelos autores corroboram conclusões recentes a favor dos procedimentos automatizados não-lineares, indicando que algoritmos de *machine learning* (em particular, *random forest*) podem superar os métodos tradicionais de previsão em termos de erro quadrático

[1] Ver os trabalhos de Breiman (1996), Bühlmann et al. (2002) e Lee e Yang (2006).

médio.

Contudo, surgiram alguns trabalhos na literatura internacional de previsão macroeconômica que voltaram o seu foco para a construção de índices de sentimentos[2] a partir das informações textuais contidas nos relatórios publicados pelo Banco Central da Inglaterra (JONES et al., 2019);(CLEMENTS; READE, 2020), que são usados como preditores para a taxa de inflação, Produto Interno Bruto (PIB) e outros indicadores macroeconômicos. Esses indicadores de polaridade do Banco Central podem ser usados como preditores diretos para as variáveis macroeconômicas do próximo período. Também podem ser utilizados em modelos mais gerais (por exemplo, modelos VAR, modelos de *machine learning*, modelos de *deep learning* e etc) para melhorar o desempenho de tais modelos de previsão.

De acordo com Jones et al. (2019) a avaliação das previsões econômicas concentrou-se tipicamente na qualidade das previsões numéricas. No entanto, esse foco nas previsões quantitativas negligencia a quantidade substancial de texto que freqüentemente as acompanha, principalmente nas publicações dos bancos centrais. O texto é incluído para fornecer contexto e nuances e pode ser avaliado usando análise textual. Além disso, as previsões numéricas nem sempre estão disponíveis ao público, mas o texto dos relatórios, as atas das reuniões e os discursos podem revelar informações sobre a avaliação das condições econômicas atuais e futuras. Ainda segundo Jones et al. (2019) a literatura estabeleceu o valor da análise textual, bem como uma metodologia geral para converter texto em *scores* quantitativos que avaliam principalmente as polaridades dos textos. De acordo com Gentzkow et al. (2019), as informações codificadas no texto são um complemento rico para os tipos de dados mais estruturados tradicionalmente usados na pesquisa empírica. De fato, nos últimos anos, ocorreu um uso intenso de dados textuais em diferentes áreas de pesquisa.

Existe um corpo de literatura em rápida expansão sobre o uso de informações textuais, como as narrativas sutis do tipo que aparecem nos relatórios de inflação e atas dos bancos centrais. Um artigo-chave é o de Stekler e Symington (2016) que investiga as "narrativas"que constituem as atas do *Federal Open Market Committee* (FOMC) entre 2006 e 2010. Seu estudo quantifica as declarações qualitativas das atas do FOMC sobre o atual e o futuro tendências da economia e compara os índices resultantes às previsões do *Greenbook* e da pesquisa de previsões profissionais dos EUA.

Dossani (2019) analisa como o tom das entrevistas coletivas do Banco Central dos Estados Unidos afeta os prêmios de risco no mercado de câmbio. Ele mede o tom como a diferença entre o número de frases *hawkish* e *dovish* feitas durante uma conferência de imprensa. Ele usou quatro contratos futuros de moeda negociados na *Chicago Mercantile Exchange* (CME) e descobriu que a aversão ao risco implícita aumenta quando os Bancos Centrais são *hawkish* e diminui quando os Bancos Centrais estão *dovish*.

[2] De acordo com Medhat et al. (2014) a análise de sentimento é o estudo computacional das opiniões, atitudes e emoções contidas em textos escritos. Em geral, a mineração de opinião ajuda a coletar informações sobre os aspectos positivos e negativos de um tópico específico.

Goldfarb et al. (2005) desenvolveram um procedimento de pontuação que eles usaram para criar um índice de previsões qualitativas feitas durante a Grande Depressão com base em uma análise textual de artigos de jornais contemporâneos. Lundquist e Stekler (2012) aplicaram o procedimento de Goldfarb et al. (2005) no estudo de avaliações qualitativas feitas por economistas durante a Grande Recessão. Mais recentemente, Stekler e Symington (2016) usaram o procedimento para examinar as atas das reuniões do FOMC, e Mathy e Stekler (2018) ajustaram o procedimento de pontuação para incorporar perspectivas mais negativas. Eles também propuseram uma metodologia de avaliação que envolvia o uso de coeficientes de correlação para comparar as pontuações com uma série de *benchmarks*. Catalfamo et al. (2018) incorporou essas duas inovações em seu exame dos noticiários franceses da economia dos EUA durante a Grande Recessão.

Jones et al. (2019) realizaram análises qualitativas via análise de sentimentos dos textos dos relatórios de inflação do Banco Central da Inglaterra no período de 2005-2014. Também construíram índices de polaridade para fora da amostra. Em seguida, compararam as pontuações com os dados de crescimento do produto em tempo real e com as projeções quantitativas correspondentes publicadas pelo Banco Central. Concluíram que a evolução geral da economia do Reino Unido foi representada com precisão no texto do Relatório de Inflação. Além disso, as regressões de eficiência sugeriram que há informações no texto que poderiam melhorar os relatórios quantitativos do Banco da Inglaterra e as previsões para um trimestre à frente.

Clements e Reade (2020) analisaram as narrativas que acompanham as previsões numéricas nos relatórios trimestrais de inflação do Banco da Inglaterra, para o período de 1997–2018. O trabalho se concentrou em saber se as narrativas contêm informações úteis sobre o curso futuro das principais variáveis macroeconômicas além das previsões pontuais, em termos de se as narrativas podem ser usadas para aprimorar a precisão das previsões numéricas. Também foi considerado se as narrativas são capazes de prever mudanças futuras nas previsões numéricas. Os autores concluíram que uma medida de sentimento derivada das narrativas pode prever os erros nas previsões numéricas de crescimento do produto, mas não da inflação. Não encontraram evidências de que mudanças passadas no sentimento prevejam mudanças subsequentes nas previsões pontuais do crescimento do produto ou da inflação, mas descobriram que os ajustes nas previsões numéricas do crescimento do produto têm um elemento sistemático.

Até o momento, essa literatura para países emergentes como o Brasil é escassa. Portanto, o presente trabalho busca preencher essa lacuna da literatura. Nesse contexto, o trabalho tem como objetivo principal construir índices de sentimentos via análise de sentimentos a partir da Ata do Copom e do Relatório de Inflação (RI) produzidos pelo Banco Central do Brasil (BCB). Este artigo busca verificar se as narrativas na ata do Copom e no RI contêm informações textuais úteis que podem ser usadas para melhorar a precisão das previsões de indicadores macroeconômicos, tais como a taxa de inflação e o crescimento do PIB. Então, para realizar esses objetivos iremos realizar dois exercícios. No primeiro vamos fazer previsões para a taxa de inflação e para o crescimento do PIB via modelos tradicionais e modelos de *machine learning*.

Nesse caso, pretendemos saber se a inclusão de índices de sentimentos nesses modelos é capaz de melhorar a acurácia da previsão. Vamos comparar o desempenho das nossas previsões com o desempenho da média do mercado, representada pelo Focus. O segundo exercício vai relacionar diretamente as informações textuais da ata do Copom e do RI com as previsões do Focus. Neste contexto, vamos verificar se as informações textuais são capazes de melhorar as previsões do Focus.

Assim, o primeiro passo será construir séries temporais que representem a polaridade (sentimento) nos textos da ata do Copom e do Relatório de Inflação por métodos de análise textual. Nessa etapa, usamos o tradicional dicionário financeiro de Loughran e McDonald (2016) com léxicos fixos, que é amplamente utilizado e difundido nos trabalhos que analisam textos financeiros. Adicionalmente, empregamos a abordagem de Lima et al. (2019) que permite que o conteúdo do dicionário seja variante ao longo do tempo.

Após a obtenção das séries de sentimentos textuais, vamos verificar se esses novos indicadores são capazes melhorar o desempenho de modelos de previsão multivariados. Os modelos multivariados utlizados nesse exercício são o Modelo de Fatores (FM), LASSO, *Random Forest* (RF) e *Support Vector Machines* (SVM). Como *benchmark* usaremos as previsões fornecidas pelo Boletim Focus. As previsões serão realizadas para cada período à frente fora da amostra (y_{t+h}) e levará apenas em consideração as informações disponíveis até o momento X_t. Também, de forma similar a Jones et al. (2019), vamos testar a eficiência da previsão do Focus pela versão estentida de equação de Mincer e Zarnowitz (1969). Neste caso, será possível saber se as previsões realizadas pelo Focus podem ser melhoradas com a incorporação de todas as informações e nuances contidas nas narrativas do BCB.

Assim, o presente trabalho possui algumas contribuições. A primeira é incorporar na comparação de desempenho das previsões o uso de um modelo *machine learning* que nenhum trabalho mencionado anteriormente utilizou, neste caso o SVM. A segunda contribuição é usar um método de dicionário variante no tempo, pois todos os trabalhos da literatura citados anteriormente usam dicionários fixos. A terceira é de preencher algumas lacunas na literatura brasileira de previsão macroeconômica por meio de uma abordagem alternativa e inovadora. Por fim, a quarta contribuição é mostrar aos agentes que os textos dos relatórios produzidos pelas instituições, principalmente os provenientes do Banco Central, contém informações textuais revelantes para o exercício de previsão que não podem ser ignoradas e devem ser incorporadas no arcabouço metodológico dos principais agentes que realizam previsão de variáveis macroeconômicas.

Os resultados obtidos mostram de forma geral que os modelos de previsão que usam as séries de sentimentos apresentam os menores erros de previsão, e em alguns casos apresentando erros de previsão inferiores ao Focus. As melhores previsões foram obtidas com os modelos que usaram as séries de sentimentos provenientes do dicionário variante no tempo. Esse fato ocorreu porque esse tipo de dicionário é capaz de incorporar novos termos que aparecem nos relatórios. Assim, as séries de sentimentos construídas a partir do dicionário variante conseguiram captar,

por exemplo, a pandemia do Covid-19, pois os termos relacionados a pandemia surgiram em 2019 e não constam nos dicionários fixos, então, o sentimento dos relatório é mais realista com o dicionário variante no tempo.

Por fim, a equação de Mincer e Zarnowitz (1969) para a eficiência das previsões do Focus mostrou que os índices de sentimentos são capazes de prever os erros das previsões do Focus para o PIB, ou seja, na média o mercado para as previsões do PIB não leva em consideração a totalidade das informações e nuances textuais contidas nos relatórios. Assim, as previsões do crescimento do PIB realizadas pelo Focus podem ser melhoradas com os índices de sentimento. Mas, isso não foi constatado para a inflação. Esse fato está em conformidade com os resultados anteriormente encontrados, pois os erros de previsão do Focus para a inflação é bem menor do que para o PIB. Ao mesmo tempo, nossas previsões com sentimentos conseguiram superar o Focus apenas nas previsões realizadas para o PIB.

O presente artigo além da introdução é dividido em quatro partes. A primeira trata dos dados e da amostragem usada no artigo. A segunda apresenta a metodologia empregada para a obtenção dos resultados. A terceira ilustra os principais resultados. Por fim, a quarta discute as principais conclusões e limitações do trabalho.

1.2 Dados

1.2.1 Ata do Copom e Relatório de Inflação

De acordo com Filho e Rocha (2010) a ata do Copom é um dos principais instrumentos de comunicação do Banco Central do Brasil, apresenta projeções econômicas para o cenário nacional e internacional, controle da inflação, decisões a respeito dos juros etc. É através dela que a autoridade monetária explica os procedimentos utilizados para a tomada de decisão de política monetária com o objetivo de tornar a comunicação mais transparente e manter sob controle as expectativas.

Assim como no trabalho de Silva et al. (2019) que também usaram as atas do Copom para um exercício de análise textual, é necessário levar em conta uma mudança referente a periodicidade de publicação, durante o período 2000 a 2005 as reuniões ocorriam de forma mensal assim como as publicações, a partir do ano de 2006 as reuniões passaram a acontecer a cada quarenta e cinco dias, sendo divulgadas oito atas ao ano. Outro ponto que merece ser destacado é o fato de que em 2002 foram publicadas treze atas ao invés de doze. A série de polaridade após ser estimada precisou ser convertida em dados trimestrais com o objetivo de que toda a amostra passasse a ter a mesma frequência.

De acordo com o BCB o Relatório de Inflação apresenta as diretrizes das políticas adotadas pelo Copom, considerações acerca da evolução recente do cenário econômico e projeções para a inflação. As projeções são apresentadas em cenários com condicionantes para algumas

variáveis econômicas. O Comitê de Política Monetária (Copom) utiliza um conjunto amplo de modelos e cenários para orientar suas decisões de política monetária. Ao expor alguns desses cenários, o Copom procura dar maior transparência às decisões de política monetária, contribuindo para sua eficácia no controle da inflação, que é seu objetivo principal.

A frequência do RI é trimestral, entretanto, o relatório possui um "*gap*"entre os anos de 2012 e 2015. Nesses anos o Copom disponibiliza apenas o sumário executivo do RI, que é um resumo do documento. Portanto, optamos por usar o sumário executivo como *proxy* do RI durante os anos de 2012 e 2015.

Assim, escolhemos para a compor a amostra com janela temporal com início no primeiro trimestre de 2005 até o segundo trimestre de 2020, totalizando 62 observações. Foi possível obter todas as atas do Copom e os Relatórios de Inflação por *web scraping*, em que importamos e trabalhamos com os documentos no formato *Portable Document Format* (PDF) na versão em língua inglesa.

1.2.2 Variáveis Macroeconômicas

Para o exercício de predição foram utilizadas um grande conjunto de variáveis macroeconômicas e financeiras como preditores. As variáveis de interesse a serem previstas são a taxa inflação e a taxa de crescimento do PIB. Para a taxa de inflação optamos por escolher o índice IPCA que é o índice de inflação oficial adotado pelo Copom para fins de política monetária. A frequência e periodicidade das variáveis são idênticas aos dos relatórios da inflação na seção anterior. A Tabela 1 e a Tabela 2 ilustram todas as variáveis. Os preditores são os mesmos escolhidos no trabalho de Araujo e Gaglianone (2020).

Tabela 1 – Lista de variáveis macroeconômicas e financeiras

Séries	Categoria	Nome	Fonte	Unidade
1	Inflação	IPCA	IBGE	%
2	Inflação	IPCA - Preços de mercado	IBGE	%
3	Inflação	IPCA - Preços monitorados	IBGE	%
4	Inflação	IPCA - Negociáveis	BCB	%
5	Inflação	IPCA - Não-negociáveis	BCB	%
6	Inflação	IPC-Fipe	Fipe	%
7	Inflação	IPC-Br	FGV	%
8	Inflação	IPA-DI - Atacado	FGV	%
9	Inflação	IGP-DI	FGV	%
10	Inflação	IGP-M	FGV	%
11	Inflação	IGP-10	FGV	%
12	Inflação	INCC	FGV	%
13	Inflação	IPC-Br - Núcleo	FGV	%
14	Inflação	IPCA - Núcleo (EX0 exclusão)	BCB	%
15	Inflação	IPCA - Núcleo (EX1 exclusão)	BCB	%
16	Inflação	IPCA - Núcleo (dupla ponderação)	BCB	%
17	Inflação	IPCA - Núcleo (médias aparadas sem suavização)	BCB	%
18	Inflação	IPCA - Núcleo (Break Even, 1 ano)	Anbima	%
19	Inflação	IPCA - Núcleo (Break Even, 2 ano)	Anbima	%
20	Inflação	IPCA - Núcleo (Break Even, 5 ano)	Anbima	%
21	Taxa de juros	Selic	BCB	%
22	Taxa de juros	TJLP	BCB	%
23	Taxa de juros	Tesouro prefixado (1 ano)	Anbima	%
24	Taxa de juros	Tesouro prefixado (2 ano)	Anbima	%
25	Taxa de juros	Tesouro prefixado (5 ano)	Anbima	%
26	Taxa de juros	Pré-DI Swap (1 ano)	BCB	%
27	Taxa de juros	Pré-DI Swap Focus (1 ano)	BCB	%
28	Taxa de juros	Tesouro IPCA (1 ano)	Anbima	%
29	Taxa de juros	Tesouro IPCA (2 ano)	Anbima	%
30	Taxa de juros	Tesouro IPCA (5 ano)	Anbima	%
31	Moeda	Base monetária	BCB	R$ mil
32	Moeda	Oferta de moeda	BCB	R$ mil
33	Moeda	Depósitos (conta corrente)	BCB	R$ mil
34	Moeda	Depósitos (poupança)	BCB	R$ mil
35	Moeda	M1	BCB	R$ mil
36	Moeda	M2	BCB	R$ mil
37	Moeda	M3	BCB	R$ mil
38	Moeda	M4	BCB	R$ mil
39	Setor bancário	Spread de crédito (recursos livros, taxa Selic)	BCB	Pontos base
40	Setor bancário	Índice de inadimplência da carteira de crédito	BCB	%
41	Setor bancário	Proporção dos empréstimos	BCB	Unidade
42	Setor bancário	Proporção das reservas	BCB	Unidade
43	Setor bancário	Crescimento das operações de crédito	BCB	R$ milhões
44	Mercado de capitais	IPOs (acumulados em 12 meses)	BCB	R$ milhões
45	Mercado de capitais	Patrimônio líquido dos fundos de ações	BCB	R$ milhões
46	Mercado de capitais	Patrimônio líquido dos fundos de investimento financeiro	BCB	R$ milhões
47	Mercado de capitais	Ibovespa	Reuters	Índice
48	Mercado de capitais	MSCI países emergentes (US$)	Reuters	Índice
49	Mercado de capitais	MSCI países desenvolvimentos (US$)	Reuters	Índice
50	Prêmio de risco	EMBI	Reuters	Pontos base
51	Prêmio de risco	EMBI (composto - média de 16 países)	Reuters	Pontos base
52	Prêmio de risco	CDS (5 anos)	Reuters	Pontos base
53	Taxa de câmbio	Taxa de câmbio nominal (R$/US$)	BCB	Unidades
54	Taxa de câmbio	Taxa de câmbio real efetiva (IPA - 13 moedas)	Funcex	Índice
55	Economia mundial	U.S. dollar index (média geométrica de 6 em relação ao US$)	Reuters	Índice
56	Economia mundial	U.S. Treasury 2 years (Treasury nominal interest rates)	Reuters	%
57	Economia mundial	U.S. Treasury 10 years (Treasury nominal interest rates)	Reuters	%
58	Economia mundial	U.S. Treasury 5 years TIPS (Treasury Inflation Protected Securities)	Reuters	%
59	Economia mundial	CRB all commodities index	Reuters	Índice
60	Economia mundial	Oil price (WTI, OklahomaUSA)	Reuters	US$/barril
61	Economia mundial	VIX CBOE volatility index (30 day expected volatility of the S&P500)	Reuters	Índice

Fonte: Araujo e Gaglianone (2020)

Tabela 2 – Lista de variáveis macroeconômicas e financeiras - continuação

Séries	Categoria	Nome	Fonte	Unidade
62	Setor externo	Índice de preços de importação	Funcex	Índice
63	Setor externo	Índice de quantidade de importação	Funcex	Índice
64	Setor externo	Índice de preços de exportação	Funcex	Índice
65	Setor externo	Índice de quantidade de exportação	Funcex	Índice
66	Setor externo	Importação (FOB, total)	MDIC/Secex	US$
67	Setor externo	Exportação (FOB, total)	MDIC/Secex	US$
68	Setor externo	Exportação (FOB, bens primários)	MDIC/Secex	US$
69	Setor externo	Reservas internacionais (total)	BCB	US$ milhões
70	Setor externo	Conta corrente (líquida)	BCB	US$ milhões
71	Setor externo	Conta corrente (acumulada em 12 meses como proporção do PIB)	BCB	%
72	Setor externo	Investimento estrangeiro direto (acumulado em 12 meses)	BCB	US$ milhões
73	Setor externo	Investimento estrangeiro em portfólio (acumulado em 12 meses)	BCB	US$ milhões
74	Atividade econômica	IBC-Br	BCB	Índice
75	Atividade econômica	PIB (acumulado em 12 meses, preços de mercado)	BCB	R$ milhões
76	Atividade econômica	Índice de confiança do consumidor	Fecomercio	Índice
77	Trabalho	Taxa de desemprego	IBGE	%
78	Trabalho	Índice de empregados registrados (comércio por atacado e varejo)	MTE	Índice
79	Trabalho	Índice de empregados registrados (setor de construção)	MTE	Índice
80	Trabalho	Horas trabalhadas na produção (São Paulo)	Fiesp	Índice
81	Trabalho	Salário real (industria, São Paulo)	Fiesp	Índice
82	Indústria	Produção industrial (total)	IBGE	Índice
83	Indústria	Produção industrial (extração mineral)	IBGE	Índice
84	Indústria	Produção industrial (industria manufatureira)	IBGE	Índice
85	Indústria	Produção industrial (bens de capital)	IBGE	Índice
86	Indústria	Produção industrial (bens intermediários)	IBGE	Índice
87	Indústria	Produção industrial (bens de consumo)	IBGE	Índice
88	Indústria	Produção industrial (bens duráveis)	IBGE	Índice
89	Indústria	Produção industrial (bens semi-duráveis e não duráveis)	IBGE	Índice
90	Indústria	Utilização da capacidade instalada (São Paulo)	Fiesp	%
91	Indústria	Utilização da capacidade instalada (indústria manufatureira)	FGV	%
92	Indústria	Produção de aço	BCB	Índice
93	Indústria	Produção de veículos (total)	BCB	Unidades
94	Indústria	Produção de veículos (veículos de passeio e veículos comerciais leves)	BCB	Unidades
95	Indústria	Produção de caminhões	BCB	Unidades
96	Indústria	Produção de ônibus	BCB	Unidades
97	Indústria	Produção de máquinas agrícolas (total)	BCB	Unidades
98	Vendas	Índice de volume de vendas no setor de varejo (total)	BCB	Índice
99	Vendas	Índice de volume de vendas no setor de varejo (combustíveis e lubrificantes)	BCB	Índice
100	Vendas	Índice de volume de vendas no setor de varejo (hiperm. superm, alim. beb. tab.)	BCB	Índice
101	Vendas	Índice de volume de vendas no setor de varejo (têxtil, vestuário e calçados)	BCB	Índice
102	Vendas	Índice de volume de vendas no setor de varejo (móveis e itens brancos)	BCB	Índice
103	Vendas	Índice de volume de vendas no setor de varejo (veículos e motos, peças de reposição)	BCB	Índice
104	Vendas	Índice de volume de vendas no setor de varejo (hipermercados e supermercados)	BCB	Índice
105	Vendas	Vendas de veículos (total)	BCB	Unidades
106	Vendas	Vendas de veículos (domésticos)	BCB	Unidades
107	Energia	Consumo de energia elétrica (comercial)	Eletrobras	GWh
108	Energia	Consumo de energia elétrica (residencial)	Eletrobras	GWh
109	Energia	Consumo de energia elétrica (industrial)	Eletrobras	GWh
110	Energia	Consumo de energia elétrica (outros)	Eletrobras	GWh
111	Energia	Consumo de energia elétrica (total)	Eletrobras	GWh
112	Setor Público	Resultado primário do setor público consolidado (fluxos mensais atuais)	BCB	R$ milhões
113	Setor Público	Resultado primário do setor público consolidado (fluxos acumulados em 12 meses)	BCB	R$ milhões
114	Setor Público	Resultado primário do setor público consolidado (fluxos acumulados em 12 meses, % PIB)	BCB	%
115	Setor Público	Dívida pública líquida (total, governo federal e banco central,% PIB)	BCB	%
116	Setor Público	Dívida pública líquida (interna, governo federal e banco central,% PIB)	BCB	%
117	Setor Público	Dívida pública líquida (externa, governo federal e banco central,% PIB)	BCB	%
118	Setor Público	Dívida pública líquida (total, setor público consolidado, saldos em reais)	BCB	R$ milhões
119	Setor Público	Dívida pública líquida (interna, setor público consolidado, saldos em reais)	BCB	R$ milhões
120	Setor Público	Dívida pública líquida (externa, setor público consolidado, saldos em reais)	BCB	R$ milhões

Fonte: Araujo e Gaglianone (2020)

As variáveis possuem frequência trimestral iniciando no primeiro trimestre de 2005 até o segundo trimestre de 2020. Os preditores foram obtidos nos sites do Instituto Brasileiro de Geografia e Estatística (IBGE), BCB, Fundação Getúlio Vargas (FGV), Fundação Instituto de Pesquisas Econômicas (Fipe), Associação Brasileira das Entidades dos Mercados Financeiro e de Capitais (Anbima), *Reuters*, Fundação Centro de Estudos do Comércio Exterior (Funcex), Se-

cretaria de Comércio Exterior (SECEX), Fecomércio, Ministério do Trabalho (MTE), Federação das Indústrias do Estado de São Paulo (Fiesp) e Eletrobrás.

1.3 Metodologia

Esta seção descreve os métodos usados neste artigo para as previsões. Assim, como Garcia et al. (2017) consideramos uma abordagem direta de previsão em que a variável a ser prevista está *h* períodos à frente, y_{t+h}, é modelada em função de um conjunto de preditores medido no tempo *t*, como:

$$y_{t+h} = T\left(x_t\right) + u_{t+h} \tag{1.1}$$

em que $T(x_t)$ é um mapeamento de um conjunto de *q* preditores, u_{t+h} é o erro de previsão e $x_t = (x_{1t}, \ldots, x_{qt})' \in \mathbb{X} \subseteq \mathbb{R}^q$ pode incluir preditores fracamente exógenos, valores defasados da variável de interesse e vários fatores calculados a partir de um grande número de covariáveis potenciais. É importante ressaltar que x_t contém apenas variáveis observadas e disponíveis no momento *t*. Observe que a consideração de modelos de previsão direta para cada horizonte evita a necessidade de estimar um modelo para a evolução de x_t.

Também como em Garcia et al. (2017) para a maioria dos métodos considerados neste artigo, o mapeamento T(·) é linear, de modo que:

$$y_{t+h} = \beta' x_t + u_{t+h} \tag{1.2}$$

em que $\beta \in \mathbb{R}^q$ é um vetor de parâmetros desconhecidos.

1.3.1 Mensurando os Índices de Sentimento

Nesta seção será apresentada a metodologia de construção dos índices de sentimentos (S) a partir dos textos das atas do Copom e do RI. Cada S_t visa capturar algumas das informações da narrativa no relatório no momento *t*, para cada documento em nossa amostra. Essa medida transforma milhares de palavras em um único número. Para obter cada série de sentimento S_t usamos duas abordagens: uma que mensura os sentimentos a partir de dicionários com léxicos fixos e outra que usa modelos de *machine learning* para construir um dicionário variante no tempo.

Antes de executar a análise lexicográfica nos documentos, realizamos uma série de transformações no texto original. O texto é primeiro dividido em uma sequência de *substrings* (*tokens*) cujos caracteres são todos transformados em minúsculas. Removemos *stop words* em inglês e limpamos o texto usando o pacote *tolower* do R.

De acordo com Shapiro et al. (2020) existem duas metodologias gerais para quantificar o sentimento no texto. A primeira é conhecida como metodologia lexical. Esta abordagem se baseia em listas predefinidas de palavras, chamadas de léxicos ou dicionários, com cada palavra atribuída uma pontuação para a emoção de interesse. Geralmente, essas pontuações são simplesmente 1, 0 e -1 para positivo, neutro e negativo, mas alguns léxicos têm mais de três categorias. As aplicações típicas desta abordagem medem o conteúdo emocional de um determinado *corpus* de texto com base na prevalência de palavras negativas vs. positivas no *corpus*. Esses métodos de correspondência de palavras são chamados de métodos de *bag-of-words* (BOW) devido as características contextuais de cada palavra, como sua ordem no texto, classe gramatical, coocorrência com outras palavras e outras características contextuais específicas ao texto em que a palavra aparece, são ignorados.

Dentre esse tipo de método destaca-se o dicionário criado por Loughran e McDonald (2011) (LM). Os autores contruíram listas de palavras negativas e positivas que são selecionadas para serem apropriadas ao texto financeiro. Eles mostram que seus dicionários são superiores para classificar textos econômicos e financeiros a outros dicionários, por exemplo, o de Apel e Grimaldi (2012) e o *Harvard Psychosociological Dictionary*, que tende a categorizar incorretamente palavras neutras em um contexto financeiro/econômico (por exemplo, impostos, custos, capital, despesa, responsabilidade, risco, excesso e depreciação). Existem 2.355 palavras negativas e 354 palavras positivas nos dicionários LM. Portanto, para a construção dos índices de sentimentos via abordagem de dicionários fixos usamos o dicionário de LM.

Shapiro et al. (2020) afirma que a segunda abordagem, mais incipiente, emprega técnicas de *machine learning* para construir modelos complexos para prever probabilisticamente o sentimento de um determinado conjunto de texto. Umas das aplicações dos modelos *machine learning* é na construção de dicionários variantes no tempo. Lima et al. (2019) usaram essa abordagem para criar um método de dicionário com termos variantes.

De acordo com Lima et al. (2019) a suposição de um dicionário invariável no tempo não parece ser realista em documentos que introduzem novas palavras ao longo do tempo ou se o vocabulário usado em períodos de recessão difere do usado em períodos de expansões econômicas. Os autores ressaltam que mesmo se o vocabulário fosse constante ao longo do tempo, o poder preditivo de algumas palavras pode variar, ou seja, a relevância das palavras se alteram ao longo do tempo, mas a literatura existente não explica esse efeito e, portanto, os preditores resultantes não refletem as informações textuais mais preditivas encontradas nos documentos em um determinado momento. Portanto, aplicamos para a construção dos índices de sentimentos via dicionários variantes no tempo usamos a abordagem desenvolvida por Lima et al. (2019).

Assim, utilizando a metodologia proposta pelos autores para construir o dicionário variante no tempo, primeiramente criamos um vetor de séries temporais, X_t, em que cada elemento do vetor mostra observações em série temporal da frequência em que cada palavra (ou

combinação de palavras) aparece na ata do Copom e no RI até o tempo *t*. Portanto, esta etapa transforma as palavras em valores numéricos sem usar um dicionário pré-especificado (fixo). Essa representação numérica é de alta dimensão e esparsa; portanto, a redução da dimensionalidade deve ser empregada na próxima etapa. Na segunda etapa, usamos um algoritmo de *machine learning* supervisionado para selecionar as séries temporais mais preditivas (palavras) $X_t^* \subset X_t$.

O modelo de *elastic net* foi escolhido para realizar a segunda etapa:

$$y_{t+h} = W_t'\boldsymbol{\beta}_h + X_t'\phi_h + \epsilon_{t+h} \tag{1.3}$$

em que $h \geq 0$ é o horizonte de previsão, $\widehat{\beta}_h$ e $\widehat{\phi}_h$ são estimadas minimizando a seguinte função objetivo:

$$\min_{\beta_h,\phi_h} \sum_t \left(y_{t+h} - W_t'\beta_h - X_t'\phi_h\right)^2 + \lambda_1 \left\|\phi_h\right\|_{\ell_1} + \lambda_2 \left\|\phi_h\right\|_{\ell_2} \tag{1.4}$$

em que W_t é um vetor *k x 1* de preditores pré-determinados, como defasagens de y_t bem como preditores tradicionais de dados estruturados e $\|\cdot\|_{\ell_1}$ e $\|\cdot\|_{\ell_2}$ são a norma ℓ_1 e ℓ_2, respectivamente. Então, a partir da seleção das palavras com maior poder preditivo, temos para cada período *t* um conjunto de palavras que servem como dicionário léxico para a obtenção da série de sentimentos *S*. No presente artigo, vamos aplicar esse método na inflação e no PIB.

Por fim, ambas abordagens de dicionário, calculam o índice de sentimento pela diferença entre palavras positivas e negativas, dividida pela soma de palavras positivas e negativas, como foi proposto por Hubert e Labondance (2018):

$$S_t = \frac{\text{PalavrasPositivas}_{\text{ t}} - \text{PalavrasNegativas}_{\text{ t}}}{\text{PalavrasPositivas}_{\text{ t}} + \text{PalavrasNegativas}_{\text{ t}}} \tag{1.5}$$

Portanto, obtemos a medida de sentimentos, *S*, que varia entre -1 e 1. Sendo -1 o tom mais pessimista e 1 o tom mais otimista.

1.3.2 Modelos de Previsão

1.3.2.1 ARMA

Um dos modelos estatísticos mais comuns usados para previsão de séries temporais é o Modelo Autoregressivo de Média Móvel (ARMA), que pressupõe que observações futuras sejam guiadas principalmente por observações recentes. A informação que geralmente exibe comportamento persistente, é amplamente consistente com essa suposição. O melhor modelo para a taxa de inflação y_t, em nossa amostra, em que a representação mais simples é o AR (1), descrito a seguir:

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t \tag{1.6}$$

em que os parâmetros estimados $[\widehat{\alpha}; \widehat{\beta}]'$ podem ser calculados usando uma amostra com t = 1,...,T observações:

$$y_{t+h} = \widehat{\beta}^h y_T + \sum_{i=0}^{h-1} \widehat{\alpha}\widehat{\beta}^i \tag{1.7}$$

em que y_{t+h} é a previsão h-passos à frente.

1.3.3 Modelo de Fatores

A idéia de que variações de tempo em um grande número de variáveis podem ser resumidas por um pequeno número de fatores é empiricamente atraente e é empregada em um grande número de estudos em economia e finanças (FORNI et al., 2000); (STOCK; WATSON, 2002). Considere, $x_{i,t}$ serem os dados observados para o i-ésimo unidade de cross-section no tempo t, para i = 1,...,N e t = 1,...,T, e considere a seguinte representação fatorial dos dados:

$$x_{i,t} = \lambda_i' F_t + e_{i,t} \tag{1.8}$$

em que F_t é um vetor de fatores comuns, λ_i é um vetor de cargas fatoriais associados com F_t e $e_{i,t}$ é o componente indiossincrático de $x_{i,t}$. Note que F_t, λ_i e $e_{i,t}$ são desconhecidos, pois apenas $x_{i,t}$ é observável. Aqui, estimamos os fatores e respectivas cargas usando a Análise de Componentes Principais (PCA), que é uma técnica bem estabelecida para redução de dimensão em séries temporais. O número de componentes é determinado pelo critério de Bai e Ng (2002). Após a estimativa do PCA dos fatores comuns F_t, empregamos a abordagem de previsão direta para modelar a taxa de inflação e o crescimento do PIB no tempo t + h, da seguinte forma:

$$y_{t+h} = \beta_h F_t + \varepsilon_{t+h} \tag{1.9}$$

Portanto, a previsão de informações da abordagem do modelo de fator direto, y_{t+h}, usando uma amostra de t = 1,...,T observações, é dado por:

$$y_{t+h} = \widehat{\beta_h}\widehat{F_T}, \qquad \text{for } h = 1, \ldots, H \tag{1.10}$$

em que $\widehat{\beta_h}$ e $\widehat{F_T}$ são os parâmetros a serem estimados.

1.3.3.1 LASSO

Neste tipo de método de encolhimento a ideia é reduzir os parâmetros que correspondem a variáveis irrelevantes a zero. Sob algumas condições, é possível estimar modelos com mais variáveis que observações.

Entre os métodos de encolhimento, o LASSO, introduzido por Tibshirani (1996) recebeu atenção especial. Foi demonstrado que o LASSO pode lidar com mais variáveis que observações, e o subconjunto correto de variáveis relevantes pode ser selecionado (EFRON et al., 2004); (MEINSHAUSEN et al., 2009); (ZHAO; YU, 2006).

O estimador LASSO é definido como:

$$\hat{\boldsymbol{\beta}} = \arg\min_{\hat{\beta}} \left[\sum_{t=1}^{T} \left(y_{t+h} - \boldsymbol{\beta}' \boldsymbol{x}_t \right)^2 + \lambda \sum_{j=1}^{k} |\beta_j| \right] \tag{1.11}$$

em que λ controla a quantidade de encolhimento e é determinado por técnicas orientadas a dados, como validação cruzada ou uso de critérios de informação. O modelo funciona mesmo quando o número de variáveis aumenta mais rapidamente que o número de observações e quando os erros são não gaussianos e heterocedásticos.

1.3.3.2 Random Forest

A metodologia de *random forest* foi proposta inicialmente por Breiman (2001) como uma maneira de reduzir a variação de árvores de regressão e baseia-se na agregação de *bootstrap* (*bagging*) de árvores de regressão construídas aleatoriamente.

De acordo com Garcia et al. (2017) uma árvore de regressão é um modelo não paramétrico baseado no particionamento binário recursivo do espaço covariável $\mathbb{X}$, em que a função T(·) é uma soma de modelos locais, cada um dos quais é determinado em $K \in \mathbb{N}$ regiões diferentes (partições) de $\mathbb{X}$. O modelo geralmente é exibido em um gráfico que possui o formato de uma árvore de decisão binária com nós $N \in \mathbb{N}$ pai (ou divididos) e nós $K \in \mathbb{N}$ terminais (também chamados de folhas) e que cresce a partir do nó raiz para os nós do terminal. Geralmente, as partições são definidas por um conjunto de hiperplanos, cada qual ortogonal ao eixo de uma determinada variável preditora, chamada de variável dividida. Portanto, condicional ao conhecimento das sub-regiões, a relação entre y_{t+h} e x_t na Equação 1.1 é aproximado por um modelo constante por partes, em que cada folha (ou nó terminal) representa um regime distinto.

Segundo Garcia et al. (2017) é possível representar matematicamente um modelo complexo de árvore de regressão introduzindo a seguinte notação. O nó raiz está na posição 0 e um nó pai na posição *j* gera nós filhos a esquerda e a direita nas posições $2j+1$ e $2j+2$, respectivamente. Todo nó pai possui uma variável dividida associada $x_{s_j t} \in \mathbf{x}_t$, onde $s_j \in \mathrm{S} = \{1, 2, \ldots, q\}$. Além disso, se permitirmos que $\mathbf{J}$ e $\mathbf{T}$ sejam os conjuntos de índices dos nós pais e terminal, respectivamente, uma arquitetura em árvore pode ser determinada completamente a partir de $\mathbf{J}$ e $\mathbf{T}$.

O modelo de previsão baseado em árvores de regressão pode ser representado matematicamente como:

$$y_{t+h} = H_{\mathbb{JT}}\left(\mathbf{x}_t;\psi\right) + u_{t+h} = \sum_{i\in\mathbb{T}} \beta_i B_{\mathbb{J}i}\left(\mathbf{x}_t;\boldsymbol{\theta}_i\right) + u_{t+h} \tag{1.12}$$

em que:

$$\begin{aligned} B_{\mathbb{J}\mathrm{i}}\left(\mathbf{x}_t;\boldsymbol{\theta}_i\right) = &\prod_{j\in\mathbb{J}} I\left(x_{s_j,t};c_j\right)^{\frac{n_{i,j}\left(1+n_{i,j}\right)}{2}} \\ &\times\left[1-I\left(x_{s_j,t};c_j\right)\right]^{(1-n_{i,j})(1+n_{i,j})} \end{aligned} \tag{1.13}$$

$$I\left(x_{s_j,t};c_j\right) = \begin{cases} 1 & \text{if } x_{s_j,t} \leq c_j \\ 0 & \text{caso contrário} \end{cases} \tag{1.14}$$

$$n_{i,j} = \begin{cases} -1 & \text{se o caminho para a folha } i \text{ não inclui o nó pai} \\ 0 & \text{se o caminho para a folha } i \text{ inclui} \\ & \text{o nó filho direito do nó pai } j \\ 1 & \text{se o caminho para a folha } i \text{ inclui} \\ & \text{o nó filho esquerdo do nó pai } j \end{cases}$$

Seja $\mathbb{J}_i$ o subconjunto de $\mathbb{J}$ que contém os índices dos nós pais que formam o caminho para a folha *i*; então, θ_i é o vetor que contém todos os parâmetros c_k, de modo que $k \in \mathbb{J}_i$, $i \in \mathbb{T}$. Observe que $\sum_{j\in\mathbb{J}} B_{\mathbb{J}i}\left(\mathbf{x}_t;\boldsymbol{\theta}_j\right) = 1, \forall \mathbf{x}_t \in \mathbb{R}^{q+1}$.

Uma *random forest* é uma coleção de árvores de regressão, cada uma especificada em uma subamostra de inicialização dos dados originais. Suponha que haja subamostras com inicialização *B* e denote a árvore de regressão estimada para cada uma das subamostras por $H_{\mathbb{J}_b\mathbb{T}_b}\left(\cdot;\psi_b\right)$. A previsão final é definida como:

$$\widehat{y}_{t+h} = \frac{1}{B}\sum_{b=1}^{B} H_{\mathbb{J}_b\mathbb{T}_b}\left(x_t;\psi_b\right) \tag{1.15}$$

As florestas aleatórias podem lidar com um número muito grande de variáveis explicativas, e o modelo previsto é altamente não linear. É importante notar que as amostras de autoinicialização são calculadas usando autoinicializações em bloco, pois estamos lidando com séries temporais.

1.3.3.3 Support Vector Machine

Desde que o SVM foi introduzido a partir da teoria de aprendizagem estatística por Vapnik (1995), uma série de estudos foi anunciada sobre sua teoria e aplicações. Comparado com a maioria das outras técnicas de aprendizado, o SVM leva a aumentar o desempenho em reconhecimento de padrões, estimativa de regressão, previsão de séries temporais financeiras, dentre outras aplicações aplicações. A seguinte breve descrição de SVM se concentra inteiramente

no problema de reconhecimento de padrões no campo de classificação. A explicação detalhada e as provas de SVM podem estar contidas nos livros de (VAPNIK, 1995) e (VAPNIK, 1999).

De acordo com Shin et al. (2005) o SVM produz um classificador binário, os chamados hiperplanos de separação ótimos, por meio do mapeamento extremamente não linear dos vetores de entrada no espaço de recursos de alta dimensão. O SVM constrói um modelo linear para estimar a função de decisão usando limites de classe não lineares com base em vetores de suporte. Se os dados forem separados linearmente, o SVM treina máquinas lineares para um hiperplano ideal que separa os dados sem erro e na distância máxima entre o hiperplano e os pontos de treinamento mais próximos. Os pontos de treinamento mais próximos do hiperplano de separação ideal são chamados de vetores de suporte. Todos os outros exemplos de treinamento são irrelevantes para determinar os limites das classes binárias. Em casos gerais em que os dados não são separados linearmente, o SVM usa máquinas não lineares para encontrar um hiperplano que minimiza o número de erros do conjunto de treinamento.

1.3.4 Acurácia da Previsão

Testes formais de capacidade preditiva podem ser feitos usando abordagens como as popularizadas por Diebold e Mariano (2002) para verificar se as diferenças no *Root Mean Square Deviation* (RMSE) refletem diferenças estatisticamente significativas entre as previsões. Assim, como no trabalho de Clements e Reade (2020), empregamos um teste baseado em regressão Diebold e Mariano (2002), onde construímos o termo e_t^{DM} como:

$$e_t^{DM} = L\left(y_t - \hat{y}_{t|t-h}\right) - L\left(y_t - \hat{y}_{t|t-h}^{EM}\right) \tag{1.16}$$

em que $\hat{y}_{t|t-h}^{EM}$ é a previsão de referência, em nosso caso, a referência são as previsões do Focus. Convencionalmente, a função de perda *L* é a perda de erro ao quadrado, isto é, $L(e) = e^2$. O desempenho igual da previsão implica $E\left(e_t^{DM}\right) = 0$, e esse teste pode ser implementado usando o modelo de regressão:

$$e_t^{DM} = \alpha + u_t \tag{1.17}$$

com hipótse nula de precisão igual, $H_0 : \alpha = 0$. A significância de α implica uma diferença no desempenho da previsão. Se $\alpha > 0$, isso implica que a previsão do *benchmark* é 'melhor' do que a previsão realizada, enquanto $\alpha < 0$ implica o contrário, ou seja, que a previsão realizada é diferente em relação à previsão do *benchmark*.

1.3.5 Eficiência da Previsão

Nessa etapa vamos comparar as projeções realizadas pelo mercado, representado pelo Focus, e as informações qualitativas contidas no texto da ata do Copom e do Relatório de Inflação,

regredindo os erros de previsão do Focus em função dos índices de sentimentos, e vamos repetir o mesmo procedimento as os erros de previsão do Focus para um trimestre à frente. Assim como em Jones et al. (2019), para este exercício de comparação usaremos uma versão extendida de uma regressão de Mincer e Zarnowitz (1969). Se as previsões do Focus contêm todos das informações relevantes incluídas no texto, então os erros de previsão do Focus não devem ser previstos pelos índices de sentimentos. A equação pode ser adaptada da seguinte maneira:

$$e_{t+h}^{Focus} = \alpha + \beta S_t + u_t \tag{1.18}$$

em que e_{t+h}^{Focus} é o erro de previsão do Focus que é definido como a diferença entre os valores realizados e os valores previstos pelo Focus, S_t o índice de sentimento e u_t é o resíduo. Testando a hipótese nula de que o coeficiente do índice de sentimento é igual a zero nos permite determinar se as previsões do Focus foram ou não melhoradas com a incorporação de informações de o texto.

1.4 Resultados

1.4.1 Índices de Sentimento

No total foram construídos três índices a partir da ata do Copom e três provenientes do RI, totalizando seis indicadores de sentimento. Os índices e suas respectivas definições podem ser visualizados na Tabela 3.

Tabela 3 – Definição dos índices de sentimento

Índice de Sentimento	Definição
CM_Index	Índice de sentimento da ata do Copom usando o dicionário fixo
IR_Index	Índice de sentimento do Relatório de Inflação usando o dicionário fixo
IPCA_CM_Index	Índice de sentimento da ata do Copom usando o dicionário variante selecionando as palavras mais preditivas para o IPCA
IPCA_IR_Index	Índice de sentimento do Relatório de Inflação usando o dicionário variante selecionando as palavras mais preditivas para o IPCA
GDP_CM_Index	Índice de sentimento da ata do Copom usando o dicionário variante selecionando as palavras mais preditivas para o PIB
GDP_IR_Index	Índice de sentimento do Relatório de Inflação usando o dicionário variante selecionando as palavras mais preditivas para o PIB

As variáveis CM_Index e IR_Index foram construídas utilizando o dicionário de Loughran e McDonald (2011), portanto são índices de sentimentos advindos de um rol de palavras pré-selecionadas que não sofrem alteração com o tempo. Já os indicadores IPCA_CM_Index e IPCA_IR_Index são criadas por um conjunto de palavras que variam a cada período *t*, ou seja, para cada relatório o dicionário consiste das palavras mais preditivas para o IPCA que são selecionadas via *machine learning*. O mesmo ocorre para os índices GDP_CM_Index e GDP_IR_Index, a diferença reside na variável de resposta, neste caso, o PIB. A principal van-

tagem dos índices de sentimentos provenientes do dicionário variante é que elas são capazes de representar um sentimento mais realista, pois podem mensurar de forma mais precisa o sentimento do documento quando ocorrem evento extremos, como por exemplo, a pandemia de Covid-19.

A Figura 13 e Figura 14 em anexo, ilustram bem esse ponto. É possível ver os 10 coeficientes mais positivos e os 10 mais negativos para o IPCA na ata do Copom e RI. Na ata do Copom para o IPCA as palavras "budget", "certain"e "gas"são os termos mais positivos e "coronavirus", "covid"e "basket"as palavras mais negativas. No RI as três palavras mais positivas são "vaccines", "arabia"e "fuels", já as mais negativas são "election", "cigarretes"e "surprises", mas também aparecem termos que representam cenários de desastre ambiental como "brumadinho"e o cenário atual da Covid-19 como "epidemic". Para o PIB também verificamos que os termos "coronavirus", "covid", "infections"e "vaccines"são os mais preditivos. No caso dos índices de dicionário fixo é impossível essas séries captarem com precisão a pandemia de Covid-19, pois palavras como "covid"e "coronavirus"não estão presentes em tal dicionário.

Quando analisamos o tamanho dos documentos percebemos que eles variam com o tempo, de modo que o número de palavras é decrescente ao longo dos trimestres, como é evidente a partir da Figura 11. A ata do Copom no primeiro trimestre de 2005 continha cerca de 5.328 palavras e no segundo trimestre de 2020 uma quantidade de 1.797 palavras. O número de palavras na ata do Copom foi drásticamente reduzido a partir do ano de 2014. Já o RI no primeiro trimestre de 2005 possuia 57.091 palavras e 36.664 no fim da amostra. Ainda sobre o RI, podemos ver pela Figura 11 que entre os anos de 2012 e 2015 existe uma redução extrema na quantidade de palavras. Como já foi explicado na seção de base de dados, isso ocorreu porque durante esses anos o BCB não produziu o RI, apenas disponibilizou um sumário executivo sobre o documento.

As palavras com maior frequência na ata do Copom e no RI podem ser visualizadas na Figura 12, que mostra a nuvem de palavras de cada documento. Palavras como "growth", "goods", "rate", "credit", "industrial", "exchange", "expectation"e "sales"são os termos mais encontrados em ambos documentos. As nuvens de palavras também ajudam a mostrar que apesar do RI ser um documento com um número de palavras muito superior em relação à ata do Copom, os termos principais que aparecem nesses documentos são similares.

A Tabela 21 mostra as correlações entre os índices de sentimentos com o IPCA e o crescimento do PIB. É possível perceber que os índices de sentimento de dicionário fixo, CM_Index e IR_Index, possuem baixas correlações com as variáveis macroeconômicas. Esse fato não é totalmente anormal, pois trabalhos similares como os de Catalfamo et al. (2018) e Mathy e Stekler (2018) também encontraram baixas correlações com dados dos Estados Unidos.

Quando observamos os índices de sentimento de dicionário variante as correlações aumentam de forma significativa. Também podemos perceber que os índices que usam o crescimento do PIB na sua construção apresentam correlações com o PIB maiores do que a correlação

do IPCA índices que usam o IPCA. Além disso, quando olhamos para os índices de dicionários variante, vemos que os provenientes da ata do Copom produzem correlações mais elevadas.

A maior correlação de todas foi do GDP_CM_Index com o crescimento do PIB (0,75). Esse valor é inferior ao encontrado nos trabalhos de Jones et al. (2019) e Clements e Reade (2020) que fizeram com dados do Banco Central da Inglaterra e usaram um dicionário fixo. Contudo, os autores não usaram o índice de sentimento "bruto", antes de os usar como preditores eles os "melhoraram"para ficarem mais próximos ao PIB e a inflação via dimensionamento e escalonamento. Então, se ele utilizassem os índices de dicionário fixo orginais, eles provavelmente iriam obter correlações muito inferiores.

Outro fato interessante é que os índices de sentimento variante no tempo possuem correlações ainda maiores com o IPCA e o PIB de um trimestre a frente. Isso mostra que o sentimento textual dos relatórios do BCB podem conter informações que já antecipam os movimentos do crescimento do PIB e da inflação.

Figura 1 – Índices de sentimento com dicionário fixo e variante no tempo e IPCA

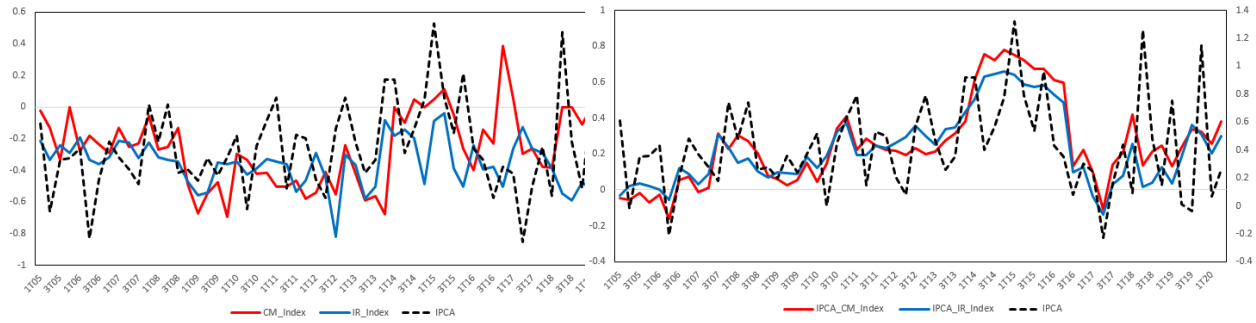
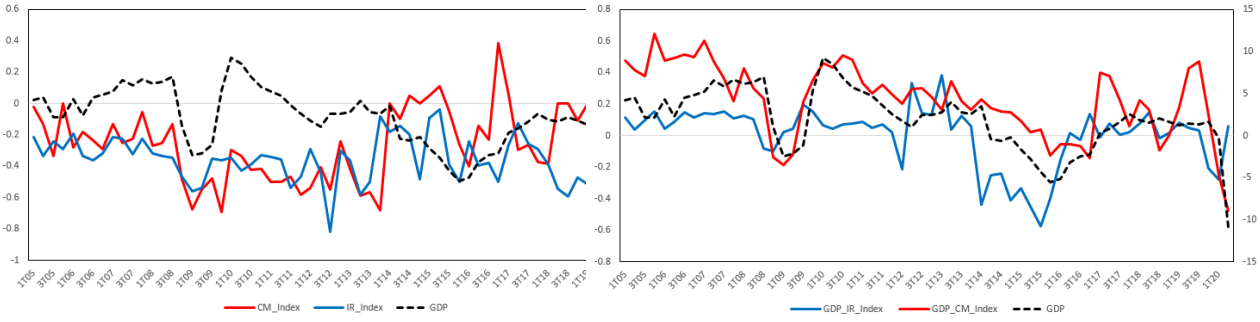


Figura 2 – Índices de sentimento com dicionário fixo e variante no tempo e PIB



A Figura 1 e Figura 2 ilustram graficamente o movimento dos índices de sentimento juntamente com o crescimento do PIB e do IPCA ao longo do tempo. Apesar das escalas diferentes podemos ver que os índices de sentimento de dicionário variante se ajustam melhor as variáveis macroeconômicas. Então, a análise gráfica confirma visualmente o que já foi constatado nas correlações.

Os índices IPCA_CM_Index e IPCA_IR_Index seguem com maior precisão os picos e vales do IPCA em relação aos indicadores CM_Index e IR_Index e como já foi mencionado, devido a baixa correlação entre os índices de sentimento de dicionário fixo. A mesma conclusão pode ser adotada para a análise do crescimento do PIB. Entretanto, CM_Index e IR_Index acompanham mais de perto os moviemntos do PIB do que o IPCA. É visível que esses índices conseguem captar a forte queda no produto devido à crise do Subprime em 2008 e também conseguem acompanhar a queda do PIB por problemas fiscais em 2015, mas falham em captar a abrupta redução do PIB no primeiro e segundo trimestre de 2020 proveniente da pandemia de Covid-19. Esse fato já era esperado, pois o dicionário financeiro de Loughran e McDonald (2011) possui termos capazes de detectar crises financeiras e fiscais, mas são incapazes de mensurar diretamente crises como a do Coronavírus. CM_Index ainda consegue acompanhar a queda do PIB de forma tardia a partir do segundo trimestre de 2020, pois o Coronavírus já tinha atingido de forma significativa a economia brasileira, dessa forma os documentos passaram a ter um sentimento pessimista devido ao fato de termos negativos aparecerem posteriormente com mais frequência.

Quando comparamos os índices de sentimento pela ótica do relatório de origem, vemos que de forma geral as séries de sentimentos provenientes da ata do Copom são mais otimistas em relação as séries do RI, inclusive IR_Index não possui nenhum valor positivo. Além disso, podemos notar que os índices de origem da ata do Copom possuem fortes picos de otimismo após 2014, já os índices de sentimento do RI se mostram mais pessimistas. Isso deve ocorrer devido a função e estrutura distinta entre os documentos, pois a ata do Copom é um documento com natureza mais comunicativa e objetiva, já o RI é um documento mais analítico. Então, como o BCB usa a ata do Copom para comunicar a decisão de política monetária o documento tende a ser mais otimista com o objetivo de ancorar as expectativas do mercado e notamos que a partir de 2014 o tom da ata do Copom passou ter um padrão mais otimista. Os índices construídos por Jones et al. (2019) e Clements e Reade (2020) que usam o RI do Banco Central da Inglaterra também seguem a linha dos nossos índices de origem do RI, ou seja, tendem a ter um tom mais pessimista.

1.4.2 Previsões e Acurácia

Para o nosso exercício de previsão usamos cinco tipo de modelos de previsão: ARMA, Modelo de Fatores, LASSO, *Random Forest* e SVM. O nosso *benchmark* são as previsões médias realizadas pelo mercado, representadas pelo Focus. Com o objetivo verificar se a inclusão dos índices de sentimento em modelos de previsão é capaz de melhorar a acurácia das previsões para o IPCA e o PIB estimamos os modelos multivariados com diferentes categorias. O que vai definir a categoria são os tipos índices de sentimento. Neste caso, a primeira categoria inclui todos os índices de sentimento (com dicionário fixo e dicionário variante), a segunda usa apenas os índices de dicionário variante, a terceira utiliza apenas os índices de dicionário fixo e por fim a quarta categoria é de modelos sem nenhum índice de sentimento. A Tabela 4 a seguir ilustra as

diferentes categorias de modelos.

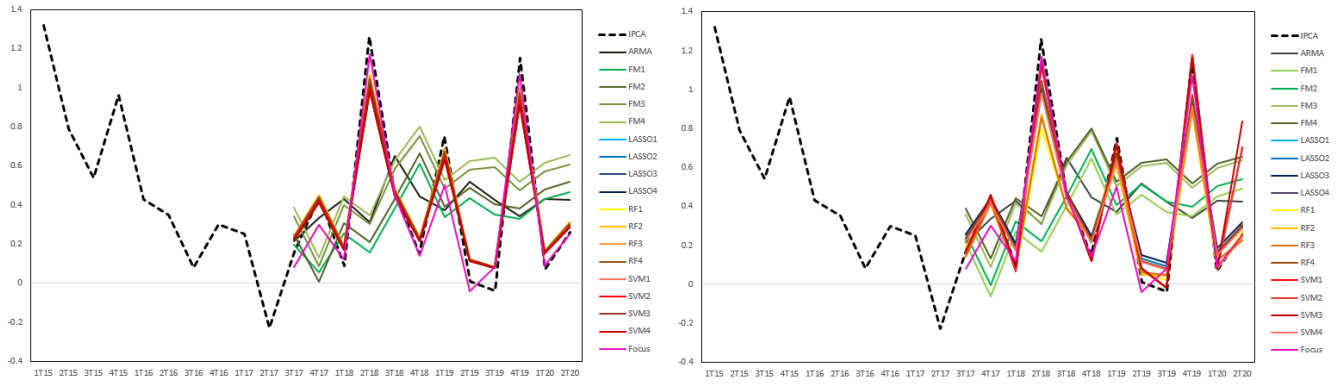
Tabela 4 – Definição das variáveis

Modelo	Definição
LASSO1	Modelo LASSO com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante + índices de sentimento com dicionário fixo
LASSO2	Modelo LASSO com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante
LASSO3	Modelo LASSO com todas as variáveis macroeconômicas + índices de sentimento com dicionário fixo
LASSO4	Modelo LASSO com todas as variáveis macroeconômicas
RF1	Modelo Random Forest com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante + índices de sentimento com dicionário fixo
RF2	Modelo Random Forest com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante
RF3	Modelo Random Forest com todas as variáveis macroeconômicas + índices de sentimento com dicionário fixo
RF4	Modelo Random Forest com todas as variáveis macroeconômicas
SVM1	Modelo Suport Vector Machines com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante + índices de sentimento com dicionário fixo
SVM2	Modelo Suport Vector Machines com todas as variáveis macroeconômicas + índices de sentimento com dicionário variante
SVM3	Modelo Suport Vector Machines com todas as variáveis macroeconômicas + índices de sentimento com dicionário fixo
SVM4	Modelo Suport Vector Machines com todas as variáveis macroeconômicas
FM1	Modelo de Fatores com todas variáveis macroeconômicas + índices de sentimento com dicionário variante + índices de sentimento com dicionário fixo
FM2	Modelo de Fatores com todas variáveis macroeconômicas + índices de sentimento com dicionário variante
FM3	Modelo de Fatores com todas variáveis macroeconômicas + índices de sentimento com dicionário fixo
FM4	Modelo de Fatores com todas variáveis macroeconômicas

A Figura 3 e Figura 4[3] ilustram as previsões do IPCA e o crescimento do PIB, respectivamente, para fora da amostra e considerando h = 1, ou seja, um trimestre a frente. A amostra total vai do primeiro trimestre de 2005 até o segundo trimestre de 2020, totalizando 62 observações. 80% da amostra total foi utilizada para compor a amostra de treinamento dos modelos e os 20% para compor a amostra de validação. Em ambas figuras percebemos que as previsões do Focus ficam muito próximos aos dados realizados e em alguns trimestres os modelos de *machine learning* conseguem superar o Focus.

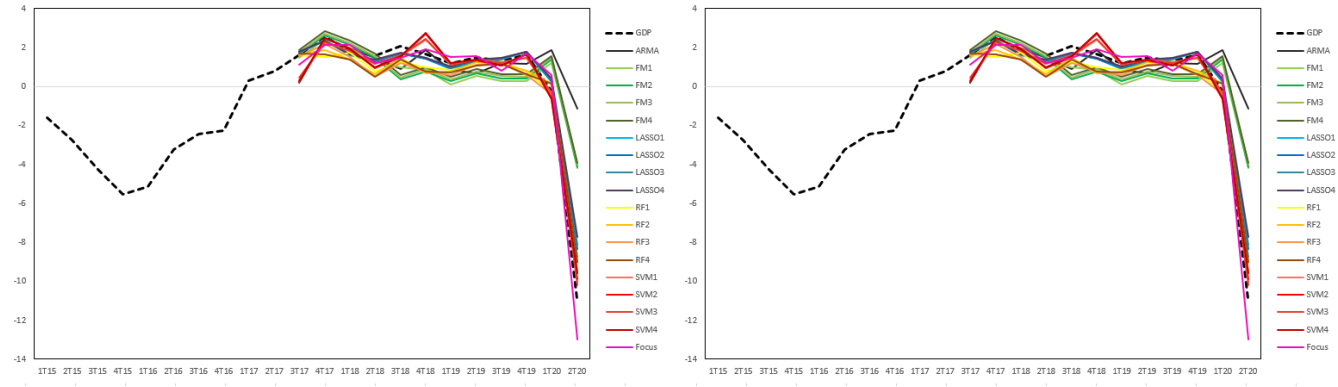
[3] Nas Figura 3 e Figura 4 optamos para fins de melhor visualização exibir as séries a partir do primeiro trimestre de 2015.

Figura 3 – Previsões do IPCA para h = 1 - ata do Copom e Relatório de Inflação



Copom Minutes Inflation Report

Figura 4 – Previsões do PIB para h = 1 - ata do Copom e Relatório de Inflação



Copom Minutes Inflation Report

A Tabela 5 e Tabela 6 ilustram o erro de previsão, medido pelo *Mean Square Deviation* (MSE) de cada modelo para até quatro trimestres à frente, juntamente com o teste de Diebold-Mariano (DM). Neste caso, para o teste de DM consideramos o Focus como referência, de modo que a hipótese nula é de previsão igual ao do Focus e a hipótese alternativa é previsão diferente do Focus.

Foi possível constatar alguns fatos. O primeiro é que assim como reportado nos trabalhos de ((MEDEIROS; MENDES, 2016), (MEDEIROS et al., 2019), (ARAUJO; GAGLIANONE, 2020)) os modelos de *machine learning* apresentaram um desempnho satisfatório em termos de acurácia para um período à frente, pois as previsões ficaram próximas às previsões do Focus e alguns até superando o *benckmark*, entretanto, para mais períodos à frente as previsões já começam a se afastar do Focus. Ou seja, em previsões de curto prazo de até um trimestre à frente, conseguimos nos aproximar e até superar o Focus, mas para trimestres mais à frente nossas previsões ficam longe do Focus. O segundo é que os MSEs reportados para o IPCA são inferiores em relação aos MSEs para o PIB. Então, como já era esperado, tanto para nós como para a média do mercado é mais fácil prever a inflação do que o crescimento do produto.

O terceiro é que os modelos que incluíram os índices de sentimento obtiveram um desempenho melhor do que os modelos que não incorporam as séries de polaridade. Os melhores modelos são os que usam ao mesmo tempo os índices de sentimento de dicionário fixo e dicionário variante (categoria 1) seguidos pelos modelos que consideram apenas os índices de sentimento de dicionário variante. Outro fato é que os modelos que usaram as séries de sentimentos oriundas da ata do Copom obteram um menor MSE, tanto para o IPCA como para o crescimento do PIB.

Tabela 5 – Mean Squared Error (MSE) e teste de Diebold-Mariano para as previsões do IPCA

	Copom				Relatório da Inflação			
Modelo	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
ARMA	0,2147***	0,2404***	0,2833***	0,3222***	0,2147***	0,2404***	0,2833***	0,3222***
FM1	0,2550***	0,2707***	0,2908***	0,3587***	0,2583***	0,2894***	0,3003***	0,3597***
FM2	0,2500***	0,2916***	0,3130***	0,3693***	0,2541***	0,2917***	0,3237***	0,3612***
FM3	0,2649***	0,3034***	0,3317***	0,3713***	0,2741***	0,3372***	0,3904***	0,4481***
FM4	0,2745***	0,3341***	0,3850***	0,4399***	0,2745***	0,3341***	0,3850***	0,4399***
LASSO1	0,0124	0,0130*	0,0142**	0,0151**	0,0147*	0,0176***	0,0195***	0,2310***
LASSO2	0,0126	0,0164***	0,0180***	0,0191***	0,0131	0,0192***	0,0199***	0,2050***
LASSO3	0,0143*	0,0203***	0,0222***	0,0236***	0,0151**	0,0195***	0,0211***	0,0236***
LASSO4	0,0196***	0,0204***	0,0223***	0,0235***	0,0196***	0,0204***	0,0223***	0,0255***
RF1	0,0190***	0,0259***	0,0263***	0,0291***	0,0208***	0,0242***	0,0264***	0,0268***
RF2	0,0202***	0,0267***	0,0279***	0,0317***	0,0218***	0,0257***	0,0279***	0,0345***
RF3	0,0204***	0,0241***	0,0260***	0,0320***	0,0215***	0,0285***	0,0299***	0,0336***
RF4	0,0231***	0,0259***	0,0280***	0,0332***	0,0231**	0,0259***	0,0280***	0,0332***
SVM1	0,0160***	0,0188***	0,0224***	0,0262***	0,0168***	0,0183***	0,0229***	0,0266***
SVM2	0,0156***	0,0161***	0,0199***	0,0230***	0,0175***	0,0189***	0,0231***	0,0280***
SVM3	0,0173***	0,0183**	0,0222***	0,0259***	0,0196***	0,0311***	0,0364***	0,0401***
SVM4	0,0224***	0,0300***	0,0348***	0,0388***	0,0224***	0,0300***	0,0348***	0,0388***
Focus	0,0102	0,0126	0,0130	0,0140	0,0102	0,0126	0,0130	0,0140

Tabela 6 – Mean Squared Error (MSE) e teste de Diebold-Mariano para as previsões do PIB

	Copom				Relatório da Inflação			
Modelo	h = 1	h = 2	h = 3	h = 4	h = 1	h = 2	h = 3	h = 4
ARMA	9,5203***	16,7311***	18,5564***	20,0584***	9,5203***	16,7311***	18,5564***	20,0584***
FM1	4,9172***	5,8518***	6,5147**	6,7372***	5,0423***	5,9473***	6,5682***	6,7749***
FM2	5,0715***	5,8099***	6,7179***	8,1984***	5,1795***	5,9147***	6,7597***	8,2218***
FM3	5,1721***	5,8951***	6,9611***	8,4422***	5,2792***	5,9792***	7,0146***	8,4620***
FM4	5,3333***	6,0396***	7,1901***	8,6615***	5,3333***	6,0396***	7,1901***	8,6615***
LASSO1	0,6455***	0,7017***	0,7613***	0,8332***	0,6614***	0,7864***	0,8207***	1,0336***
LASSO2	0,7778***	1,0629***	1,1532***	1,2622***	0,6844***	0,8448***	0,8991***	1,0093***
LASSO3	0,6594***	0,9342***	1,0136***	1,1094***	0,6962***	0,9621***	1,0517***	1,1483***
LASSO4	0,8564***	1,3453***	1,4595***	1,5975***	0,8564***	1,3453***	1,4595***	1,5975***
RF1	0,4359	0,6776***	0,7161***	0,8406***	0,4708	0,7752***	0,8179***	0,9265***
RF2	0,4591	0,5727***	0,7945***	0,9457***	0,4824	0,6832***	0,7225***	1,1748***
RF3	0,4933	0,6627***	0,7390***	0,8564***	0,5026	0,6584***	0,8925***	1,0644***
RF4	0,5524**	0,6832***	0,7985***	1,1748***	0,5524**	0,6832***	0,7985***	1,1748***
SVM1	0,4404	0,5472**	0,7406***	1,0012***	0,4634	0,5534**	0,7902***	0,9431***
SVM2	0,4892	0,6563***	0,8178***	0,8609***	0,4949	0,7499***	0,8495***	1,0112***
SVM3	0,4939	0,6829***	0,8320***	0,9143***	0,5042	0,7794***	0,9307***	1,1990***
SVM4	0,5361**	0,7001***	0,8917***	1,0632***	0,5361**	0,7001***	0,8917***	1,0632***
Focus	0,4670	0,5469	0,6321	0,7444	0,4670	0,5469	0,6321	0,7444

A Tabela 6 ilustra que para a previsão do IPCA o modelo LASSO se destacou, tanto com os índices de sentimento da ata do Copom como para os do RI. Para a inflação apenas esse modelo conseguiu não rejeitar a hipótese nula do teste de DM de igualdade com o Focus, os demais modelos rejeitaram a hipótese. Já para as previsões do PIB os modelos RF e SVM apresentaram um desempenho melhor, iclusive alguns deles conseguiram superar a acurácia do Focus, tais como o RF1 e RF2 com os índices de sentimento da ata do Copom e SVM1 em

ambos documentos. Todos esses modelos usam séries de polaridade, então, para o crescimento do produto só foi possível superar a média do mercado com a inclusão de índices de sentimento nas previsões.

Apesar de Medeiros et al. (2019) e Araujo e Gaglianone (2020) também terem obtido sucesso em superar o Focus com previsões do IPCA para um período à frente sem usar índices de sentinnto, vale ressaltar que os autores usaram um frequência mensal e o presente artigo usou uma frequência trimestral. Além disso, como foi constatado no presente trabalho, a inclusão de informações textuais da ata do Copom e do Relatório de Inflação só foi fundamental para superar o Focus nas previsões do crescimento do PIB. Para as previsões do IPCA as informações textuais foram capazes de melhorar o desempenho dos nossos modelos, mas não possibilitou ganhar da média do mercado. Esse fato não é estranho, pois como já mencionamos, é mais fácil obter menores erros de previsão com a inflação do que com o PIB, então, com o IPCA existe uma menor margem para a melhoria no desempenho com o uso de índices de sentimento.

Adicionalmente, verificamos se os melhores modelos de previsão ilustrados anteriormente são diferentes estatisticamente entre si. Então, como pode ser visto na Tabela 22 e Tabela 23, realizamos novamente o teste de DM para os três melhores modelos de previsão para cada horizonte de previsão. De forma geral é possível ver que a hipótese nula de igualdade de acurária entre os modelos não pode ser rejeitada. Então, não existe diferença estatística no desempenho dos melhores modelos de cada horizonte, relatório e variável macroeconômica.

1.4.3 Regressões de Eficiência

Comparamos as previsões da média do mercado para o IPCA e para o PIB com as informações qualitativas provenientes dos textos da ata do Copom e do RI, regredindo os erros de previsão de curto prazo do Focus nos índices de sentimento e os erros de previsão do Focus de um trimestre à frente com os índices de sentimento, usando uma versão estentidida de uam regressão de Mincer e Zarnowitz (1969). Se as previsões do Focus contiverem todas as informações relevantes que estão incluídas no texto da ata do Copom e do RI, então os erros de previsão não devem ser previstos pelas índices de sentimento dos textos.

Tabela 7 – Regressão de Mincer–Zarnowitz para a inflação

Erro de previsão em (t)				
	Copom_Index	IR_Index	IPCA_CM_Index	IPCA_IR_Index
Constante	-0,024	-0,037	-0,005	-0,015
	(0,016)	(0,030)	(0,017)	(0,018)
Coeficiente do	0,054	0,070	0,024	0,015
Índice de Sentimento	(0,046)	(0,077)	(0,050)	(0,059)
R^2	0,023	0,013	0,040	0,011
Erro de previsão em (t+1)				
	Copom_Index	IR_Index	IPCA_CM_Index	IPCA_IR_Index
Constante	-0,013	-0,016	-0,009	-0,013
	(0,016)	(0,031)	(0,017)	(0,018)
Coeficiente do	0,005	0,012	0,011	0,006
Índice de Sentimento	(0,047)	(0,080)	(0,051)	(0,060)
R^2	0,018	0,038	0,084	0,165

Tabela 8 – Regressão de Mincer–Zarnowitz para o PIB

Erro de previsão em (t)				
	CM_Index	IR_Index	GDP_CM_Index	GDP_IR_Index
Constante	-0,243	-0,497	-0,088	-0,388**
	(0,186)	(0,361)	(0,181)	(0,130)
Coeficiente do	0,624	0,281	1,367*	1,577**
Índice de Sentimento	(0,533)	(0,908)	0,564	(0,553)
R^2	0,022	0,016	0,089	0,079
Erro de previsão em (t+1)				
	CM_Index	IR_Index	GDP_CM_Index	GDP_IR_Index
Constante	-0,330	-0,368	-0,079	-0,389**
	(0,189)	(0,369)	(0,198)	(0,122)
Coeficiente do	0,274	0,076	1,345*	2,604***
Índice de Sentimento	(0,545)	(0,940)	(0,623)	(0,650)
R^2	0,043	0,011	0,073	0,214

Os resultados das regressões mostram que todos os coeficientes dos índices de sentimento não rejeitam a hipótese nula de que são zero para os erros de previsão do IPCA, ou seja, as informações textuais da ata do Copom e do RI não afetam os erros de previsão do Focus para a inflação. Então, não há evidências de que as informações textuais desse documentos possam melhorar as previsões realizadas pelo Focus para a inflação. O mesmo foi constatado para as previsões do IPCA pelo Focus para um trimestre à frente. Esses resultados estão em conformidade ao que foi encontrado por Clements e Reade (2020) que também não encontraram um coeficiente estatisticamente significativo do índice de sentimento na explicação dos erros de previsão da inflação do Banco Central da Inglaterra.

Já para as previsões do crescimento do PIB encontramos eviências de que as informações textuais contêm informações relevantes que podem melhorar as previsões do Focus, pois os erros de previsão em tempo real e para um trimestre à frente do PIB são explicados pelos índices de sentimento. Contudo, esse resultado só é válido para os índices de sentimento de dicionário variante (GDP_CM_Index e GDP_IR_Index) e os maiores coeficientes e mais significativos foram os do GDP_IR_Index. Jones et al. (2019) encontraram significância estatística do coeficiente do índice de sentimento apenas para os erros previsão em tempo real, já os erros de previsão para um trimestre à frente do crescimento do produto não apresentaram relação com o índice de sentimento. Já Clements e Reade (2020) obtiveram resultados similares em termos de significância estatística do coeficiente do índice de sentimento, tanto para o erro de previsão no mesmo período da polaridade como para os erros de previsão em *t+1*. Assim como no presente artigo, Clements e Reade (2020) encontraram sinais positivos dos coeficientes, enquanto Jones et al. (2019) estimaram sinais negativos.

Como os erros de previsão são definidos como valores realizados menos previstos, isso significa que as previsões do Focus ao longo deste período foram excessivamente pessimistas, e que o tom mais positivo do texto é uma avaliação mais precisa que proporciona uma redução nos erros de previsão. Portanto, esses resultados mostram que levar em consideração as informações textuais da ata do Copom e do RI são úteis para obter uma imagem mais completa das condições econômicas atuais e futuras.

1.5 Considerações Finais

A proposta do presente artigo foi de verificar se as informações textuais contidas na ata do Copom e no Relatório de Inflação publicados pelo Banco Central do Brasil são capazes de melhorar as previsões de variáveis macroeconômicas, especificamente, a taxa de inflação e o crescimento do PIB.

A partir dos resultados obtidos, podemos concluir alguns pontos. Constatamos que o uso de uma abordagem de análise textual via *machine learning* para selecionar as palavras que vão compor o dicionário, produz índices de sentimentos que captam de forma mais realista o sentimento existente nos textos das publicações do BCB, além de produzir índices mais correlacionados com as variáveis macroeconômicas. As correlações entre os índices de sentimento e as variáveis macroeconômicas revelam que os indicadores de sentimento, neste caso os de dicionário variante, que além de apresentarem correlações elevadas em tempo real, também possuem uma elevada correlação com as variáveis macroeconômicas para um trimestre à frente. Isso revela que as informações textuais dos documentos do BCB podem conter informações relevantes sobre o futuro do IPCA e, principalmente, do crescimento do PIB.

Também foi possível verificar que os modelos de *machine learning* proporcionam um elevado grau de acurácia em relação os modelos tradicionais como ARMA e Modelo de Fatores. Outro ponto interessante é que os modelos que incorporaram índices de sentimentos como preditores obteram os menores MSEs e que os melhores modelos foram os que usaram ao mesmo tempo os índices de dicionário variante e os dicionário fixo, seguidos pelos modelos que utilizaram os índices provenientes de dicionário variante e os piores modelos foram os que não consideraram nenhuma variável de sentimento no seu conjunto de preditores. Além disso, o teste de Diebold-Mariano ilustrou que as previsões de curto prazo para um trimestre à frente dos modelos de machine learning ficaram próximas às previsões realizadas pelo Focus, e para o crescimento do PIB, algumas de nossas previsões foram capazes de superar o Focus.

Um resultado relevante que foi obtido foi que os índices de sentimento de dicionário variante oriundos tanto da ata do Copom como do RI conseguem explicar os erros de previsão do Focus em tempo real para o crescimento do PIB. O mesmo foi encontrado para os erros de previsão para um trimestre à frente. Esse fato significa que o texto da ata do Compom e do RI podem reduzir o erro de previsão do Focus e, neste caso, melhorar as previsões do Focus. Essa análise também aponta que o coeficiente dos índices de sentimento foi positivo, indicando que as previsões do Focus para o PIB foram excessivamente pessimistar durante a janela temporal da amostra. Já para o IPCA os mesmo resultado não foi encontrado pois os coeficientes estimandos não são estatisticamente significativos. Esses resultados fazem sentido, pois o Focus apresentou menor MSE nas previsões para o IPCA e maiores para o PIB. Além disso, só conseguimos superar o Focus com o uso de índices de sentimento no crescimento do PIB.

Por fim, podemos conluir que as informações textuais da ata do Copom e do RI conseguiram melhorar as previsões de variáveis macroeconômicas, principalmente previsões para o PIB e

que é importante o mercado passe a considerar tais informações textuais em suas previsões de curto prazo. Além disso, tais informações podem ser úteis para mostrar as condições futuras da economia.

Entretanto, o presente trabalho ainda pode ser considerado em estágio inicial. Dessa forma, o artigo possui algumas limitações. Uma delas é deixar de fora do exercício de previsão alguns algoritmos de *machine learning* mais recentes, principalmente, o LASSO adaptativo. Também foram deixados de fora alguns modelos multivariados relevantes na literatura, como os modelos *threshold* e o modelo Midas. Outra limitação advém do uso de apenas um tipo de dicionário variante no tempo, pois existem outros métodos importantes, como o VADER e os dicionários de Shapiro e Wilson (2019) que usam ML para a construção do dicionário de palavras. Um exercício interessante que poderá ser realizado no futuro é a comparação do desempenho das predições antes e depois do período da pandemia de Covid-19.

Parte II

Risco Bancário

2 *Machine learning* e Análise de Sentimento: Projetando o Risco de Insolvência Bancária

2.1 Introdução

A falência de uma instituição bancária pode gerar efeitos devastadores em uma economia e no sistema financeiro de um país. Atualmente, vários tipos de modelos de previsão são empregados com o objetivo de prever o risco de falência dos bancos, fornecendo informações aos reguladores para que estes sejam capazes de tomar alguma decisão de forma antecipada, evitando ou minimizando os efeitos negativos sobre o resto do sistema financeiro.

Dentro deste contexto, este trabalho pretende construir uma nova métrica de classificação de risco de falência dos principais bancos de capital aberto negociados na Brasil, Bolsa, Balcão (B3). Essa nova medida de risco bancário será comparada com o Z-score que é uma métrica tradicional e amplamente utilizada na literatura. Após a construção da variável de risco, via técnicas de clusterização, esta será projetada por um conjunto de modelos de predição com o intuito de saber qual deles oferece a melhor acurácia. Além disso, o trabalho também busca verificar se o sentimento do gestor da instituição bancária é uma variável relevante na predição da nova métrica construída.

A importância deste trabalho está ligada ao fato que crises financeiras sempre têm consequências catastróficas, em especial, a crise *Subprime*, iniciada em 2008 com o estouro da bolha do mercado imobiliário americano. Essa crise gerou múltiplas consequências na economia global, mostrando, entre outras questões, que os problemas financeiros das instituições bancárias vão além dos problemas sociais e econômicos e são capazes de afetar agentes em todo o mundo.

Barbosa (2017) diz que diante da forte repercussão financeira adversa gerada, comportamentos de insolvência bancária passaram a ganhar cada vez mais destaque na literatura, uma vez que tanto os investidores como os donos de depósitos tendem a perder a confiança nessas instituições inadimplentes, o que pode contaminar os demais bancos presentes no mercado e, no longo prazo, resultar em uma crise bancária. Essa última denota em consequências ainda mais severas que vão desde a paralisação da oferta de crédito para empresas e famílias até a fuga de capitais daquele país.

De acordo com Lepetit e Strobel (2013) a insolvência de uma instituição bancária ocorre quando as perdas incorridas por essa instituição não podem ser cobertas pelos seus recursos próprios. A partir desse fato a literatura tem desenvolvido medidas que buscam mensurar o risco

de insolvência, entre essas, destacam-se, sistema CAMELS[1] e o Z-score, em que o último indica a distância em que o banco se encontra de um comportamento de insolvência (Suss e Treitel (2019), Vieira et al. (2020), Viswanathan et al. (2020)).

Além de mensurar o risco bancário, a literatura também passou a se preocupar em prever e antecipar a falência destas instituições. Um conjunto de técnicas foram desenvolvidas ao longo dos anos na tentativa de fornecer aos analistas e tomadores de decisão métodos eficazes de previsão do risco de falência bancário com base em vários índices financeiros e modelos matemáticos, com esses modelos incluindo regressões lineares e logísticas, *splines* de regressão adaptativa multivariada, análise de sobrevivência, programação linear e quadrática e programação de critérios múltiplos, conforme visto em Karels e Prakash (1987), Ezzamel et al. (1987) e Ravi et al. (2008). Segundo Huang e Yen (2019) muitas dessas técnicas são tipicamente baseadas nas suposições de separabilidade linear e normalidade multivariada e, de fato, na independência das variáveis explicativas. No entanto, essas condições são frequentemente violadas em situações da vida real.

Com o aumento expressivo do número de dados e informações alguns autores começaram a empregar técnicas de *machine learning* (ML) para a predição do risco bancário[2]. De acordo com Huang e Yen (2019) as técnicas de ML têm a capacidade de extrair informações significativas de dados não estruturados, ao mesmo tempo que lidam com a não linearidade de maneira eficaz. No entanto, a aplicação de técnicas avançadas de ML à previsão financeira ainda é uma área relativamente nova para os pesquisadores explorarem.

Paule-Vianez et al. (2019) afirmam que as variáveis utilizadas na maioria dos estudos para prever o risco de falência das intituições financeiras têm sido os índices financeiros, especialmente os índices classificados em capital, ativos, gestão, resultados, liquidez e sensibilidade (Sistema CAMELS) e algumas variáveis econômicas[3]. Outros autores buscaram incluir novas variáveis na explicação do risco de insolvência bancário. Uma delas é o sentimento que o gestor da instituição bancária transmite por meio dos relatórios trimestrais e comunicações ao mercado. A ideia é captar por meio do sentimento textual, uma relação direta entre o risco de insolvência e o tom pessimista, buscando aumentar a capacidade preditiva dos modelos.

Essa discussão pode ser encontrada em Gupta et al. (2016), em que os autores ressaltam que o sentimento textual é capaz de predizer a falência dos bancos de capital aberto, considerando ainda que o tom textual otimista na comunicação dos gestores tem um maior poder preditivo para essas instituições, identificando que os bancos insolventes apresentam sentimentos mais positivos do que os seus pares não falidos.

A literatura de previsão estabeleceu o valor da análise textual, bem como uma metodolo-

[1] CAMELS significa "Adequação de capital, qualidade de ativos, gerenciamento, ganhos, liquidez e sensibilidade".

[2] Ver os trabalhos de Sun e Li (2012), Erdogan (2013), Kim et al. (2016), Chou et al. (2017), Xia et al. (2017), Hsu (2019), Suss e Treitel (2019), dentre outros.

[3] Ver Cole e Gunther (1998), González-Hermosillo (1999), Kumar e Ravi (2007), Curry et al. (2007), Rosa e Gartner (2017), Constantin et al. (2018), dentre outros.

gia geral para converter texto em *scores* quantitativos que avaliam principalmente as polaridades dos textos. De acordo com Gentzkow et al. (2019), as informações codificadas no texto são um complemento rico para os tipos de dados mais estruturados tradicionalmente usados na pesquisa empírica. De fato, nos últimos anos, ocorreu um uso intenso de dados textuais em diferentes áreas de pesquisa.

Assim, este trabalho busca contribuir com a literatura descrita acima em alguns pontos, sendo eles: o primeiro é a construção de uma nova métrica de risco de insolvência bancária a partir do agrupamento de *cluster* por meio da técnica *k-means* que consiste em um método de ML não supervisionado e que permite classificar uma base de dados através de agrupamentos que minimizam o erro quadrado. Isso nos permite modelar o risco de insolvência em vez de falência total. Essa abordagem tem uma série de vantagens principais, sendo a mais importante o alinhamento às necessidades práticas dos órgãos reguladores que procuram intervir muito antes do fracasso, ou seja, da falência total do banco.

Em segundo lugar, o trabalho contribui com a literatura, indo além das técnicas de modelagem convencionais, utilizando métodos da literatura de ML ao lado de abordagens mais tradicionais. De acordo com Suss e Treitel (2019) abordagens convencionais, como modelos de regressão logística, são incapazes de levar em conta interações complexas e não linearidades, tendendo a ter um desempenho pior do que suas contrapartes de aprendizado de máquina mais flexíveis. Neste artigo, comparamos a regressão logística (LOGIT), com cinco modelos de ML: *Naive Bayes* (NB), *Random Forest* (RF), *Adaptive Boosting* ou AdaBoost (AB), *Support Vector Machines* (SVM) e *Decision Trees* (DT).

A terceira é verificar se o tom textual dos relatórios trimestrais dos bancos é capaz de melhorar a acurácia na previsão do risco de falência bancária. Além disso, o presente trabalho utiliza-se de um dicionário variante no tempo na construção da variável de sentimento bancário. Até o momento na literatura de previsão de risco bancário, os escassos trabalhos que aplicam alguma variável de sentimento bancário utilizam um dicionário fixo. Portanto, o uso de um dicionário variante no tempo é algo inédito nesta discussão.

Os resultados indicam que a classificação de risco bancária foi capaz de classificar 17% da amostra no grupo de maior risco (1), enquanto 83% da amostra ficou no grupo de menor risco de falência (0). Utilizando a métrica do Z-score verificamos que 65% da amostra faz parte do grupo de baixo risco e 35% da amostra no grupo de elevado risco. Desse modo, o algoritmo *k-means* é mais rigoroso em classificar um banco na categoria de alto risco. Na sequência utilizamos os dados já descritos para projetar o risco de insolvência bancária. Os resultados desta etapa mostraram que o modelo de árvore de decisão apresentou o melhor desempenho para a amostra de teste. Além disso, constatou-se que a inclusão de variáveis de sentimento bancário é capaz de melhorar o desempenho dos modelos de previsão, principalmente, quando o sentimento bancário é construído a partir de um dicionário variante no tempo.

O presente artigo é dividido em quatro seções. A primeira é a introdução apresentada

acima. A segunda apresenta a metodologia aplicada para a construção da variável de risco de insolvência bancário e das variáveis de sentimento bancário. Nessa seção também são mostrados os modelos de previsão. A terceira ilustra os principais resultados obtidos. Por fim, a quarta indica as conclusões finais, discutindo as principais contribuições e limitações do atigo.

2.2 Metodologia

A estratégia empírica da pesquisa foi realizada em quatro etapas: a primeira foi usar o algoritmo *k-means* para realizar os agrupamentos (*clusters*) das instituições bancárias utilizadas na amostra e assim criar uma nova forma de classificação de risco de insolvência bancária. Após a construção da série de *cluster*, esta passou a ser considerada a variável dependente.

O segundo passo foi a construção das métricas de sentimento bancário. Uma variável foi criada a partir de um dicionário fixo e a outra por meio de um dicionário variante no tempo, com o intuito de verificar se o sentimento do gestor da instituição financeira contido nos relatórios trimestrais é capaz de melhorar a predição do risco de insolvência. A terceira etapa foi empregar técnicas estatísticas de ML supervisionada para prever a classificação das instituições bancárias. A ideia é identificar o modelo preditivo mais robusto para projetar a variável de risco de insolvência construída. Por fim, foram calculadas as acurácias e as taxas de falsos negativos e falsos positivos dos modelos.

2.2.1 Risco de Insolvência Bancária

A construção da nova variável usada como *proxy* para o risco de insolvência bancária foi realizada por meio da clusterização dos dados dos bancos escolhidos para o estudo. Para compor a amostra, foram escolhidos 12 bancos de capital aberto com ações negociadas na B3 que podem ser visualizados na Tabela 9.

Tabela 9 – Lista dos bancos

Instituição Bancária	Definição
Banco ABC	ABC
Banrisul	BANRS
Banco do Brasil	BB
Bradesco	BRA
Banco BRB	BRB
BTG Pactual	BTG
Banco Indusval	IND
Itáu	ITA
Banco Mercantil	MER
Banco Panamericano	PAN
Banco Pine	PIN
Santander	SAN

A janela temporal tem início no quarto trimestre de 2012 e termina no primeiro trimestre de 2021. A frequência dos dados é trimestral. Assim como no trabalho de Damasceno et al. (2021), a variável escolhida para a criação do *cluster* foi o desvio padrão móvel de doze trimestres do retorno sobre os ativos da empresa (σROA) que é usado no cálculo do Z-score. Além disso, o novo indicador de risco será comparado com o Z-score.

Similarmente a pesquisa de Vieira et al. (2020), o Z-score foi mensurado através das médias móveis e do desvio padrão móvel considerando doze trimestres (três anos), em que os valores dos onze trimestres anteriores e o do trimestre contemporâneo foram utilizados para essa mensuração, conforme apresentado na equação abaixo:

$$Z_{score\{i,t\}} = \frac{\mu ROA_{i,t} + \mu ETS_{i,t}}{\sigma ROA_{i,t}} \tag{2.1}$$

Em que:
$Z_{score\{i,t\}}$ = risco de insolvência do banco i, no período t;
$\mu ROA_{i,t}$ = média móvel de doze trimestres do retorno sobre os ativos do banco i, no periodo t;
$\mu ETS_{i,t}$ = média móvel de doze trimestres da razão entre o patrimônio líquido e o ativo total para a firma i, no periodo t;
$\sigma ROA_{i,t}$ = desvio padrão móvel de doze trimestres do retorno sobre os ativos da empresa i, no periodo t.

Com o intuito de categorizar os bancos que apresentam um maior risco de insolvência foi utilizado como ponto de corte o primeiro quartil da variável Z-score, o qual foi calculado para cada trimestre analisado. A escolha desse limiar se deu uma vez que os valores menores do Z-score denotam bancos que apresentam maior probabilidade de insolvência, assim os bancos identificados no primeiro quartil receberam o valor 1, já os demais foram categorizados com o valor 0.

Para a criação da nova medida de risco, primeiramente, as variáveis financeiras dos bancos sofreram um processo de normalização com o intuito de melhorar a clusterização. Em seguida, foi utilizado o algoritmo *k-means*, em que foi definida a existência de dois *clusters*, um para os bancos com alta probabilidade de falência (1) e outro para os bancos com baixa probabilidade de falência (0). Além disso, diferentemente do trabalho de Damasceno et al. (2021), na presente pesquisa optou-se por clusterizar os bancos um trimestre por vez e não calcular os *clusters* com todos os trimestres de uma só vez. Esse caminho foi escolhido, pois caso contrário a clusterização pelo *k-means* ficaria distorcida.

Após a obtenção da variável de risco, ela foi comparada com a categorização do Z-score mencionada no parágrafo anterior e, posteriormente, utilizada como variável dependente dos modelos de predição.

2.2.2 Agrupamento por cluster - K-means

De acordo com Varella e Quadrelli (2017) o algoritmo *k-means* usa uma forma simples e fácil de classificar um conjunto de dados por meio de um número de *cluster*, que deve ser fixado antes da execução do algoritmo.

Segundo Fortuna e Maturo (2019) o conceito principal é definir k centroides, um para cada cluster pré-definido, esses centroides devem ser definidos por conta da separação entre os resultados. Desta forma a escolha deve ser feita para colocá-los distantes um dos outros. A seguir deve-se fazer com que cada ponto pertença a um conjunto de dados e vinculá-lo ao centroide mais próximo. O autor ainda informa que o algoritmo *k-means* tem como objetivo minimizar uma função objetivo, neste caso específico uma função quadrática de erro que pode ser vista na equação abaixo.

$$J = \sum_{j=1}^{k} \sum_{i=1}^{x} \left\| x^{(j)}{}_i - c_j \right\|^2 \tag{2.2}$$

O processo termina sempre com um resultado de separação, embora este não necessariamente seja um resultado ótimo. O algoritmo *k-means* é sensível aos centros de agrupamento selecionados aleatoriamente, por conta disso, é necessária sua definição assertiva e que deve ser tomada como uma premissa.

2.2.3 Logit

Provencher et al. (2002) diz que a regressão logística é uma das técnicas mais utilizadas para a área de análise de risco de crédito. Apresenta como característica que a difere da regressão linear discriminante a possibilidade de identificação de crescimento não linear do *default* sob um formato de uma função sigmoide, com crescimento acelerado, gerando maior acurácia preditiva em muitos casos. Para esses autores, este método é bastante utilizado em situações em que a variável dependente assume valores dicotômicos, como é o caso dos problemas de classificação de risco de uma empresa. De acordo com Silva et al. (2016) a sua execução consiste em estimar a probabilidade de ocorrência de um evento com base em um conjunto de variáveis. Em sua forma funcional a regressão logística pode ser representada por:

$$P_a = \frac{1}{1 + e^{-(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n)}} \tag{2.3}$$

Em que P_a representa a probabilidade de uma entidade assumir um dado valor (normalmente expresso em termos de uma variável dicotômica), $\beta_1 \ldots \beta_n$, representam os coeficientes das variáveis (ou características) e $\mathrm{x}_1 \ldots \mathrm{x}_n$ são as variáveis explicativas ou características.

2.2.4 Naive Bayes

Silva et al. (2016) fala que dentre os modelos supervisionados de ML, as redes bayesianas consistem em um dos métodos mais utilizados e que se configuram como uma classe de modelos estatísticos que apresentam resultados significativos para lidar com eventos de elevada incerteza, sendo aplicado o Teorema de Bayes para identificar comportamentos relacionados à dependência probabilística condicional.

O autor ainda mostra que especificamente esse modelo utiliza um grupo de variáveis aleatórias (atributos) retratadas em um grafo estatístico por meio de nós e arcos, os quais são definidos em função de uma relação de precedência (condicional), ou seja, refletem a probabilidade de ocorrência de um evento de interesse em estudo, em virtude da ocorrência de um outro evento tido como condicional, obtendo dessa forma a tabela de probabilidade condicional. Por meio dessa abordagem, tem-se o algoritmo de classificação *Naive Bayes*. Assim, durante a etapa de treinamento, o algoritmo faz uso dos dados dessa subamostra para compreender quais os valores condicionais das variáveis independentes (atributos) que estão associadas às classes do modelo, para que em uma segunda etapa denominada de validação, utilizando os dados da base de teste, seja possível realizar predições sobre a classe de cada observação na amostra, fazendo uso dos valores das variáveis preditoras identificados pelo modelo na etapa de treinamento.

2.2.5 Decision Trees e Random Forest

A metodologia de *random forest* foi proposta inicialmente por Breiman (2001) como uma maneira de reduzir a variação de árvores de regressão e baseia-se na agregação de *bootstrap* (*bagging*) de árvores de regressão construídas aleatoriamente.

De acordo com Choubin et al. (2018), uma árvore de decisão pode ser considerada um aprendiz básico no campo de ML. A principal vantagem das árvores de decisão em relação aos modelos de regressão linear é que, no caso de um relacionamento altamente não linear e complexo entre os recursos e a resposta, as árvores de decisão podem superar as abordagens clássicas. Embora as árvores de decisão possam não ser muito robustas e geralmente possam fornecer menos precisão preditiva do que alguns dos outros métodos de regressão, essas desvantagens podem ser facilmente melhoradas agregando muitas árvores de decisão, usando métodos, como *bagging* e *random forest*. Esses métodos têm em comum que podem ser considerados métodos de aprendizagem por conjunto.

Segundo Garcia et al. (2017) é possível representar matematicamente um modelo complexo de árvore de regressão introduzindo a seguinte notação. O nó raiz está na posição 0 e um nó pai na posição *j* gera nós filhos a esquerda e a direita nas posições $2j+1$ e $2j+2$, respectivamente. Todo nó pai possui uma variável dividida associada $x_{s_j t} \in \mathbf{x}_t$, onde $s_j \in \mathrm{S} = \{1, 2, \ldots, q\}$. Além disso, se permitirmos que $\mathbf{J}$ e $\mathbf{T}$ sejam os conjuntos de índices dos nós pais e terminal, respectivamente, uma arquitetura em árvore pode ser determinada completamente a partir de $\mathbf{J}$ e

$\mathbf{T}$.

O modelo de previsão baseado em árvores de regressão pode ser representado matematicamente como:

$$y_{t+h} = H_{\mathbb{JT}}\left(\mathbf{x}_t; \psi\right) + u_{t+h} = \sum_{i \in \mathbb{T}} \beta_i B_{\mathbb{J}i}\left(\mathbf{x}_t; \boldsymbol{\theta}_i\right) + u_{t+h} \tag{2.4}$$

em que:

$$\begin{aligned} B_{\mathbb{J}\mathrm{i}}\left(\mathbf{x}_t; \boldsymbol{\theta}_i\right) = & \prod_{j \in \mathbb{J}} I\left(x_{s_j,t}; c_j\right)^{\frac{n_{i,j}\left(1+n_{i,j}\right)}{2}} \\ & \times \left[1 - I\left(x_{s_j,t}; c_j\right)\right]^{(1-n_{i,j})(1+n_{i,j})} \end{aligned} \tag{2.5}$$

$$I\left(x_{s_j,t}; c_j\right) = \begin{cases} 1 & \text{if } x_{s_j,t} \leq c_j \\ 0 & \text{caso contrário} \end{cases} \tag{2.6}$$

$$n_{i,j} = \begin{cases} -1 & \text{se o caminho para a folha } i \text{ não inclui o nó pai} \\ 0 & \text{se o caminho para a folha } i \text{ inclui} \\ & \text{o nó filho direito do nó pai } j \\ 1 & \text{se o caminho para a folha } i \text{ inclui} \\ & \text{o nó filho esquerdo do nó pai } j \end{cases}$$

Seja $\mathbb{J}_i$ o subconjunto de $\mathbb{J}$ que contém os índices dos nós pais que formam o caminho para a folha *i*; então, θ_i é o vetor que contém todos os parâmetros c_k, de modo que $k \in \mathbb{J}_i$, $i \in \mathbb{T}$. Observe que $\sum_{j \in \mathbb{J}} B_{\mathbb{J}i}\left(\mathbf{x}_t; \boldsymbol{\theta}_j\right) = 1, \forall \mathbf{x}_t \in \mathbb{R}^{q+1}$.

Uma *random forest* é uma coleção de árvores de regressão, cada uma especificada em uma subamostra de inicialização dos dados originais. Suponha que haja subamostras com inicialização *B* e denote a árvore de regressão estimada para cada uma das subamostras por $H_{\mathbb{J}_b\mathbb{T}_b}\left(\cdot; \psi_b\right)$. A previsão final é definida como:

$$\widehat{y}_{t+h} = \frac{1}{B} \sum_{b=1}^{B} H_{\mathbb{J}_b\mathbb{T}_b}\left(x_t; \psi_b\right) \tag{2.7}$$

As florestas aleatórias podem lidar com um número muito grande de variáveis explicativas, e o modelo previsto é altamente não linear. É importante notar que as amostras de autoinicialização são calculadas usando autoinicializações em bloco, pois estamos lidando com séries temporais.

2.2.6 Support Vector Machine (SVM)

Desde que o SVM foi introduzido a partir da teoria de aprendizagem estatística por Vapnik (1995), uma série de estudos foi anunciada sobre sua teoria e aplicações. Comparado

com a maioria das outras técnicas de ML, o SVM aumenta o desempenho em reconhecimento de padrões, estimativa de regressão, previsão de séries temporais financeiras, dentre outras aplicações. Ressalta-se que a breve descrição de SVM se concentra inteiramente no problema de reconhecimento de padrões no campo de classificação. A explicação detalhada e as provas de SVM podem ser verificadas nos livros de Vapnik (1995) e Vapnik (1999).

De acordo com Shin et al. (2005), o SVM produz um classificador binário, os chamados hiperplanos de separação ótimos, por meio do mapeamento não linear dos vetores de entrada no espaço de recursos de alta dimensão. O SVM constrói um modelo linear para estimar a função de decisão usando limites de classes não lineares com base em vetores de suporte. Se os dados forem separados linearmente, o SVM treina máquinas lineares para um hiperplano ideal que separa os dados sem erro e na distância máxima entre o hiperplano e os pontos de treinamento mais próximos. Os pontos de treinamento mais próximos do hiperplano de separação ideal são chamados de vetores de suporte. Todos os outros exemplos de treinamento são irrelevantes para determinar os limites das classes binárias. Em casos gerais em que os dados não são separados linearmente, o SVM usa máquinas não lineares para encontrar um hiperplano que minimiza o número de erros do conjunto de treinamento.

2.2.7 Adaptive Boosting (AdaBoost)

O AdaBoost é um método de ML desenvolvido por Freund et al. (1999). De acordo com Taherkhani et al. (2020) o Adaboost combina todos os classificadores fracos para criar um classificador forte. Ele pode ser aplicado em combinação com vários outros algoritmos de aprendizagem para melhorar o desempenho. Os resultados dos aprendizes fracos são associados a uma acumulação ponderada que representa os resultados finais do classificador ponderado. O Adaboost é adaptativo, uma vez que os aprendizes fracos posteriores são ajustados para favorecer instâncias que foram classificadas erroneamente pelos classificadores anteriores.

Os autores também dizem que quando o AdaBoost está em processo de treinamento, ele escolhe as funções ideais para aumentar o desempenho de predição do modelo, reduzir a dimensão e possivelmente encurtar o tempo de execução, já que não há necessidade de calcular recursos irrelevantes.

2.2.8 Mensurando os Índices de Sentimento

Nesta seção será apresentada a metodologia de construção dos índices de sentimento bancário a partir dos Formulários de Informações Trimestrais (ITR) e Demonstrações Financeiras Padronizadas (DFP) dessas organizações exigidos pela Comissão de Valores Mobiliários CVM. Dessa maneira, para construção das variáveis de sentimento textual, inicialmente foi obtida a classificação setorial disponibilizada no site da B3, a qual apresenta a lista dos bancos de capital aberto no país. Posteriormente, foram recolhidos os ITRs e as DFPs dessas companhias por meio do endereço eletrônico da CVM, coletando dessa forma os relatórios dos três primeiros

trimestres de cada ano analisado, mediante os documentos ITR e o último trimestre por sua vez, correspondente aos relatórios DFP.

Assim como no trabalho de Damasceno et al. (2021), após a obtenção dos documentos ITR e DFP, que fazem parte da base de dados da pesquisa, iniciou-se a etapa de transformação dos arquivos do formato original, em *Portable Document Format* (PDF), para texto separado por tabulações (TXT). Posteriormente, estes arquivos gerados na etapa anterior passaram por um processo de pré-ajuste da amostra, sendo removidos os espaços duplos, pontuações, números, bem como os stopwords, que se configuram como uma lista de preposições, pronomes, conjunções e formas verbais que não apresentam relevância explicativa em documentos textuais.

Para explorar as informações textuais presentes nos relatórios ITR e DFP foi empregada a Análise de Processamento Natural, mediante algoritmos de leitura automatizada escritos em linguagem R. Além disso, foi aplicada a técnica do *vector space model*, que considera as palavras presentes nos textos como vetores, os quais serão utilizados para a estimação do peso de cada palavra em um determinado documento de acordo com a frequência das mesmas, conforme apresentado na equação a seguir.

$$P_{i,j} \begin{cases} \frac{(1+\log(Tf_{i,j}))}{(1+\log(a_j))} \times \log \frac{N}{df_i}, & \text{se} \quad Tf_{i,j} \geq 1 \\ 0, & \text{se} \quad Tf_{i,j} = 0 \end{cases} \tag{2.8}$$

Em que $P_{i,j}$ consiste no peso da palavra i no documento j, por sua vez, $Tf_{i,j}$ compreende a totalidade de ocorrências de uma determinada palavra i em um relatório j, a_j diz respeito à média de frequências das palavras de um documento financeiro, N é o total de relatórios da amostra, e, df_i é o total de relatórios administrativos com ao menos uma ocorrência da palavra i.

A aplicação de ponderações com logaritmos foi realizada objetivando minimizar a atuação de palavras de alta frequência (*outliers*) nos documentos da base de dados, evitando que aqueles apresentem um peso maior na estimação. Acerca desse assunto, Loughran e McDonald (2011) argumentam que tal prática reduz a interferência de *outliers*, comprovando a eficácia desse método após examinarem relatórios 10-K, produzidos por firmas norte-americanas.

De acordo com Shapiro et al. (2020) existem duas metodologias gerais para quantificar o sentimento no texto. A primeira é conhecida como metodologia lexical. Esta abordagem se baseia em listas predefinidas de palavras, chamadas de léxicos ou dicionários, com cada palavra atribuída uma pontuação para a emoção de interesse. Geralmente, essas pontuações são simplesmente -1, 0 e 1 para negativo, neutro e positivo, mas alguns léxicos têm mais de três categorias. As aplicações típicas desta abordagem medem o conteúdo emocional de um determinado corpus de texto com base na prevalência de palavras negativas versus positivas no corpus. Esses métodos de correspondência de palavras são chamados de métodos de *bag-of-words* devido as características contextuais de cada palavra, como sua ordem no texto, classe gramatical, coocorrência com outras palavras e outras características contextuais específicas ao texto em que a palavra aparece são ignorados. Shapiro et al. (2020) afirma que a segunda

abordagem, mais incipiente, emprega técnicas de ML para construir modelos complexos para prever probabilisticamente o sentimento de um determinado conjunto de texto.

Nesse ponto precisamos enfatizar que vamos utilizar duas abordagens distintas para a construção do sentimento do gestor. Em uma delas será adotada a estimação do sentimento textual com dicionário fixo, tendo como base o algoritmo de leitura desenvolvido por Machado et al. (2019). Nele são desconsiderados os termos que não estavam associados, no referido dicionário, a nenhuma das duas tipologias de sentimento.

Em um segundo momento será utilizada a abordagem com dicionários variantes no tempo. Essa discussão está disponível em Lima et al. (2019) e, conforme destacado por esses, a suposição de um dicionário invariável no tempo não parece ser realista em documentos que introduzem novas palavras ao longo do tempo ou se o vocabulário usado em períodos de recessão difere do usado em períodos de expansões econômicas. Os autores ressaltam que mesmo se o vocabulário fosse constante ao longo do tempo, o poder preditivo de algumas palavras pode variar, ou seja, a relevância das palavras se alteram, mas a literatura existente não explica esse efeito e, portanto, os preditores resultantes não refletem as informações textuais mais preditivas encontradas nos documentos em um determinado momento. Com base nisso, aplicamos para a construção do sentimento bancário de dicionário variante no tempo a abordagem desenvolvida por Lima et al. (2019).

Assim, utilizando a metodologia proposta pelos autores para construir o dicionário variante no tempo, primeiramente criamos um vetor, $X_{i,t}$, em que cada elemento do vetor mostra observações em série temporal da frequência em que cada palavra (ou combinação de palavras) aparece nos relatórios de cada banco *i* até o tempo *t*. Portanto, esta etapa transforma as palavras em valores numéricos sem usar um dicionário pré-especificado (fixo). Essa representação numérica é de alta dimensão e esparsa; portanto, a redução da dimensionalidade deve ser empregada na próxima etapa. Na segunda etapa, usamos o ML para selecionar as palavras mais preditivas $X^*_{i,t} \subset X_{i,t}$.

O modelo de *elastic net* foi escolhido para realizar a segunda etapa:

$$y_{i,t+h} = W'_{i,t}\boldsymbol{\beta}_h + X'_{i,t}\phi_h + \epsilon_{i,t+h} \tag{2.9}$$

em que $h \geq 0$ é o horizonte de previsão, $\widehat{\beta}_h$ e $\widehat{\phi}_h$ são estimadas minimizando a seguinte função objetivo:

$$\min_{\beta_h,\phi_h} \sum_t \left(y_{i,t+h} - W'_{i,t}\beta_h - X'_t\phi_h\right)^2 + \lambda_1 \left\|\phi_h\right\|_{\ell_1} + \lambda_2 \left\|\phi_h\right\|_{\ell_2} \tag{2.10}$$

em que W_t é um vetor *k x 1* de preditores pré-determinados, como defasagens de y_t bem como preditores tradicionais de dados estruturados e $\|\cdot\|_{\ell_1}$ e $\|\cdot\|_{\ell_2}$ são a norma ℓ_1 e ℓ_2, respectivamente. Então, a partir da seleção das palavras com maior poder preditivo, temos para

cada período *t* um conjunto de palavras que servem como dicionário léxico para a obtenção da série de sentimento bancário. No presente artigo, vamos aplicar esse método para a nova variável de risco bancário. Assim, o algoritmo vai selecionar em cada relatório ITR no período *t* de cada banco *i* as palavras que mais explicam mudanças no risco de insolvência e a partir desse conjunto de palavras, ou seja, dicionário variante no tempo, são geradas as séries de sentimento bancário para cada instituição.

Por fim, ambas abordagens de dicionário, calculam o índice de sentimento pela diferença entre palavras positivas e negativas, dividida pela soma de palavras positivas e negativas, como foi proposto por Hubert e Labondance (2018):

$$SB_t = \frac{\text{PalavrasPositivas}_{\,t} - \text{PalavrasNegativas}_{\,t}}{\text{PalavrasPositivas}_{\,t} + \text{PalavrasNegativas}_{\,t}} \tag{2.11}$$

Portanto, obtemos a medida de sentimento bancário, *SB*, que varia entre -1 e 1.

2.2.9 Preditores

Para prever os risco de insolvência bancária usamos dados macroeconômicos e financeiros dos bancos como preditores. A janela temporal e frequência dos dados é idêntica aos da Seção 2.2.1. A Tabela 10 apresenta as variáveis utilizadas como preditores.

Tabela 10 – Preditores

Variável	Operacionalização	Fonte
Sentimento Textual Bancário de Dicionário Fixo (SBF)	Tom do sentimento textual de cada relatório da amostra com dicionário fixo	Resultados da pesquisa
Sentimento Textual Bancário de Dicionário Variante (SBV)	Tom do sentimento textual de cada relatório da amostra com dicionário variante	Resultados da pesquisa
Tamanho (TAM)	Razão entre depósitos totais e ativo total	Economática e Resultados da pesquisa
Retorno sobre os ativos (ROA)	Razão entre o lucro operacional e ativo total	Economática e Resultados da pesquisa
Capitalização (ETS)	Razão entre patrimônio líquido e ativo total	Economática e Resultados da pesquisa
Produto Interno Bruto (PIB)	Variação percentual do PIB real em relação ao trimestre anterior	IBGE
Índice Nacional de Preços ao Consumidor Amplo (IPCA)	Variação percentual do IPCA em relação ao trimestre anterior	IBGE
Ibovespa (IBOV)	Variação percentual do Ibovespa em relação ao trimestre anterior	Anbima
Ciclo Econômico (CICLO)	Ciclo do PIB real trimestral obtido pelo filtro HP	Resultados da pesquisa

A utilização de uma variável que mede o tom do sentimento textual dos relatórios trimestrais já existe na literatura. Damasceno et al. (2021) construiu uma variável de sentimento bancário com o dicionário fixo de Machado et al. (2019) para os bancos brasileiros de capital

negociados na B3. Entretanto, como já foi mencionado na seção anterior, um dicionário fixo ao longo do tempo não consegue captar a alteração da importância das palavras ao longo do tempo e, principalmente, não é capaz de captar o surgimento de novos termos importantes, por exemplo, o termo "Covid-19"surgiu em 2020 nos relatórios dos bancos e ele não é captado pelo dicionário de Machado et al. (2019). Então, na presente pesquisa também construímos uma variavel de sentimento bancário com um dicionário variante no tempo a partir da abordagem de Lima et al. (2019).

Com relação as variáveis financeiras, essas foram escolhidas de acordo com a literatura já consolidada, entre esses, destaca-se: Rosa e Gartner (2017), Vieira et al. (2020). No caso das variáveis macroeconômicas, o PIB já é amplamente utilizado na literatura. Assim como Suss e Treitel (2019), adicionamos a inflação (IPCA) e um índice acionário (IBOV). Suss e Treitel (2019) também recomendaram que trabalhos futuros sobre o tema incorporassem ao conjunto de preditores uma variável *proxy* para captar os efeitos do ciclo econômico. Para atender esta sugestão incluímos o ciclo econômico (Ciclo) e esse foi obtido a partir da aplicação do filtro Hodrick-Prescott na série do PIB.

2.3 Resultados

2.3.1 Estatística Descritiva

Nessa seção será ilustrada a estatística descritiva das variáveis usadas como preditores para a variável dependente risco bancário. A Tabela 11 mostra os valores obtidos.

Tabela 11 – Estatística Descritiva

	Observações	Média	Desv. Pad.	Mediana	Min	Max	Assimetria
LIQ	396	0,17	0,37	0	0	1	1,71
TAM	396	18,22	2,10	17,66	14,60	21,47	0,20
ETS	396	0,10	0,04	0,10	0,03	0,29	1,52
ROA	396	0,002	0,009	0,003	-0,04	0,05	-0,75
SBF	396	-0,002	0,09	-0,01	-0,18	0,55	1,92
SBV	396	0,13	0,31	0,16	-0,72	0,76	-0,35
PIB	396	1,44	3,42	2,30	-7,32	10,70	-0,21
IBOV	396	0,21	7,82	0,30	-29,9	16,97	-1,37
IPCA	396	0,52	0,40	0,44	-0,23	1,35	0,35
CICLO	396	190,00	46170,17	-1593,71	-192664	81402,12	-1,57

Com relação as variáveis de sentimento bancário, pode-se observar por meio das medidas de tendência central que durante o período de análise os relatórios financeiros ITR e DFP desses bancos, apresentaram de maneira geral um tom de neutralidade, dados que os valores para as duas variáveis (SBF e SBV) se situam próximo ao zero.

Ao observar os valores mínimos e máximos é possível notar a presença de bancos que apresentam em seus relatórios sentimentos textuais distintos, denotando o comportamento hete-

rogêneo por parte dessas firmas financeiras, no que se refere ao tom textual dos seus documentos, principalmente, para o sentimento bancário com dicionário variante. Esses resultados para o sentimento bancário com dicionário fixo estão em conformidade ao que Damasceno et al. (2021) encontrou em seu trabalho. Para o sentimento bancário com dicionário variante, esses resultados até o momento não existem na literatura.

Para as variáveis financeiras dos bancos, os resultados da estatística descritiva ficaram em linha com o obtido por Damasceno et al. (2021). A variável TAM que os seus valores correspondem ao logaritmo natural dos ativos totais de cada banco mostra que em média os bancos escolhidos, apresentaram o valor de 18,22, e mediana de 17,66.

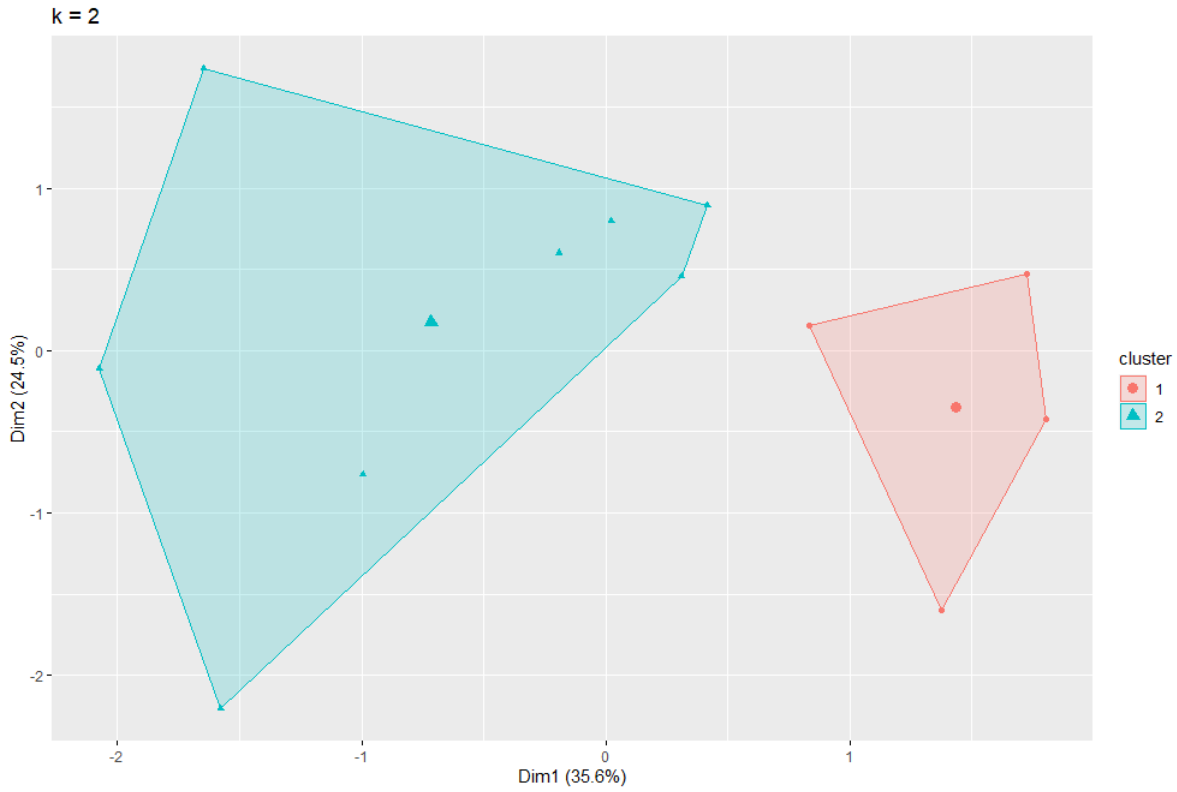
Em relação à capitalização dos bancos (ETS), os valores observados ilustram em média um valor de 0,17. Isso significa que essas companhias operam de maneira geral com certa alavancagem financeira, pois o patrimônio líuqido representa 17% dos ativos totais dessas instituições, valor bem acima do encontrado por Damasceno et al. (2021) de 0,103%. A rentabilidade dos bancos retratada por meio do retorno sobre os ativos (ROA) demonstra que essas empresas obtêm em média retornos de 0,002%, para cada trimestre, ou seja, a rentabilidade na média para o período foi praticamente nula.

Já entre as variáveis macroeconômicas vale destacar o PIB que apresentou um média de crescimento elevada durante o período de análise com valor de 1,44%. Entretanto, vale destacar que essa taxa foi obtida devido aos primeiros anos da amostra, já que nos últimos anos a taxa de crescimento do PIB foi baixa.

2.3.2 Clusterização e Risco Bancário

Para a realização do agrupamento dos *clusters* foi considerado o desvio padrão móvel do retorno sobre os ativos (σROA), empregado para a execução do algoritmo *k-means*. O Z-score foi usado como parâmetro de comparação com os resultados obtidos pelo método de clusterização. Para tal, foi criada uma variável *dummy* que classificou o grupo de maior risco (1), definido pelo primeiro quartil do Z-score, e o grupo de menor risco (0) é representado pelos demais valores. A Figura 5 mostra um exemplo do processo de clusterização pelo *k-means*, neste caso, do primeiro trimestre de 2021.

Figura 5 – Clusterização formada a partir do desvio padrão móvel do retorno sobre os ativos (σROA)



Fonte: Resultados da pesquisa.

Nota: O software classifica dois clusters entre 1 e 2, mas é equivalente aos valores 0 e 1 usados no presente artigo.

Ao compararmos os resultados da classificação de risco bancário, definida pelo *k-means*, com o Z-score é possível perceber que há uma distinção na separação dos grupos quanto a métrica adotada, conforme descrito na Tabela 12. Esse achado também foi reportado por Damasceno et al. (2021), em que esses encontraram similaridade entre Z-score e a clusterização apenas para o grupo de menor risco.

Tabela 12 – Comparação do agrupamento dos clusters com a classificação do Z-score - 1T21

	Total de observações com o Z-score	Total de observações com σROA	Observações comuns quanto ao risco	Percentual de similaridade
Grupo de maior risco (1)	139	69	25	17.99%
Grupo de menor risco (0)	257	327	163	63.42%
Total de observações	396	396	188	-

Como pode ser visto na Tabela 11, o método *k-means* é mais rigoroso quanto a classificação dos bancos em uma situação de maior risco. Além disso, verificamos que o método de agrupamento de *clusters* classificou 63,42% das observações de menor risco rotuladas pelo Z-score. A maior diferença acontece quando é considerado do grupo de elevado risco. O Z-score

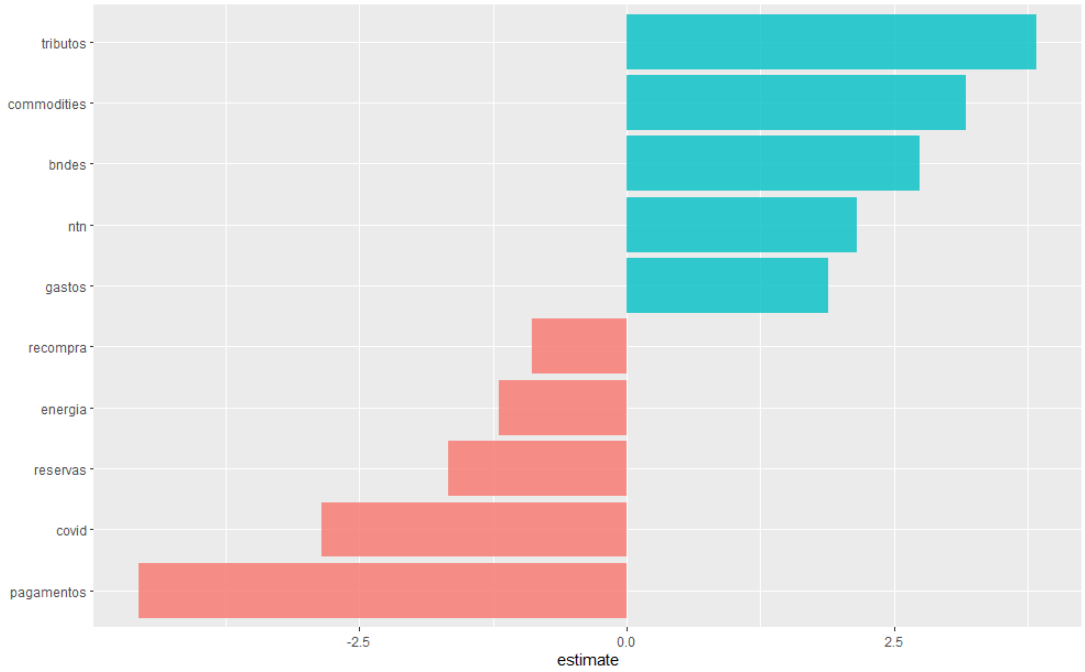
categorizou 139 observações no grupo de maior risco. No entanto, pelo método de *clusters*, houve *matching* com apenas 17,99% das observações. Ressalta-se ainda, que ambos os percentuais foram obtidos após a comparação das mesmas observações pelas duas medidas de classificação empregadas. Com a categorização da *dummy* do Z-score 65% da amostra faz parte do grupo de baixo risco e 35% da amostra no grupo de elevado risco. Com a variável de cluster 83% da amostra faz parte do grupo de menor risco, enquanto, 17% do grupo de maior risco. Esse resultado sugere que o método de classificação pelo agrupamento é mais rigoroso, quanto a categorização do grupo de maior risco, do que o Z-score.

2.3.3 Análise do Sentimento Bancário

Nesta seção serão apresentados os índices de sentimento bancário, construídos com o dicionário variante no tempo, dos doze bancos utilizados na amostra. Como foi mencionado na seção 2.2.8, foi usado para a criação das séries de sentimento variante no tempo a abordagem proposta por (LIMA et al., 2019). Assim, para cada instituição bancária o algoritmo de *Elastic Net* seleciona as palavras mais preditivas para a nova variável de risco em cada ITR. A partir disso a série de tom textual é gerada. As palavras consideradas mais preditivas para o risco de falência varia de acordo com os relatórios de cada instituição bancária.

A Figura 6, ilustrativamente, apresenta os 5 principais termos positivos e negativos mais preditivos para o risco de insolvência encontrados nos relatórios do Banco ABC. Neste caso, o dicionário variante no tempo encontrou que termos como "títulos", "gastos", "covid"e "pagamentos"contidos nos relatórios trimestrais do Banco ABC são as palavras que possuem um maior efeito no risco bancário.

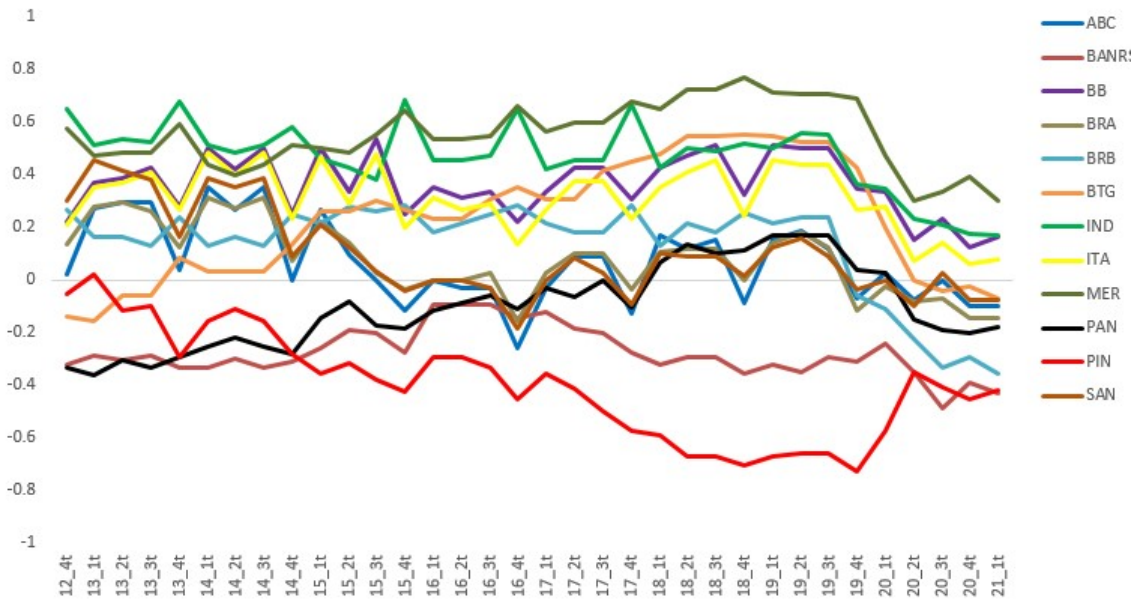
Figura 6 – Coeficientes mais positivos e negativos na determinação do sentimento bancário - Banco ABC - 1T21



Fonte: Resultados da pesquisa.

Após o algoritmo *Elastic Net* encontrar as palavras mais relevantes na explicação do risco bancário no trimestre t, o valor do tom do sentimento é obtido. No exemplo da Figura 6, dado esse conjunto de palavras, o valor do tom do Banco ABC no primeiro trimestre de 2021 foi de -0,10, ou seja, um tom pessimista.

Figura 7 – Séries de sentimento bancário com dicionário variante no tempo



Fonte: Resultados da pesquisa.

Já a Figura 7 ilustra as séries de sentimento bancário de cada instituição bancária. É possivel notar que ao longo do tempo o tom dos ITRs do Banco Mercantil foi o mais otimista ao longo do tempo, enquanto o Banco Pine apresentou o tom mais pessimista durante o período de análise. A Figura 7 também mostra que a partir do segundo trimestre de 2020 o tom do sentimento textual de praticamente todos os bancos obteve uma forte queda, ou seja, o tom dos relatórios passou a ser mais pessimista após o início da pandemia da Covid-19, iniciada durante esse período.

2.3.4 Análise dos Modelos de Previsão

Assim como no trabalho de Damasceno et al. (2021) para o cálculo da acurácia em um primeiro momento é necessário fazer uso da matriz de confusão, sendo essa considerada como uma ferramenta útil durante a etapa de avaliação de modelos de classificação, ao fornecer os dados necessários para a mensuração das medidas de especificidade e sensibilidade, as quais posteriormente são utilizadas para o cálculo da acurácia dos modelos. A sensibilidade pode ser definida como:

$$\text{sensibilidade} = \frac{TP}{TP + FN} \tag{2.12}$$

Em que: sensibilidade = Taxa de verdadeiros positivos, indicando o percentual de instituições bancárias que possuem elevado risco de falência futura, e que foram classificadas corretamente; TP = Verdadeiro positivo, indica o número de casos positivos (maior risco de insolvência) que foram corretamente identificados; FN = Falso negativo, indica o número de casos positivos (maior risco de insolvência) que foram classificados de forma incorreta como casos negativos (menor risco de insolvência).

Já a especificidade é definida pela seguinte expressão:

$$\text{especificidade} = \frac{TN}{TN + FP} \tag{2.13}$$

especificidade = Taxa de verdadeiros negativos, indica a proporção de bancos que apresentam menor risco de insolvência e que foram classificados corretamente; TN = Verdadeiro negativo, indica o número de casos negativos (menor risco de insolvência) que foram corretamente identificados; FP = Falso positivo, indica o número de casos negativos (menor risco de insolvência) que foram classificados de forma incorreta como casos positivos (maior risco de insolvência).

Finalmente, a acurácia é mensurada como:

$$\text{Acurácia} = \frac{TP + TN}{TP + TN + FP + FN} \tag{2.14}$$

Foram escolhidos seis modelos de previsão, conforme descrito na Tabela 13. Como mencionado anteriormente, além de encontrar o modelo com maior poder de predição do risco

bancário, também vamos verificar se as variáveis de sentimento bancário são capazes de melhorar a acurácia dos modelos, ou seja, se o tom contido nos relatórios trimestrais dos bancos é uma informação relevante no monitoramento do risco de falência de tais instituições.

Tabela 13 – Definição dos modelos de previsão

Sigla	Definição	Sigla	Definição
NB1	Modelo Naive Bayes s/ sentimento bancário	NB2	Modelo Naive Bayes com SBF
NB3	Modelo Naive Bayes com SBV	NB4	Modelo Naive Bayes com SBF e SBV
LOGIT1	Modelo Logit s/ sentimento bancário	LOGIT2	Modelo Logit com SBF
LOGIT3	Modelo Logit com SBV	LOGIT4	Modelo Logit com SBF e SBV
SVM1	Modelo SVM s/ sentimento bancário	SVM2	Modelo SVM com SBF
SVM3	Modelo SVM com SBV	SVM4	Modelo SVM com SBF e SBV
RF1	Modelo Random Forest s/ sentimento bancário	RF2	Modelo Random Forest com SBF
RF3	Modelo Random Forest com SBV	RF4	Modelo Random Forest com SBF e SBV
AB1	Modelo AdaBoost s/ sentimento bancário	AB2	Modelo AdaBoost com SBF
AB3	Modelo AdaBoost com SBV	AB4	Modelo AdaBoost com SBF e SBV
DT1	Modelo Árvore de Decisão s/ sentimento bancário	DT2	Modelo Árvore de Decisão com SBF
DT3	Modelo Árvore de Decisão com SBV	DT4	Modelo Árvore de Decisão com SBF e SBV

Portanto, como é possível ver na Tabela 14, para cada modelo foram testadas quatro especificações. A primeira é a estimação com todos os preditores menos as duas variáveis de sentimento bancário (SBF e SBV). A segunda é a estimação com todos os preditores menos SBV. A terceira é a estimação com todos os preditores menos SBF. A quarta é a estimação com todos os preditores.

Assim como nos trabalhos de Rosa e Gartner (2017) e Damasceno et al. (2021), foi aplicada a etapa de *cross-validation* para evitar a ocorrência de *over-fitting*. Com relação a divisão da amostra entre treino e teste, optou-se por colocar 80% da amostra para o treino dos modelos e 20% da amostra para o teste e mensuração da acurácia.

As acurácias das previsões dos modelos podem ser vista na Tabela 14. De forma geral, o modelo que atingiu a maior taxa de acurácia foi o DT (especificação 3), com acurácia de 95% para a amostra de teste. Já os modelos SVM e LOGIT apresentaram a pior acurácia, com valores de 85%. Vale destacar que esses dois últimos modelos, assim como o modelo NB, não apresentaram alteração em suas acurácias com as modificações nos preditores. Esse fato ocorreu porque no SVM e LOGIT em todas as 4 especificações só previram valores 0, ou seja, eles não foram capazes de prever valores 1. No caso do modelo NB, nas quatro especificações ele só conseguiu prever apenas um valor para alto risco.

Tabela 14 – Acurácia das previsões dos modelos

Modelo/Especificação	1	2	3	4
NB	0,8625	0,8625	0,8625	0,8625
RF	0,9125	0,8875	0,9125	0,925
SVM	0,8500	0,8500	0,8500	0,8500
LOGIT	0,8500	0,8500	0,8500	0,8500
AB	0,9250	0,9125	0,9375	0,9125
DT	0,9250	0,9250	0,9500	0,9375

Também destacamos que com excessão dos três modelos citados no parágrafo anterior, a inclusão do sentimento bancário proporciona ganhos de desempenho dos modelos, pois em nenhum deles a especificação do tipo 1 (modelos sem variáveis de sentimento bancário) conseguiu superar a acurácia das especificações com variáveis de sentimento bancário. Pode-se também perceber que modelos com a especificação 3 (modelos com SBV e sem SBF), com excessão do RF, obtiveram os melhores desempenhos.

Quando comparado com os resultados obtidos pelos estudos com temática similiar, a presente pesquisa conseguiu um desempenho superior aos trabalhos de Climent et al. (2019), Suss e Treitel (2019), Huang e Yen (2019) Damasceno et al. (2021). Contudo, Viswanathan et al. (2020) obtiveram um desempenho superior a 95%, neste caso, os autores reportam uma acurácia de 95,93% com o modelo *random forest*.

Os resultados obtidos mostram alguns pontos que merecem destaque. O primeiro deles é que os modelos de ML do tipo *ensemble* obtiveram acurácias superiores ao modelo tradicional (modelo logit). Esses resultados convergem ao encontrado pela literatura, com excessão do trabalho de Damasceno et al. (2021) que encontrou um melhor desempenho com o modelo logit, pois o autor reporta que o modelo *random forest* foi descartado por apresentar *over-fitting*.

O segundo é que a inclusão do sentimento bancário como preditor é capaz de melhorar o desempenho dos modelos de predição. O terceiro é que entre as variáveis de sentimento bancário, a que aplicou um dicionário variante no tempo conseguiu um resultado melhor em termos de acurácia. Cabe ressaltar que esses resultados ainda não foram reportados na literatura por se tratar de uma discussão inédita.

2.3.5 Taxas de Falsos Positivos e Falsos Negativos

A questão da identificação do melhor modelo para prever o risco de insolvência bancária, exposta na seção anterior, é de grande relevância para os tomadores de decisão de um sistema de alerta preventivo. Contudo, isso não é suficiente para um sistema de acompanhamento do risco bancário. De acordo com Suss e Treitel (2019) duas métricas de desempenho diferentes são relevantes para os tomadores de decisão de um sistema de alerta preventivo: as taxas de erro de falso negativo (FN) e falso positivo (FP).

A taxa FN representa a proporção de bancos que efetivamente possuem alto risco, no entanto, o algoritmo classifica como de baixo risco, é provavelmente a mais importante das duas métricas do ponto de vista regulatório. Um sistema de alerta preventivo que não aciona o alarme quando deveria, especialmente para instituições grandes e sistemicamente importantes, pode ter consequências negativas para toda a economia. Obviamente, para minimizar a taxa de FN, podemos apenas diminuir o limite de probabilidade prevista pelo qual classificamos os bancos de baixo e alto risco. Já a taxa FP é a proporção de bancos de baixo risco e que são classificados erroneamente como de alto risco. A taxa FN é calculada como $\frac{FN}{FN+TP}$ e taxa FP é obtida por $\frac{FP}{FP+TN}$.

Tabela 15 – Taxas de falsos positivos e falsos negativos

	Taxa de FN	Taxa de FP
NB	14,10%	16,45%
RF	5,71%	4,35%
SVM	15,00%	16,45%
LOGIT	15,00%	16,45%
AB	4,35%	2,94%
DT	2,94%	2,94%

É possivel ver na Tabela 15 que o modelo DT obteve a menor taxa de FN, mostrando que apenas 2,94% das previsões foram classificadas de baixo risco quando possuim risco alto. Comparando esse resultado com o mesmo limiar de 25% do trabalho de Suss e Treitel (2019), o presente artigo conseguiu obter taxas de FN próximas ao encontrado pelos autores. O menor valor que os autores encontraram foi de 2,2% com o modelo *random forest*. Com relação as taxas de FP, os modelos do presente trabalho conseguiram taxas mais baixas quando comparada ao encontrado por Suss e Treitel (2019).

2.4 Considerações Finais

Esse artigo tem três contribuições para a literatura que trata de risco de insolvência bancária, sendo elas: utilização da técnica não-supervisionada de *cluster* para classificar, conforme o risco de insolvência, os bancos negociados na B3; construção dos índices de sentimento bancário a partir dos relatórios trimestrais (ITR) e Demonstrações Financeiras Padronizadas (DFP) dessas organizações exigidos pela CVM, tendo como base a abordagem com dicionários variantes no tempo; por fim, comparamos várias técnicas de aprendizado de máquina e estatísticas clássicas, implementando um procedimento rigoroso de validação cruzada aleatória de bloco duplo para avaliar o desempenho da previsão fora da amostra.

Na análise de sentimento verificamos que, ao longo do tempo, o tom dos ITRs do Banco Mercantil foi o mais otimista, enquanto o Banco Pine apresentou o tom mais pessimista durante o período de análise. Outro dado interessante é que o índice gerado foi capaz de captar os efeitos negativos da crise sanitária da Covid-19, presente nos relatórios dos bancos. Em outras palavras, a partir do segundo trimestre de 2020 o tom do sentimento textual de praticamente todos os bancos obteve uma forte queda, ou seja, o tom dos relatórios passou a ser mais pessimista.

Os resultados obtidos mostraram que o algoritmo de árvore de decisão é superior na previsão dos dados da amostra de teste de classificação, embora os algoritmos do tipo *ensemble* (*random forest* e AdaBoost) tenham se aproximado em termos de acurácia. Também constatamos que a inclusão do sentimento bancário como preditor promove ganhos de acurácia, principalmente, se for utilizado o dicionário variante no tempo. Além disso, constatamos que o modelo de árvore de regressão possui o menor custo em termos de falsos negativos e falsos positivos.

Do ponto de vista prático, apresentamos uma medida alternativa para classificação de risco, desenvolvemos um indicador capaz de captar o sentimento do gestor do banco quanto as condições de mercado e, por fim, projetamos o risco de insolvência bancária por meio de diferentes técnicas de aprendizagem de máquina. A fusão destas etapas auxilia os gestores, reguladores e supervisores a antecipar comportamentos atípicos ou que envolvem maior risco de insolvência das instituições financeiras.

Cabe ressaltar que este artigo tratou, exclusivamente, os bancos de capital aberto negociados na B3 no Brasil, devido o nível de regulação e a quantidade de informações disponíveis para estes. Portanto, estudos futuros podem expandir o escopo para os demais bancos do país. Também podem ser construídas variáveis de sentimento bancário com o uso de outros dicionários. Dada a janela amostral escolhida em que houve a pandemia da Covid-19, em 2020, então, trabalhos futuros podem verificar os resultados antes e depois da pandemia e mensurar os resultados nesses dois ambientes.

Parte III

Preferências do Banco Central

3 O Banco Central do Brasil reage ao Sentimento da Política Fiscal? Uma análise a partir de processamento de linguagem natural

3.1 Introdução

Após o artigo seminal de Taylor (1993)[1], diversos autores buscaram estimar, para diferentes economias, funções de reação com o objetivo de capturar e entender o comportamento dos bancos centrais. A maioria desses trabalhos obtiveram estimações a partir de modelos de equações simples. Mais recentemente, alguns trabalhos incorporaram na literatura de funções de reação os modelos DSGE. Este é o caso de Smets e Wouters (2007), Lubik e Schorfheide (2007), e Finocchiaro e Heideken (2013), dentre outros[2].

Silva e Besarria (2018) enfatizam que a literatura caminhou no sentido de introduzir elementos que pudessem refletir mais fielmente a maneira como os bancos centrais reagem a alterações no cenário econômico. Uma delas tem sido permitir a suavização da taxa de juros, com a introdução do nível passado da taxa de juros na função de reação, outra tem sido considerar expectativas de inflação, ao invés da inflação passada como na regra original. De acordo com os autores, esta última alteração foi implementada com o objetivo de refletir o comportamento *foward-looking* dos bancos centrais quando da tomada de decisões e, em particular, o reconhecimento de que a política monetária afeta a economia com um certa defasagem.

Alguns outros trabalhos que surgiram na literatura inseriram variáveis adicionais na função de reação. Uma destas variáveis tem sido a inserção de variáveis fiscais na função de reação do Banco Central. Autores como Canzoneri et al. (2001) e Kumhof et al. (2010) adicionaram variáveis fiscais, como a dívida pública e o superávit primário, além das medidas tradicionais. A inclusão de variáveis fiscais na função de reação do BCB torna possível a verificação de como a condução da política monetária é afetada pelo desempenho da política fiscal. No Brasil, nos últimos anos, esse debate ganhou notoriedade em que uma grande quantidade

[1] Segundo Nobrega et al. (2020) o trabalho de Taylor (1993) trouxe ao debate a questão fundamental entre regras de política e o discricionarismo na consecução da política monetária, atentando para os possíveis ganhos de credibilidade obtidos junto aos agentes econômicos em virtude do fato de o governo seguir regras claras no combate à inflação. A regra de Taylor, conforme inicialmente proposta, descreve a reação do Banco Central em relação à inflação através de seus desvios em relação à meta preestabelecida e dos ciclos de negócios. Posteriormente, Clarida et al. (1998) propõem uma versão prospectiva da regra original, na qual a reação ocorria em virtude das expectativas inflacionárias.

[2] Em relação aos estudos aplicados à economia brasileira ver os trabalhos de Minella et al. (2003), Holland et al. (2005), Aragón e Medeiros (2013), Lopes et al. (2012), Medeiros et al. (2018), dentre outros.

de estudos buscaram testar a existência de dominância de uma das políticas e como acontece a interação entre ambas[3].

Franco (2021) fala que há certa resistência em admitir uma "natureza fiscal"nos juros, sobretudo entre parlamentares e suas respectivas assessorias, pois tais agentes têm uma propensão a negar de maneira sistemática a existência de restrição orçamentária, escassez de recursos e, em última instância, a existência de um "problema fiscal", fenômeno acima designado como negacionismo fiscal. Seria como admitir que o orçamento termina por determinar a taxa de juros. Franco (2021) enfatiza que a política fiscal está presente nas expectativas que movem as fórmulas e os modelos que orientam o regime de metas, e os dirigentes do BCB não se furtam a enfatizar a importância da politica fiscal e, genericamente, das reformas, de tempos em tempos.

Entretanto, até o momento, nenhum trabalho testou a inclusão de variáveis relacionadas ao sentimento da autoridade fiscal na função de reação do Banco Central. A proposta do presente capítulo é investigar se o Banco Central do Brasil tem reagido ao tom das comunicações da autoridade fiscal. Vamos usar duas estratégias para atingir os objetivos. A primeira estima funções de reação do banco central utilizando equações simples, com a inclusão do tom (sentimento) da política fiscal, como um dos argumentos da equação em um modelo de Mínimos Quadrados Ordinários (MQO) e um modelo de Método Generalizado de Momento (GMM). Para a estimação das funções de reação a segunda estratégia se utiliza de um modelo Dinâmico Estocástico de Equilíbrio Geral (DSGE).

A inclusão do sentimento da política fiscal é capaz de ilustrar como a política monetária se comporta ao saber das perspectivas da autoridade fiscal sobre a situação fiscal do país. Se o coeficiente do SPF for significativo estatisticamente na função de reação do BCB, então, pode-se sugerir que a percepção da autoridade fiscal acerca do cenário fiscal captada pelo tom de suas publicações é uma variável relevante na tomada de decisão sobre a taxa de juros pela autoridade monetária e, possivelmente, que a política fiscal domine as ações da política monetária. Assim, a principal contribuição do artigo é incluir um índice de polaridade que mensure o sentimento do gestor de política fiscal sobre a conjuntura fiscal em uma função de reação do Banco Central e assim fornecendo uma abordagem alternativa para testar a hipótese de dominância fiscal. A própria criação da variável de sentimento da política fiscal também pode ser considerada como uma contribuição, pois até o momento nenhum trabalho para o Brasil criou tal variável a partir de *text mining* das publicações do Tesouro Nacional.

A variável que mede o sentimento da política fiscal foi criada a partir do uso de *text mining* e análise de polaridade nos relatórios fiscais produzidos pelo BCB. Para a realilzação do processo de *text mining* primeiramente usamos um método tradicional de léxico, neste caso, o dicionário de Loughran e McDonald (2011). Também empregamos o método de dicionário variante no tempo de Lima et al. (2019) que utiliza *machine learning* para a construção do

[3] Ver os trabalhos de Issler e Lima (2000), Schymura (2015), Tanner e Ramos (2003), Fialho e Portugal (2005), Ázara (2006), Aguiar (2007), Gadelha e Divino (2008), Junior (2010), Ornellas (2011), Araujo e Besarria (2014), Ferreira et al. (2015) e Nobrega et al. (2020), dentre outros.

dicionário. Após a construção dessas novas variáveis, elas foram inseridas na função de reação do banco central.

Os resultados mostram que a inclusão do índice de sentimento fiscal proveniente de dicionário fixo não apresentou relevância na função de reação do BCB nos modelos MQO e GMM. Entretanto, o índice criado a partir de um dicionário variante no tempo apresenta um grande impacto na função de reação da autoridade monetária. Além disso, no modelo DSGE a inclusão do sentimento da política fiscal aumenta de forma significativa a densidade marginal quando comparada com um modelo básico sem o índice. Portanto, os resultados indicam que existem indícios da ocorrência de uma dominância da política fiscal nas ações de condução da política monetária.

Além desta introdução, o artigo apresenta quatro outras seções. Na seção 2 é apresentada uma descrição da metodologia utilizada, apresentando a função de reação do Banco Central e o modelo DSGE. A seção 3 discute os principais resultados obtidos. Por fim, a seção 4 relata as principais conclusões e limitações do artigo.

3.2 Metodologia

3.2.1 Função de Reação do Banco Central

Taylor (1993) argumentou que o comportamento do Banco Central americano poderia ser descrito por uma regra simples, que associava mudanças na taxa de juros a desvios da inflação e do produto de seu potencial. No presente artigo, em particular, a regra de política monetária é do tipo *foward-looking* e será modificada com a inclusão do sentimento da política fiscal e pode ser descrita como:

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1-\rho_r)\left[r_\pi E_t \widehat{\pi}_{t+1} + r_y \hat{y}_t + r_s \Delta \hat{s}_t\right] + e_t. \tag{3.1}$$

ou em sua forma estimada

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + \Gamma_\pi \widehat{\pi}_{t+1} + \Gamma_y \hat{y}_t + \Gamma_s \Delta \hat{s}_t + e_t \tag{3.2}$$

em que as variáveis com circunflexo estão na forma log-desvio do filtro de Hodrick-Prescott (HP); r_t é a taxa de juros nominal; π_{t+1} é a taxa de inflação; y_t é o produto real da economia; Δs_t é a variação do sentimento da política fiscal; e_t é choque que captura os componentes não sistemáticos na regra de política monetária; e $\Gamma_\pi = (1-\rho_r)\, r_\pi, \Gamma_y = (1-\rho_r)\, r_y, \Gamma_s = (1-\rho_r)\, r_s$.

A regra de política monetária apresentada acima foi estimada por MQO e GMM. De acordo com Silva e Besarria (2018) em relação ao método MQO, destaca-se que esse pode gerar estimativas viesadas e inconsistentes na presença de endogeneidade. Assim, o método GMM passa a ser escolhido como alternativa ao MQO. Sobre o GMM, os autores ainda destacam que a

adequação da inferência estatística, gerada a partir desse método, está ligada à exogeneidade e relevância dos instrumentos adotados. Outro ponto é que a eficiência dos estimadores está diretamente ligada a análise de identificação da seleção das variáveis instrumentais. Assim, os critérios de Andrews (1999) foram usados para selecionar de forma adequada o conjunto de instrumentos, já a hipótese de sobre-identificação foi tratada a partir do teste J.

Para a estimação a partir do GMM usamos como instrumento as variáveis rendimento real, selic, investimento consumo das famílias, expectativa de inflação. Também foram utilizadas as defasagens das séries de sentimento fiscal e PIB como variáveis instrumentais. Para estimar a regra de política monetária para o Brasil foram utilizadas dados trimestrais no período do primeiro trimestre de 2003 ao segundo trimestre de 2021. A série de PIB foi tratada na forma de desvio da tendência estimada pelo filtro HP. Em relação à série de inflação, essa foi analisada na forma de desvio de meta de inflação.

3.2.2 Procedimento de Estimação Textual

Nesta seção será apresentada a metodologia de construção dos índices de sentimentos (S) a partir dos textos dos Relatórios Mensais da Dívida Pública Federal[4]. A divulgação dos relatórios foi iniciada em novembro do ano 2000, na língua portuguesa, e em março de 2003 na língua inglesa. No presente trabalho, optou-se por utilizar a versão do relatório divulgada na língua inglesa, devido, principalmente, ao fato de o mais notório e aceito dicionário utilizado em análise de sentimento ser elaborado nesta língua, proposto por Loughran e McDonald (2011).

Após realizada a coleta dos relatórios da dívida no website do Tesouro Nacional por *web scraping*, algumas etapas de tratamento do conjunto de documentos foram realizadas com o intuito de extrair o máximo de informações possíveis do *corpus* linguístico, minimizando, assim, a perda de informações decorrentes da manipulação da amostra. Antes de executar a análise lexicográfica nos documentos, realizamos uma série de transformações no texto original. O texto é primeiro dividido em uma sequência de *substrings* (*tokens*) cujos caracteres são todos transformados em minúsculas. Removemos *stop words* em inglês e limpamos o texto usando o pacote *tolower* do R.

Cada S_t visa capturar algumas das informações da narrativa no relatório no momento t, para cada documento em nossa amostra. Essa medida transforma milhares de palavras em um único número. Para obter cada série de sentimento da política fiscal S_t usamos duas abordagens: uma que mensura os sentimentos a partir de dicionários com léxicos fixos e outra que usa modelos de *machine learning* para construir um dicionário variante no tempo.

De acordo com Shapiro et al. (2020) existem duas metodologias gerais para quantificar o sentimento no texto. A primeira é conhecida como metodologia lexical. Esta abordagem

[4] O relatório apresenta informações sobre emissões, resgates, estoque, perfil de vencimentos e custo médio, dentre outras, para a Dívida Pública Federal, nela incluídas as dívidas internas e externa de responsabilidade do Tesouro Nacional.

se baseia em listas predefinidas de palavras, chamadas de léxicos ou dicionários, com cada palavra atribuída uma pontuação para a emoção de interesse. Geralmente, essas pontuações são simplesmente 1, 0 e -1 para positivo, neutro e negativo, mas alguns léxicos têm mais de três categorias. As aplicações típicas desta abordagem medem o conteúdo emocional de um determinado corpus de texto com base na prevalência de palavras negativas vs. positivas no corpus. Esses métodos de correspondência de palavras são chamados de métodos de *bag-of-words* (BOW) devido as características contextuais de cada palavra, como sua ordem no texto, classe gramatical, coocorrência com outras palavras e outras características contextuais específicas ao texto em que a palavra aparece, são ignorados.

Dentre esse tipo de método destaca-se o dicionário criado por Loughran e McDonald (2011) (LM). Os autores construíram listas de palavras negativas e positivas que são selecionadas para serem apropriadas ao texto financeiro. Eles mostram que seus dicionários são superiores para classificar textos econômicos e financeiros a outros dicionários, por exemplo, o de Apel e Grimaldi (2012) e o *Harvard Psychosociological Dictionary*, que tende a categorizar incorretamente palavras neutras em um contexto financeiro/econômico (por exemplo, impostos, custos, capital, despesa, responsabilidade, risco, excesso e depreciação). Existem 2.355 palavras negativas e 354 palavras positivas nos dicionários LM. Portanto, para a construção dos índices de sentimentos via abordagem de dicionários fixos usamos o dicionário de LM.

Shapiro et al. (2020) afirma que a segunda abordagem, mais incipiente, emprega técnicas de aprendizado de máquina (ML) para construir modelos complexos para prever probabilisticamente o sentimento de um determinado conjunto de texto. Umas das aplicações dos modelos ML é na construção de dicionários variantes no tempo. Lima et al. (2019) usaram essa abordagem para criar um método de dicionário variante.

De acordo com Lima et al. (2019) a suposição de um dicionário invariável no tempo não parece ser realista em documentos que introduzem novas palavras ao longo do tempo ou se o vocabulário usado em períodos de recessão difere do usado em períodos de expansões econômicas. Os autores ressaltam que mesmo se o vocabulário fosse constante ao longo do tempo, o poder preditivo de algumas palavras pode variar, ou seja, a relevância das palavras se alteram ao longo do tempo, mas a literatura existente não explica esse efeito e, portanto, os preditores resultantes não refletem as informações textuais mais preditivas encontradas nos documentos em um determinado momento. E no atual momento de pandemia do Covid-19 o uso de um dicionário variante no tempo é fundamental em que novos termos começam a se tornar relevantes nas comunicações das autoridades monetárias e fiscais. Portanto, aplicamos para a construção dos índices de sentimentos via dicionários variantes no tempo usamos a abordagem desenvolvida por Lima et al. (2019).

Assim, utilizando a metodologia proposta pelos autores para construir o dicionário variante no tempo, primeiramente criamos um vetor de séries temporais, X_t, em que cada elemento do vetor mostra observações em série temporal da frequência em que cada palavra (ou

combinação de palavras) aparece no relatório mensal da dívida até o tempo *t*. Portanto, esta etapa transforma as palavras em valores numéricos sem usar um dicionário pré-especificado (fixo). Essa representação numérica é de alta dimensão e esparsa; portanto, a redução da dimensionalidade deve ser empregada na próxima etapa. Na segunda etapa, usamos o *supervised machine learning* para selecionar as séries temporais mais preditivas (palavras) $X_t^* \subset X_t$.

O modelo de *elastic net* foi escolhido para realizar a segunda etapa:

$$y_{t+h} = W_t'\boldsymbol{\beta}_h + X_t'\phi_h + \epsilon_{t+h} \tag{3.3}$$

em que $h \geq 0$ é o horizonte de previsão, $\widehat{\beta}_h$ e $\widehat{\phi}_h$ são estimadas minimizando a seguinte função objetivo:

$$\min_{\beta_h,\phi_h} \sum_t \left(y_{t+h} - W_t'\beta_h - X_t'\phi_h\right)^2 + \lambda_1 \left\|\phi_h\right\|_{\ell_1} + \lambda_2 \left\|\phi_h\right\|_{\ell_2} \tag{3.4}$$

em que W_t é um vetor *k x 1* de preditores pré-determinados, como defasagens de y_t bem como preditores tradicionais de dados estruturados e $\|\cdot\|_{\ell_1}$ e $\|\cdot\|_{\ell_2}$ são a norma ℓ_1 e ℓ_2, respectivamente. Então, a partir da seleção das palavras com maior poder preditivo, temos para cada período *t* um conjunto de palavras que servem como dicionário léxico para a obtenção da série de sentimentos *S*. No presente artigo, a nossa variável de resposta para a construção do dicionário variante será a DBGG como proporção do PIB.

Por fim, ambas abordagens de dicionário, calculam o índice de sentimento pela diferença entre palavras positivas e negativas, dividida pela soma de palavras positivas e negativas, como foi proposto por Hubert e Labondance (2018):

$$S_t = \frac{\text{Palavras Positivas}_{\text{t}} - \text{Palavras Negativas}_{\text{t}}}{\text{Palavras Positivas}_{\text{t}} + \text{Palavras Negativas}_{\text{t}}} \tag{3.5}$$

Portanto, obtemos a medida de sentimentos, *S*, que varia entre -1 e 1.

3.2.3 Modelo DSGE

Para a presente seção utilizamos o modelo base do trabalho de Jesus et al. (2020). A alteração que o presente artigo traz nesse modelo é a criação de outra versão para a regra de taxa de juros. Então, serão estimados dois modelos DSGE. O primeiro é o modelo base com a regra de juros tradicional como no trabalho de Jesus et al. (2020) (modelo básico). O segundo com uma regra de taxa de juros modificada com a criação de uma variável de sentimento fiscal dentro do modelo e inserida na regra de Taylor.

3.2.4 Famílias

3.2.4.1 Famílias Pacientes

A utilidade das famílias pacientes depende positivamente do seu nível total de consumo, do seu estoque de imóveis e do seu tempo disponível para o lazer. Assim, as famílias pacientes possuem a seguinte função utilidade:

$$U_t(C_t^{'}, L_t^{'}, H_t^{'}) = \sum_{t=0}^{\infty}(\beta^{'})^t[alogC_t^{'} + log(1 - L_{p,t}^{'} - L_{g,t}^{'}) + jlogH_t^{'}] \tag{3.6}$$

em que, $\beta^{'} \in (0, 1)$ é o fator de desconto das famílias pacientes, $C^{'}$ é o consumo total das famílias pacientes, de modo que, $(C_t^{'} = C_{p,t}^{'} + \mu_p C_{G,t})$; $C_{G,t}$ é o consumo governamental que provê bens e serviços para as famílias e este é determinado de forma exógena; $C_{p,t}^{'}$ representa a parte privada do consumo da família rica e; μ_p é o parâmetro que mede o grau de substituição entre o consumo privado e o consumo em bens/serviços públicos das famílias pacientes; $L_p^{'}$ são as horas trabalhadas das famílias pacientes nas firmas e $L_g^{'}$ horas trabalhadas das famílias pacientes no setor público; $H^{'}$ são os imóveis em poder das famílias pacientes; a simboliza um choque de preferências no consumo das famílias e; j um choque de preferências nos imóveis.

A forma funcional da função utilidade[5] das famílias segue a forma de Iacoviello (2005), em que, é incluído um bem durável, sendo este representado pelos imóveis que estão em posse das famílias. A expressão $(C_t^{'} = C_{p,t}^{'} + \mu_p C_{G,t})$ é utilizada por Barro (1981), Aschauer (1985), Christiano e Eichenbaum (1992) e McGrattan (1994), dentre outros. Neste caso, é definido que o consumo total das famílias é uma combinação linear entre o consumo privado e o consumo público.

Então, as famílias pacientes se deparam com a seguinte restrição orçamentária:

$$(1+\tau^c)C_t^{'}+I_t^{'}+R_{t-1}b_{t-1}^{'}/\pi_t+q_t\Delta H_t^{'} = (1-\tau^l)(w_{p,t}^{'}L_{p,t}^{'}+w_{g,t}^{'}L_{g,t}^{'})+(1-\tau^k)R_t^k K_{t-1}^{'}+b_t^{'}+TR_t \tag{3.7}$$

em que, π_t é a taxa de inflação bruta $(\pi_t = P_t/P_{t-1})$; P_t é o nível geral de preços; R_t é a taxa de juros nominal da economia; $b_t^{'}$ é a quantidade de títulos públicos em posse das famílias pacientes $(b_t^{'} = B_t^{'}/P_t)$; R_t^k é a taxa de remuneração do capital em termos nominais; q_t é o preço real dos imóveis $(q_t = Q_t/P_t)$; $w_{p,t}^{'}$ é o salário real proveniente do setor privado da família paciente $(w_{p,t}^{'} = W_{p,t}^{'}/P_t)$; $w_{g,t}^{'}$ é o salário real proveniente do setor público da família paciente $(w_{g,t}^{'} = W_{g,t}^{'}/P_t)$; $I_t^{'}$ é o nível de investimento realizado pelas famílias ricas; TR_t é a transferência governamental para as famílias; τ^c, τ^l e τ^k são as alíquotas que incidem sobre o consumo, sobre a renda do trabalho e sobre a remuneração do capital, respectivamente.

[5] Assim como Iacoviello (2005), será suposto que a habitação e o consumo sejam separáveis. Bernanke (1984) estudou o comportamento conjunto do consumo de bens duráveis e não-duráveis e encontrou que a separabilidade entre os dois tipos de bens pode ser considerada uma boa aproximação.

O processo de acumulação do capital pode ser ilustrado por meio da seguinte expressão:

$$K_t^{'} = (1 - \delta_k)K_{t-1}^{'} + I_t^{'} \tag{3.8}$$

sendo, δ_k a taxa de depreciação do capital físico.

$\epsilon_{k,t}$ é o custo de ajustamento do capital, que pode ser representado pela seguinte expressão:

$$\epsilon_{k,t} = \frac{\psi_k}{2\delta_k}\left(\frac{I_t^{'}}{K_{t-1}^{'}} - \delta_k\right)^2 \tag{3.9}$$

3.2.4.2 Famílias Impacientes

Da mesma maneira que as famílias pacientes, a utilidade das famílias impacientes depende de forma positiva do seu nível de consumo, estoque de imóveis e do tempo disponibilizado para o lazer. Então, a função utilidade das famílias impacientes possui a seguinte forma funcional:

$$U_t(C_t^{''}, L_t^{''}, H_t^{''}) = \sum_{t=0}^{\infty}(\beta^{''})^t[alogC_t^{''} + log(1 - L_{p,t}^{''} - L_{g,t}^{''}) + jlogH_t^{''}] \tag{3.10}$$

em que, $\beta^{''} \in$ (0,1) é o fator de desconto das famílias impacientes e $\beta^{''} < \beta^{'}$; $C^{''}$ é o consumo das famílias impacientes, de modo que, ($C_t^{''} = C_{p,t}^{''} + \mu_i C_{G,t}$); $C_{p,t}^{''}$ é o consumo privado das famílias impacientes; $L^{''}$ são as horas trabalhadas pelas famílias impacientes e; μ_i é o parâmetro que mede o grau de substituição entre o consumo privado e o consumo em bens/serviços públicos das famílias impacientes.

Então, as famílias impacientes se deparam com a seguinte restrição orçamentária:

$$(1+\tau^c)C_t^{''} + R_{t-1}b_{t-1}^{''}/\pi_t + (1-\tau^q)q_t\Delta H_t^{''} = (1-\tau^l)(w_{p,t}^{''}L_{p,t}^{''} + w_{g,t}^{''}L_{g,t}^{''}) + b_t^{''} + TR_t \tag{3.11}$$

sendo, $(1 + \tau^c)C_t^{''}$ o nível de consumo total das famílias impacientes; $R_{t-1}b_{t-1}^{''}/\pi_t$ é a rentabilidade em termos reais dos títulos públicos das famílias impacientes; $q_t\Delta H_t^{''}$ é a rentabilidade dos imóveis das famílias impacientes; τ^q é a alíquota do subsídio imobiliário concedido pelo governo para as famílias pobres e; $(1 - \tau^l)(w_{p,t}^{''}L_{p,t}^{''} + w_{g,t}^{''}L_{g,t}^{''})$ representa a renda do trabalho das famílias impacientes.

Além disso, assim como em Iacoviello (2005), as famílias impacientes possuem restrição ao mercado de títulos que pode ser representada pela seguinte equação:

$$R_t b_t^{''} = \pi_{t+1}m_w(w_{p,t+1}^{''}L_{p,t}^{''} + w_{g,t+1}^{''}L_{g,t}^{''}) + \pi_{t+1}m_q(q_{t+1}H_t^{''}) \tag{3.12}$$

em que m_w representa a fração da renda do salário das famílias impacientes dado como garantia para detenção de empréstimos e m_q é a parcela do valor dos imóveis também dados como garantia, esses termos m_w e m_q também são conhecidos como LTV. Se $m_w = 0$ e $m_q = 0$, então as famílias pobres são excluídas do mercado financeiro.

3.2.5 Firmas

3.2.5.1 Firmas Atacadistas

As firmas atacadistas utilizam como insumo o trabalho fornecido pelas famílias pacientes e impacientes ($L_p^{'}, L_g^{'}, L_p^{''}, L_g^{''}$), o capital privado fornecido pelas famílias pacientes ($K^{'}$) e o capital público (K^g). Para esse caso, o setor público auxilia o setor produtivo privado fornecendo infraestrutura por meio do capital público e esse evolui a partir do seguinte processo:

$$K_t^g = (1 - \delta_k)K_{t-1}^g + I_t^g \tag{3.13}$$

Assim, como em Torres (2015), a função de produção das firmas atacadistas é do tipo Cobb-Douglas e pode ser representada pela seguinte expressão:

$$Y_{j,t} = A_t(K_t^{'})^{\alpha}(L_t^{'})^{\eta\gamma}(L_t^{''})^{(1-\eta)\gamma}(K_t^g L_t^g)^{1-\alpha-\gamma} \tag{3.14}$$

em que $L_t^g = L_{g,t}^{'} + L_{g,t}^{''}$.

O preço do bem intermediário pode ser definido como:

$$P_{j,t} = \left(\frac{\psi}{\psi - 1}\right) cm_t \tag{3.15}$$

podemos definir, $\frac{\psi}{\psi-1}$ como *mark-up*, que representa a diferença entre o preço e o custo marginal de produção das firmas atacadistas.

3.2.5.2 Firmas Varejistas

O produto final da economia é produzido pelas firmas varejistas por meio da seguinte tecnologia:

$$Y_t = \left[\int_0^1 Y_{j,t}^{\frac{\psi-1}{\psi}} dj\right]^{\frac{\psi}{\psi-1}} \tag{3.16}$$

com, $Y_{j,t}$ sendo o produto intermediário; ψ é a elasticidade de substituição entre os bens intermediários e $\psi > 1$. Esse método de agregação do bem intermediário é chamado de agregação de Dixit-Stiglitz. Esse parâmetro representa o *mark-up* no mercado de bens.

As firmas varejistas maximizam o lucro sujeito a função de produção, dado o preço dos bens intermediários, $P_{j,t}$, e o preço do bem final P_t.

A demanda pelo bem intermediário j é uma função decrescente do seu preço relativo e uma função crescente em relação à produção do bem final.

O preço do bem final é representado por:

$$P_t = \left[\int_0^1 P_{j,t}^{1-\psi} dj \right]^{\frac{1}{1-\psi}} \tag{3.17}$$

O tipo de rigidez inserido no presente modelo é a chamada rigidez nominal, assim como encontrado em Goodfriend e King (1997), Chari et al. (2000), Kiley (2003), Huang et al. (2004) e Dib (2003), dentre outros. O papel da rigidez de preços na modelagem macroeconômica é de ser uma forma de aumentar a persistência no produto e na inflação.

3.2.6 Governo

3.2.6.1 Política Fiscal

O papel da autoridade fiscal na economia é arrecadar tributos e emitir títulos para financiar o seu investimento público e o seu gasto. A arrecadação tributária do governo (T_t) é composta dos impostos que incidem sobre o consumo das famílias (τ^c), sobre a renda do trabalho (τ^l) e sobre os rendimentos do capital (τ^k). Assim como em Cavalcanti et al. (2018), a restrição orçamentária do governo é definida da seguinte forma:

$$d_t = R_{t-1} d_{t-1} - SP_t \tag{3.18}$$

em que, SP_t representa o superávit primário real do governo; d_t é o valor real da dívida pública ($d_t = \frac{D_t}{P_t} = R_t b_t$). A variável b_t é a quantidade total de títulos públicos ($b_t = b_t^{'} + b_t^{''}$).

O superávit primário é dado pela diferença entre a arrecadação total e a despesa total do governo durante o mesmo período:

$$SP_t = T_t - I_t^g - G_t \tag{3.19}$$

em que:

$$G_t = C_{G,t} + TR_t + \tau^q q_t \Delta H_t^{''} \tag{3.20}$$

$$T_t = \tau^l (w_{p,t}^{'} L_{p,t}^{'} + w_{g,t}^{'} L_{g,t}^{'} + w_{p,t}^{''} L_{p,t}^{''} + w_{g,t}^{''} L_{g,t}^{''}) + \tau^k R_t^k K_t^{'} + \tau^c (C_t^{'} + C_t^{''}) \tag{3.21}$$

sendo, $\tau^l (w_{p,t}^{'} L_{p,t}^{'} + w_{g,t}^{'} L_{g,t}^{'} + w_{p,t}^{''} L_{p,t}^{''} + w_{g,t}^{''} L_{g,t}^{''})$ a receita do governo proveniente da rendas das famílias; $\tau^k R_t^k K_t^{'}$ é a receita do governo advinda da tributação da remuneração do capital físico em posse da famílias pacientes; $\tau^c (C_t^{'} + C_t^{''})$ é a receita do governo obtida a partir do

consumo total das famílias, consequentemente, o consumo governamental também é tributado. O investimento público (I_t^g) é considerado como um choque exógeno.

A variável que busca representar o sentimento da política fiscal será definida no modelo a partir da situação da conjuntura fiscal. Neste caso, no modelo um aumento/redução na dívida do governo provoca uma polaridade negativa/positiva na variável que mede o sentimento da autoridade fiscal[6]. Neste caso, a variável na prática irá mensurar a negatividade líquida da autoridade e podemos definir a variável como:

$$S_t = d_t - d_{t-1} + e_{SF,t} \tag{3.22}$$

em que se a diferença da dívida pública for positiva/negaiva teremos uma negatividade líquida maior/menor, resultando em um sentimento da política fiscal mais pessimista/otimista. $e_{SF,t}$ representa um choque proveniente de mudanças abruptas nos cenário fiscal.

3.2.6.2 Política Monetária

A autoridade monetária adota uma meta para a inflação e determina a taxa de juros por meio de uma regra proposta por Taylor (1993). Como indicado, será considerado duas versões da regra de taxa de juros. Uma versão básica, onde o banco central olha apenas para o nível passado da taxa de juros, e para os desvios da evolução futura e do PIB, a qual chamaremos de "regra básica", enquanto na segunda versão, incluiremos os desvios do preço dos imóveis de seu nível de estado estacionário. Chamaremos esta versão da regra de instrumento de versão ampliada.

A versão básica pode ser representada pela seguinte expressão:

$$\hat{R}_t = \phi_R \hat{R}_{t-1} + (1-\phi_R)[\phi_\pi(E_t(\pi_{t+p}) - \bar{\pi}_t) + \phi_Y E_t(\hat{Y}_{t+z})] + e_{R,t} \tag{3.23}$$

Esta regra especifica que a taxa de juros nominal corrente depende de um componente inercial ou defasado $(\hat{R}_{t-1})$; desvio da inflação esperada da meta definida pela autoridade monetária; o hiato do produto, representado pelo desvio do produto em relação ao seu valor de estado estacionário; e um choque *i.i.d.* da política monetária, $e_{R,t}$. Os subscritos p e z são números inteiros que assumem qualquer valor.

A versão expandida pode ser representada por:

$$\hat{R}_t = \phi_R \hat{R}_{t-1} + (1-\phi_R)[\phi_\pi(E_t(\pi_{t+p}) - \bar{\pi}_t) + \phi_Y E_t(\hat{Y}_{t+z}) + \phi_s E_t(S_{t+z})] + e_{R,t} \tag{3.24}$$

em que S_t é a variável que mede o sentimento fiscal.

[6] Vale ressaltar que escolhemos a dívida pública pois esta é a variável fiscal que já leva em consideração o desempenho das demais.

3.2.7 Choques

O modelo inclui seis choques exógenos: um choque de produtividade agregada, $e_{A,t}$; um choque de política monetária, $e_{R,t}$; um choque de investimento público, $e_{I,t}$; um choque salarial do setor público, $e_{w,t}$; um choque de horas de trabalho no setor público, $e_{l,t}$, um choque de transferências governamentais, $e_{tr,t}$ e um choque de política fiscal, $e_{SF,t}$. Supõe-se que todos os choques são $iid(0, \sigma_l)$ em que $l = A, R, I_g, w_g, l_g, TR$. Os processos estocásticos que definir a evolução de A_t, $e_{R,t}$, $I_{g,t}$, $w_{g,t}$, $L_{g,t}$, TR_t e $e_{SF,t}$ são dados pelas seguintes expressões:

$$A_t = \rho_A A_{t-1} + e_{A,t} \tag{3.25}$$

$$e_{R,t} = \rho_e e_{R,t-1} + \epsilon_{R,t} \tag{3.26}$$

$$I_{g,t} = \rho_I I_{g,t-1} + e_{I,t} \tag{3.27}$$

$$w_{g,t} = \rho_w w_{g,t-1} + e_{w,t} \tag{3.28}$$

$$L_{g,t} = \rho_L L_{g,t-1} + e_{L,t} \tag{3.29}$$

$$TR_t = \rho_{tr} TR_{t-1} + e_{tr,t} \tag{3.30}$$

$$e_{SF,t} = \rho_{sf} e_{SF,t-1} + e_{sf,t} \tag{3.31}$$

3.2.8 Equilíbrio

O equilíbrio nesta economia é caracterizado por um conjunto de alocações de famílias, empresas e governo. Dada a política fiscal adotada pelo governo,

$$\{\tau^c, \tau^l, \tau^k, \tau^q, \gamma_d, \gamma_Y, C_{g,t}, w_{g,t}, TR_t, I_{g,t}\}_{t=0}^{\infty}$$

, o equilíbrio competitivo é caracterizado por um conjunto de decisões das famílias

$$\{C_t^{'}, C_{p,t}^{'}, C_t^{''}, C_{p,t}^{''}, L_{p,t}^{'}, L_{p,t}^{''}, H_t^{'}, H_t^{''}, b_t^{'}, b_t^{''}, I_{p,t}\}_{t=0}^{\infty}$$

; por uma sequência de oferta de capital privado e público $\{K_{p,t}, K_{g,t}\}_{t=0}^{\infty}$; por uma sequência de preços $\{P_t, P_{j,t}, R_{k,t}, q_t, w_{p,t}\}_{t=0}^{\infty}$; pela taxa de juros nominal da economia $\{R_t\}_{t=0}^{\infty}$, que seja compatível com a solução do problema do consumidor, com o problema dos atacadistas e varejistas, com as restrições orçamentárias do governo e com a regra de política monetária. O equilíbrio

também é caracterizado por uma sequência de choques $\{e_{A,t}, \epsilon_{R,t}, e_{I,t}, e_{w,t}, e_{L,t}, e_{tr,t}\}_{t=0}^{\infty}$. Além disso, as seguintes condições são satisfeitas:

$$Y_t = I_{p,t} + C'_{p,t} + C''_{p,t} + G_t$$
$$H'_t + H''_t = 1$$

3.2.9 Estimação Bayesiana

Nesta seção, discutimos nossa metodologia para estimar e avaliar os modelos. A solução do modelo DSGE foi obtida a partir de uma aproximação de Taylor de primeira ordem das condições de equilíbrio em torno do valor de regime permanente não estocástico. Dada a solução do modelo como estado do espaço e vetor de variáveis observáveis, os modelos foram estimados por meio de técnicas bayesianas. Em particular, um algoritmo Metropolis – Hastings, que é um método Monte Carlo Markov Chain (MCMC), foi empregado para obter a distribuição de probabilidade posterior dos parâmetros. Duas sequências independentes foram geradas, cada uma consistindo em 400.000 retiradas, usando o algoritmo Metropolis-Hastings. A aceitação média ao longo das duas cadeias foi em torno de 40%, e a convergência foi avaliada com os métodos propostos por Brooks e Gelman (1998). As primeiras 180.000 retiradas foram descartadas para garantir a independência das condições iniciais. As estatísticas de interesse foram então calculadas com base na distribuição conjunta da probabilidade ergódica posterior dos parâmetros estruturais.

Para a estimação, três variáveis foram utilizadas para cada trimestre: PIB real, taxa de juros nominal e consumo das famílias. Essas variáveis foram escolhidas por serem as variáveis endógenas mais relevantes. As variáveis foram utilizadas em logaritmo natural e ajustadas sazonalmente. O componente cíclico das variáveis foi obtido do filtro Hodrick-Prescott e tem período trimestral iniciando no primeiro trimestre de 2003 e finalizando no segundo trimestre de 2021. O modelo foi estimado usando Dynare dentro do software Matlab.

3.2.10 Calibração e Distribuições a Priori

Alguns parâmetros foram mantidos fixos durante o processo de estimativa, enquanto outros foram estimados. Para os parâmetros que foram mantidos fixos, optamos por usar valores da literatura relacionada (Christiano e Eichenbaum (1992), Lim e McNelis (2008), Silva et al. (2014), Cavalcanti et al. (2018), Wesselbaum (2017)). Tabela 24 em Apêndice (A) apresenta uma breve descrição desses parâmetros.

Para os parâmetros estimados, optou-se por utilizar uma distribuição a priori semelhante às utilizadas na literatura relacionada. Para os parâmetros que indicam o grau de substituição entre o consumo privado e o consumo de bens e serviços públicos, μ_p e μ_i, usamos uma distribuição beta anterior com uma média de 0,50, que consiste no valor encontrado para o Brasil por Ferreira

e Nascimento (2005), Santana et al. (2012) e Bezerra et al. ()[7] com um desvio padrão de 0,02 para ambos. Para os parâmetros da regra de Taylor, usamos uma distribuição a priori e valores para os hiperparâmetros que comumente aparecem na literatura (SMETS; WOUTERS, 2003). O parâmetro que rege a resposta do banco central às mudanças nos preços, ϕ_π, foi definido em 1,5, satisfazendo o princípio de Taylor. Para o coeficiente que mede a resposta do banco central ao hiato do produto, ϕ_Y, usamos uma distribuição normal a priori com uma média de 0,125 (CARVALHO et al., 2013).

Para o parâmetro que indica a participação do capital físico na função de produção, semelhante a Cavalcanti et al. (2018), adotamos uma distribuição normal a priori com média de 0,30 e desvio padrão de 0,05 . Por fim, para todos os parâmetros autorregressivos, empregamos uma distribuição beta a priori com média de 0,95 e desvio padrão de 0,02.

3.3 Resultados

3.3.1 Sentimento da Política Fiscal

Como mencionado nas seções anteriores, foram construídos dois índices de sentimento da política fiscal, sendo uma a partir do dicionário de léxico fixo (SFF) de Loughran e McDonald (2011) e outro via dicionário com palavras variantes no tempo (SFV) de Lima et al. (2019).

Os coeficientes de correlação entres os índices de sentimento e as variáveis da política fiscal e política monetária podem ser visualizados na Tabela 16. O índice SFF apresentou correlação positiva com a dívida pública e negativa com o superávit primário e a taxa de juros selic. Esse fato não parece ser intuitivo, pois espera-se que o sentimento da política fiscal reflita o cenário fiscal, ou seja, um contexto fiscal mais positivo (menor dívida e maior superávit) deve gerar uma comunicação com um tom mais otimista por parte da autoridade fiscal. Obviamente, nada impede da autoridade fiscal em um momento de piora das contas públicas ter um tom mais otimista com o propósito de mitigar a reação negativa dos agentes. Contudo, essa ideia é pouco plausível dada a natureza do Relatório Mensal da Dívida que é um documento técnico de natureza mais analítica.

Tabela 16 – Coeficientes de correlação

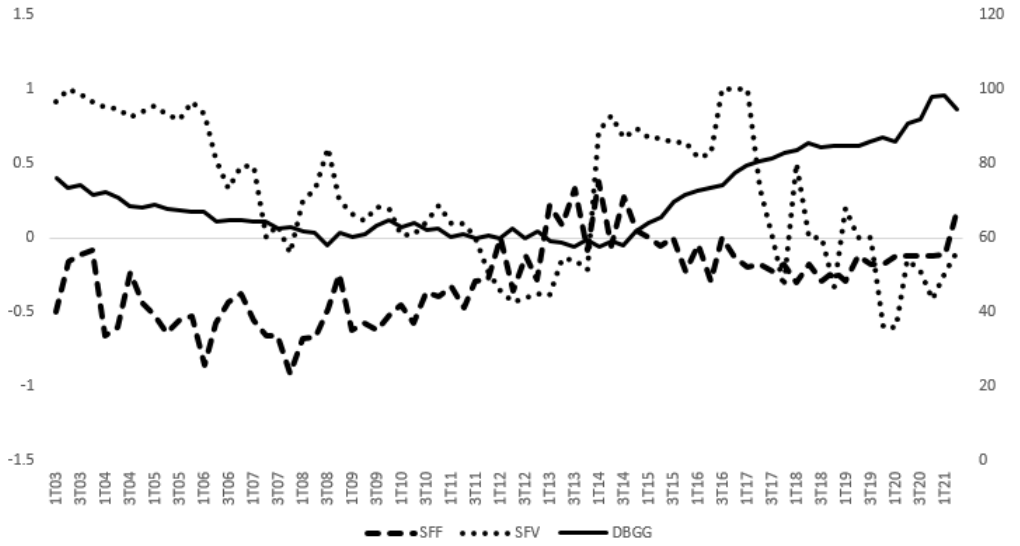
	DBGG/PIB	SUP/PIB	Selic
SFF	0,19	-0,46	-0,18
SFV	-0,17	0,27	0,62

Já o SFV obteve correlações mais intuitivas, em que este possui uma correlação negativa com a dívida pública e positiva com o superávit primário. Esse fato ocorreu pois o SFV utiliza um

[7] Este valor encontrado para o Brasil pode ser considerado conservador. Bailey (1971) e Aschauer (1985) encontraram valores entre 0,23 e 0,42 para os Estados Unidos.

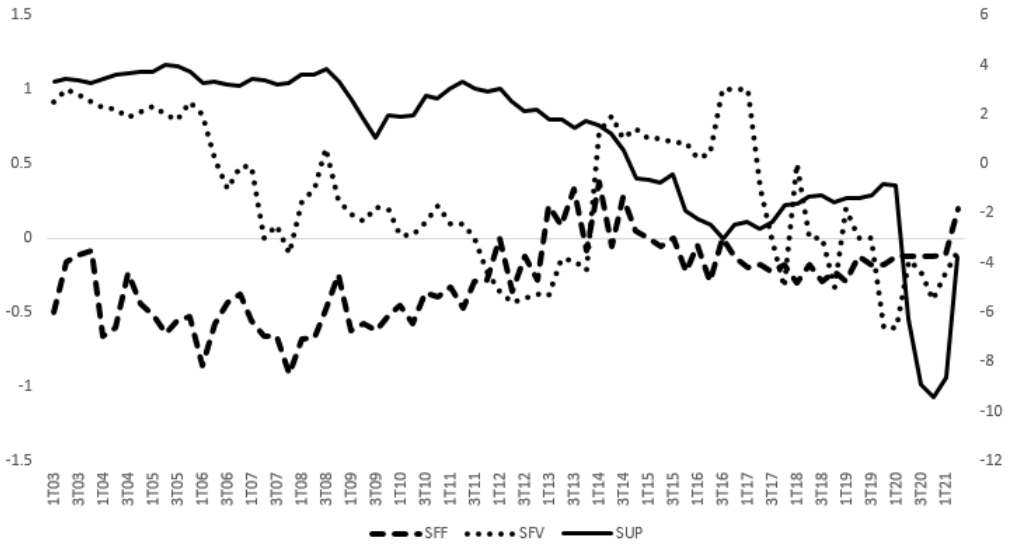
dicionário variante no tempo, que dada a sua natureza, tende a produzir um índice de sentimento mais realista em relação a um dicionário puramente financeiro com rol de palavras fixas e com o mesmo peso das palavras ao longo do tempo. Outro fato que chama a atenção é a forte correlação positiva com a taxa selic com valor de 0,62%. Isso é interessante porque o SFV usa como variável de resposta a dívida pública.

Figura 8 – Trajetória trimestral dos índices de sentimento da política fiscal e da DBGG como proporção do PIB



A Figura 8 e a Figura 9 apresentam a trajetória trimestral dos índices de sentimento juntamente com a DBGG como proporção do PIB e do superávit primário acumulado nos últimos doze meses como proporção do PIB, respectivamente. Primeiramente, é possível observar que na maior parte do tempo o SFV foi mais otimista que o SFF.

Figura 9 – Trajetória trimestral dos índices de sentimento da política fiscal e do superávit primário como proporção do PIB



Também é possível perceber que graficamente (assim também como em termos de correlação) os índices de sentimento se ajustam mais ao superávit primário do que a dívida pública. Isso pode ser um indicativo de que a autoridade fiscal mostra um maior foco no superávit primário em sua comunicação.

3.3.2 Estimações da Função de Reação

A Tabela 17 apresenta as estimações dos parâmetros na forma reduzida da função de reação do Banco Central por MQO e GMM. As colunas 2 e 3 apresentam as estimativas da função de reação por MQO, enquanto que as colunas 4, 5 e 6 apresentam as estimativas da função de reação obtidas a partir do método GMM, sendo que a primeira inclui os índices de sentimento fiscal na função de reação e a segunda restringe a resposta dos juros as variações nos índices de sentimento.

Assim como no trabalho de Silva e Besarria (2018), de forma geral, os resultados obtidos a partir da estimativa da função de reação, independentemente do método, evidenciam um elevado grau de suavização na dinâmica da taxa de juros, indicando que o Banco Central faz mudanças de forma gradual na taxa de juros. Quanto ao coeficiente relacionado às expectativas de inflação, percebe-se que esse é estatisticamente significante e maior que a unidade, indicando que o Banco Central satisfaz o princípio de Taylor, aumentando a taxa de juros real em resposta aos desvios da inflação esperada.

Tabela 17 – Estimação dos parâmetros na forma reduzida da função de reação

	MQO_1	MQO_2	GMM_{IR1}	GMM_{IR2}	GMM_R
$Selic_{t-1}$	0,6831	0,6522	0,6844	0,6923	0,7828
	[0,0544]	[0,0474]	[0,0852]	[0,0942]	[0,0964]
Hiato da Inflação	0,4150	0,4123	0,4280	0,4170	0,5340
	[0,0709]	[0,0689]	[0,1095]	[0,1011]	[0,1197]
Hiato do PIB	0,0121	0,0095	0,0123	0,0119	0,0088
	[0,0019]	[0,0016]	[0,0025]	[0,0023]	[0,0071]
SFF	0,5311	-	0,5553	-	-
	[0,7097]	-	[0,7243]	-	-
SFV	-	1,8569	-	1,8734	-
	-	[0,4872]	-	[0,5302]	-

Nota: Os termos entre colchetes representam os desvios padrão dos coeficientes estimados.

Tabela 18 – Estimação dos parâmetros na forma estrutural da função de reação

	MQO_1	MQO_2	GMM_{IR1}	GMM_{IR2}	GMM_R
$Selic_{t-1}$	0,6831	0,6522	0,6844	0,6923	0,7828
	-	-	-	-	-
Hiato da Inflação	1,3096	1,1855	1,2753	1,3552	2,4585
	-	-	-	-	-
Hiato do PIB	0,0382	0,0273	0,0367	0,0387	0,0405
	-	-	-	-	-
SFF	1,6759	-	1.6546	-	-
	-	-	-	-	-
SFV		5,3390	-	6,0884	-
	-	-	-	-	-

Em relação aos índices de sentimento, as estimativas do MQO e GMM indicam que a variável SFF não possui significância estatística. Já a variável SFV apresentou sinal positivo nos modelos, então, quando o tom das comunicações do Tesouro Nacional é otimista o BCB aumenta a taxa de juros e quando o tom é mais pessimista a autoridade monetária reduz a taxa de juros. Vale também ressaltar que o peso atribuído às variações no sentimento da política fiscal foi superior ao peso relacionado às expectativas de inflação.

Esses resultados indicam que o sentimento da política fiscal afeta de forma significativa a condução da política monetária, mais especificamente com o índice SFV. Vale lembrar que esse índice de sentimento usa em sua construção a dívida pública como variável dependente, e o SFF usa um dicionário financeiro fixo. Então, por esses resultados fica evidente que a conjuntura fiscal (representado pelo índice de sentimento) é uma variável relevante na condução da política monetária. Portanto, tais resultados apontam para a ocorrência de uma possível situação de dominância fiscal.

A despeito desses resultados, a estimação de equações simples apresentam alguns problemas. Como destacado por Lubik e Schorfheide (2007), Finocchiaro e Heideken (2013) e Silva e Besarria (2018), estas estimações sofrem com a presença de endogeneidade quando estimadas por MQO e podem apresentar viés quando estimadas por GMM, em função do tamanho da amostra, viés relacionado ao uso de estágios nas estimações por GMM em dois estágios e GMM iterativo, que é proporcional ao número de condições de momentos em modelos de variáveis instrumentais. Além disso, na prática, encontrar bons instrumentos para implementar o método GMM é não trivial. Instrumentos inválidos ou fracos representam um sério desafio para a boa inferência e podem comprometer as estimativas.

3.3.3 Estimações dos Modelos DSGE

Esta subseção apresenta os resultados da estimação dos modelos DSGE. A Tabela 19 apresenta os valores médios, os desvios padrão e os valores correspondentes aos limites inferiores (MDP inf) e superiores (MDP Sup) do intervalo de credibilidade de 95% de Máxima Densidade a Posteriori (MDP) dos parâmetros estimados utilizando a técnica de inferência Bayesiana para os dois tipos de modelos estimados.

Tabela 19 – Resultados da Estimação Bayesiana

	Priori			Modelo sem Sentimento da Política Fiscal			Modelo com Sentimento da Política Fiscal		
	Dist. a Priori	Média a Priori	Desv. Padrão	Média a Posteriori	MDP Inf.	MDP Sup.	Média a Posteriori	MDP Inf.	MDP Sup.
ρ_A	Beta	0,95	0,02	0,9918	0,9874	0,9964	0,9913	0,9866	0,9958
ρ_r	Beta	0,95	0,02	0,9528	0,9229	0,9799	0,9942	0,9907	0,9979
ρ_{tr}	Beta	0,95	0,02	0,9494	0,9212	0,9822	0,9364	0,9027	0,9643
ρ_I	Beta	0,95	0,02	0,9985	0,9978	0,9994	0,9985	0,9977	0,9995
ρ_L	Beta	0,95	0,02	0,9470	0,9147	0,9732	0,9475	0,9139	0,9836
ρ_w	Beta	0,95	0,02	0,9578	0,9362	0,9850	0,9598	0,9432	0,9769
σ_A	Gama Reversa	0,01	2	0,0118	0,0079	0,0153	0,0233	0,0185	0,0279
σ_r	Gama Reversa	0,01	2	0,0091	0,0023	0,0169	0,0384	0,0316	0,0449
σ_{tr}	Gama Reversa	0,01	2	0,0221	0,0173	0,0266	0,0097	0,0022	0,0188
σ_I	Gama Reversa	0,01	2	0,0089	0,0022	0,0168	0,0120	0,0081	0,0159
σ_L	Gama Reversa	0,01	2	0,0084	0,0023	0,0157	0,0085	0,0023	0,0157
σ_w	Gama Reversa	0,01	2	0,0093	0,0023	0,0173	0,0108	0,0022	0,0210
σ_{sf}	Gama Reversa	0,01	2	-	-	-	0,0089	0,0023	0,0166
μ_p	Beta	0,50	0,02	0,3058	0,2768	0,3360	0,3245	0,2946	0,3490
μ_i	Beta	0,50	0,02	0,3021	0,2697	0,3401	0,3027	0,2788	0,3266
ϕ_R	Beta	0,80	0,10	0,8436	0,7591	0,9459	0,8281	0,7183	0,9335
ϕ_π	Normal	1,50	0,50	1,4539	1,0432	1,9279	0,7083	0,1919	1,2302
ϕ_Y	Normal	0,125	0,05	0,0729	0,0293	0,1153	0,0719	0,0160	0,1268
ϕ_S	Normal	0,300	0,05	-	-	-	0,3118	0,2569	0,3759
α	Normal	0,30	0,05	0,1526	0,0947	0,2040	0,1846	0,1299	0,2388

Assim como encontrado no trabalho de Silva e Besarria (2018), observa-se que os parâmetros estimados sofrem pouca alteração entre os dois modelos com médias posteriores muito próximas entre os dois modelos[8]. Os resultados das estimações revelam que os dados brasileiros são pouco informativos quanto à quantidade de trabalho das famílias pacientes na produção do bem intermediário. Resultado semelhante foi identificado por Finocchiaro e Heideken (2013) nas estimações deste parâmetro para o Reino Unido e Japão, onde o valor da média a posteriori foi exatamente igual ao valor da priori. Para os parâmetros que definem o máximo de empréstimo que famílias impacientes e empreendedores podem tomar, os valores foram mais baixos do que a média a priori. Estas estimativas talvez reflitam o fato de que famílias e empresas no país sejam mais restritas no acesso ao crédito do que no caso dos países desenvolvidos. Os parâmetros dos choques de preferência e tecnologia foram mais altos do que a média a priori, revelando que estes choques são mais persistentes do que a hipótese inicial como definida pelos hiperparâmetros da distribuição a priori. Finocchiaro e Heideken (2013) obtêm resultados semelhantes nas estimações dos processos dos choques para os Estados Unidos, Reino Unido e Japão.

Em relação aos parâmetros da função de reação do banco central, o parâmetro que mede a resposta do banco central a mudanças na expectativa de inflação foi positivo e maior que a unidade, satisfazendo o princípio de Taylor. De mesmo modo, o parâmetro que mede a resposta a desvios do produto foi positivo. Ambos os parâmetros sugerem o comportamento de um banco central operando em um regime de metas de inflação flexível, atribuindo peso tanto à inflação quanto ao lado real da economia. Em relação ao sentimento da política fiscal, a média do parâmetro que reflete a resposta do banco central, ϕ_S, foi positiva e significativa. Portanto, assim como no caso dos modelo de equação simples, há indicações de que o BCB considerou explicitamente o sentimento da política fiscal, ou seja, a conjuntura fiscal em sua função de reação no período de análise.

[8] A Figura 15, Figura 16, Figura 17 e Figura 18 no Apêndice B apresentam as distribuições a priori e a posteriori dos parâmetros.

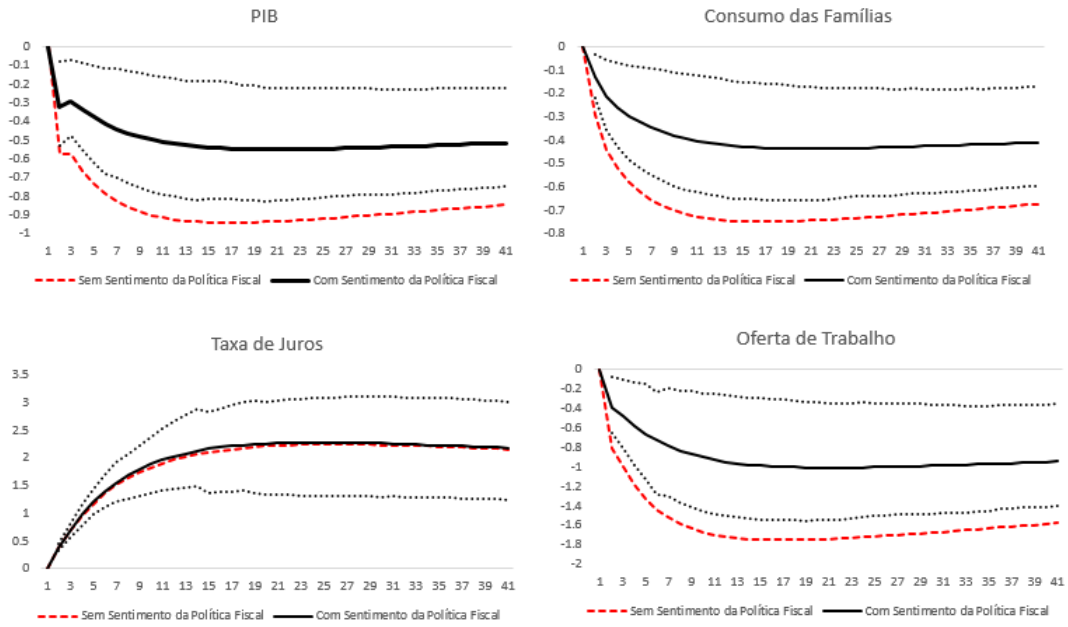
De acordo com Silva e Besarria (2018) uma ferramenta conveniente na análise Bayesiana é o uso das estimativas na comparação entre modelos alternativos. Uma destas formas é utilizar a densidade marginal dos dados associada com cada modelo e compará-las entre si e, por conseguinte, escolher o modelo que é melhor suportado pelos dados. Uma das formas de se obter a densidade marginal dos dados, a partir da distribuição conjunta posterior, é utilizar o estimador de Geweke (1999), o estimador de média harmônica modificado. A Tabela 20 apresenta os valores para a densidade marginal dos dados (em log) computada utilizando este estimador.

Tabela 20 – Comparação entre os Modelos

Especificação	Densidade Marginal dos Dados	Log Fator Bayers
Sem sentimento da política fiscal	147.6836	0
Com sentimento da política fiscal	220.4311	72.7475

A partir dos critérios de avaliação dos modelos criado por Kass e Raftery (1995) encontramos alguma evidência favorável ao modelo com o sentimento da política fiscal. Portanto, com base nas estimações, pode-se dizer de que há evidência, embora limitada, de que o BCB considerou explicitamente na sua decisão de taxa de juros o comportamento da política fiscal.

Figura 10 – Funções de Resposta à Impulso a um choque de um desvio padrão na Taxa Nominal de Juros



Nota: As linhas pontilhadas representam um intervalo de credibilidade de 68% para o caso com sentimento da política fiscal.

Uma análise das funções de resposta (Figura 10 à impulso indica que a inclusão do sentimento da política fiscal na função de reação alteram a transmissão da política monetária sobre o PIB, o consumo das famílias e a oferta de trabalho, a exceção foi a taxa de juros. Também é possível perceber que o aumento na taxa de juros trouxe efeitos recessivos típicos na economia,

mostrando que o choque positivo de juros promoveu uma redução no consumo, oferta de trabalho e demanda agregada, sendo esses efeitos observados, independentemente, de o Banco Central incluir ou não o sentimento da política fiscal na função de reação.

Os resultados obtidos a partir das estimações dos modelos DSGE, embora limitados, vão em linha com o encontrado nas estimaçõs dos modelos MQO e GMM. Ou seja, os resultados indicam que o sentimento da política fiscal é uma variável relevante na tomada de decisão da autoridade monetária sobre a taxa de jurps. Então, essa nova abordagem de investigação indica que existe uma grande possibilidade da ocorrência do fenômeno da dominância fiscal no período analisado. Esses resultados convergem aos encontrados nos trabalhos de Ázara (2006), Junior et al. (2021), Ornellas e Portugal (2011) e Nobrega et al. (2020).

3.4 Considerações Finais

A conjuntura da situção fiscal recente no Brasil levanta a questão se, e em que medida, o Banco Central do Brasil tem reagido a este cenário durante o perído de 2003 a 2021. Este artigo investiga esta problemática através de duas estratégias principais. A primeira consiste na estimação de funções de reação do banco central utilizando equações simples, com a inclusão do sentimento da política fisca como um dos argumentos da equação. A segunda desenvolve um modelo Dinâmico Estocástico de Equilíbrio Geral (DSGE) e usa o modelo para produzir inferências sobre o comportamento da política monetária diante do sentimento da autoridade fiscal.

Os resultados indicam que o banco central incorporou explicitamente, nas suas decisões de política monetária, o comportamento do sentimento da autoridade fiscal. Ao reagir com um sinal positivo ao tom da comunicação da autoridade fiscal, o banco central brasileiro pode ter assumido uma postura de aumentar/reduzir a taxa de juros quando o cenário fiscal foi mais otimista/negativo, dando indícios de uma possível dominância fiscal.

Apesar dos resultados serem promissores, futuras versões precisam incorporar algumas questões a fim de tornar os resultados obtidos mais robustos. Uma primeira questão é a de construir uma variável de sentimento fiscal por meio de dicionários em português para testar se existe grandes alterações nos resultados. Então, para futuros trabalhos no tema, seria interessante o uso do dicionário de Machado et al. (2019). Outro ponto é que além do dicionário de Lima et al. (2019), existem outros dicionários que usam machine learning, então é importante também testar essa outras alternativas para fornecer mais robustez aos resultados.

Também é recomendado que as próximas pesquisas além de testar a inclusão do sentimento da autoridade fiscal também verifiquem a incerteza da política fiscal. No presente trabalho discutimos se o sentimento ou tom da política fiscal afeta as decisões de taxa de juros da autoridade monetária, mas talvez, a incerteza da política também possui relevância ou até mais do que o sentimento. Portanto, recomendamos a inclusão de algum índice de incerteza da política fiscal,

principalmente o Índice de Incerteza Macroeconômica (IIM) criado por Besarria et al. (2021) que é um índice de incerteza construído a partir de *text mining* do Relatório Mensal da Dívida Pública Federal.

Referências

AGUIAR, M. T. de. *Dominância Fiscal e a Regra de Reação Fiscal: Uma Análise Empírica para o Brasil*. 73 p. Dissertação (Mestrado) — Universidade de São Paulo - USP, 2007. Disponível em: <http://www.teses.usp.br/teses/disponiveis/12/12140/tde-19102007-124240/pt-br.php>.

ANDREWS, D. W. Consistent moment selection procedures for generalized method of moments estimation. *Econometrica*, Wiley Online Library, v. 67, n. 3, p. 543–563, 1999.

APEL, M.; GRIMALDI, M. The information content of central bank minutes. *Riksbank Research Paper Series*, n. 92, 2012.

ARAGÓN, E. K. d. S. B.; MEDEIROS, G. B. de. Testing asymmetries in central bank preferences in a small open economy: A study for brazil. *Economía*, Elsevier, v. 14, n. 2, p. 61–76, 2013.

ARAUJO, G. S.; GAGLIANONE, W. P. Machine learning methods for inflation forecasting in brazil: new contenders versus classical models. 2020.

ARAUJO, J. M.; BESARRIA, C. da N. Relações de dominância entre as políticas fiscal e monetária: uma análise para economia brasileira no período de 2003 a 2009. *Revista de Economia*, v. 40, n. 1, 2014.

ASCHAUER, D. Fiscal Policy and Aggregate Demand. *American Economic Review*, v. 75, n. 1, p. 117–127, 1985. Disponível em: <http://www.jstor.org/stable/1812707>.

ÁZARA, A. d. *Dominância fiscal e suas implicações sobre a política monetária no Brasil: uma análise do período 1999-2005*. Tese (Doutorado), 2006.

ÁZARA, A. de. *Dominância Fiscal e Suas Implicações Sobre a Política Monetária no Brasil: Uma Análise do Período 1999-2005*. Dissertação (Mestrado) — Fundação Getúlio Vargas - FGV, 2006. Disponível em: <http://bibliotecadigital.fgv.br/dspace/handle/10438/2022>.

BAI, J.; NG, S. Determining the number of factors in approximate factor models. *Econometrica*, Wiley Online Library, v. 70, n. 1, p. 191–221, 2002.

BAI, J.; NG, S. Forecasting economic time series using targeted predictors. *Journal of Econometrics*, Elsevier, v. 146, n. 2, p. 304–317, 2008.

BAILEY, M. *National Income and the Price Level: Aggregate Supply and Aggregate Demand Model*. New York: Mcgraw-Hill, 1971. Disponível em: <http://www.edb.gov.hk/attachment/en/curriculum-development/kla/pshe/references-and-resources/economics/eco_asad_booklet-12e[1].pdf>.

BAŃBURA, M.; GIANNONE, D.; REICHLIN, L. Large bayesian vector auto regressions. *Journal of applied Econometrics*, Wiley Online Library, v. 25, n. 1, p. 71–92, 2010.

BARBOSA, J. H. d. F. Early warning system para distress bancário no brasil. 2017.

BARBOSA, R. B.; FERREIRA, R. T.; SILVA, T. M. d. Previsão de variáveis macroeconômicas brasileiras usando modelos de séries temporais de alta dimensão. *Estudos Econômicos (São Paulo)*, SciELO Brasil, v. 50, n. 1, p. 67–98, 2020.

BARRO, R. Output Effects of Government Purchases. *Journal of Political Economy*, v. 89, n. 6, p. 1086–1121, 1981. Disponível em: <https://dash.harvard.edu/bitstream/handle/1/3451294/barro_outputeffects.pdf?sequence=4>.

BERNANKE, B. Permanent Income, Liquidity, and Expenditure on Automobiles: Evidence from Panel Data. *The Quarterly Journal of Economics*, v. 99, n. 3, p. 587–614, 1984. Disponível em: <https://www.jstor.org/stable/1885966?seq=1#page_scan_tab_contents>.

BESARRIA, C. C. da N. et al. Incerteza macroeconômica e seus efeitos fiscais: Uma análise a partir de precessamento natural e modelos dinâmicos estocásticos de equilíbrio gerla (dsge). *CADERNOS DE FINANÇAS PÚBLICAS*, v. 21, n. 1, 2021.

BEZERRA, A. et al.

BOX, G. E. et al. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.

BREIMAN, L. Bagging predictors. *Machine learning*, Springer, v. 24, n. 2, p. 123–140, 1996.

BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, n. 1, p. 5–32, 2001.

BROOKS, S.; GELMAN, A. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, v. 7, n. 4, p. 434–455, 1998. Disponível em: <https://www.sciencedirect.com/science/article/pii/0304393283900600?via\%3Dihub>.

BÜHLMANN, P.; YU, B. et al. Analyzing bagging. *The Annals of Statistics*, Institute of Mathematical Statistics, v. 30, n. 4, p. 927–961, 2002.

CANZONERI, M. B.; CUMBY, R. E.; DIBA, B. T. Is the price level determined by the needs of fiscal solvency? *American Economic Review*, v. 91, n. 5, p. 1221–1238, 2001.

CARRIERO, A.; KAPETANIOS, G.; MARCELLINO, M. Forecasting large datasets with bayesian reduced rank multivariate models. *Journal of Applied Econometrics*, Wiley Online Library, v. 26, n. 5, p. 735–761, 2011.

CARVALHO, D.; SILVA, M.; SILVA, I. Efeitos dos choques fiscais sobre o mercado de trabalho brasileiro. *Revista Brasileira de Economia*, v. 67, n. 2, p. 177–200, 2013. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-71402013000200002&lng=pt&nrm=iso&tlng=en>.

CATALFAMO, E. et al. French nowcasts of the us economy during the great recession: A textual analysis. *Research Program on Forecasting Working Paper*, n. 2018-001, 2018.

CAVALCANTI, A. et al. The Macroeconomic Effects of Monetary Policy Shocks under Fiscal Rules Constrained by Public Debt Sustainability. *Economic Modelling*, v. 71, n. 1, p. 184–201, 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0264999317302468>.

CHARI, V.; KEHOE, P.; MCGRATTAN, E. Sticky-Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem? *Econometrica*, v. 68, n. 5, p. 1151–1179, 2000. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1111/ecta.2000.68.issue-5/issuetoc>.

CHOU, C.-H.; HSIEH, S.-C.; QIU, C.-J. Hybrid genetic algorithm and fuzzy clustering for bankruptcy prediction. *Applied Soft Computing*, Elsevier, v. 56, p. 298–316, 2017.

CHOUBIN, B. et al. Precipitation forecasting using classification and regression trees (cart) model: a comparative study of different approaches. *Environmental earth sciences*, Springer, v. 77, n. 8, p. 1–13, 2018.

CHRISTIANO, L.; EICHENBAUM, M. Current Real Business Cycle Theories and Aggregate Labor Market Fluctuations. *American Economic Review*, v. 82, n. 3, p. 430–450, 1992. Disponível em: <https://www.jstor.org/stable/2117314>.

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.

CLEMENTS, M. P.; READE, J. J. Forecasting and forecast narratives: The bank of england inflation reports. *International Journal of Forecasting*, Elsevier, 2020.

CLIMENT, F.; MOMPARLER, A.; CARMONA, P. Anticipating bank distress in the eurozone: An extreme gradient boosting approach. *Journal of Business Research*, Elsevier, v. 101, p. 885–896, 2019.

COLE, R. A.; GUNTHER, J. W. Predicting bank failures: A comparison of on-and off-site monitoring systems. *Journal of Financial Services Research*, Springer, v. 13, n. 2, p. 103–117, 1998.

CONSTANTIN, A.; PELTONEN, T. A.; SARLIN, P. Network linkages to predict bank distress. *Journal of Financial Stability*, Elsevier, v. 35, p. 226–241, 2018.

CURRY, T. J.; ELMER, P. J.; FISSEL, G. S. Equity market data, bank failures and market efficiency. *Journal of Economics and Business*, Elsevier, v. 59, n. 6, p. 536–559, 2007.

DAMASCENO, P. I. d. S. et al. Risco de insolvência e sentimento textual bancário: uma análise dos bancos de capital aberto no brasil. Universidade Federal da Paraíba, 2021.

DIB, A. An Estimated Canadian DSGE Model with Nominal and Real Rigidities. *The Canadian Journal of Economics*, v. 36, n. 4, p. 949–972, 2003. Disponível em: <http://www.jstor.org/stable/3131808>.

DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business & economic statistics*, Taylor & Francis, v. 20, n. 1, p. 134–144, 2002.

DOSSANI, A. Central bank tone and currency risk premia. *Available at SSRN 3304785*, 2019.

EFRON, B. et al. Least angle regression. *The Annals of statistics*, Institute of Mathematical Statistics, v. 32, n. 2, p. 407–499, 2004.

EICKMEIER, S.; ZIEGLER, C. How successful are dynamic factor models at forecasting output and inflation? a meta-analytic approach. *Journal of Forecasting*, Wiley Online Library, v. 27, n. 3, p. 237–265, 2008.

ERDOGAN, B. E. Prediction of bankruptcy using support vector machines: an application to bank bankruptcy. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 83, n. 8, p. 1543–1555, 2013.

EZZAMEL, M.; MAR-MOLINERO, C.; BEECH, A. On the distributional properties of financial ratios. *Journal of Business Finance & Accounting*, Wiley Online Library, v. 14, n. 4, p. 463–481, 1987.

FERREIRA, L. A. M. et al. Dominância fiscal ou dominância monetária no brasil: uma análise do regime de metas de inflação. Universidade Federal de Uberlândia, 2015.

FERREIRA, P.; NASCIMENTO, L. Welfare and Growth Effects of Alternative Fiscal Rules for Infrastructure Investment in Brazil. *Ensaios Econômicos*, n. 604, p. 1–65, 2005. Disponível em: <https://bibliotecadigital.fgv.br/dspace/bitstream/handle/10438/422/1996.pdf>.

FIALHO, M. L.; PORTUGAL, M. S. Monetary and fiscal policy interactions in brazil: an application of the fiscal theory of the price level. *Estudos Econômicos (São Paulo)*, SciELO Brasil, v. 35, n. 4, p. 657–685, 2005.

FILHO, A. E. C.; ROCHA, F. Como o mercado de juros futuros reage à comunicação do banco central? *Economia aplicada*, SciELO Brasil, v. 14, n. 3, p. 265–292, 2010.

FINOCCHIARO, D.; HEIDEKEN, V. V. Do central banks react to house prices? *Journal of Money, Credit and Banking*, v. 45, n. 8, p. 1659–1693, 2013. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/jmcb.12065>.

FORNI, M. et al. The generalized dynamic-factor model: Identification and estimation. *Review of Economics and statistics*, MIT Press, v. 82, n. 4, p. 540–554, 2000.

FORNI, M. et al. Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, Elsevier, v. 50, n. 6, p. 1243–1255, 2003.

FORTUNA, F.; MATURO, F. K-means clustering of item characteristic curves and item information curves via functional principal component analysis. *Quality & Quantity*, Springer, v. 53, n. 5, p. 2291–2304, 2019.

FRANCO, G. *Lições Amargas: Uma História Provisória da Atualidade*. [S.l.]: História Real, 2021.

FREUND, Y.; SCHAPIRE, R.; ABE, N. A short introduction to boosting. *Journal-Japanese Society For Artificial Intelligence*, JAPANESE SOC ARTIFICIAL INTELL, v. 14, n. 771-780, p. 1612, 1999.

GADELHA, S. R. d. B.; DIVINO, J. A. Dominância fiscal ou dominância monetária no brasil? uma análise de causalidade. *Economia Aplicada*, SciELO Brasil, v. 12, p. 659–675, 2008.

GARCIA, M. G.; MEDEIROS, M. C.; VASCONCELOS, G. F. Real-time inflation forecasting with high-dimensional models: The case of brazil. *International Journal of Forecasting*, Elsevier, v. 33, n. 3, p. 679–693, 2017.

GENTZKOW, M.; KELLY, B.; TADDY, M. Text as data. *Journal of Economic Literature*, v. 57, n. 3, p. 535–74, 2019.

GEWEKE, J. Using simulation methods for bayesian econometric models: inference, development, and communication. *Econometric reviews*, Taylor & Francis, v. 18, n. 1, p. 1–73, 1999.

GOLDFARB, R. S.; STEKLER, H. O.; DAVID, J. Methodological issues in forecasting: Insights from the egregious business forecast errors of late 1930. *Journal of Economic Methodology*, Taylor & Francis, v. 12, n. 4, p. 517–542, 2005.

GONZÁLEZ-HERMOSILLO, M. B. *Determinants of ex-ante banking system distress: A macro-micro empirical exploration of some recent episodes*. [S.l.]: International Monetary Fund, 1999.

GOODFRIEND, M.; KING, R. The New Neoclassical Synthesis and the Role of Monetary Policy. *NBER Working Paper*, v. 12, p. 231–296, 1997. Disponível em: <http://www.nber.org/chapters/c11040.pdf>.

GUPTA, A.; SIMAAN, M.; ZAKI, M. When positive sentiment is not so positive: Textual analytics and bank failures. *Available at SSRN 2773939*, 2016.

HOLLAND, M. et al. Monetary and exchange rate policy in brazil after inflation targeting. *Encontro Nacional de Economia*, v. 33, 2005.

HSU, M.-F. A fusion mechanism for management decision and risk analysis. *Cybernetics and Systems*, Taylor & Francis, v. 50, n. 6, p. 497–515, 2019.

HUANG, K.; LIU, Z.; PHANEUF, L. Why Does the Cyclical Behavior of Real Wages Change Over Time? *American Economic Review*, v. 94, n. 4, p. 836–856, 2004. Disponível em: <https://www.jstor.org/stable/3592795?seq=1#page_scan_tab_contents>.

HUANG, Y.-P.; YEN, M.-F. A new perspective of performance comparison among machine learning algorithms for financial distress prediction. *Applied Soft Computing*, Elsevier, v. 83, p. 105663, 2019.

HUBERT, P.; LABONDANCE, F. Central bank sentiment. *URL: https://www. nbp. pl/badania/seminaria/14xi2018. pdf. Working paper*, 2018.

IACOVIELLO, M. House Prices, Borrowing Constraints, and Monetary Policy in the Business Cycle. *American Economic Review*, v. 95, n. 3, p. 739–764, 2005. Disponível em: <https://www2.bc.edu/matteo-iacoviello/research_files/AER_2005.pdf>.

INOUE, A.; KILIAN, L. How useful is bagging in forecasting economic time series? a case study of us consumer price inflation. *Journal of the American Statistical Association*, Taylor & Francis, v. 103, n. 482, p. 511–522, 2008.

ISSLER, J. V.; LIMA, L. R. Public debt sustainability and endogenous seigniorage in brazil: time-series evidence from 1947–1992. *Journal of development Economics*, Elsevier, v. 62, n. 1, p. 131–147, 2000.

JESUS, D. P. de; BESARRIA, C. da N.; MAIA, S. F. The macroeconomic effects of monetary policy shocks under fiscal constrained: An analysis using a dsge model. *Journal of Economic Studies*, Emerald Publishing Limited, 2020.

JONES, J. T.; SINCLAIR, T. M.; STEKLER, H. O. A textual analysis of bank of england growth forecasts. *International Journal of Forecasting*, Elsevier, 2019.

JUNIOR, C. J. C.; GARCIA-CINTADO, A. C.; JUNIOR, K. M. A modern approach to monetary and fiscal policy. *International Review of Economics Education*, Elsevier, p. 100232, 2021.

JUNIOR, K. M. Há dominância fiscal na economia brasileira? uma análise empírica para o período do governo lula. *Indicadores Econômicos FEE*, v. 38, n. 1, 2010.

KADIYALA, K. R.; KARLSSON, S. Numerical methods for estimation and inference in bayesian var-models. *Journal of Applied Econometrics*, Wiley Online Library, v. 12, n. 2, p. 99–132, 1997.

KARELS, G. V.; PRAKASH, A. J. Multivariate normality and forecasting of business bankruptcy. *Journal of Business Finance & Accounting*, Wiley Online Library, v. 14, n. 4, p. 573–593, 1987.

KASS, R.; RAFTERY, A. Bayes factors journal of the american statistical association. 90, 773-795, 1995.

KILEY, M. How Should Unemployment Benefits Respond to the Business Cycle? *Topics in Economic Analysis & Policy*, v. 3, n. 1, 2003. Disponível em: <https://www.degruyter.com/view/j/bejeap.2003.3.issue-1/bejeap.2003.3.1.1066/bejeap.2003.3.1.1066.xml>.

KIM, Y. J.; BAIK, B.; CHO, S. Detecting financial misstatements with fraud intention using multi-class cost-sensitive learning. *Expert Systems with Applications*, Elsevier, v. 62, p. 32–43, 2016.

KOOP, G. M. Forecasting with medium and large bayesian vars. *Journal of Applied Econometrics*, Wiley Online Library, v. 28, n. 2, p. 177–203, 2013.

KUMAR, P. R.; RAVI, V. Bankruptcy prediction in banks and firms via statistical and intelligent techniques–a review. *European journal of operational research*, Elsevier, v. 180, n. 1, p. 1–28, 2007.

KUMHOF, M.; NUNES, R.; YAKADINA, I. Simple monetary rules under fiscal dominance. *Journal of Money, Credit and Banking*, Wiley Online Library, v. 42, n. 1, p. 63–92, 2010.

LEE, T.-H.; YANG, Y. Bagging binary and quantile predictors for time series. *Journal of econometrics*, Elsevier, v. 135, n. 1-2, p. 465–497, 2006.

LEPETIT, L.; STROBEL, F. Bank insolvency risk and time-varying z-score measures. *Journal of International Financial Markets, Institutions and Money*, Elsevier, v. 25, p. 73–87, 2013.

LIM, G.; MCNELIS, P. *Econometric Methods*. [S.l.]: Cambridge: The MIT Press, 2008.

LIMA, L. R.; GODEIRO, L.; MOHSIN, M. Time-varying dictionary and the predictive power of fed minutes. *Available at SSRN 3312483*, 2019.

LOPES, K. C. et al. Preferências assimétricas variantes no tempo na função perda do banco central do brasil. Universidade Federal da Paraíba, 2012.

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.

LOUGHRAN, T.; MCDONALD, B. Textual analysis in accounting and finance: A survey. *Journal of Accounting Research*, Wiley Online Library, v. 54, n. 4, p. 1187–1230, 2016.

LUBIK, T. A.; SCHORFHEIDE, F. Do central banks respond to exchange rate movements? a structural investigation. *Journal of Monetary Economics*, Elsevier, v. 54, n. 4, p. 1069–1087, 2007.

LUNDQUIST, K.; STEKLER, H. O. Interpreting the performance of business economists during the great recession. *Business Economics*, Springer, v. 47, n. 2, p. 148–154, 2012.

MACHADO, M. A. V. et al. Índice de sentimento textual: uma análise empírica do impacto das notícias sobre risco sistemático. *Revista Contemporânea de Contabilidade*, v. 16, n. 40, p. 24–42, 2019.

MATHY, G.; STEKLER, H. Was the deflation of the depression anticipated? an inference using real-time data. *Journal of Economic Methodology*, Taylor & Francis, v. 25, n. 2, p. 117–125, 2018.

MCGRATTAN, E. The Macroeconomic Effects of Distortionary Taxation. *Journal of Monetary Economics*, v. 33, n. 3, p. 573–601, 1994. Disponível em: <https://www.sciencedirect.com/science/article/pii/0304393294900442>.

MEDEIROS, M. C.; MENDES, E. F. 1-regularization of high-dimensional time-series models with non-gaussian and heteroskedastic errors. *Journal of econometrics*, Elsevier, v. 191, n. 1, p. 255–271, 2016.

MEDEIROS, M. C. et al. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics*, Taylor & Francis, p. 1–22, 2019.

MEDEIROS, R. K. d. et al. Mudanças nas preferências do banco central: um estudo empírico para o brasil. Universidade Federal da Paraíba, 2018.

MEDHAT, W.; HASSAN, A.; KORASHY, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams engineering journal*, Elsevier, v. 5, n. 4, p. 1093–1113, 2014.

MEINSHAUSEN, N.; YU, B. et al. Lasso-type recovery of sparse representations for high-dimensional data. *The annals of statistics*, Institute of Mathematical Statistics, v. 37, n. 1, p. 246–270, 2009.

MINCER, J. A.; ZARNOWITZ, V. The evaluation of economic forecasts. In: *Economic forecasts and expectations: Analysis of forecasting behavior and performance*. [S.l.]: NBER, 1969. p. 3–46.

MINELLA, A. et al. Inflation targeting in brazil: constructing credibility under exchange rate volatility. *Journal of international Money and Finance*, Elsevier, v. 22, n. 7, p. 1015–1040, 2003.

MOL, C. D.; GIANNONE, D.; REICHLIN, L. Forecasting using a large number of predictors: Is bayesian shrinkage a valid alternative to principal components? *Journal of Econometrics*, Elsevier, v. 146, n. 2, p. 318–328, 2008.

NOBREGA, W. C. L.; MAIA, S. F.; BESARRIA, C. da N. Interação entre as políticas fiscal e monetária: uma análise sobre o regime de dominância vigente na economia brasileira. *Análise Econômica*, v. 38, n. 75, 2020.

ORNELLAS, R.; PORTUGAL, M. S. Fiscal and Monetary Policy Interaction in Brazil. *XXXIII Encontro Brasileiro de Econometria, 2011, Foz do Iguaçú. Anais do XXXIII Encontro Brasileiro de Econometria. Rio de Janeiro: Sociedade Brasileira de Econometria (SBE)*, 2011. Disponível em: <http://www.bcu.gub.uy/Comunicaciones/Jornadas%20de%20Economa/t_portugal_marcelo_2011_.pdf>.

ORNELLAS, R. da S. Fiscal and monetary interaction in brazil. In: *33º Meeting of the Brazilian Econometric Society*. [S.l.: s.n.], 2011.

PAULE-VIANEZ, J.; GUTIÉRREZ-FERNÁNDEZ, M.; COCA-PÉREZ, J. L. Prediction of financial distress in the spanish banking system: An application using artificial neural networks. *Applied Economic Analysis*, Emerald Publishing Limited, 2019.

PROVENCHER, B.; BAERENKLAU, K. A.; BISHOP, R. C. A finite mixture logit model of recreational angling with serially correlated random utility. *American Journal of Agricultural Economics*, Wiley Online Library, v. 84, n. 4, p. 1066–1075, 2002.

RAVI, V. et al. Soft computing system for bank performance prediction. *Applied soft computing*, Elsevier, v. 8, n. 1, p. 305–315, 2008.

ROSA, P. S.; GARTNER, I. R. Financial distress in brazilian banks: an early warning model. *Revista Contabilidade & Finanças*, SciELO Brasil, v. 29, p. 312–331, 2017.

SANTANA, P.; CAVALCANTI, T.; PAES, N. Impactos de Longo Prazo de Reformas Fiscais Sobre a Economia Brasileira. *Revista Brasileira de Economia*, v. 66, n. 2, p. 247–269, 2012. Disponível em: <http://www.scielo.br/pdf/rbe/v66n2/a06v66n2.pdf>.

SCHYMURA, L. G. A sombra da dominância fiscal e a reação do sistema político. *Revista Conjuntura Econômica*, v. 69, n. 11, p. 8–10, 2015.

SHAPIRO, A. H.; SUDHOF, M.; WILSON, D. J. Measuring news sentiment. *Journal of Econometrics*, Elsevier, 2020.

SHAPIRO, A. H.; WILSON, D. Taking the fed at its word: A new approach to estimating central bank objectives using text analysis. In: FEDERAL RESERVE BANK OF SAN FRANCISCO. [S.l.], 2019.

SHIN, K.-S.; LEE, T. S.; KIM, H.-j. An application of support vector machines in bankruptcy prediction model. *Expert systems with applications*, Elsevier, v. 28, n. 1, p. 127–135, 2005.

SILVA, M. E. A. d.; BESARRIA, C. d. N. Política monetária e preços dos imóveis no brasil: Uma análise a partir de um modelo dsge. *Revista Brasileira de Economia*, SciELO Brasil, v. 72, n. 1, p. 117–143, 2018.

SILVA, P. H. N.; BESARRIA, C. d. N.; SILVA, M. D. d. O. P. da. *Mensurando o sentimento de incerteza da política econômica*. São Paulo: ANPEC - Associação Nacional dos Centros de Pós-Graduação em Economia, 2019. 20 p. (Anais do 47º Encontro Nacional de Economia). Disponível em: <https://www.anpec.org.br/encontro/2019/submissao/files{_}I/i4-6d0fea421a2127dcd80097f52edc812f.>

SILVA, R. A.; RIBEIRO, E. S.; MATIAS, A. B. Aprendizagem estatística aplicada à previsão de default de crédito. *Revista de Finanças Aplicadas*, v. 7, n. 2, p. 1–19, 2016.

SILVA, W.; PAES, N.; OSPINA, R. A Substituição da Contribuição Patronal para o Faturamento: Efeitos Macroeonômicos, sobre a Progressividade e Distribuição de Renda no Brasil. *Revista Brasileira de Economia*, v. 68, n. 4, p. 517–545, 2014. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/rbe/article/view/14269>.

SIMS, C. A. Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, JSTOR, p. 1–48, 1980.

SMETS, F.; WOUTERS, R. An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, v. 5, n. 1, p. 1123–1175, 2003. Disponível em: <https://academic.oup.com/jeea/article-abstract/1/5/1123/2280815?redirectedFrom=fulltext>.

SMETS, F.; WOUTERS, R. Shocks and Frictions in US Business Cycles: a Bayesian DSGE Approach. *American Economic Review*, v. 97, n. 3, p. 586–606, 2007. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/aer.97.3.586>.

STEKLER, H.; SYMINGTON, H. Evaluating qualitative forecasts: The fomc minutes, 2006–2010. *International Journal of Forecasting*, Elsevier, v. 32, n. 2, p. 559–570, 2016.

STOCK, J. H.; WATSON, M. W. Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, Taylor & Francis, v. 97, n. 460, p. 1167–1179, 2002.

SUN, J.; LI, H. Financial distress prediction using support vector machines: Ensemble vs. individual. *Applied Soft Computing*, Elsevier, v. 12, n. 8, p. 2254–2265, 2012.

SUSS, J.; TREITEL, H. Predicting bank distress in the uk with machine learning. Bank of England Working Paper, 2019.

TAHERKHANI, A.; COSMA, G.; MCGINNITY, T. M. Adaboost-cnn: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning. *Neurocomputing*, Elsevier, v. 404, p. 351–366, 2020.

TANNER, E.; RAMOS, A. M. Fiscal sustainability and monetary versus fiscal dominance: Evidence from brazil, 1991–2000. *Applied Economics*, Taylor & Francis, v. 35, n. 7, p. 859–873, 2003.

TAYLOR, J. Discretion Versus Policy Rules in Practice . *Carnegie-Rochester Conference Series on Public Policy*, v. 39, p. 195–214, 1993. Disponível em: <http://www.nviegi.net/teaching/taylor2.pdf>.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, Wiley Online Library, v. 58, n. 1, p. 267–288, 1996.

TORRES, J. *Introduction to Dynamic Macroeconomic General Equilibrium Models*. Vernon Art & Science: Incorporated, 2015.

VAPNIK, V. *The nature of statistical learning theory*. [S.l.]: Springer science & business media, 1995.

VAPNIK, V. N. An overview of statistical learning theory. *IEEE transactions on neural networks*, IEEE, v. 10, n. 5, p. 988–999, 1999.

VARELLA, J. L.; QUADRELLI, G. Redes neurais e análise de potência. *Revista de Tecnologia Aplicada*, v. 6, n. 3, 2017.

VIEIRA, C. A. M.; SILVA, R. R. da; FLORÊNCIO, D. B. Complexidade e risco dos conglomerados financeiros operantes no brasil complexity and risk in brazilian banking holding companies. *Revista BASE–v*, v. 17, n. 2, 2020.

VISWANATHAN, P.; SRINIVASAN, S.; HARIHARAN, N. Predicting financial health of banks for investor guidance using machine learning algorithms. *Journal of Emerging Market Finance*, SAGE Publications Sage India: New Delhi, India, v. 19, n. 2, p. 226–261, 2020.

WESSELBAUM, D. Expectation shocks and fiscal rules. *International Economics and Economic Policy*, v. 14, p. 1–21, 2017. Disponível em: <https://link.springer.com/article/10.1007/s10368-017-0389-z>.

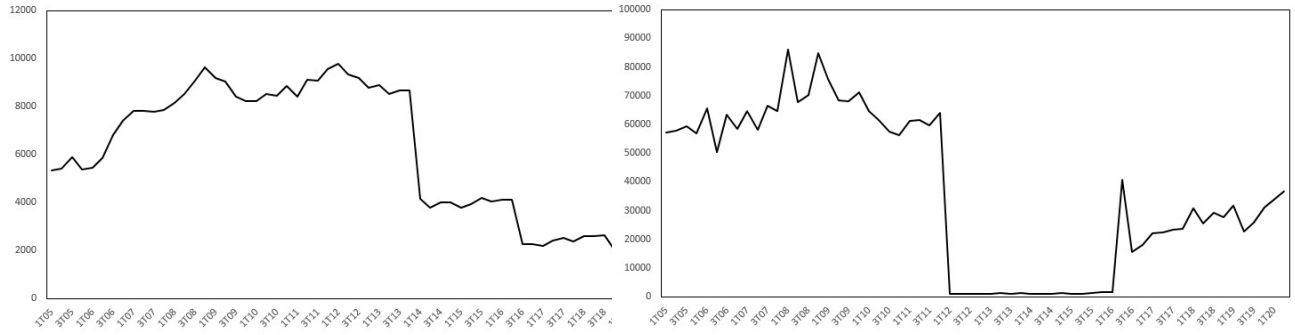
XIA, Y. et al. A boosted decision tree approach using bayesian hyper-parameter optimization for credit scoring. *Expert Systems with Applications*, Elsevier, v. 78, p. 225–241, 2017.

ZHAO, P.; YU, B. On model selection consistency of lasso. *Journal of Machine learning research*, v. 7, n. Nov, p. 2541–2563, 2006.

Apêndices

APÊNDICE A – Capítulo 1

Figura 11 – Quantidade de palavras na ata do Copom e no Relatório da Inflação

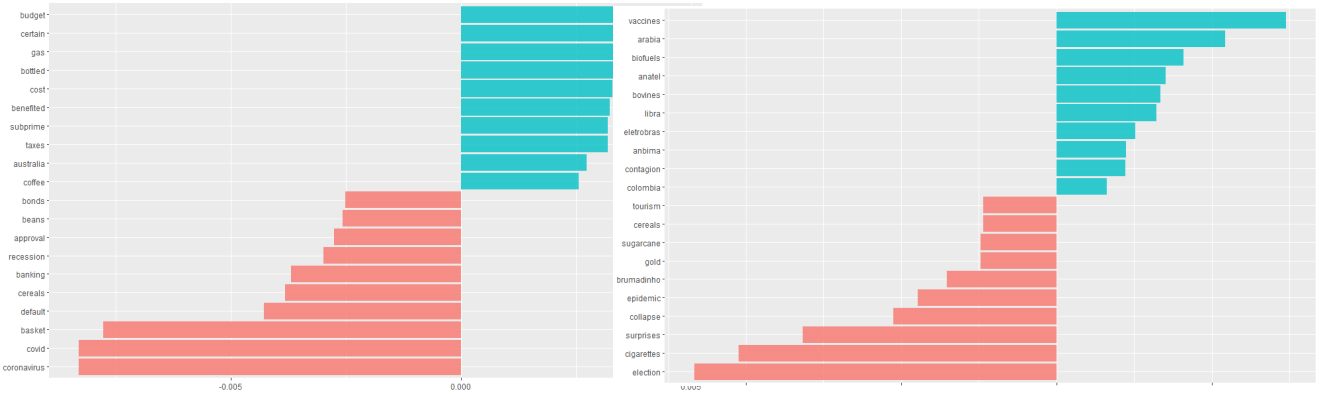


Copom Minutes | Inflation Report

Figura 12 – Nuvem de palavras - ata do Copom e Relatório de Inflação

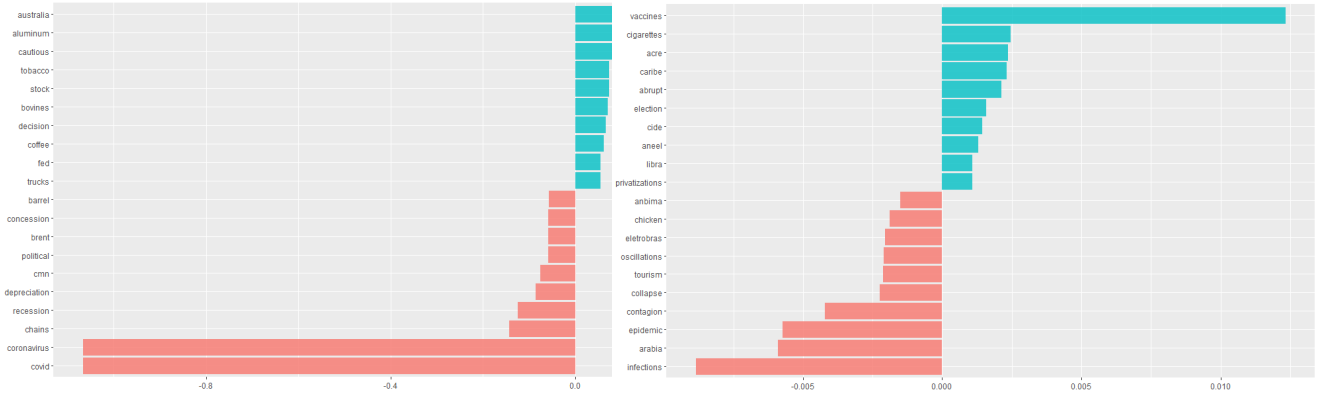
Copom Minutes | Inflation Report

Figura 13 – Os 10 coeficientes mais positivos e negativos da ata do Copom e do Relatório da Inflação para o IPCA



Copom Minutes | Inflation Report

Figura 14 – Os 10 coeficientes mais positivos e negativos da ata do Copom e do Relatório da Inflação para o PIB



Copom Minutes Inflation Report

Tabela 21 – Correlação entre os índices de sentimento e as variáveis macroeconômicas

	$IPCA_t$	$IPCA_{t+1}$	PIB_t	PIB_{t+1}
CM_Index	0,17	-0,08	-0,15	-0,39
IR_Index	0,15	-0,05	0,18	0,13
IPCA_CM_Index	0,45	0,52	-	-
IPCA_IR_Index	0,35	0,57	-	-
GDP_CM_Index	-	-	0,75	0,75
GDP_IR_Index	-	-	0,41	0,71

Tabela 22 – Teste de Diebold-Mariano para os três melhores modelos de previsão do IPCA

Ata do Copom				Relatório da Inflação			
h = 1	LASSO1	LASSO2	LASSO3	h = 1	LASSO2	LASSO1	LASSO3
LASSO1	-	0,9803	0,3423	LASSO2	-	0,6823	0,3095
LASSO2	0,9803	-	0,3206	LASSO1	0,6823	-	0,588
LASSO3	0,3423	0,3206	-	LASSO3	0,3095	0,588	-
h = 2	LASSO1	SVM2	LASSO2	h = 2	LASSO1	SVM1	SVM2
LASSO1	-	0,2543	0,2296	LASSO1	-	0,7853	0,5341
SVM2	0,2543	-	0,9383	SVM1	0,7853	-	0,8787
LASSO2	0,2296	0,9383	-	SVM2	0,5341	0,8787	-
h = 3	LASSO1	LASSO2	SVM2	h = 3	LASSO1	LASSO2	LASSO3
LASSO1	-	0,1294	0,0623	LASSO1	-	0,9004	0,6923
LASSO2	0,1294	-	0,4732	LASSO2	0,9004	-	0,7299
SVM2	0,0623	0,4732	-	LASSO3	0,6923	0,7299	-
h = 4	LASSO1	LASSO2	SVM2	h = 4	LASSO2	LASSO1	LASSO3
LASSO1	-	0,1113	0,0023	LASSO2	-	0,6811	0,6166
LASSO2	0,1113	-	0,253	LASSO1	0,6811	-	0,9177
SVM2	0,0023	0,253	-	LASSO3	0,6166	0,9177	-

Tabela 23 – Teste de Diebold-Mariano para os três melhores modelos de previsão do PIB

Ata do Copom				Relatório da Inflação			
h = 1	RF1	SVM1	RF2	h = 1	SVM1	RF1	RF2
RF1	-	0,9793	0,7632	SVM1	-	0,8994	0,7471
SVM1	0,9793	-	0,8182	RF1	0,8994	-	0,8802
RF2	0,7632	0,8182	-	RF2	0,7471	0,8802	-
h = 2	SVM1	RF2	SVM2	h = 2	SVM1	RF3	SVM4
SVM1	-	0,7871	0,0020	SVM1	-	0,0114	0,0000
RF2	0,7871	-	0,0398	RF3	0,0114	-	0,3612
SVM2	0,0020	0,0398	-	SVM4	0,0000	0,3612	-
h = 3	RF1	RF3	SVM1	h = 3	RF2	SVM1	RF4
RF1	-	0,8509	0,8033	RF2	-	0,0555	0,0398
RF3	0,8509	-	0,9596	SVM1	0,0555	-	0,9932
SVM1	0,8033	0,9596	-	RF4	0,0398	0,9932	-
h = 4	LASSO1	RF1	RF3	h = 4	RF1	SVM1	LASSO2
LASSO1	-	0,9625	0,8472	RF1	-	0,8733	0,2154
RF1	0,9625	-	0,9483	SVM1	0,8733	-	0,2994
RF3	0,8472	0,9483	-	LASSO2	0,2154	0,2994	-

APÊNDICE B – Capítulo 2

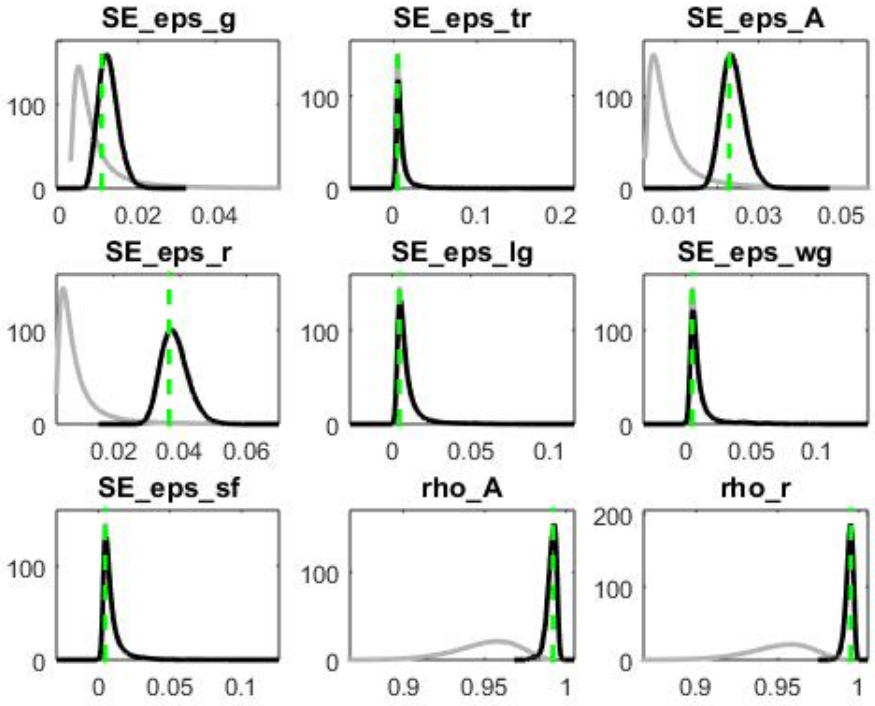
APÊNDICE C – Capítulo 3

Tabela 24 – Calibração dos Parâmetros

Parâmetros	**Descrição**	**Valor**
θ	Fator de rigidez dos preços	0,85
δ_k	Taxa de depreciação do capital físico	0,02
ψ_k	Ajuste de capital físico	2,00
ψ	Elasticidade de substituição entre bens intermediários	6,00
m_w	Proporção do salário usada como garantia	0,90
m_q	Proporção do valor do imóvel usado como garantia	0,85
τ^c	Taxa de imposto sobre o consumo doméstico	0,2313
τ^l	Taxa de imposto sobre a renda do trabalho	0,1713
τ^k	Taxa de imposto sobre a renda de capital	0,1441
$\beta^{'}$	Fator de desconto das famílias dos pacientes	0,99
$\beta^{''}$	Fator de desconto de famílias impacientes	0,94

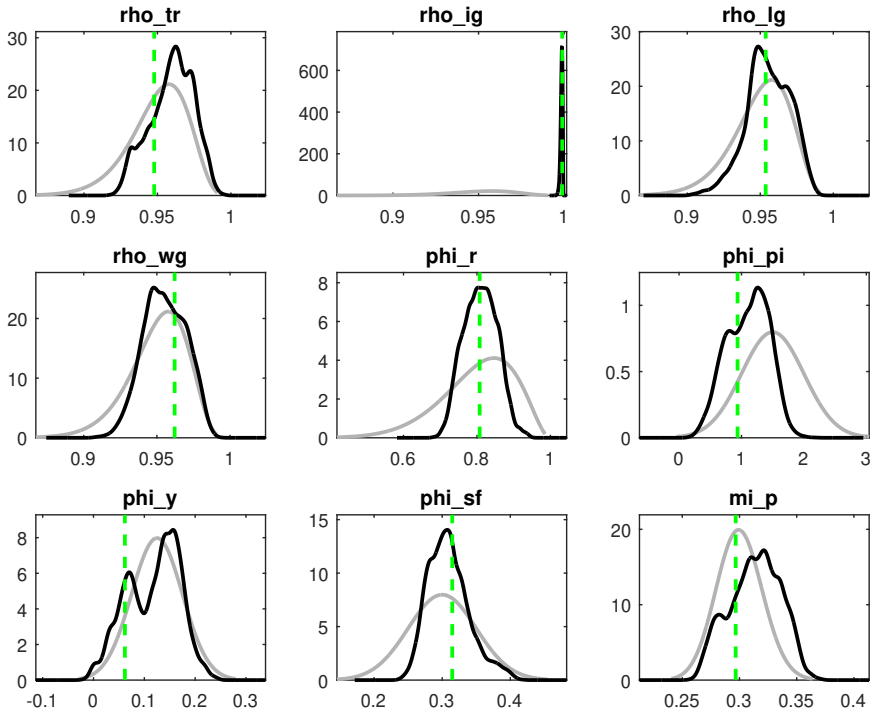
Appendix B - Distribuição a Priori e a Posteriori dos Parâmetros

Figura 15 – Distribuições a Priori e a Posteriori: Modelo com sentimento da política fiscal



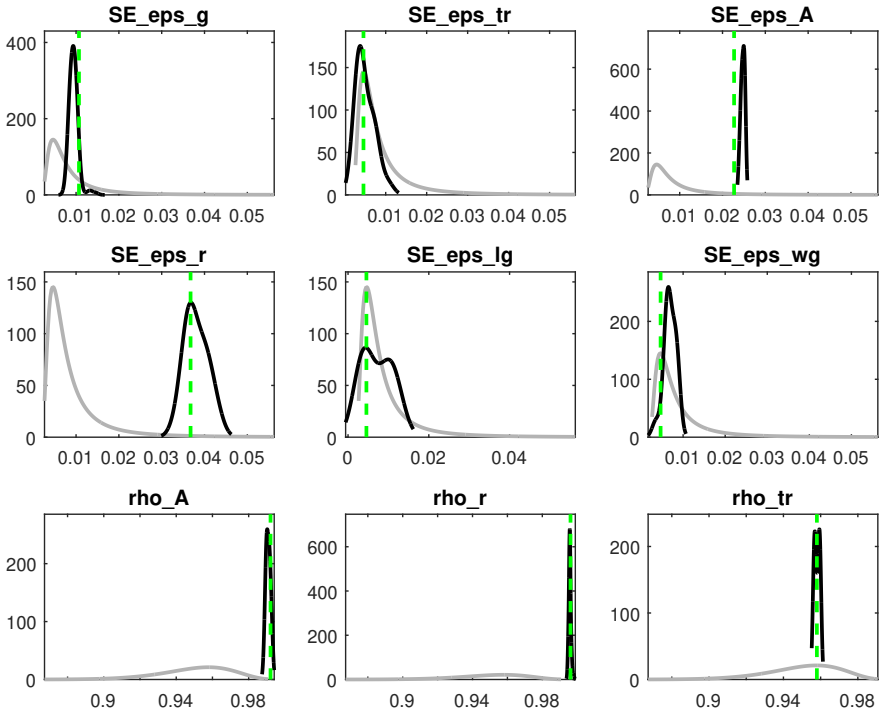
Nota: A linha tracejada representa a moda posterior obtida no processo de otimização, enquanto a curva em preto representa a distribuição a posteriori e a curva mais clara a distribuição a priori dos parâmetros

Figura 16 – Distribuições a Priori e a Posteriori: Modelo com sentimento da política fiscal



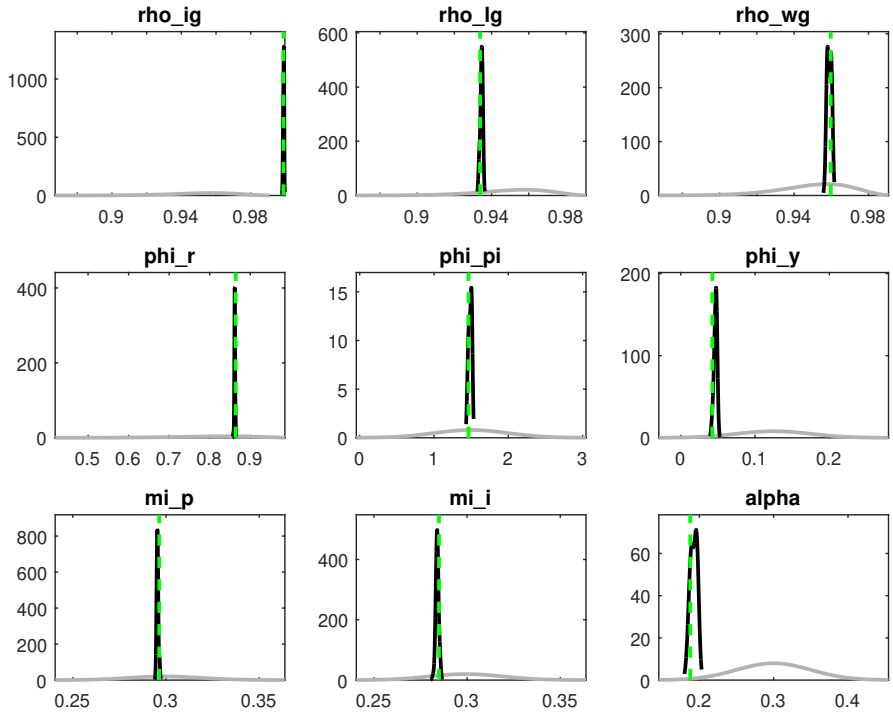
Nota: A linha tracejada representa a moda posterior obtida no processo de otimização, enquanto a curva em preto representa a distribuição a posteriori e a curva mais clara a distribuição a priori dos parâmetros

Figura 17 – Distribuições a Priori e a Posteriori: Modelo sem sentimento da política fiscal



Nota: A linha tracejada representa a moda posterior obtida no processo de otimização, enquanto a curva em preto representa a distribuição a posteriori e a curva mais clara a distribuição a priori dos parâmetros

Figura 18 – Distribuições a Priori e a Posteriori: Modelo sem sentimento da política fiscal



Nota: A linha tracejada representa a moda posterior obtida no processo de otimização, enquanto a curva em preto representa a distribuição a posteriori e a curva mais clara a distribuição a priori dos parâmetros