

PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

Faculdade de Economia, Administração, Contabilidade e Atuária

Curso de Ciências Econômicas

Alberto Freire da Silva

RA00231490

**COMUNICAÇÃO DO BANCO CENTRAL: UMA ANÁLISE QUALITATIVA COM
MACHINE LEARNING**

São Paulo/SP

2023



ALBERTO FREIRE DA SILVA

**COMUNICAÇÃO DO BANCO CENTRAL: UMA ANÁLISE QUALITATIVA COM
MACHINE LEARNING**

Monografia submetida à apreciação de Banca Examinadora do Departamento de Economia, como exigência parcial para a obtenção do grau de bacharel em ciências econômicas.

Orientador: Prof. Dr. Carlos Eduardo Ferreira de Carvalho

São Paulo

2023

O autor desta obra autoriza sua publicação eletrônica na Biblioteca Digital da PUC-SP.

Este trabalho é somente para uso privado de atividades de pesquisa e ensino. Não é autorizada sua reprodução para quaisquer fins lucrativos. Esta reserva de direitos abrange a todos os dados do documento bem como seu conteúdo. Na utilização ou citação de partes do documento é obrigatório mencionar nome da pessoa autora do trabalho e demais itens da referência bibliográfica.

Ficha Catalográfica

Silva, Alberto Freire da .

COMUNICAÇÃO DO BANCO CENTRAL: UMA ANÁLISE QUALITATIVA COM MACHINE LEARNING, 2023

35 p.

Trabalho de Conclusão de Curso de Graduação em Ciências Econômicas

Orientador: Carlos Eduardo Ferreira de Carvalho.

1. Política Monetária 2. Macroeconomia, 3. Machine Learning. Pontifícia Universidade Católica de São Paulo.
Faculdade de Economia, Administração, Contabilidade e Atuária

AGRADECIMENTOS

Primeiramente, agradeço A Deus por todas a benções concedidas e àquelas ainda por vir. Em segundo lugar, aos meus pais – Cátia e Marcos – pelos ensinamentos, paciência e amor. Em terceiro, mas em mesmo grau de importância, aos meus avós maternos – Sebastião e Maria – e aos meus avós paternos – Paulo Roberto e Regina – e ao meu falecido avô Alberto.

Aos amigos de sala de aula, em especial ao Bruno, e a todos os parceiros de ritmo da Bateria FEA PUC, obrigado por fazerem esses 5 anos inesquecíveis.

À PUC e aos professores da FEA, obrigado por todos os ensinamentos e pela contribuição ao meu desenvolvimento intelectual. Ao professor Carlos Eduardo, por todas as conversas, que tangenciavam desde eleições presidenciais até desenvolvimento econômico.

RESUMO

O text mining, ferramenta já consolidada em diversas áreas das ciências sociais aplicadas, destaca-se como um recurso interessante na análise qualitativa de textos de diversos gêneros, desde discursos presidenciais até capítulos de obras fictícias. No âmbito da política monetária, essa abordagem demonstra um potencial significativo para aprimorar a qualidade das análises dos comunicados dos bancos centrais, uma vez que tem a capacidade de mitigar a influência da subjetividade na interpretação.

Este trabalho direciona seu foco para a exploração desses recursos, aplicando-os especificamente às atas do COPOM (Comitê de Política Monetária), abrangendo desde a gestão de Henrique Meirelles até parte da gestão de Campos Neto. O objetivo é demonstrar como tais ferramentas podem contribuir de maneira significativa para a compreensão de temas sensíveis, tais como o cenário corrente e prospectivo, balanço de riscos, e o estado da natureza e os objetivos subjacentes da política monetária.

Palavras chave: Macroeconomia; Política Monetária; Machine Learning, Comunicação, Text Mining, História econômica



LISTA DE GRÁFICOS

Figura 1 – Tokenização de um texto	24
Figura 2 – Processo de tidyzing de um texto.....	25
Figura 3 – Mudança na Selic em t-1 e tom da ata do Copom em t.....	28
Figura 4 – Tf-idf aplicado às atas de 2021	30
Figura 5 – Modelagem de tópicos e seus termos.....	31
Figura 6 – Mapa de calor dos temas para cada presidente	32

SUMÁRIO

1 INTRODUÇÃO	8
1.1 OBJETIVOS	10
1.1.1 Objetivo geral	10
1.1.2 Objetivos específicos	10
1.3 METODOLOGIA	13
2 MACROECONOMIA E O PAPEL DOS BANCOS CENTRAIS	17
2.1.1 CRIAÇÃO DO BCB E REGIME ATUAL	17
2.1.2 REVOLUÇÃO DAS EXPECTATIVAS RACIONAIS	20
2.1.2 NOVOS MODELOS DE POLÍTICA ECONOMICA	22
3.0 TEXT MINING	23
3.1 O PROCESSO BÁSICO	23
3.2 STOP WORDS	25
3.3 STEMMING	25
4. FERRAMENTAS APLICADAS.....	26
4.1 ANÁLISE DE SENTIMENTO	26
4.2 TERM FREQUENCY.....	29
4.3 TOPIC MODELLING	30
Bibliografia	34

1 INTRODUÇÃO

A razão de escolha desse trabalho é a importância adquirida pela política monetária no tempo recente. O mercado financeiro acompanha a decisão de taxa de juros do BCB, FED e BCE impreterivelmente, em especial em períodos de dinâmicas inflacionárias desafiadoras como em 2022, no qual os Bancos Centrais adquiriram papel fundamental no rumo que a economia mundial tomará.

No momento em que esta introdução está sendo escrita, a Zona do Euro acumula 9.1% de inflação acumulada em 12 meses, 33% para energia, em meio a chegada do inverno europeu e maiores sanções ao comércio do gás natural russo (EUROSTAT STATISTICS, 2022). Nesse cenário, o BCE tenta de toda a forma comunicar seu compromisso em parar a dinâmica inflacionária a despeito do risco de jogar a economia europeia em estagflação aos moldes da crise dos anos 70, e no dia 8 de setembro aumentou a taxa básica de juro em 75 p.p, o maior aumento da história e a maior taxa desde 2011 (EUROPEAN CENTRAL BANK, 2022). Paralelo a isso, o Banco Central do Brasil, por meio de seu presidente, Roberto Campos Neto, levou volatilidade ao mercado ao indicar que o ciclo de aperto monetário não acabou – abrindo a possibilidade de um ajuste residual (ESTADÃO CONTEÚDO, 2022) – apesar de, no início de setembro de 2022, o juro real estimado *ex-ante* estar em 8% e as expectativas de inflação estarem em queda pela 13ª semana seguida, como previsto na 248ª reunião do COPOM (BANCO CENTRAL DO BRASIL, 2022). Por fim, após o período de busca de um “pouso forçado” e maior leniência com a dinâmica inflacionária, o Federal Reserve, através de seus economistas, reafirma continuamente seu compromisso em levar a inflação para a meta. (POWELL, 2022)

Nesse sentido, a comunicação do Banco Central é essencial para que a inflação seja entregue na meta definida. O consenso macroeconômico integra em seus modelos as expectativas dos agentes em relação à inflação futura, e sua rigidez (ou não) pode dificultar (facilitar) sobremaneira o trabalho da política monetária nos custos incorridos pela sociedade em termos de perda de poder de compra e atividade econômica:

Proposição da ineficácia da política monetária: Se as expectativas são racionais e a taxa de inflação é perfeitamente antecipada (não existe surpresa inflacionária) a política monetária não influencia o nível de atividade. Já o componente estrutural da política monetária afeta a taxa de inflação de longo prazo. (Licha, 2014, p. 32)

Em um de seus *boxes* do Relatórios de Inflação, material publicado pelo BCB após decisões de política monetária, se afirma e analisa o papel das expectativas na eficácia da política monetária, tal qual afirmado pela teoria:

A inflação depende de diversos fatores, alguns mais relacionados à política monetária do que outros. Entre os mais dependentes estão as expectativas de inflação. Expectativas ancoradas, significando sua manutenção em torno da meta em qualquer cenário, inclusive após a ocorrência de choques relevantes, tornam menos custosa a ação do Banco Central no combate a pressões inflacionárias. Por esse motivo a ancoragem das expectativas é vista como elemento chave para a política monetária. (Banco Central do Brasil, 2021, p.1)

As formas de se estimar a qualidade dessa comunicação acontece principalmente por meio do boletim FOCUS, divulgado semanalmente pelo BCB, agregação da mediana de expectativas das instituições financeiras sobre as principais variáveis econômicas, como definido por (BANCO CENTRAL DO BRASIL). Além disso, análises qualitativas do “teor” das atas (duras ou serenas quanto à tolerância do banco central à inflação desviando da meta) e os cenários considerados são essenciais. Novamente, ambas seguem a teórica macroeconômica corrente:

Um compromisso supõe que o banco central se compromete a seguir certas regras de política monetária. Ele deve indicar seu comportamento futuro através orientando as expectativas do mercado. A comunicação do banco central com o público, em relação a suas intenções, cumpre um papel central neste contexto. Existem várias alternativas para realizar esta orientação futura. Uma alternativa é recolher e publicar as expectativas de inflação do mercado. Outra prática utilizada pelos bancos centrais para administrar as expectativas inflacionárias é publicar a trajetória futura da taxa de política dada a informação corrente. (Licha, 2014, p. 90)

Hoje já existe um esforço relevante na academia e no mercado para antecipar esses movimentos e uma técnica já usada por economistas do Federal Reserve é o *text mining*, uma forma de agregar as informações contidas nas atas do BC americano, como em (SHAPIRO e WILSON, 2021), (BENCHIMOL, KAZINNIK e SAADON, 2022) e (BHOLAT, HANSEN, *et al.*, 2015). No Brasil, o uso desse ferramental já é realizado há pelo menos dez anos:

Carvalho, Cordeiro e Vargas (2013) constroem um índice de comunicação que utiliza o método de pesquisa do Google para identificar se a sinalização das atas do COPOM está associada a uma postura mais “*hawkish*” ou “*dovish*”. Os resultados encontrados sugerem que o conteúdo das atas é relevante para explicar mudanças nas taxas de juros de diferentes prazos de maturação. Chague *et al.*

(2013) constroem um índice - denominado “fator de otimismo” - que sumariza as informações contidas nas atas do COPOM de acordo com predeterminados grupos semânticos, baseados no dicionário Harvard IV. Os resultados sugerem que taxas de juros de longo prazo são sensíveis ao grau de otimismo do Banco Central: maior otimismo da autoridade monetária implica em menores taxas de juros.

Pelos trabalhos resenhados é possível verificar que a literatura sobre comunicação mostrou profusão nas últimas décadas. Os resultados encontrados corroboram a ideia de que maior transparência - através da divulgação sistemática de documentos - está associada a maior eficiência da política monetária. Em particular, como notam Jansen (2011b) e Bulír, Cihák e Jansen (2012), existem consideráveis evidências de que o aumento da quantidade de comunicação dos bancos centrais está associado à maior previsibilidade da política monetária. Entretanto, não apenas a quantidade de comunicação por parte dos Bancos Centrais é importante, mas a qualidade do que é comunicado é crucial para torná-los mais transparentes. Esse último aspecto parece ser um dos pontos a serem explorados em futuras pesquisas. (Wilher, 2015, p.46)

1.1 OBJETIVOS

1.1.1 Objetivo geral

De acordo com (FILHO e FILHO, 2015), um tema deve 1) fugir do senso comum; 2) haver possibilidade de ser pesquisado; 3) ser relevante; 4) ser prazeroso ao autor em ser pesquisado. Por outro lado, Gil (2017) dá ênfase em 1) ser suscetível de solução, 2) delimitado a uma dimensão viável, e 3) deve ser ético.

Tendo em vista essas orientações, principalmente ao considerar, como demonstrado na introdução, a importância da comunicação do banco central e a viabilidade dessa análise, já que a ferramenta já é utilizada, o objetivo será analisar a comunicação do Banco Central do Brasil no período entre as gestões Henrique Meirelles e Campos Neto; no caso da última, até o fim de 2022. Ou seja, o período contemplado será de 2004 até 2022

1.1.2 Objetivos específicos

Visando cumprir a especificação e delimitação, os objetivos específicos são:

Estimar de maneira qualitativa e automatizada a comunicação do Banco Central do Brasil aos agentes econômicos;

Usar o insumo da teoria econômica e da história para enriquecer a análise;

Contribuir para a intersecção entre a ciência de dados e a economia aplicada

1.2 JUSTIFICATIVA

De acordo com (FILHO e FILHO, 2015):

Uma hipótese é um enunciado formal das relações esperadas entre, pelo menos, uma variável dependente e uma independente. É possível que, numa hipótese, estejam outras variáveis, como as intervenientes, por exemplo. (Filho & Filho, 2015, p.24)

Portanto, esta monografia está delimitada em analisar, a partir da análise das atas do COPOM – Comitê de Política Monetária – com *text frequency*, *sentiment analysis*, *topic modelling*, a comunicação do Banco Central do Brasil frente aos agentes econômicos, cujo objetivo é cumprir seu mandato de inflação na meta e estabilidade macroeconômica¹. A oportunidade deste trabalho, portanto, é auferir a forma pela qual a autoridade monetária usou um de suas ferramentas mais importantes: a comunicação. “um crítico componente da política monetária” (IMF'S MONETARY AND MACROPRUDENTIAL POLICIES DIVISION (MCMMP), 2022). É importante para a ciência econômica tal qual dito por (FILHO e FILHO, 2015):

Uma pergunta de pesquisa é todo problema cuja resposta seja de interesse da comunidade científica. É a comunidade científica que julgará a proposta que estará contida no projeto de investigação. Se interessar a essa comunidade que sejam identificadas as questões do ambiente organizacional levadas em consideração no desenvolvimento organizacional, o pesquisador poderá se propor a enfrentar tal desafio. (Filho & Filho, 2015, p.15)

Tal qual definido por (Gil, 2017), a hipótese proposta nesta monografia é aplicável, vai analisar a atuação do Banco Central do Brasil em gerir a política monetária, e a ferramenta usada será – como indicado pela macroeconomia atual – a ancoragem de

¹ “Desde a introdução do sistema de metas para a inflação no Brasil em 1999, a condução da política monetária pelo Banco Central do Brasil (BCB) tem se tornado mais transparente. Um elemento chave desta crescente transparência tem sido o esforço empreendido pelo BCB para comunicar sua visão sobre as perspectivas da economia, os elementos principais que caracterizam tais perspectivas e as possíveis conseqüências para a política monetária. Este esforço vem sendo feito através de um aumento da quantidade e qualidade da informação divulgada ao público, a fim de explicar as decisões de política monetária. Uma comunicação clara e frequente desempenha um papel de extrema importância para a credibilidade das ações do BCB, ancorando as expectativas de inflação do mercado e auxiliando na estabilidade do ambiente macroeconômico.” (JANOT e MOTA, 2012)

expectativas através dos dados disponíveis para o público: atas do COPOM. Nesse sentido, se relaciona com a uma teoria, a macroeconômica:

“Em muitas pesquisas sociais, este critério não é considerado. Entretanto, as hipóteses elaboradas sem qualquer vinculação às teorias existentes não possibilitam a generalização de seus resultados” (Gil, 2017, p.17).

Ou seja, se relaciona às técnicas disponíveis (*text mining* nas atas do COPOM) e tem referências empíricas na literatura especializada, como (SHAPIRO e WILSON, 2021):

“Nem sempre uma hipótese teoricamente bem elaborada pode ser testada empiricamente. É necessário que haja técnicas adequadas para a coleta dos dados exigidos para seu teste. Por essa razão, recomenda-se aos pesquisadores o exame de relatórios de pesquisa sobre o assunto a ser investigado, com vistas ao conhecimento das técnicas utilizadas.” (Gil, 2017, p.20)

É específica, pois analisa o papel do BC na condução da política monetária; e é conceitualmente clara

“Por essa razão, são preferíveis as hipóteses que especificam o que de fato se pretende verificar.” (Gil, 2017, p.20)

Por fim, é conceitualmente clara em seu objetivo e ferramentas para análise:

Os conceitos contidos na hipótese, particularmente os referentes a variáveis, precisam estar claramente definidos. Devem-se preferir as definições operacionais, isto é, aquelas que indicam as operações particulares que possibilitam o esclarecimento do conceito. Por exemplo, uma hipótese pode-se referir ao nível de religiosidade, que será definido operacionalmente pela frequência aos cultos religiosos. (Gil, 2017, p.20)

A escolha pelo período 2004 e 2022 decorre, primeiramente, pelo fato de ser um período menos conturbado da economia brasileira, sem as grandes turbulências internas e externas que o país enfrentou, principalmente com a eleição de Lula a seu primeiro mandato como presidente. A partir de 2004, com a reafirmação do tripé macroeconômico e do regime de metas, há um ambiente mais estável; apesar da pandemia de COVID-19 se contrapor a esse objetivo, por conta do amadurecimento institucional ao qual o país experimentou nesses 16 anos, essa crise é um objeto de estudo que demonstra a evolução do próprio Banco Central em sua comunicação.

1.3 METODOLOGIA

O propósito é fazer uma análise qualitativa da comunicação do BCB, em especial no conteúdo contido nas atas de suas reuniões. Uma questão a ser apontada nesse contexto é a capacidade das ferramentas de *sentiment analysis*, *term frequency* e *topic modelling* têm para isso. A primeira define, a partir de um “dicionário” pré-definido, emoções definidas para cada termo contido em um texto. O léxico NRC, por exemplo, tem algumas definições possíveis, como “raiva”, “confiança” e “tristeza”; o AFINN, por outro lado, tem um *score* que varia entre +5 e -5 em um espectro de positivo e negativo:

All three lexicons are based on unigrams, i.e., single words. These lexicons contain many English words and the words are assigned scores for positive/negative sentiment, and also possibly emotions like joy, anger, sadness, and so forth. The NRC lexicon categorizes words in a binary fashion (“yes”/“no”) into categories of positive, negative, anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. The Bing lexicon categorizes words in a binary fashion into positive and negative categories. The AFINN lexicon assigns words with a score that runs between -5 and 5, with negative scores indicating negative sentiment and positive scores indicating positive sentiment. All of this information is tabulated in the sentiments dataset, and tidytext provides the function `get_sentiments()` to get specific sentiment lexicons without the columns that are not used in that lexicon. (Sigle & Robinson, 2017, p.24)

Nesse sentido, dá para se definir o quão *hawkish* ou *dovish* o BC é a partir da análise das atas; em outras palavras, o quão duro ou suave é em sua comunicação. No jargão do mercado financeiro, os termos medem a “sensibilidade dos *policymakers* são à riscos inflacionários”:

A useful ranking of central bank preferences is based on their in- inflation preferences – or how sensitive policymakers are to increases in inflation risks. Borrowing terms used to describe individuals’ appetites for war, those that worry relatively more about inflation risks are called inflation hawks whereas those that worry relatively less about inflation risks are called Inflation Doves. Ordering individual central bank committee members according to their preferences along a single common dimension, “how much do I care about inflation risks relative to output and unemployment risks,” is important as it allows researchers to compare preferences across actors and over time and to test theories about committee decision-making. (Baerg & Lowe, 2020, p.3)

Nesse sentido, o *topic modelling* é interessante pois consegue, com ferramental matemático, extrair de cada texto o tema mais presente em cada um. Por exemplo, em um típico texto na coluna de política da Folha de São Paulo, as palavras mais presentes

costumam ser “congresso” e “presidente”, enquanto uma análise teatral terá termos relacionados á prática; o papel dessa ferramenta é fazer essa divisão.

Latent Dirichlet allocation is one of the most common algorithms for topic modeling. Without diving into the math behind the model, we can understand it as being guided by two principles: Every document is a mixture of topics We imagine that each document may contain words from several topics in particular proportions. For example, in a two-topic model we could say “Document 1 is 90% topic A and 10% topic B, while Document 2 is 30% topic A and 70% topic B.” Every topic is a mixture of words For example, we could imagine a two-topic model of American news, with one topic for “politics” and one for “entertainment.” The most common words in the politics topic might be “President,” “Congress,” and “government,” while the entertainment topic may be made up of words such as “movies,” “television,” and “actor.” Importantly, words can be shared between topics; a word like “budget” might appear in both equally. LDA is a mathematical method for estimating both of these at the same time: finding the mixture of words that is associated with each topic, while also determining the mixture of topics that describes each document. (Sigle & Robinson, 2017, p.97-98)

Por fim, o *term frequency* extrai de um determinado texto os termos mais frequentes em cada um, excluindo os chamados *stop words*, aquelas não uteis em análises (como artigos, preposições e pronomes, como “a”, “que”, “o” e etc.). Para o fim desta monografia, conseguimos capturar aspectos conjunturais de maneira mais precisa como se, por exemplo, na recente guerra da Ucrânia, a qual poderia ter impactos relevantes em preço de commodities e na dinamica inflacionaria global, a ata numero 245 tratou; enquanto na anterior a enfase era na reabertura da economia, preço de petróleo e a taxa de cambio do real em relação ao dólar. É importante ressaltar, entretanto, que a ferramenta trata o texto como um capítulo, ou seja, a análise demanda que todo o livro seja analisado para a determinação dos termos mais relevantes em cada um, ponderando palavras frequentes com menor peso e as novas com relevancia maior; assim, aquelas “inéditas” definem qual seria a tonica do capítulo.

A central question in text mining and natural language processing is how to quantify what a document is about. Can we do this by looking at the words that make up the document? One measure of how important a word may be is its term frequency (tf), how frequently a word occurs in a document, as we examined in Chapter 1. There are words in a document, however, that occur many times but may not be important; in English, these are probably words like “the,” “is,” “of,” and so forth. We might take the approach of adding words like these to a

list of stop words and removing them before analysis, but it is possible that some of these words might be more important in some documents than others. A list of stop words is not a very sophisticated approach to adjusting term frequency for commonly used words. Another approach is to look at a term's inverse document frequency (idf), which decreases the weight for commonly used words and increases the weight for words that are not used very much in a collection of documents. This can be combined with term frequency to calculate a term's tf-idf (the two quantities multiplied together), the frequency of a term adjusted for how rarely it is used. (Sigle & Robinson, 2017, p.40)

Com isso se obtém uma ferramenta poderosa: qualitativamente ter em mãos a preocupação do Banco Central em cada ata, dividido em grande medida entre juros, inflação, atividade e expectativas. Consegue capturar, nesse sentido, a história de cada reunião de política monetária ocorrido desde o início da divulgação das atas.

Longe de ser algo fora da realidade na análise econômica tradicional, pesquisadores do Federal Reserve já usam o *text mining*:

In this paper, we propose a new approach to estimating central bank objectives. Our approach relies on the assumption that the central bank's preferences regarding desired levels of inflation and economic activity are embedded in the words spoken by central bank policymakers at internal meetings. Specifically, we demonstrate that the sentiment expressed by policymakers can be used to directly estimate the central bank's loss function, including the implicit inflation target. This approach could be used internally and externally to study the preferences of any central bank that has transcripts or detailed summaries of their policymaking deliberations. (Shapiro & Wilson, 2021, p.2)

Esta pesquisa, portanto, será do modelo explicativa, com uso de métodos qualitativos e quantitativos para análise, pesquisa bibliográfica e experimental, com pesquisa ex-post-facto de dados secundários, tal qual orientado por (FILHO e FILHO, 2015)

Como parte importante deste trabalho é a informação disponível no Boletim FOCUS do BCB, compilação da mediana de expectativas das principais instituições financeiras sobre as principais variáveis econômicas, bem como o uso eventual de dados financeiros como o juro DI, dado de alta frequência que consegue capturar a visão de momento do mercado quanto aos movimentos da taxa Selic futura.

As possíveis limitações residem no isolamento de variáveis endógenas: melhora de indicadores econômicos e consenso quanto ao compromisso do BC em manter a estabilidade

de preços – visto que os dois tem uma relação clara empiricamente mais difícil de ser quantificada, como na equação elaborada em (Wilher, 2015, p.20):

$$\pi_{t+1}^e = E[\pi_{t+1}|I_t]$$

Desse modo, as expectativas de inflação no período subsequente é a esperança matemática, ou “(...) *the subjective probability distribution of outcomes tend to be distributed, for the same information set, about the ‘objective’ probability distributions of outcomes*” (Muth, 1961, p. 361), e toda a informação disponível, que pode ser tanto os sinais em discursos do presidente da instituição ou a opinião de um especialista em produtos agrícolas sobre a pressão decorrente de uma seca, por exemplo.

O capítulo 2 servirá como introdução para o consenso macroeconômico atual sobre bancos centrais e determinantes da taxa de inflação, bem como às ferramentas de *text mining* e sua contribuição para a economia

O capítulo 3 será mais técnica do text mining, com descrição e exemplos de como o processo funciona.

Por fim, o capítulo 4 mostrará o uso dessas ferramentas quando aplicadas ao nosso objeto de estudo, as atas de política monetária do BCB no período de 2004 até 2022.

2 MACROECONOMIA E O PAPEL DOS BANCOS CENTRAIS

Discutimos anteriormente um pouco da literatura relacionada ao papel do Banco Central na economia moderna. A partir do uso da literatura relacionada ao assunto, trataremos de forma sucinta aspectos históricos no e também técnicos do *mainstream*. Em relação ao primeiro.

2.1.1 CRIAÇÃO DO BCB E REGIME ATUAL

A criação do Banco Central do Brasil, o qual suas funções se confundiam com a do Banco do Brasil e a antiga SUMOC, ocorreu em meio ao Plano de Ação Econômica do Governo (PAEG), em 1964, um esforço do Estado em promover reformas estruturais no sistema financeiro, tributário e no mercado de trabalho do Brasil (GIAMBIAGI, 2011). Dentre os objetivos da modernização da política monetária, dentre outros, era limitar a expansão de crédito privado, o qual era muito relacionado à expansão dos meios de pagamento, que gerava inflacionária na economia; importante ressaltar, entretanto, a preocupação com a atividade produtiva, de forma a austeridade não prejudicá-la:

As metas do Paeg para a inflação indicavam uma estratégia assumidamente gradualista. O Plano não se propôs a eliminar o processo inflacionário em curto espaço de tempo, mas apenas a atenuá-lo ao longo de três anos, admitindo ainda uma inflação de dois dígitos (10%) no terceiro ano. A opção pelo gradualismo foi justificada no Plano com base no argumento de que havia a necessidade de uma “inflação corretiva” e de evitar-se uma grave crise de estabilização. (GIAMBIAGI, 2011, p.52)

As funções da instituição, tal qual contida em (BANCO CENTRAL DO BRASIL), é:

1. Inflação baixa e estável
2. Sistema financeiro seguro e eficiente
3. Banco do governo
4. Banco dos bancos
5. Emissor do dinheiro

Além disso, aspecto relevante a ser ressaltado é o regime pelo qual a política monetária opera no Brasil, o regime de metas de inflação, que ocorreu após as sucessivas crises internacionais ocorridas no fim dos anos 90 e, tal qual ocorrida no início da ditadura militar, foi uma tentativa de modernizar o regime monetário. Ele começa a ser operacionalizado após a institucionalização do regime de câmbio flutuante, e define ao BCB a responsabilidade e a independência operacional de conduzir a política monetária de forma a alcançar uma meta inflacionária. No interim de sua implementação está a deterioração da credibilidade do Estado, seja do lado fiscal ou monetário, em manter a estabilidade da moeda após a crise russa de 1999 apesar do aumento da taxa de juro em 6 pontos percentuais e a negociação de feita com o FMI de ajuda financeira. Por isso, a necessidade de uma âncora de credibilidade e ancoragem de expectativas era necessária até mesmo para a continuidade do sucesso do Plano Real. De acordo com os formuladores, havia a necessidade de se ancorar as expectativas de médio prazo, já que, por ora, a guinada no juro já faria esse trabalho. A ferramenta utilizada foi o regime de metas, factível por apresentar certa rigidez mas trazer discricionariedade suficiente para atuação do BC em choques no produto ou na inflação. Nesse sentido, foi adotado um processo gradualista pois o estabelecimento imediato de uma âncora inflacionária em meio a uma crise cambial poderia ter o efeito contrário, portanto foi anunciada a institucionalização do regime apenas em junho, três meses depois:

“A fully discretionary monetary policy without an explicit anchor would not tackle the inevitable uncertainties, especially during the transition period. It was natural to opt for a more rigid system, one that would represent a definite, strong commitment but that could also offer some indication of the future path of the economy; one that would allow enough flexibility for policymaking but that could also effectively anchor the public expectations. The authorities therefore decided to set up an inflation-targeting framework. However, the immediate announcement of a numerical inflation target was out of the question. The exchange rate was still overshooting, making any realistic assessment of pass-through effects all but impossible. Thus, setting a specific target value would not enhance credibility; on the contrary, it could even bring the new policy regime into disrepute. The Central Bank took a gradualist approach: it made clear that monetary policy would aim at keeping

inflation under control, but a formal inflation targeting framework would be in place only by the end of June.” (BOGDANSKI, FREITAS, *et al.*, 2001, p.12)

Os princípios teóricos a embasar este diagnóstico se relacionam ao novo consenso macroeconômico dos anos 90:

- A prioridade da política monetária é a estabilidade de preços;
- Estabilizar a taxa de inflação permite estabilizar o nível de atividade;
- O instrumento de política monetária deve ser a taxa de juros para empréstimos interbancários. Os agregados monetários são instáveis e difíceis de serem controlados e por esse motivo não podem ser utilizados como instrumentos de política monetária.⁵⁴ Já as taxas de juros básicas podem ser controladas pelos bancos centrais;
- A política fiscal é coadjuvante da política monetária para estabilizar a taxa de inflação e o nível de atividade;
- A credibilidade do banco central é importante para ter uma inflação baixa;
- É fundamental para a política monetária ter objetivos e procedimentos transparentes para a determinação da taxa de juros de política e, em especial, regras para enfrentar os choques diversos. (LICHA, 2014, p.62)

Isso fica claro num documento feito pelo FMI em discussões feitas com técnicos do BCB relacionado à implementação do regime de metas no Brasil, o qual já havia sido implementado em outros países:

“Low and stable inflation was singled out as the primary long-run objective of monetary policy, and inflation targeting was regarded as an effective framework for guiding monetary policy. In particular, inflation targeting was seen as providing a nominal anchor both for monetary policy and inflation expectations, making this anchor identical to the long-run objective of monetary policy; providing more transparency and accountability to the design and implementation of monetary policy; facilitating its communication, understanding, and assessment; and providing effective policy guidance by focusing policymakers’ attention on the long-run consequences of short-term policy actions.” (Brazil – Selected Issues and Statistical Appendix – International Monetary Fund, July 16th, 1999, *in* BOGDANSKI, TOMBINI e WERLANG, 2001)

Entre outros pontos, para os fins desta monografia os seguintes pontos de implementação são relevante, como presente em (BOGDANSKI, TOMBINI e WERLANG, 2001):

1. As metas de inflação são conhecidas dois anos antes do ano em questão;
2. A meta será definida baseada na variação entre janeiro e dezembro do referido ano, baseado num índice de preços conhecido, excetuando-se a possibilidade do uso de núcleos (excluindo a possibilidade da exclusão de itens mais voláteis, por exemplo)
3. No caso de desvio da inflação à meta, o presidente do BCB envia carta aberta para o Ministro da Fazenda explicando as razões para tal, as medidas a serem adotadas para convergência, e o período necessário para elas terem efeito;
4. O BCB publicará relatório trimestral sobre a performance de regime de metas, os resultados das medidas monetárias e as perspectivas inflacionárias.

Dentre as justificativas para o ponto (2), ao contrário do caso dos EUA, por exemplo, que tem dois índices de preços – um para produtor, outro para consumidor – se tornou importante no regime de metas por conta da existência de vários índices de preço no Brasil, o que poderia custar à credibilidade da política monetária principalmente num momento tão conturbado quando o fim dos anos 90.

Em relação aos outros, fica claro o peso significativo dado às expectativas na condução de política monetária, a ser tratado posteriormente.

2.1.2 REVOLUÇÃO DAS EXPECTATIVAS RACIONAIS

A questão das expectativas é tratada desde Keynes em sua Teoria Geral, um conceito revolucionário na época e até hoje muito relevante. Lá, entretanto, ela é restringida à definição da condição necessária para que ocorram investimentos na economia: a eficiência marginal do capital deve ser maior que a taxa de juros a remunerá-lo; quando não o é, se opta pela aplicação financeira.

Em 1976, Robert Lucas revoluciona a macroeconomia com sua crítica, que basicamente criticava a forma como os parâmetros de política econômica eram estimados. Isso decorre de as expectativas dos agentes econômicos serem racionais e usarem das informações disponíveis para preverem o futuro; ou seja, não são estáticas e usam apenas o passado. Em linguagem matemática, como formulado em (LICHA, 2014), suponha produto como $Y =$

$a + v\pi$, na qual a é o parâmetro da estrutura econômica e v o instrumento da política. A crítica de Lucas define que o parâmetro não é independente do instrumento, ou seja, exige endogeneidade entre as variáveis e a suposição de um a constante incorre em omissão de variável; como ela é correlacionada com a variável incluída π , os parâmetros serão tendenciosos e inconsistentes. Em outras palavras:

Com a introdução da hipótese das expectativas racionais no começo da década de 1970 as decisões dos agentes econômicos passaram a depender dos instrumentos de política econômica, tornando-se variáveis endógenas. Mudanças nos instrumentos de política econômica provocam mudanças nas decisões dos agentes (que são contingentes ao estado da economia) conforme destacam a crítica de Lucas. Com a hipótese das expectativas racionais as relações econômicas manipuladas pela política econômica não são estruturais e mudam com câmbios na política. (LICHA, 2014, p.34-35)

Para ilustrar sua importância, é necessário apontar o período histórico em que ocorre. A equação macroeconômica mais importante era então a curva de Phillips, definida por:

$$\pi^t - \pi^{t-1} = (m + z) - \alpha u$$

Onde π é a inflação efetiva; π^{t-1} é a inflação do período anterior; m o *mark-up* das firmas; z fatores que afetam a decisão de salários (rigidez); α um coeficiente (constante); e u a taxa de desemprego

Naquele momento a curva de Phillips perdera sua capacidade de previsão e estimação da economia no cenário inflacionário dos anos 1970. Nela, havia uma relação constante entre o nível de preços e o desemprego; em outras palavras, a determinação do desemprego dependia do quanto o *policymaker* estaria disposto a inflacionar a economia. Nos anos 70 houve uma mudança no comportamento da inflação:

A inflação tornou-se mais persistente. Aumentou a probabilidade de a inflação alta de um ano ser seguida por uma inflação alta no seguinte. Assim, as pessoas, ao formarem suas expectativas, começaram a levar em conta a persistência da inflação. Essa mudança na formação de expectativas acabou por modificar a natureza da relação entre desemprego e inflação. (BLANCHARD, 2017, p. 175)

Nesse sentido, as mudanças decorrentes da crítica de Lucas foram: (1) as expectativas são racionais e os instrumentos de política afetam as decisões dos agentes e a estrutura da economia, (2) os modelos econométricos, como no caso da curva de Phillips, apresentam

erro de especificação. (LICHA, 2014, p.35). Nesse sentido, por conta de (1), a economia não era um sistema estático cujos parâmetros poderiam ser descobertos e definiam como ela funciona: com (2), a economia se torna um sistema dinâmico onde a única constante é sua mudança.

Além disso, cria-se o conceito de produto potencial, no qual é definido que a atividade econômica flutua ao redor de um nível que, salvo choques de oferta positivos (ganhos de produtividade) na economia, os desvios em relação à ela geram aumento (diminuição) do nível de preços quando o produto efetivo é maior (menor). A taxa natural de desemprego é aquela que mantém a taxa de inflação constante, a chamada NAIRU (*non-accelerating inflation rate of unemployment*), como definido em (BLANCHARD, 2017). O efeito da política monetária no produto acontece apenas quando $\pi^{efetiva} \neq \pi^{esperada}$, pois considerando preços e salários flexíveis e informação perfeita, os choques de demanda decorrentes da política monetária são antecipados e não tem efeito na economia. Entretanto, “o componente estrutural da política monetária afeta a taxa de inflação de longo prazo” (LICHA, 2014, p.33). Nesse modelo o papel estabilizador do Estado é impossibilitado. Um aspecto relevante a ser considerado é a rigidez de preço, pois ela determina a resposta da atividade à expansão monetária, ou seja: quanto mais rigidez nominal há na economia, maior o poder do instrumento monetário em estimular a atividade, pois os agentes não podem ajustar sua renda (ou lucro).

Por fim, na nova curva de Phillips o desvio da inflação afeta o desvio do desemprego em relação ao seu potencial, dado um coeficiente, ou seja, “o produto agregado é maior que o produto potencial se o nível geral dos preços (e em consequência, a taxa de inflação) é maior que o nível geral esperado de preços (taxa de inflação esperada) (LICHA, 2014, p.45) :

$$\pi - \pi^e = \alpha(u_t - u_n)$$

2.1.2 NOVOS MODELOS DE POLÍTICA ECONOMICA

O resultado dessa mudança de paradigma teve efeitos concretos na ciência econômica e não se limitou exclusivamente ao debate acadêmico. No cenário agora posto, temos em mente a hipótese de expectativas adaptativas e os desdobramentos, como o aprendizado adaptativo². Entre outras coisas, uma das discussões mais importantes é o regime pelo qual a política

² Os agentes aprendem a fazer previsão com os dados observados e seus modelos no decorrer do tempo, ou seja, agem como econométristas que buscam entender a estrutura econômica.

monetária deve seguir seu curso: de forma discricionária, com instrumentos de política determinados *ex-post*; e compromisso com decisões já conhecidas. Esses dois regimes se referem à função de reação do BC a choques de demanda e oferta, bem como a seu conservadorismo em relação à desvio da inflação à meta e viés de estímulo ao produto. Em termos matemáticos, é definido de acordo com a função de perda social (FPS), como definido por LICHA (2014, p.48).

$$L = w\pi^2 + (Y - kY_p)^2,$$

Na qual π é a inflação e w é a tolerância do BC à seu desvio (o BC é menos tolerante à inflação quanto maior o k); Y e Y_p se referem ao produto e seu potencial, respectivamente, enquanto k , se maior que 1, define o viés inflacionário do formulador da política.

Quando o regime é o de regra (compromisso), os agentes conseguem prever a inflação, ou seja, $\pi = \pi^e$, pois sabem exatamente como será a política em casos de choques. Para o regime discricionário, π^e é determinado ex-ante e a partir disso se decide o instrumento de política, caracterizando um jogo sequencial entre o público e o BC. Nesse caso existe um viés inflacionário na formação das expectativas, da qual acabam sendo validadas e tomadas em consideração para o instrumento de política. Dessa forma, a função de perda acaba sendo definida de forma direta pelo viés inflacionário e de forma inversa ao grau de conservadorismo do BC, afinal, expectativas altistas aumentam o custo de desinflação em termos de perda na atividade econômica; apenas um BC totalmente conservador pode ancorar perfeitamente as expectativas em regime discricionário.

3.0 TEXT MINING

3.1 O PROCESSO BÁSICO

A quarta revolução industrial é tida como a era da informação, na qual o poder da *internet* e dos computadores definem a tônica dos avanços tecnológicos do mundo. Entre várias outras coisas, essa facilidade criou o ramo da programação e o esforço coletivo de facilitar e aprimorar análises qualitativas que, apesar de funcionarem, eram demasiadamente mecânicas. O *text mining* é uma ferramenta na qual se tenta, através do poder da programação, aprimorar a análise textual em diversos sentidos, seja pela quantidade de palavras, análise de sentimentos, e distinção de tópicos ou palavras. Há uma forma definida de organização, base para todos os processos realizados posteriormente

Em primeiro lugar, o formato é o *tidy*, definido como

1. Cada variável é uma coluna
2. Cada observação é uma linha

Assim, o formato de análise textual é uma tabela com um *token* por linha, sendo ele uma unidade de texto – normalmente apenas uma palavra, mas também pode ser um parágrafo inteiro. O processo de *tokenização*, após a extração do texto em si, é a segunda parte técnica. Para facilitar esse processo, o pacote *tidytext*, desenvolvido por (SILGE e ROBINSON, 2016), é referência nesse sentido. No exemplo abaixo se mostra o formato. Uma frase foi dividida em duas linhas, o que facilmente poderia ser dois parágrafos ou até duas páginas de uma ata do copom. O processo de tokenização, como dito anteriormente, cria uma tabela com cada palavra do texto e o divide por uma métrica, nesse caso, a linha em que cada uma aparece:

```
## # A tibble: 2 × 2
##   line text
##   <int> <chr>
## 1     1 1 O Banco Central do Brasil
## 2     2 2 é presidido por Campos Neto
```

```
## # A tibble: 10 × 2
##   line word
##   <int> <chr>
## 1     1 1 o
## 2     2 1 banco
## 3     3 1 central
## 4     4 1 do
## 5     5 1 brasil
## 6     6 2 é
## 7     7 2 presidido
## 8     8 2 por
## 9     9 2 campos
## 10    10 2 neto
```

Figura 1 – *Tokenização* de um texto

Em suma, o processo básico é feito da seguinte forma, com o detalhe de a obtenção da série de atas do COPOM ser feito via *webscraping*, uma automatização do processo de coleta de dados.:

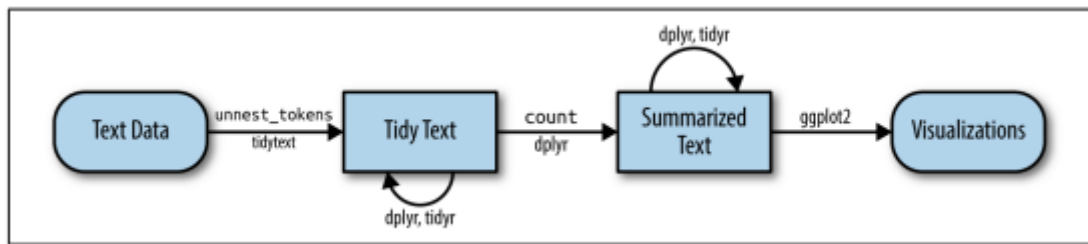


Figura 2 – Processo de *tidying* de um texto

3.2 STOP WORDS

Outro aspecto importante é a remoção das *stop words*, artigos, preposições e outros termos que nada dizem sobre o texto, mas são importantes gramaticalmente; na imagem anterior, “o” “é” “por” são exemplos possíveis. Ou seja, o texto – como presente na ata – é assim:

"15. For the set of regulated prices inflation, the projected change is 6.8% in 2016, unchanged relative to the April Copom meeting. Among other factors, this projection considers average increase of 19.7% in the water and sewage tariffs, of 13.6% in the prices of health plans and reduction of 3.5% in electricity prices. For 2017, an increase of 5.0% is considered, unchanged relative to the value projected at the last three meetings. The items for which there are more information available were projected individually and, for the others, the projections are based on endogenous determination models for regulated prices, which compute, among other factors, seasonal components, market prices inflation and the IGP (General Price Index) change."

E fica assim

"15. For set regulated prices inflation, projected change 6.8% 2016, unchanged relative April Copom meeting. Among factors, projection considers average increase 19.7% water sewage tariffs, 13.6% prices health plans reduction 3.5% electricity prices. For 2017, increase 5.0% considered, unchanged relative value projected last three meetings. The items information available projected individually, others, projections based endogenous determination models regulated prices, compute, among factors, seasonal components, market prices inflation IGP (General Price Index) change."

3.3 STEMMING

Por último, mas não menos importante, é o processo de stemming. Ele é importante principalmente para a questão do topic modelling, um modelo cujo objetivo é extrair – a partir de um conjunto de textos – os tópicos presentes neles. Considerado uma “normalização” dos termos, ele extrai o radical das palavras, de forma a não haver repetições desnecessárias, que atrapalham esse processo. “inflation” e “inflationary”, por exemplo, se referem ao mesmo tema (tópico) que é a inflação. Nesse sentido, não há utilidade em ter ambos no conjunto de palavras a serem utilizadas, logo, se faz o uso de “inflat”, sem sufixos

e prefixos, e o efeito desejado é alcançado da mesma maneira. Usando o texto anterior, o texto após o processo de stemming fica na seguinte maneira:

"15. For the set of regul price inflati, the project chang is 6.8% in 2016, unchang relat to the April Copom meeting. Among other factor, this project consid averag increas of 19.7% in the water and sewag tariffs, of 13.6% in the pric of health plan and reduct of 3.5% in electr pric. For 2017, an increas of 5.0% is consider, unchang relat to the valu project at the last three meetings. The item for which there are more inform avail were project individu and, for the others, the project are base on endogen determin model for regul prices, which compute, among other factors, season components, market price inflat and the IGP (Gener Price Index) chang."

Um ponto adotado por este trabalho foi a remoção da seção “Sumário dos dados analisados pelo Copom” – presente nas atas entre 2004 e 2016 – como adotado por (ASSOCIADOS, 2019), por ser uma parte do texto cujo efeito na comunicação da política monetária é nulo, por ser mera descrição da evolução dos dados econômicos da reunião anterior correspondente à cada ata. Dessa forma, tentou-se resgatar o que posteriormente se formalizou a partir de 2016: as seções “Avaliação da conjuntura econômica e do cenário básico do Copom”; “Riscos em torno do cenário básico para a inflação”; “Discussão sobre a condução da política monetária”; e “Decisão de política monetária”, como recomendado por (BLINDER, EHRMANN, *et al.*, 2008).³

4. FERRAMENTAS APLICADAS

4.1 ANÁLISE DE SENTIMENTO

Na linha do dito anteriormente, a intenção do *text mining* é facilitar análises qualitativas. Nesse sentido, o processo de análise de sentimento mais simples é a atribuição de notas para cada termo, e no âmbito deste trabalho é o que mais faz sentido. O sentimento é definido pela somatório dos *scores* de todos os termos presente no texto

$$Sentimento = \sum_{1}^{n} S_n$$

A outra forma é a atribuição de qualidade para cada um -como tristeza, raiva, medo, entre outros - como feito no léxico NRC. No caso do Bing, há a divisão apenas entre positivo e

³ “Central banks communicate about at least four different aspects of monetary policy: their overall objectives and strategy, the motives behind a particular policy decision, the economic outlook, and future monetary policy decisions. Central banks’ objectives and strategies tend to be more stable, so the corresponding signals show less variability over time than signals about the other three items.”

negativo; para o AFINN, o *score* é no intervalo entre -5 e 5. Em suma, é uma mensuração completamente subjetiva e arbitrária. O léxico referência para textos financeiros é o Loughran–McDonald, mas para o propósito deste trabalho será usado a metodologia presente em (APEL e GRIMALDI, 2014), onde são extraídos não apenas termos específicos, mas também suas variações: *increasing*, por exemplo, variação de “*increase*”, é computado; assim como “*higher*” e outros adjetivos. Além disso, são computados também termos compostos, como “*inflation expectations*”, e não apenas “*inflation*”, bem como “*disinflation*”. Por fim, em relação à temas relacionados à política fiscal, é filtrado termos correlatos como “*debt-to-gdp*”, e não apenas “*fiscal*” ou “*debt*”.

Por fim, se tem uma lista de adjetivos – os mais presentes no conjunto de atas, desde 2004 até 2022– e cada um tem uma pontuação “*Hawk*” ou “*dove*” baseado em cada assunto específico. “*contract*”, por exemplo, é *dove* para uma inflação menor que o mês anterior, mas *hawk* quando se trata de piora no resultado primário do governo. No caso deste trabalho, a conta do índice de sentimentos considera a proporção de palavras “*hawk*” e “*dove*” no texto, como em (Ibid.):

$$Net\ Index = \left[\left(\frac{\#hawk}{\#hawk + \#dove} \right) - \left(\frac{\#dove}{\#hawk + \#dove} \right) \right] + 1_4$$

. Essa especificação é feita individualmente para os temas cenário, inflação, setor externo, fiscal, e atividade. Cada termo é atribuído de maneira a fazer mais sentido para cada tópico.

⁴ (Apel & Grimaldi, 2014, p.10)

Hawk	Dove
Worse*	contract*
deteriora*	eas*
increas*	decelerat*
gr*	weak*
expand*	cool*
heat*	converg*
accelerat*	low*
strength*	decreas*
strong*	reduc*
fast*	slow*
high*	improv*
converg*	diverg*
better*	worse*
surplus	deficit*
improv*	deteriorat*
deficit	surplus
inconsistent ³	consistent

Figura 3 - Tabela de termos para a análise de sentimentos

Nesse sentido, este é o resultado obtido:

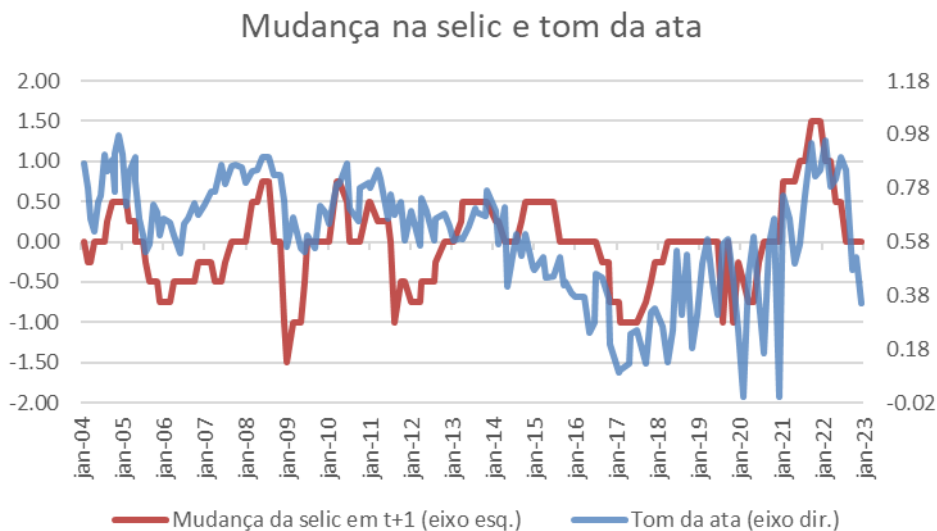


Figura 4 – Mudança na Selic em t-1 e tom da ata do Copom em t

No gráfico, observamos uma correlação clara, principalmente até 2014. Isso se prova a partir da melhora (piora) do tom das atas com a queda (alta) na taxa Selic no encontro

posterior. A partir de 2015, a tendência de queda no tom das atas pode antecipar quedas consecutivas na taxa Selic, assim como as altas, a partir da tendência de alta nos textos.

4.2 TERM FREQUENCY

Quando se analisa as atas do Copom, uma das tarefas do economista é capturar a ênfase do BC em um determinado aspecto da conjuntura econômica, como incerteza fiscal decorrente de uma expansão de gastos, ou de uma seca que pode pressionar o preço de produtos primários. No início do ano, após a questão da PEC dos precatórios no fim de 2021, o ciclo de aperto monetário parecia estar terminando até que a guerra da Ucrânia começou e trouxe diversas incertezas em meio a um cenário global já nebuloso. A política de lockdown na China, o aperto na oferta global de petróleo e commodities já estavam dados, mas a adição da incerteza quanto aos impactos da guerra para o resto do mundo, principalmente pela questão dependência do gás natural russo para a Europa e relevância da Ucrânia em commodities, pressionou tanto o cenário do BCB quanto do FED, com o segundo sendo forçado a começar o processo de aperto monetário.

A automatização desse processo ocorre pelo *inverse document frequency* (idf), que diminui o peso daquelas palavras com maior ocorrência numa série de textos e aumenta as outras. Combinado com a estatística tf-idf, é criada uma maneira de se mensurar a importância de termos numa coleção de documentos, como no caso de capítulos de um livro ou nas atas do Copom.

The statistic tf-idf is intended to measure how important a word is to a document in a collection (or corpus) of documents, for example, to one novel in a collection of novels or to one website in a collection of websites. (SIGLE e ROBINSON, 2017, p.40)

Em termos matemáticos, é definida da seguinte maneira:

$$idf(\text{termo}) = \ln \frac{n_{\text{capítulos}}}{n_{\text{capítulos que contém o termo}}}$$

Quando realizamos esse processo e removemos as *stop words*, uma análise feita desde março de 2021 fica assim:

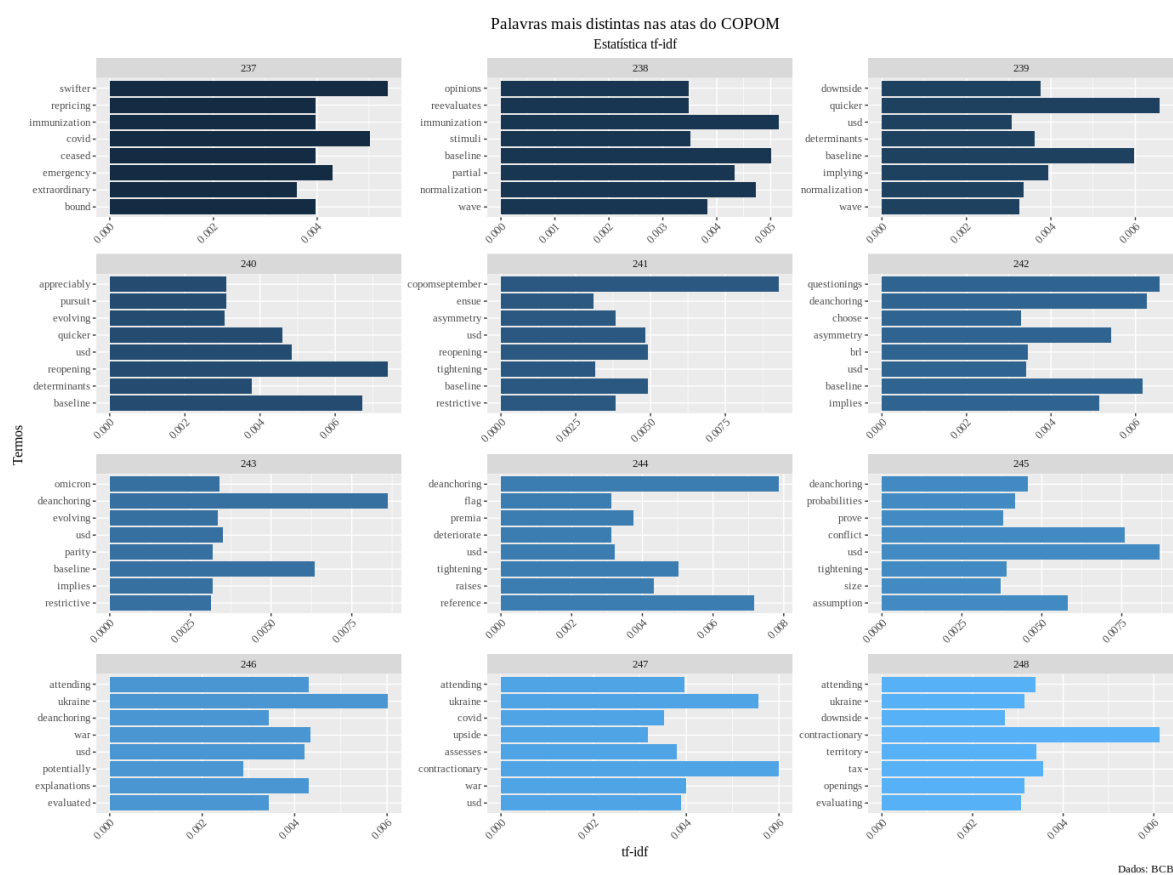


Figura 5 – Tf-idf aplicado às atas de 2021

Tendo esses dados em mãos, dá para traçar uma linha do tempo das atas. No início, ênfase em termos relacionados à covid, como reabertura, imunização, normalização e ondas de infecção. A partir da 241ª ata, sinalização para a reunião posteriores e câmbio. Daí em diante, aparece questões relacionadas ao risco fiscal da PEC dos precatórios, desancoragem de expectativas (deterioração, premio de risco). Em março, ata pós início da guerra da Ucrânia, termos relacionados à guerra tomam protagonismo (Ucrânia, potencialmente, guerra, câmbio), além de algum resíduo de COVID por conta do lockdown adotado na china. Na última reunião, a de agosto, já aparece a sinalização de ajuste residual no juro, além da avaliação de risco da autoridade monetária quanto à redução de impostos aprovada pelo governo federal.

4.3 TOPIC MODELLING

Por fim, o *topic modelling* é uma ferramenta cujo objetivo é estimar a proporção de tópicos presentes em um texto. Para o nosso caso, um exemplo simples seria estimar o quanto uma determinada ata do copom trata de atividade econômica, inflação e expectativas. Para fazer

essa mensuração, para cada palavra é associada uma característica qualitativa. Nesse sentido, é estimada a probabilidade de encontrarmos, conforme a leitura do texto é realizada, cada um dos tópicos determinados.

Abaixo observamos 5 tópicos (definidos arbitrariamente, tendo em vista que já nesse número há sobreposição de termos), e suas palavras associadas. No tema 1, fica claro a ênfase em dados de atividade, dados os termos “industri”, “year”. No tema 2, “scenario”, “fiscal”, “target” mostram preocupação com a evolução do cenário econômico. No tema 3, “market”, “expect”, “meet”, “remain” apontam para preocupação com as expectativas dos agentes. No quarto tema, “copom”, “rate”, “monetari”, “polici” apontam para a decisão do BC no tocante à política monetária, com suas justificativas. Por fim, o último tópico trata especificamente da inflação, inferido a partir dos termos “price”, “inflat”, “rate” e “market”.

Atividade	Cenário e Balanço de Riscos	Expectativas	Decisão de Política Monetária	Inflação
increas	inflat	price	econom	price
year	polici	inflat	copom	consid
industri	monetari	rate	inflat	chang
last	copom	market	rate	accord
respect	scenario	contin	risk	inflat
oper	project	remain	monetari	growth
data	meet	meet	scenario	adjust
averag	process	growth	polici	relat
may	fiscal	expect	level	rate
accord	target	also	expect	market

Figura 6 – Modelagem de tópicos e seus termos

Quando projetamos esse recurso num gráfico de calor, no qual os valores (*gamma*, na linguagem do modelo) atribuídos para cada tópico para os respectivos textos, temos a imagem abaixo. Nela, observamos que no mandato de Henrique Meirelles, o tema mais preponderante foi sobre atividade, que se justifica, afinal foi o período de crescimento econômico mais vultoso neste século. Posteriormente, a gestão Tombini deu ênfase para inflação e em seguida para o cenário e balanço de riscos; também faz sentido, afinal um dos pontos sensíveis do período foi a inflação alta, dada principalmente por uma inflação cheia muito influenciada por medidas políticas – refletida nos preços administrados – mas aparente nos preços livres. Além disso, o BC já atentava para os riscos dados pela dita política monetária parafiscal dada pelos créditos subsidiados pelo BNDES e seu reflexo na eficácia da política monetária. Na gestão Goldfajn, a virada citada anteriormente em relação à comunicação do BC fica aparente. A gestão deu muita ênfase nas justificativas das suas

decisões, em meio à muita discussão sobre revisão da meta de inflação no período, que não ocorreu exatamente pelo BC formular sua justificativa para não fazê-lo. Por fim, a gestão Roberto Campos deu início ao chamado *forward guidance*⁵, e isso parece ter tido resultado na nossa estimativa de modelo de tópicos, onde o tema mais presente foi o de Cenário e Balanço de Riscos, seguido pelo Decisão de Política Monetária.

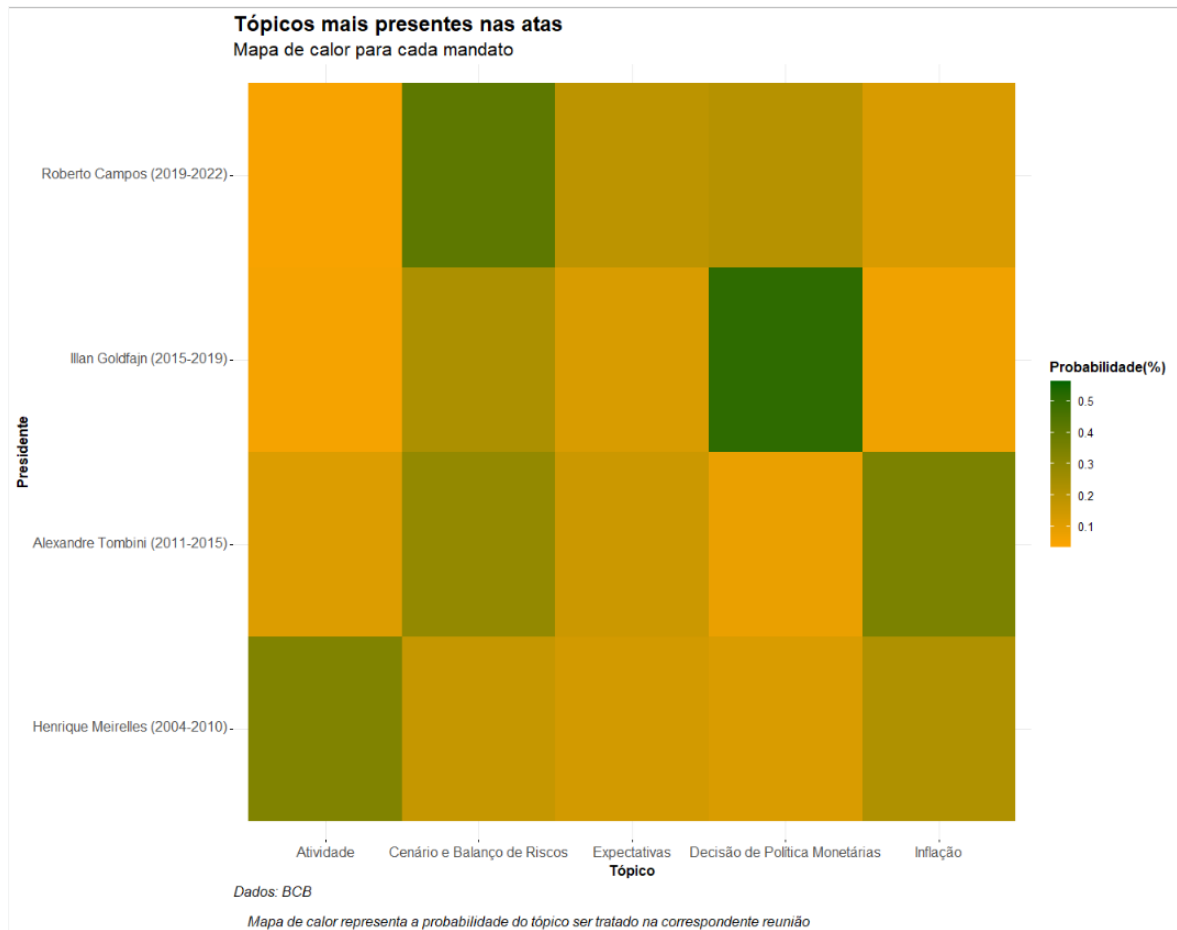


Figura 7 – Mapa de calor dos temas para cada presidente

⁵ “De forma a prover o estímulo monetário considerado adequado para o cumprimento da meta para a inflação, mas mantendo a cautela necessária por razões prudenciais, o Copom considerou a utilização de uma “prescrição futura” (isto é, um “forward guidance”) como um instrumento de política monetária adicional. O Copom discutiu as limitações no uso deste instrumento em países emergentes. Em relação aos pares desenvolvidos, países emergentes são mais suscetíveis a contágio de crises externas e possuem maiores vulnerabilidades nos fundamentos econômicos. Consequentemente, devido à maior imprevisibilidade e volatilidade, o uso de tal instrumento torna-se mais desafiador. O Comitê concluiu que, apesar dessas limitações, a prescrição futura seria a estratégia de implementação de política que atualmente apresenta a melhor relação custo benefício. A prescrição futura cumpre o papel de transmitir a visão do Comitê sobre suas ações futuras e tende a ajustar as expectativas expressadas na parte intermediária da curva de juros.” (BANCO CENTRAL DO BRASIL, 2020)

5. RESUMO

A partir deste trabalho, foi mostrado – de modo exploratório – a aplicação de técnicas modernas para a comunicação do Banco Central. Todas essas ferramentas tem parâmetros ad hoc, ou seja, podem ser definidos livremente a gosto do programador. Por outro lado, o processo de formatação desses textos segue uma regra de bolso, e tal escalabilidade dá ao economista maior espaço para aprimorar os parâmetros de sua análise.

Nesse sentido, ficou claro o poder dessas ferramentas para contar a história da economia brasileira, principalmente em cenários de estresse nos mercados e turbulências – sejam aquelas puramente econômicas (inflação em alta) ou as exógenas a ela (pandemia da COVID-19).

Como já dito, existem formas econométricas de mensurar os efeitos práticos da comunicação ao banco central, uma linha de pesquisa ainda incipiente na ciência econômica brasileira. Um desafio, por exemplo, é a consolidação de um ‘dicionário’ de termos econômico-financeiros em português.

Tudo isso posto, o *machine learning* tem grande potencial de liderar a evolução da análise econômica.

Bibliografia

APEL, M.; GRIMALDI, M. How Informative Are Central Bank Minutes? **Review of Economics**, 65, 2014. 53-76. Disponível em: <<https://www.degruyter.com/document/doi/10.1515/roe-2014-0104/html>>.

ASSOCIADOS, M. C. **Um índice de tom para a ata do Copom**. [S.l.], p. 6. 2019.

BAERG, N.; LOWE, W. A textual Taylor rule: estimating central bank preferences combining topic and scaling methods. **Political Science Research and Methods**, Janeiro 2020. 106-122.

BANCO CENTRAL DO BRASIL. **Ata da 232ª Reunião do Comitê de Política Monetária (Copom) do Banco Central do Brasil**. [S.l.], p. 5. 2020.

BANCO CENTRAL DO BRASIL. Ancoragem das expectativas de inflação e condução da política. **Relatório de Inflação**, 23, Setembro 2021. 56-62. Disponível em: <<https://www.bcb.gov.br/content/ri/relatorioinflacao/202109/ri202109p.pdf>>.

BANCO CENTRAL DO BRASIL. Atas do Comitê de Política Monetária - Copom, Brasília, Agosto 2022. Disponível em: <<https://www.bcb.gov.br/publicacoes/atascopom/03082022>>.

BANCO CENTRAL DO BRASIL. Institucional. **Acesso à informação**. Disponível em: <<https://www.bcb.gov.br/acessoinformacao/institucional>>. Acesso em: Setembro 2022.

BANCO CENTRAL DO BRASIL. Perguntas frequentes - Expectativas de mercado. **Banco Central do Brasil**. Disponível em: <<https://www.bcb.gov.br/controleinflacao/faqexpectativa>>.

BARBOSA, F. D. H.; CAMÊLO, F. D.; JOÃO, I. C. A Taxa de Juros Natural e a Regra de Taylor no Brasil: 2003–2015. **Revista Brasileira de Economia**, Dezembro 2016. Disponível em: <<https://bibliotecadigital.fgv.br/ojs/index.php/rbe/article/view/60542/63363>>.

BENCHIMOL, J.; KAZINNIK, S.; SAADON, Y. Text mining methodologies with R: An application to central bank texts. **Machine Learning with Applications**, Junho 2022.

BHOLAT, D. et al. Text mining for central banks. **Centre for Central Banking Studies**, 33, 2015. Disponível em: <<https://www.bankofengland.co.uk/-/media/boe/files/ccbs/resources/text-mining-for-central-banks>>.

BLANCHARD, O. **Macroeconomia**. 7ª edição. ed. São Paulo: Pearson Education do Brasil, 2017.

BLINDER, A. S. et al. Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. **ECB Working Paper 898**, 1 Maio 2008.

BOGDANSKI, J. et al. Inflation Targeting in Brazil: Shocks, Backward-Looking Prices, and IMF Conditionality. **Working Paper Series no. 24**, Brasília, Agosto 2001.

BOGDANSKI, J.; TOMBINI, A. A.; WERLANG, S. R. C. Implementing Inflation Targeting in Brazil. **Working Paper Series n. 1**, Brasília, Julho 2001. Disponível em: <<https://www.bcb.gov.br/pec/wps/ingl/wps01.pdf>>.

CUNHA, P. V. D. **The New Macroeconomic Matrix and Monetary Policy (draft)**. [S.l.]. 2017.

ESTADÃO CONTEÚDO. Campos Neto: “Vamos avaliar um possível ajuste final em setembro”. **InfoMoney**, Setembro 2022. Disponível em: <<https://www.infomoney.com.br/economia/campos-neto-vamos-avaliar-um-possivel-ajuste-final-em-setembro/>>.

EUROPEAN CENTRAL BANK. **Monetary policy decisions**. Frankfurt: [s.n.], 2022. Disponível em: <<https://www.ecb.europa.eu/press/pressconf/shared/pdf/ecb.ds220908~1d17d49d04.en.pdf>>.

EUROSTAT STATISTICS. **Euro area annual inflation up to 9.1%**. [S.l.]. 2022.

FILHO, E. J. M. A.; FILHO, M. C. F. **PLANEJAMENTO DA PESQUISA CIENTIFICA**. São Paulo: Atlas, 2015.

GARCIA, G. T. 2003-2017, O REGIME DE METAS INFLACIONÁRIAS E A POLÍTICA DE JUROS BRASILEIRA: ESTUDO SOBRE A EFICÁCIA DO MODELO ENTRE. **Trabalhos de Conclusão de Curso - FCLAR**, 29 Junho 2018.

GIAMBIAGI, F. **Economia Brasileira Contemporânea [recurso eletrônico]: 1945-2010**. Rio de Janeiro: Elsevier, 2011.

GIL, A. C. **Como Elaborar Projetos de Pesquisa**. Rio de Janeiro: Atlas, 2017.

IMF'S MONETARY AND MACROPRUDENTIAL POLICIES DIVISION (MCMMP). Monetary Policy Framework: Central Bank Communications. **Monetary and Capital Markets Department: Technical Assistance Handbook**, Janeiro 2022. Disponível em: <<https://www.imf.org/-/media/Files/Publications/Miscellaneous/English/2022/mcm-technical-assistance-handbook/central-bank-communications.ashx>>.

JANOT, M.; MOTA, D. E.-J. D. S. O Impacto da Comunicação do Banco Central do Brasil sobre o Mercado Financeiro. **Trabalho para Discussão**, janeiro 2012. 1-33. Disponível em: <<https://www.bcb.gov.br/pec/wps/port/TD265.pdf>>.

LICHA, A. L. **Teoria da Política Monetária - Uma abordagem a nível intermediário**. Rio de Janeiro: Alta Books, 2014.

MUTH, J. F. Rational expectations and the theory of price movements. **Econometrica**, n. 29(3), 1961. 315-335.

POWELL, J. H. **Monetary Policy and Price Stability**. Federal Reserve. Jackson Hole. 2022.

REUTERS. Insight: Brazil's Tombini - An inflation hawk, believe it or not. **Reuters**, 2013. Disponível em: <<https://www.reuters.com/article/us-brazil-economy-tombini-insight-idINBRE97M05F20130823>>.

SHAPIRO, A. H.; WILSON, D. Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis. **Federal Reserve Bank of San Francisco Working Paper 2019-02**, 2021. Disponível em: <<https://www.frbsf.org/economic-research/publications/working-papers/2019/02/>>.

SIGLE, J.; ROBINSON, D. **Text Mining with R: A Tidy Approach**. [S.l.]: CC BY-NC-SA 3.0 US, 2017.

SILGE, J.; ROBINSON, D. tidytext: Text Mining and Analysis Using Tidy Data Principles in R. **Journal of Open Source Software**, 2016.

SILVA, J. A. D. A política monetária nos governos Lula e Dilma. **Revista de Economia da UEG**, Janeiro 2017. Disponível em: <https://www.researchgate.net/profile/Alderir-Silva/publication/340224258_A_politica_monetaria_nos_governos_Lula_e_Dilma/inks/5e7d7ede92851caef4a2130e/A-politica-monetaria-nos-governos-Lula-e-Dilma.pdf>.

THE BANKER. Central Banker of the Year 2018. **The Banker Awards 2018**, 2018. Disponível em: <https://www.bcb.gov.br/pre/asimp/bcimprensa/governor_of_the_year_2018.pdf>.

VALOR. Carta de Tombini explica estouro da meta de inflação em 2015. **Valor**, Janeiro 2016. Disponível em: <<https://valor.globo.com/financas/noticia/2016/01/08/carta-de-tombini-explica-estouro-da-meta-de-inflacao-em-2015.ghtml>>.

WILHER, V. **Clareza da Comunicação do Banco Central e Expectativas de Inflação: evidências para o Brasil**. Niterói: PPGE/UFF, Faculdade de Economia, Universidade Federal Fluminense, Dissertação de Mestrado: [s.n.], 2015. Disponível em: <<http://institutomillennium.org.br/wp-content/uploads/2015/05/Disserta%C3%A7%C3%A3o-V%C3%ADtor-Wilher-Revisada1.pdf>>.