

TÓPICOS, TOM E TAXA DE JUROS: DESAGREGANDO O IMPACTO DA COMUNICAÇÃO DO COPOM NA CURVA DE JUROS*

TOPICS, TONE, AND INTEREST RATES: DISAGGREGATING THE IMPACT OF COPOM'S COMMUNICATION ON THE YIELD CURVE

Matheus dos Santos Breitenbach **
Carlos Eduardo Schonerwald da Silva***

RESUMO

Este trabalho mensura o impacto granular da comunicação do Comitê de Política Monetária (Copom) sobre a estrutura a termo da taxa de juros (ETTJ) brasileira, preenchendo uma lacuna na literatura nacional, que até então focava em índices de sentimento agregados. Para tal, utiliza as atas das reuniões de 2003 a 2025 e emprega a Modelagem de Tópicos (LDA) para identificar seis temas macroeconômicos distintos. Em seguida, constroem-se índices de Otimismo Líquido e Incerteza para cada tópico com base no dicionário financeiro de Loughran e McDonald (2011). A análise econométrica final combina LASSO para seleção de variáveis e MQO para estimação. Os resultados confirmam a hipótese de que o efeito da comunicação não é monolítico. O principal achado é que o mercado reage de formas opostas à origem da incerteza, um aumento inesperado na incerteza sobre o setor externo e fluxos de crédito (Tópico 2) associa-se a uma queda nos juros futuros, precificando uma reação expansionista do BC, enquanto a incerteza sobre os índices de preços e produção industrial (Tópico 4) eleva os juros, aumentando o prêmio de risco doméstico. Adicionalmente, o otimismo sobre a atividade econômica e crescimento (Tópico 3) reduz a volatilidade do mercado, gerando um efeito estabilizador.

Palavras-chave: Política Monetária. Comunicação. Modelagem de Tópicos. Índices de Sentimento. Estrutura a Termo da Taxa de Juros (ETTJ).

ABSTRACT

This paper measures the granular impact of the Monetary Policy Committee's (Copom) communication on the Brazilian yield curve, filling a gap in the national literature, which has so far focused on aggregated sentiment indices. To this end, it uses minutes from 2003 to 2025 and employs Topic Modeling (LDA) to identify six distinct macroeconomic topics. Subsequently, topic-specific Net Optimism and Uncertainty indices are constructed based on the Loughran and McDonald (2011) financial dictionary. The final econometric analysis combines LASSO for variable selection and OLS for estimation. The results confirm the hypothesis that the effect of communication is not monolithic. The main finding is that the market reacts in opposite ways to the source of uncertainty, an unexpected increase in uncertainty regarding the external sector and credit flows (Topic 2) is associated with a fall in future interest rates, pricing an expansionary reaction from the Central Bank, whereas uncertainty regarding price indices and industrial production (Topic 4) raises rates, increasing

* Trabalho de Conclusão de Curso apresentado, em 2025/2, ao Departamento de Economia e Relações Internacionais da Faculdade de Ciências Econômicas da Universidade Federal do Rio Grande do Sul (UFRGS), como requisito parcial para obtenção do título de Bacharel em Ciências Econômicas.

** Discente do curso de Ciências Econômicas da UFRGS (matheusbreitenbach01@gmail.com)

*** Orientador. Doutor em Economia pela University of Utah. Mestre e graduado em Economia pela Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS). Professor do Departamento de Economia e Relações Internacionais da UFRGS (carlos.schonerwald@ufrgs.br).

the domestic risk premium. Additionally, optimism regarding economic activity and growth (Topic 3) reduces market volatility, generating a stabilizing effect.

Keywords: Monetary Policy. Communication. Topic Modeling (LDA). Sentiment Indices. Yield Curve.

1 INTRODUÇÃO

O paradigma de condução da política monetária contemporânea passou por grandes transformações nas últimas décadas. Se no passado, as autoridades monetárias intencionalmente mantinham ambiguidade, hoje o cenário é distinto (Blinder *et al.*, 2008). A transparência e a comunicação clara são vistas não apenas como meios de prestação de contas, mas também como ferramentas essenciais para a eficácia da própria política. A visão moderna pressupõe que o sucesso da política não depende tanto do controle efetivo da taxa de juros de curíssimo prazo, mas sim da capacidade de influenciar as expectativas do mercado sobre a trajetória futura dessa taxa (Woodford, 2001). Em relação aos objetivos da política monetária, a estrutura a termo da taxa de juros (ETTJ) torna-se, portanto, o reflexo mais direto do sucesso (ou fracasso) dessa estratégia de comunicação. A validação empírica de que a comunicação possui um impacto autônomo sobre os preços dos ativos foi consolidada por Gürkaynak, Sack e Swanson (2005), que, ao decompor os anúncios de política monetária, demonstraram que o "fator trajetória" (a palavra) frequentemente possui um impacto superior sobre os juros de longo prazo do que a própria "ação" (a decisão de juros).

Dado que o conteúdo textual exerce um efeito próprio, o desafio metodológico deslocou-se para a mensuração desse impacto, processo que foi catalisado pelos avanços no campo de *Text as Data*. No Brasil, Chague *et al.* (2015) foram pioneiros ao construir um "Fator de Otimismo" (OF) a partir das atas do Copom, utilizando uma metodologia que combina classificações de dicionários de sentimento por meio da Análise de Componentes Principais (PCA). Seus resultados demonstraram que esse fator impactava a curva de juros brasileira. Internacionalmente, trabalhos como o de Hansen e McMahon (2016) avançaram para uma análise mais granular. Os autores utilizaram o modelo *Latent Dirichlet Allocation* (LDA) da Modelagem de Tópicos para desagregar a comunicação do *Federal Open Market Committee* (FOMC), entre comunicação sobre a situação econômica e *forward guidance*. Os autores concluíram que o mercado reage de formas distintas para cada esfera da comunicação. Embora Moreno-Pérez e Minozzo (2024) tenham aplicado LDA às atas do Copom, seu foco foi o impacto na economia real. Portanto, esta abordagem granular, que diferencia os temas, ainda não foi aplicada ao problema da ETTJ no Brasil.

Diante desta lacuna, a questão problema que norteia este estudo é: De que forma a estrutura a termo da taxa de juros (ETTJ) brasileira é impactada pelo tom da comunicação do Copom, quando este é desagregado por tópicos? O objetivo geral é, portanto, mensurar este impacto. Para alcançar esta finalidade, o estudo emprega uma metodologia híbrida que combina processamento de linguagem natural (LDA e dicionário de sentimentos) com regularização econométrica (LASSO), permitindo isolar os tópicos mais relevantes da comunicação.

A contribuição deste estudo ocorre em múltiplas frentes. Do ponto de vista empírico, o trabalho preenche a lacuna identificada ao aplicar uma análise granular de tópicos para mensurar o impacto da comunicação do Copom sobre a ETTJ brasileira, aprofundando os achados de Chague *et al.* (2015). O estudo avança na teoria ao fornecer evidências empíricas de que o efeito da comunicação não é monolítico. Metodologicamente, essa contribuição é alcançada combinando o processamento de linguagem natural com técnicas de regularização econométrica (LASSO), oferecendo um modelo robusto para testar quais índices de sentimento por tópico são os regressores mais relevantes, determinando assim quais canais temáticos

possuem maior impacto sobre a ETTJ. Por fim, esta análise oferece implicações práticas tanto para a autoridade monetária, pois evidencia como a ênfase em diferentes temas é especificada, quanto para os agentes financeiros, que podem refinar seus modelos de especificação¹.

Os principais resultados deste estudo confirmam a hipótese do efeito não uniforme, revelando que o mercado decompõe a origem do tom e reage de formas opostas. O achado mais robusto é que um aumento inesperado na incerteza sobre o Setor Externo e Fluxos de Crédito (Tópico 2) está associado a uma queda significativa nos juros futuros, sugerindo que os agentes especificam uma provável reação expansionista do Banco Central para mitigar o choque. Em contraste direto, a incerteza sobre os Índices de Preços e Produção Industrial (Tópico 4) eleva os juros, refletindo uma elevação do prêmio de risco doméstico. Adicionalmente, o estudo identifica que o otimismo sobre a Atividade Econômica e Crescimento (Tópico 3) possui um efeito estabilizador, reduzindo a volatilidade do mercado após a divulgação da ata.

2 REFERENCIAL TEÓRICO

Nesta seção, é apresentada a fundamentação teórica que serve de base para o presente estudo. A abordagem se inicia com a evolução da discussão sobre a importância da comunicação dos bancos centrais como um instrumento de política monetária, para então explorar os mecanismos teóricos da metodologia e as evidências empíricas já encontradas em trabalhos similares sobre o impacto da comunicação do Banco Central nos mercados financeiros e na economia.

2.1 PAPEL DA COMUNICAÇÃO NA POLÍTICA MONETÁRIA MODERNA

A condução da política monetária contemporânea passou por uma profunda transformação nas últimas décadas. A era em que os bancos centrais operavam sob um manto de ambiguidade intencional deu lugar a uma nova forma de operar, na qual a transparência e a comunicação clara são vistas não apenas como um dever, relacionado à prestação de contas, mas também como ferramentas essenciais para a eficácia da própria política (Blinder *et al.*, 2008). Essa ideia é encapsulada na afirmação de Bernanke (2015), ex-presidente do Federal Reserve, de que a política monetária é "98% comunicação e apenas 2% ação".

Tal mudança de comportamento foi reflexo da evolução teórica do assunto. A visão moderna sugere que a autoridade monetária comunique sua visão sobre o panorama econômico, os riscos e suas prováveis ações futuras. Ao fazer isso, a autoridade consegue ancorar as expectativas e, com isso, aumentar a potência do canal de transmissão da política monetária, moldando ativamente não somente os juros de curto prazo, mas também as expectativas do mercado (Blinder *et al.*, 2008).

A visão moderna também postula que o sucesso da política monetária não depende tanto do controle efetivo da taxa de juros de curto prazo, mas sim da capacidade de influenciar as expectativas do mercado sobre a trajetória futura dessa taxa (Woodford, 2001). A razão para tal é que as decisões econômicas dos agentes, como consumo e investimento, são influenciadas pelas taxas de juros de médio e longo prazo, as quais, por sua vez, são determinadas pelas expectativas sobre a sequência de taxas de juros futuras (Blinder *et al.*, 2008). Em relação aos objetivos da política monetária, a ETTJ torna-se, portanto, o reflexo mais direto do sucesso (ou fracasso) dessa estratégia de comunicação.

¹ Para a elaboração deste trabalho, contou-se com o auxílio da ferramenta de inteligência artificial generativa Google Gemini 2.5 Pro. O uso da ferramenta teve como finalidades: atuar como assistente de pesquisa e redação; realizar a revisão gramatical; aprimorar o estilo da escrita; auxiliar na geração dos códigos de programação empregados na análise; e contribuir para a revisão do artigo.

A materialização dessa "política pela palavra" pode ser compreendida através do conceito de *forward guidance* (orientação futura). Campbell *et al.* (2012) distinguem o *forward guidance* em duas modalidades principais. A primeira, "Délfica", consiste na divulgação de previsões e da visão do banco central sobre a economia, com o intuito de reduzir a incerteza dos agentes. A segunda, "Odisseica", implica um compromisso público com uma ação futura, buscando moldar as expectativas do presente. Um exemplo da abordagem Odisseica é as *open mouth operations* abordadas em Guthrie e Wright (2000), nas quais a comunicação, por si só, move a taxa de juros ao carregar uma "ameaça" crível de que ações serão tomadas caso o mercado não se ajuste na direção desejada.

A validação empírica de que a comunicação possui um impacto autônomo sobre os preços dos ativos foi consolidada por Gürkaynak, Sack e Swanson (2005). Neste estudo, os autores decomponeram os anúncios de política monetária em um "fator meta" (a ação) e um "fator trajetória" (*path factor*), este último associado ao conteúdo textual do comunicado sobre o futuro da política. Seus resultados demonstram que o "fator trajetória" não apenas importa, mas frequentemente possui um impacto superior sobre os juros de longo prazo do que a própria decisão de juros. Este achado fornece a base fundamental para a hipótese deste trabalho: se o conteúdo textual exerce um efeito próprio, é plausível que o tom empregado nesse texto seja um de seus principais vetores de transmissão. Portanto, investigar como o tom dos diferentes tópicos discutidos pelo Copom afeta a curva de juros é um passo natural para aprofundar a compreensão sobre como a comunicação, como instrumento da política monetária moderna, de fato funciona, e quais são seus impactos.

2.2 FUNDAMENTAÇÃO METODOLÓGICA PARA ANÁLISE DE TEXTOS

A transformação de dados textuais não estruturados em variáveis quantitativas para análise econometrônica, um campo conhecido como *Text as Data*, tornou-se uma área de intensa inovação na pesquisa econômica (Gentzkow; Kelly; Taddy, 2019). Dentre as diversas aplicações e métodos, a análise de sentimento se destaca como uma ferramenta conveniente para capturar as percepções e o tom dos agentes econômicos. A abordagem tradicional para essa tarefa é a baseada em dicionários, que consiste em contar a frequência de palavras pré-classificadas como positivas ou negativas em um determinado texto.

Para a abordagem tradicional é preciso ter um cuidado especial quanto à escolha de um dicionário representativo. Exemplo disso foi demonstrado em um dos trabalhos mais influentes na área, por Loughran e McDonald (2011). Os autores demonstraram que dicionários de propósito geral, como o Harvard IV, apresentam baixo desempenho em contextos específicos, como em seu caso, textos financeiros. Palavras como "passivo", "custo" ou "risco", frequentemente associadas a uma conotação negativa, no contexto financeiro possuem conotação neutra ou factual, e não necessariamente negativa. Esse achado ressalta a necessidade de se utilizar dicionários corretos para cada domínio.

A robustez da análise de sentimento pode ser aprimorada para além da simples contagem de palavras. Estudos como o de Shapiro, Sudhof e Wilson (2022) mostram que a combinação de múltiplos dicionários de domínio específico e a incorporação de regras linguísticas simples, como o tratamento de negações (por exemplo, a palavra "não" invertendo o sentido de um termo positivo), aumentam significativamente a acurácia da medição de sentimento.

A literatura metodológica para bancos centrais aponta para um avanço crucial, a combinação da análise de sentimento com a modelagem de tópicos não supervisionada, como o *Latent Dirichlet Allocation* (LDA) (Bholat *et al.*, 2015). Essa abordagem híbrida, central para este projeto, permite uma análise mais granular. Em vez de gerar um único índice de sentimento para um documento inteiro, é possível primeiro identificar sobre o que o banco central está discutindo, segregando em tópicos latentes como "inflação", "atividade econômica" ou

"estabilidade financeira", para então medir o tom específico empregado em cada um desses temas, como explorado nos trabalhos de Hansen e McMahon (2016) e De La Huerta Avila (2024).

Finalmente, é imperativo posicionar a abordagem escolhida (LDA + Dicionários) dentro do espectro de técnicas disponíveis. Embora a fronteira metodológica resida em arquiteturas de *deep learning*, especialmente para aquelas com arquiteturas de *Transformers* inauguradas em Vaswani *et al.* (2017), sua aplicação em estudos econométricos específicos ainda é incipiente e frequentemente associada a uma menor interpretabilidade, o problema da caixa-preta. Além disso, destacam-se a maior simplicidade, menor custo e a menor chance de *data mining* como características do método tradicional.

Dessa forma, a metodologia deste projeto se situa em uma posição estratégica: é sofisticada o suficiente para capturar nuances temáticas (via LDA), indo além das medidas agregadas, ao mesmo tempo que se mantém transparente e fundamentada em dicionários validados pela literatura, garantindo replicabilidade e clareza na interpretação dos resultados.

2.3 ESTUDOS RELACIONADOS

A constatação teórica de que a comunicação dos bancos centrais é um potente instrumento de política monetária impulsionou inúmeras pesquisas empíricas, principalmente utilizando métodos de análise de sentimentos. Um dos trabalhos pioneiros na área foi o de Apel e Blix-Grimaldi (2014). O estudo emprega a contagem de palavras com base em dicionários específicos para quantificar diferentes aspectos da linguagem usada nas atas, focando em capturar o tom da política monetária através de classificações como austero e suave. O objetivo foi verificar se essas medidas textuais contêm informações adicionais, que ajudem a prever a trajetória futura da economia e as próprias decisões de política monetária. Os principais resultados foram positivos, as medidas mostraram capacidade preditiva para futuras alterações na taxa de juros, além de ajudaram a prever variáveis macroeconômicas como o crescimento do PIB.

Para o Brasil, Chague *et al.* (2015) foram pioneiros ao construir um "Fator de Otimismo" a partir das atas do Copom, utilizando uma metodologia que combina em um único fator os diferentes sentimentos do dicionário Harvard IV, por meio da Análise de Componentes Principais (PCA). Seus resultados indicam que um aumento no otimismo da ata leva a uma queda nas taxas de juros de longo prazo, enquanto um tom pessimista eleva a volatilidade, estabelecendo uma importante evidência quantitativa do impacto da comunicação do Copom sobre a curva de juros brasileira.

A evolução desta literatura se deu na busca por uma análise mais granular, reconhecendo que a comunicação de um banco central não é monolítica. O trabalho seminal de Hansen e McMahon (2016) representa um marco metodológico ao ser um dos principais trabalhos a combinar a modelagem de tópicos, usando LDA, com a análise de sentimento para os comunicados do FOMC. Em vez de um único índice agregado, eles constroem medidas de sentimento distintas para *forward guidance* e avaliação da conjuntura econômica, e concluíram que a sinalização sobre o futuro da política monetária possui um impacto macroeconômico consideravelmente maior. Esta abordagem de decompor a comunicação em seus temas constituintes também foi feita por outros pesquisadores, como De La Huerta Avila (2024) para o Banco do México, que também aplicou o arcabouço de LDA e análise de sentimento para investigar o poder preditivo da comunicação sobre as futuras decisões de juros.

O trabalho de Moreno-Pérez e Minozzo (2024) também utiliza o LDA sobre as atas do Copom para construir índices temáticos. No entanto, o foco de sua análise e a métrica construída não reside nos efeitos nos juros, mas sim em mensurar seu impacto sobre variáveis da economia real.

Enquanto Chague *et al.* (2015) analisaram o efeito de um índice de tom agregado sobre a curva de juros, Moreno-Pérez e Minozzo (2024) investigaram o efeito de um índice de incerteza por tópico sobre a economia real, uma análise do impacto de índices de sentimentos desagregados por tópico sobre a ETTJ, até onde temos ciência, não foi explorada no Brasil. É precisamente nesta lacuna que o presente projeto se insere, buscando combinar a maior granularidade gerada pela LDA para aprofundar a análise iniciada por Chague *et al.* (2015) sobre o efeito da comunicação sobre a curva de juros.

3 METODOLOGIA

A presente seção detalha os procedimentos metodológicos adotados para mensurar o impacto da comunicação do Copom na ETTJ brasileira. A metodologia é dividida em quatro etapas principais:

- a) coleta e tratamento dos dados textuais e de mercado;
- b) modelagem de tópicos, através do modelo *Latent Dirichlet Allocation* (LDA), para identificação dos tópicos latentes nas atas;
- c) construção dos índices de sentimento por tópico com base em dicionário;
- d) modelagem econométrica para a análise de impacto.

3.1 COLETA E TRATAMENTO DE DADOS

A construção da base de dados para este estudo envolveu a coleta e o tratamento de duas fontes distintas de informação: de um lado, os dados textuais não estruturados do conteúdo textual das Atas do Copom e, de outro, os dados quantitativos de mercado referentes à ETTJ brasileira. As subseções seguintes descrevem em detalhe a origem, o período de amostragem e, crucialmente, os procedimentos de pré-processamento aplicados a cada uma dessas fontes para prepará-las para os próximos passos do trabalho.

3.1.1 Atas do Copom

Sendo este um trabalho focado na comunicação do Banco Central do Brasil (BCB), a matéria-prima de análise elaborada aqui é o conteúdo textual das próprias Atas do Copom. Para a coleta deste material, utilizou-se a API da instituição para obter tanto o link para download das atas quanto a data de divulgação de cada uma. Com o link em mãos, realizamos o download das atas em formato PDF. Optou-se por baixar as atas em sua versão em inglês, dado que o dicionário de sentimentos selecionado para esta análise (Loughran e McDonald, 2011) foi desenvolvido para este idioma. A amostra selecionada compreende todas as atas publicadas desde a 85^a reunião, de 27 de junho de 2003, até a última minuta disponível no fechamento do 1º semestre de 2025, a Ata número 271, publicada em 24 de junho de 2025. O marco inicial de 2003 se fez necessário devido à disponibilidade da série histórica para os dados da ETTJ.

Após a coleta, o corpus textual foi submetido a um rigoroso processo de pré-processamento. Inicialmente, extraímos o conteúdo de cada ata em PDF para arquivos de texto (.txt). Em seguida, realizamos uma limpeza para remover elementos considerados como ruído, como cabeçalhos, números de páginas, e rodapés. Por fim, reconstruímos a estrutura original dos parágrafos. Realizamos este processo majoritariamente por meio do uso de Expressões Regulares (regex), mas, para um subconjunto de atas, especificamente as de número 201, 207, e 208, onde a extração automatizada apresentou falhas pontuais, realizamos correções manuais para garantir a integridade do texto.

Por fim, submetemos o texto ao processo de lematização, no qual reduzimos as palavras a sua forma canônica, lema. A lematização vincula palavras com significados semelhantes em

uma única palavra, por exemplo, transformando "aumentou" e "aumentando" em "aumentar". Tal passo é fundamental para a análise, principalmente para garantir a relevância dos tópicos encontrados nas etapas futuras. Ademais, removeram-se palavras sem valor semântico (*stopwords*), números, símbolos e sinais do corpus.

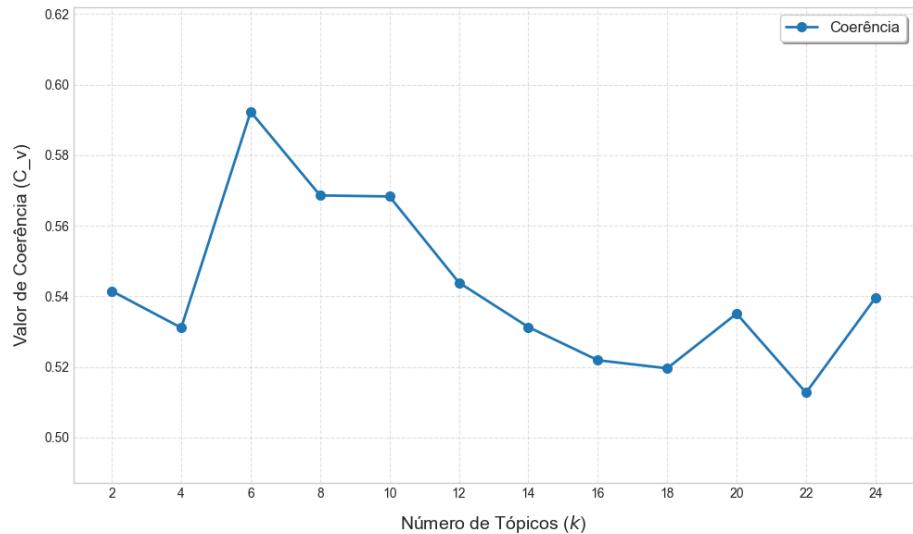
3.1.2 Vértices da curva de juros

Para a obtenção dos dados dos vértices da curva de juros, utilizou-se a base de dados da plataforma Com Dinheiro. Este estudo utiliza os vértices de 21, 63, 126, 252 e 504 dias úteis. Dessa forma, temos dados relevantes da ETTJ que representam tanto a parte de curtíssimo prazo da curva de juros quanto o prazo médio e o início do longo prazo. Ao utilizar estes vértices, obtém-se um termômetro da curva inteira, com enfoque na região de maior liquidez.

3.2 MODELAGEM DE TÓPICOS

Para identificar os temas latentes na comunicação do Copom, empregou-se a modelagem de tópicos através da técnica *Latent Dirichlet Allocation* (LDA). O LDA é um modelo de aprendizagem de máquina não supervisionado para identificar *clusters* de tópicos em conjuntos de documentos. O modelo assume que cada documento é uma mistura de vários tópicos, onde cada tópico, por sua vez, é uma distribuição de probabilidades de palavras. Usualmente, o LDA é utilizado para encontrar tópicos entre diferentes documentos de texto, porém, para a presente análise, foi utilizado com base nos parágrafos. Esta escolha diverge da utilizada em Hansen e McMahon (2016), que utilizaram, no lugar de parágrafos, as frases dos comunicados do FOMC. A abordagem escolhida foi considerada a mais adequada, pois nem sempre o tema de uma ideia se mostra explícito em uma frase por si só, sendo necessário, geralmente, as frases ao seu redor para entender o contexto. Vale ressaltar que foram removidos os parágrafos que tivessem menos de 5 lemas, remoção destes parágrafos visou evitar ruídos vindos dos nomes das seções, que se repetem em inúmeras atas.

Figura 1 - Valor de coerência em função do número de tópicos (k)



Fonte: Elaborado pelo autor (2025).

Para o modelo LDA, existe uma variável que precisa ser definida pelo pesquisador, o número de tópicos k presentes nos documentos. Essa dependência torna a definição do número ótimo de tópicos um processo chave da metodologia, pois, caso a definição seja infundada, é possível que o resultado se torne não relevante. Para evitar a decisão de k por meio de uma

escolha arbitrária, realizou-se o teste de coerência, ou *coherence score*, que serve para avaliar quantitativamente se as palavras que compõem um tópico fazem sentido juntas do ponto de vista semântico e contextual. Para encontrar esse k ótimo, rodamos o modelo LDA para todos os números de k pares de 2 até 24, e assim escolhemos aquele com o maior resultado de coerência, sendo este o modelo com k igual a 6. A figura 1 apresenta o resultado do teste de coerência para os diferentes valores de k.

Quadro 1 - Tópicos identificados pelo modelo LDA e suas principais palavras-chave

Tópico	Palavras-chave
0: Mercado de Trabalho e Atividade Real	<i>employment, increase, real, accord, rate, job</i>
1: Cenário e Projeções de Inflação	<i>rate, inflation, price, projection, meeting, scenario</i>
2: Setor Externo e Fluxos de Crédito	<i>billion, u, operation, total, credit, reach</i>
3: Atividade Econômica e Crescimento	<i>growth, economy, economic, market, demand, activity</i>
4: Índices de Preços e Produção Industrial	<i>increase, month, price, good, production, industrial</i>
5: Política Monetária e Risco	<i>inflation, monetary, policy, copom, committee, risk</i>

Fonte: Elaborado pelo autor (2025).

Com a definição do k ótimo, foi construído o modelo LDA e, assim, identificados os tópicos presentes nas atas do Copom. Após isso, de forma qualitativa, foi definido um nome para cada um desses tópicos. O quadro 1 apresenta os nomes dos tópicos e as seis palavras mais frequentes de cada tópico. Por fim, para cada parágrafo das atas, foi assinalado um tópico mais provável.

3.3 A CONSTRUÇÃO DOS ÍNDICES

Para quantificar o tom empregado em cada tópico, adotou-se uma abordagem de análise de sentimento baseada em dicionário. A escolha recaiu sobre o dicionário financeiro de Loughran e McDonald (2011), validado na literatura por sua especificidade ao contexto financeiro, superando o desempenho de dicionários de sentimento gerais nos temas de sua abordagem. Utilizamos a versão de 2024 do dicionário, que classifica palavras em algumas classes possíveis, sendo elas: *Negative, Positive, Uncertainty, Litigious, Strong_Modal, Weak_Modal, Constraining* e *Complexity*. Para nossa análise, decidiu-se construir dois índices: um de otimismo líquido e outro de incerteza, relacionados às classes *Negative, Positive* e *Uncertainty* do dicionário.

Com as palavras lematizadas e seus tópicos definidos, foi feito o processo de contagem de palavras para cada parágrafo de cada ata. Ou seja, em cada parágrafo contou-se o número de lemas totais e o número de lemas classificados como otimistas, pessimistas e de incerteza. Para tanto, foi necessário lematizar também os termos do dicionário. Com os lemas contados e os tópicos definidos, foram geradas 12 séries temporais (seis de otimismo e seis de incerteza). As definições desses índices serão melhor explicadas abaixo.

3.3.1 Índice de otimismo líquido

Para cada ata, foram agregados todos os resultados das contagens de lemas para cada tópico. Com isso, foi construído o índice de otimismo líquido "OL", que pode ser definido pela equação 1.

$$OL_k = \frac{Palavras_Positivas_k - Palavras_Negativas_k}{Palavras_Totais_k} \quad (1)$$

Estão representados na equação 1 os seguintes termos: k , que representa o tópico ($k = 0, \dots, 5$), $\text{Palavras_Positivas}_k$ é resultado da contagem de lemas que se classificam como positivos no dicionário utilizado, e de forma igual $\text{Palavras_Negativas}_k$ é o resultado da contagem de lemas que se classificam como negativos, enquanto Palavras_Totais_k representa o resultado da contagem de lemas. Com isso, o índice representa a proporção de lemas positivos líquidos presente em cada tópico em questão, podendo variar de -1 a 1.

É importante destacar que o cálculo da contagem de palavras positivas e negativas utilizadas no índice de otimismo líquido considerou uma janela de negação, na qual palavras como "não" ou "nunca", ao estarem presentes imediatamente antes de um lema relevante, invertem o sentido da contagem.

Com o Índice OL gerado, podemos construir a variável que iremos utilizar nos modelos, o índice de choque. Como o foco da análise recai sobre a surpresa da comunicação, isto é, a informação nova contida na ata que possa a vir afetar o mercado, as variáveis independentes escolhidas para os modelos foram a primeira diferença de cada índice, que pode ser representada genericamente como ($\Delta\text{Índice}_{k,t} = \text{Índice}_{k,t} - \text{Índice}_{k,t-1}$).

3.3.2 Índice de incerteza

De forma análoga, foram agregados os resultados da contagem de palavras para cada ata e tópico. Com isso, foi construído o índice de incerteza "INC", podendo ser definido pela equação 2.

$$\text{INC}_k = \frac{\text{Palavras_Incerteza}_k}{\text{Palavras_Totais}_k} \quad (2)$$

Sendo assim, este índice representa a proporção de lemas de incerteza presente em cada ata e tópico em questão, podendo variar de 0 a 1.

3.4 MODELAGEM ECONOMÉTRICA

A etapa final da metodologia consiste na estimativa do impacto dos índices de sentimento na ETTJ em seus vértices selecionados. Para esta análise, foram selecionados dois focos: a análise quanto ao nível dos juros e quanto à volatilidade, seguindo uma análise inspirada pela de Chague *et al.* (2015).

Dada a existência de, no mínimo, seis regressores nos modelos, os quais poderiam ser correlacionados elevando o risco de multicolinearidade e *overfitting*, optou-se por uma pré-seleção das variáveis que iriam compor a regressão final utilizando o *Least Absolute Shrinkage and Selection Operator* (LASSO). O LASSO é particularmente adequado para este contexto, pois realiza a seleção de variáveis de forma automática, ajustando os coeficientes dos regressores menos relevantes a zero e retendo apenas aquelas com maior poder preditivo, removendo as que podemos considerar como espúrias ou sem real poder explicativo. Sendo assim, tanto para os modelos de nível da curva quanto em relação à volatilidade, foi feita inicialmente a seleção das variáveis através do uso do LASSO. Já para a estimativa dos coeficientes, foi utilizada uma regressão via Mínimos Quadrados Ordinários (MQO). Ressalta-se que as variáveis *dummies* não foram incluídas na etapa LASSO, sendo adicionadas somente na segunda etapa, de estimação via MQO.

3.4.1 Modelos de nível

Quebramos a análise quanto ao nível em dois modelos, um para a cada índice gerado. O modelo definido quanto ao otimismo pode ser definido pela equação 3:

$$\Delta ETTJ_{a,v} = const_v + \sum_{k=0}^5 \beta_{v,k} \cdot \Delta OL_{a,k} + \epsilon_v \quad (3)$$

Explicitando os termos do modelo, na equação 3, a variável dependente, $\Delta ETTJ_{a,v}$, representa a variação da taxa de juros no dia da publicação da ata “a” para o vértice de “v” dias úteis, ou seja, o valor percentual do dia, menos o valor do dia anterior. Quanto às variáveis independentes, temos inicialmente o termo “ $const_v$ ”, que representa um coeficiente constante de intercepto específico para cada vértice, capturando o nível médio de variação não explicado pelos regressores. A principal variável explicativa é “ $\Delta OL_{a,k}$,” que quantifica a mudança no otimismo do tópico “k” entre a ata atual e a anterior. O impacto dessa mudança na taxa de juros é medido pelo coeficiente “ $\beta_{v,k}$ ”. O somatório, denotado por “ $\sum_{k=0}^5 \beta_{v,k}$ ”, agrupa o efeito de todos os “k” tópicos. Por fim, o termo “ ϵ_v ” representa o resíduo da regressão.

O modelo para o caso da incerteza foi definido de forma análoga, tendo como forma funcional o modelo descrito na equação 4.

$$\Delta ETTJ_{a,v} = const_v + \sum_{k=0}^5 \beta_{v,k} \cdot \Delta INC_{a,k} + \epsilon_v \quad (4)$$

3.4.2 Modelos de volatilidade

A etapa final da metodologia consiste na estimativa do impacto dos índices construídos na volatilidade de cada vértice estudado da ETTJ.

O processo de construção ocorreu da seguinte forma: primeiramente, foi definido que a variável dependente do modelo seria o logaritmo natural da volatilidade móvel das diferenças de cinco dias dos juros futuros. Foi escolhido no formato logaritmo a fim de mensurar de melhor forma o efeito das variáveis independentes sobre a dependente. Quanto ao período do tempo, foi escolhido cinco dias úteis pois é um período longo o suficiente para termos dados para o cálculo da volatilidade, e curto o suficiente para não adicionar resíduos desnecessários.

Para o caso do índice de otimismo, temos como forma funcional a equação 5.

$$\ln(Vol_5du_{t,v}) = const_v + \beta_0 \cdot dummy_ata_v + \sum_{k=0}^5 \beta_{v,k} \cdot \Delta OL_{a,k} + \epsilon_v \quad (5)$$

Na equação 5, utilizamos como variável dependente a volatilidade em sua forma log normal, $\ln(Vol_5du_{t,v})$, onde “t” representa o tempo, e “v” representa o vértice da ETTJ. Para as variáveis independentes, utilizamos o termo $const_v$, que representa um coeficiente constante de intercepto específico para cada maturidade. Há também uma *dummy* chamada de “*dummy_atá*” que assume o valor 1 caso seja dia de divulgação de ata e 0 caso contrário. O modelo também inclui o somatório, denotado por “ $\sum_{k=0}^5 \beta_{v,k}$ ” que agrupa o efeito de todos os “k” tópicos, onde “ $\beta_{v,k}$ ” é o coeficiente para cada vértice e tópico e “ $\Delta OL_{a,k}$ ” é a variação de índice de uma ata para a outra. Por fim, o termo ϵ_v representa o resíduo da regressão.

O modelo para a incerteza foi desenvolvido de forma análoga, sendo suas definições apresentadas a seguir, na equação 6.

$$\ln(Vol_5du_{t,v}) = const_v + \beta_0 \cdot dummy_ata_v + \sum_{k=0}^5 \beta_{v,k} \cdot \Delta INC_{a,k} + \varepsilon_v \quad (6)$$

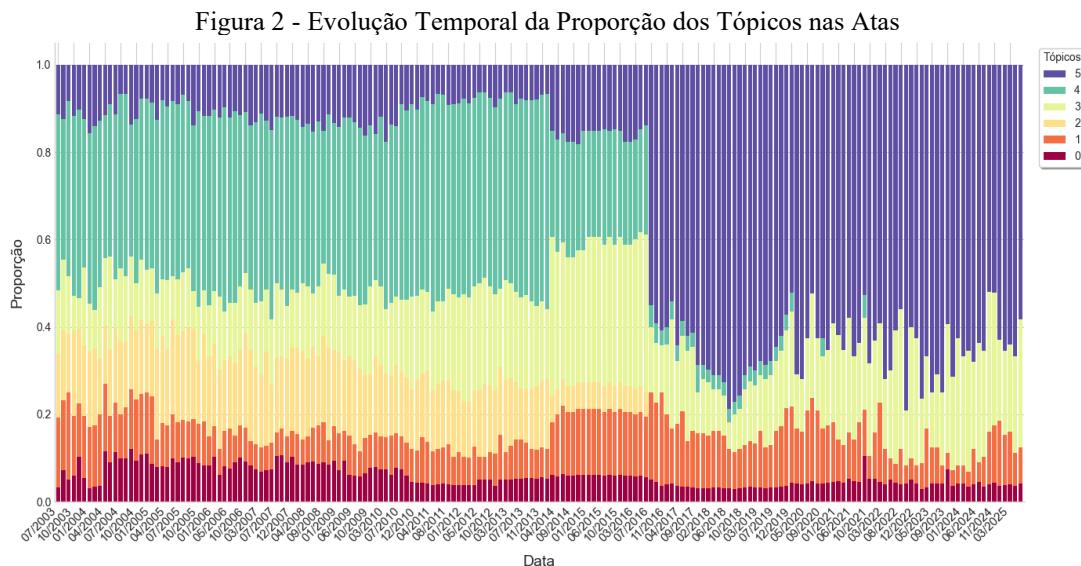
4 ANÁLISE DOS DADOS

Esta seção apresenta os resultados da pesquisa, concentrando-se em dois pontos principais. Primeiramente, realiza-se uma análise descritiva da comunicação do Copom ao longo do tempo, demonstrando como as ferramentas utilizadas no estudo foram capazes de identificar alterações no estilo de comunicação da autarquia. Em seguida, são apresentados e interpretados os resultados dos modelos econôméticos, que buscam quantificar o impacto dos índices de sentimentos, descritos anteriormente, sobre o nível e a volatilidade da ETTJ.

4.1 UM PANORAMA DA COMUNICAÇÃO

Antes de proceder à análise econômética do impacto da comunicação na ETTJ, é fundamental realizar uma análise descritiva dos dados construídos. Esta subseção oferece um panorama da evolução da comunicação do Copom ao longo da amostra (2003-2025), demonstrando como as ferramentas de modelagem de tópicos (LDA) e análise de sentimentos capturaram mudanças estruturais no comportamento da autarquia. Primeiramente, será analisada a evolução da própria composição temática das atas, seguida de uma análise visual da trajetória dos índices de Otimismo e Incerteza ao longo do tempo.

4.1.1 A evolução dos tópicos



Fonte: Elaborado pelo autor (2025).

A figura 2, que ilustra a proporção de cada um dos seis tópicos identificados nas atas do Copom desde 2003, revela uma notável mudança na composição temática ao longo do tempo. O achado mais proeminente é a presença de duas quebras estruturais na composição dos tópicos das Atas. A primeira quebra, menos acentuada, ocorre no início de 2014, em paralelo à redução do número de palavras da ata. A segunda quebra, mais expressiva, ocorre a partir de meados de 2016, período no qual se observa um aumento significativo na dominância do Tópico 5 (política monetária e risco). Em contrapartida, tópicos como o Tópico 2 (setor externo e fluxos de crédito) e o Tópico 4 (índices de preços e produção industrial) ficam no limiar do

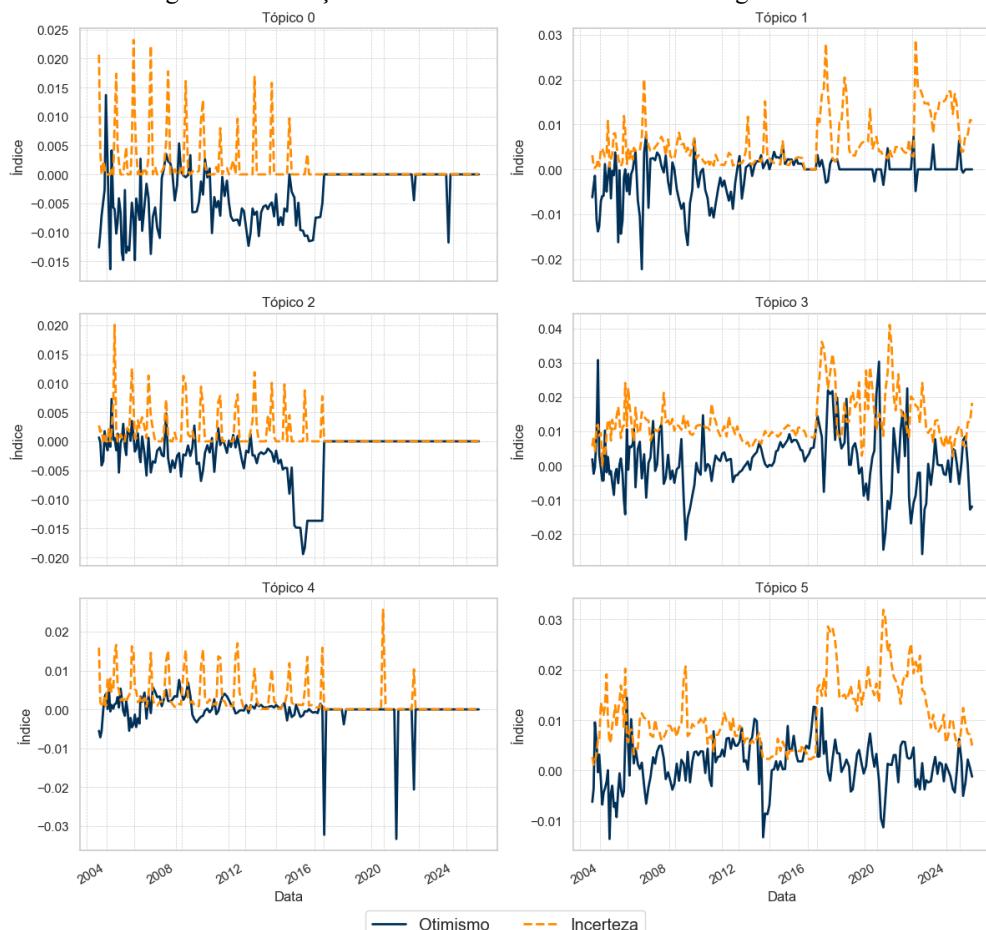
desaparecimento. Essa mudança sugere uma comunicação do Banco Central mais focada em sua deliberação principal e nos riscos associados, em detrimento de descrições detalhadas da conjuntura. O período da segunda quebra identificada é contemporâneo à troca de comando na instituição, com a saída de Alexandre Tombini e a entrada de Ilan Goldfajn. Junto a este período, a ata passou a apresentar novo formato visual, o que talvez indique uma tentativa de sinalizar uma alteração na forma de comunicação da entidade.

4.1.2 A evolução dos sentimentos

A figura 3 detalha a evolução temporal das duas famílias de índices construídos para este trabalho: Otimismo Líquido (linha sólida azul) e Incerteza (linha tracejada laranja), desagregados pelos seis tópicos identificados. Primeiramente, notam-se novamente as características dos índices, enquanto o Otimismo Líquido pode assumir valores negativos (refletindo pessimismo), o índice de Incerteza é estritamente não negativo, capturando a proporção de termos de dúvida.

Outro ponto evidente deste gráfico é a reconfirmação da quebra estrutural que ocorreu em 2016. Os Tópicos 0 (Mercado de Trabalho), 2 (Setor Externo) e 4 (Índices de Preços) apresentam flutuações de sentimento claras de 2004 até aproximadamente 2016. Após esse período, as linhas de ambos os índices (otimismo e incerteza) ficam "achatadas" em zero, com eventuais movimentos. Isso não se deveu devido ao sentimento da comunicação ter ficado neutro, mas sim porque esses tópicos deixaram de ser discutidos, como a figura 2 já apontava. Além disso, nota-se uma tendência sazonal nos índices de incerteza para esses 3 tópicos.

Figura 3 - Evolução dos Índices de Sentimentos ao Longo das Atas



Fonte: Elaborado pelo autor (2025).

Em contrapartida, os Tópicos 1 (Cenário de Inflação), 3 (Atividade Econômica) e 5 (Política Monetária e Risco) mostram atividade contínua ao longo de toda a amostra. Para esses tópicos, é possível notar uma relação inversa entre os índices: picos de Incerteza frequentemente coincidem com vales de Otimismo (azul).

Quanto ao tópico dominante dos últimos anos, o 5º (Política Monetária e Risco), nota-se um caráter visivelmente volátil em toda a série. Destaque para a incerteza (laranja), que atingiu seus maiores picos históricos no período pós-2020, refletindo o cenário desafiador de pandemia e subsequente choque inflacionário.

Tabela 1 - Desvio Padrão dos Índices de Sentimento por Tópico

Tópico	DP OL	DP INC
0	0.0043	0.0052
1	0.0047	0.0043
2	0.0026	0.0036
3	0.0087	0.0055
4	0.0057	0.0053
5	0.0045	0.0036

Fonte: Elaborado pelo autor (2025).

A tabela 1 apresenta as estatísticas descritivas dos índices de sentimento. Os desvios-padrão (DP) desta tabela são cruciais para a interpretação econômica dos coeficientes nas regressões subsequentes, pois permitem mensurar o impacto (em pontos-base) de um choque de um desvio-padrão em cada índice de sentimento.

4.2 RESULTADOS ECONOMÉTRICOS: NÍVEL

Nesta subseção, são apresentados os resultados das regressões que estimam o impacto da surpresa (variação) nos índices de sentimento sobre o nível da ETTJ no dia de publicação da ata.

4.2.1 O impacto do otimismo no nível

A tabela 2 resume os coeficientes da regressão entre a variação dos índices de otimismo e a resposta dos vértices da ETTJ. Uma análise preliminar dos resultados revela dois pontos centrais. Primeiro, o impacto da comunicação de otimismo não é uniforme para os diferentes tópicos e vértices, algo que também se repete para as próximas análises. Segundo, o poder explicativo geral do modelo é baixo.

O primeiro impacto estatisticamente significativo ocorre no Tópico 0 (Mercado de Trabalho e Atividade Real). Uma surpresa otimista de um desvio padrão em ΔOL_0 está associada a uma elevação de 0,8 pontos-base (b.p.) no vértice de 63 dias úteis (d.u.). A interpretação econômica é direta e alinhada à teoria macroeconômica, um mercado de trabalho mais aquecido que o esperado é percebido pelos agentes como um vetor de pressão inflacionária futura. Como resposta, o mercado especifica uma maior probabilidade de uma política monetária mais contracionista (ou menos expansionista) no curto a médio prazo, refletida no vértice de 63 d.u.

O segundo impacto relevante é observado no Tópico 1 (Cenário e Projeções de Inflação), que afeta o vértice de 126 d.u. em que um desvio padrão no índice afeta o vértice em -1,2 b.p. Ou seja, um aumento no otimismo do Copom sobre suas próprias projeções de inflação leva a uma queda nos juros. Este resultado está alinhado com a teoria econômica, dado que

"Otimismo" neste contexto sinaliza que o Banco Central está confiante de que a inflação está sob controle e convergindo para a meta. Diante dessa sinalização, o mercado especifica a trajetória futura da Selic para baixo.

Tabela 2 - Impacto da Variação nos Índices de Otimismo Sobre o Nível da ETTJ

Vértice	Variável Dependente: Δ Taxa DI Futuro				
	(21 d.u.)	(63 d.u.)	(126 d.u.)	(252 d.u.)	(504 d.u.)
const	0.0013 (0.0012)	-0.0004 (0.0032)	0.0063 (0.0057)	0.0056 (0.0081)	0.0111 (0.0105)
Δ OL ₀		1.7782** (0.7365)			
Δ OL ₁			-2.5656** (1.2125)		
Δ OL ₄	0.3099 (0.2045)				
R ² Ajustado	0.0070	0.0256	0.0185	0.0000	0.0000
Observações	134	134	134	134	134

*** p<0.01, ** p<0.05, * p<0.1.

Fonte: Elaborado pelo autor (2025).

Talvez mais revelador do que os coeficientes significativos seja a ausência deles. Tópicos centrais e constantemente presentes nas atas não apresentaram impacto estatístico em nenhum vértice. Isso pode sugerir que, no que tange ao otimismo, o mercado reage menos à avaliação geral da economia ou à própria deliberação da política monetária.

Por fim, é importante discutir a limitação mais evidente do modelo: o baixo poder explicativo. O R² Ajustado é consistentemente baixo em todas as regressões, com um valor máximo de 2,6% no vértice de 63 d.u. Este resultado, embora contido, é um resultado esperado, dado que a curva de juros é naturalmente influenciada por inúmeros outros fatores, e está em linha com os achados de outros pesquisadores. Chague *et al.*, 2015, por exemplo, também perceberam o baixo poder explicativo de seu índice "OF", já Hansen e McMahon (2016) chegaram à conclusão de que o seu índice "EcSit" ao ser testado em um modelo *Factor-Augmented Vector Autoregression* (FAVAR), não tem impacto significativo nos juros.

4.2.2 O impacto da incerteza no nível

A análise de impacto da incerteza, summarizada na tabela 3, apresenta resultados substancialmente mais robustos e economicamente significativos do que os encontrados no modelo de otimismo. O poder explicativo, embora ainda modesto, é visivelmente superior, com o R² Ajustado de 9,14% no vértice de 504 dias úteis. A seleção de variáveis pelo LASSO revela que o mercado reage de forma distinta à incerteza em diferentes segmentos da curva.

O principal vetor de impacto é, sem dúvida, o Tópico 2 (Setor Externo e Fluxos de Crédito). Uma surpresa na incerteza deste tópico provoca um impacto negativo e estatisticamente significativo em quase toda a extensão da curva, apresentando também um aumento na magnitude do impacto. Um choque de um desvio padrão no índice, para o vértice de 504 d.u. causa um efeito de -4,9 b.p. Este resultado, que à primeira vista parece contraintuitivo (pois incerteza geralmente se associa a aumento de prêmio de risco), expõe uma dinâmica crucial.

Teoricamente, a incerteza externa exerce dois efeitos latentes sobre a curva de juros. O primeiro efeito referente ao aumento dos juros é causado tanto devido ao aumento do prêmio

de risco quanto para se defender contra um *flight to quality*. O segundo efeito, aparentemente mais forte, é o de queda nos juros impulsionada pela especificação de uma reação expansionista do Banco Central, para mitigar os efeitos de um choque adverso. O coeficiente que começa inexistente e vai crescendo em magnitude ao longo da curva indica que a leitura do mercado quanto à influência desses dois vetores é variada ao longo da curva.

Tabela 3 - Impacto da Variação nos Índices de Incerteza Sobre o Nível da ETTJ

Vértice	Variável Dependente: Δ Taxa DI Futuro				
	(21 d.u.)	(63 d.u.)	(126 d.u.)	(252 d.u.)	(504 d.u.)
const	0.0014 (0.0012)	-0.0003 (0.0032)	0.0063 (0.0056)	0.0058 (0.0078)	0.0114 (0.0100)
Δ INC ₀			1.2100 (1.1169)	0.4819 (1.8334)	
Δ INC ₂		-1.7299* (0.8878)	-5.3966*** (1.6050)	-9.1881*** (2.3953)	-13.7118*** (3.0685)
Δ INC ₃	-0.3489 (0.2120)				
Δ INC ₄				2.5437 (1.8990)	5.2560** (2.0850)
R ² Ajustado	0.0092	0.0150	0.0484	0.0600	0.0914
Observações	134	134	134	134	134

*** p<0.01, ** p<0.05, * p<0.1.

Fonte: Elaborado pelo autor (2025).

Como contraste, o Tópico 4 (Índices de Preços e Produção Industrial) exibe a reação tradicional ao prêmio de risco. Um aumento de um desvio padrão na incerteza sobre a inflação e a indústria afeta o vértice de 504 d.u. em 2,8 b.p. Este achado de reações opostas encontra um paralelo no trabalho de Moreno-Pérez e Minozzo (2024), que também identificaram efeitos opostos da incerteza dependendo de sua fonte. Em sua análise, focada na economia real, um choque de incerteza sobre “condições econômicas gerais” elevava a inflação, enquanto um choque de incerteza sobre “inflação e política monetária” a reduzia

Em suma, os resultados da incerteza são muito mais ricos que os do otimismo. Eles mostram que o mercado é capaz de decompor a origem da incerteza: quando ela é doméstica e inflacionária (Tópico 4), os juros sobem, quando ela é externa (Tópico 2) ou ligada à atividade real (Tópico 3), os juros caem, com o mercado precificando uma resposta de afrouxamento monetário do Banco Central.

4.3 RESULTADOS ECONOMÉTRICOS: VOLATILIDADE

Esta subseção investiga como a comunicação do Copom, capturada pelos índices de Otimismo e Incerteza, afeta a volatilidade do mercado de juros futuros. A análise se divide em duas partes: o impacto do otimismo e o impacto da incerteza.

4.3.1 O impacto do otimismo na volatilidade

A tabela 4 apresenta os resultados da estimação do modelo descrito na equação 5, buscando avaliar como a “surpresa” no otimismo da comunicação (ΔOL) impacta a volatilidade dos juros futuros, medida pelo logaritmo natural da volatilidade móvel de cinco dias úteis.

Tabela 4 - Impacto da Variação nos Índices de Otimismo Sobre a Volatilidade da ETTJ

Vértice	Variável Dependente: log(Volatilidade Móvel)				
	(21 d.u.)	(63 d.u.)	(126 d.u.)	(252 d.u.)	(504 d.u.)
const	-4.5624*** (0.0129)	-3.9818*** (0.0119)	-3.3683*** (0.0109)	-2.8320*** (0.0096)	-2.4674*** (0.0088)
dummy_ata	0.2485*** (0.0705)	0.2272*** (0.0646)	0.2415*** (0.0592)	0.1799*** (0.0523)	0.1471*** (0.0480)
ΔOL_0		-13.6223 (14.8096)			-14.4752 (11.0087)
ΔOL_2			-37.1850 (24.4910)		-24.0355 (18.2054)
ΔOL_3	-14.4879* (8.0294)	-13.2227* (7.3759)	-14.7746** (6.7212)	-13.5062** (5.9417)	-13.7283** (5.4828)
ΔOL_5	-17.1322 (15.5625)	-15.0901 (14.3636)			-13.2869 (10.6772)
R ² Ajustado	0.0026	0.0028	0.0035	0.0027	0.0030
Observações	5524	5524	5524	5524	5524

*** p<0.01, ** p<0.05, * p<0.1.

Fonte: Elaborado pelo autor (2025).

Primeiramente, observa-se o efeito da “dummy_ata”, que apresenta um valor positivo e altamente significativo. Tal efeito também foi identificado por Chague *et al.* (2015), os quais pontuaram que para uma regressão linear simples, a publicação da ata tem efeito de redução de volatilidade, mas ao adicionar o índice “OF” observou que o sinal do coeficiente ficou positivo para alguns vértices. Isso sugere que, embora o conteúdo textual da ata possa reduzir a volatilidade, o próprio evento, isoladamente, gera um efeito que eleva a volatilidade. Ainda sobre o “dummy_ata”, a magnitude decrescente do coeficiente com o aumento da maturidade sugere que o impacto imediato da comunicação na incerteza do mercado é mais pronunciado nos vértices mais curtos.

Quanto ao impacto dos índices de otimismo, a seleção de variáveis via LASSO reteve diferentes tópicos para diferentes partes da curva, mas o resultado mais consistente e robusto é sem dúvidas os do Tópico 3 (Atividade Econômica e Crescimento). O coeficiente ΔOL_3 é negativo e estatisticamente significativo em todas as maturidades.

A interpretação econômica é que um aumento inesperado no otimismo do Copom sobre a atividade econômica tende a reduzir a volatilidade do mercado de juros. Uma possível explicação é que uma visão mais otimista sobre o cenário de crescimento reduz a incerteza dos agentes sobre a trajetória futura da economia e, consequentemente, sobre a necessidade de reações abruptas da política monetária, acalmando o mercado.

Os demais tópicos selecionados pelo LASSO (ΔOL_0 , ΔOL_2 , ΔOL_5) apresentaram coeficientes também negativos, mas não foram estatisticamente significativos na regressão MQO final. Isso sugere que, embora o LASSO tenha identificado alguma relevância preditiva inicial, esses tópicos não possuem um impacto robusto e discernível na volatilidade.

Finalmente, assim como nos modelos de nível, o poder explicativo geral dos modelos de volatilidade para o otimismo é extremamente baixo, com R^2 Ajustado não atingindo nem mesmo 1% para qualquer vértice. Isso indica que a variação no otimismo da ata, embora com

um efeito detectável vindo do Tópico 3, explica uma fração ínfima da volatilidade diária dos juros.

4.3.2 O impacto da incerteza na volatilidade

A tabela 5 mostra os resultados da estimação do modelo da equação 6, que avalia o impacto da “surpresa” na incerteza da comunicação (ΔINC) sobre a volatilidade dos juros futuros.

Tabela 5 - Impacto da Variação nos Índices de Incerteza Sobre a Volatilidade da ETTJ

Vértice	Variável Dependente: log(Volatilidade Móvel)				
	(21 d.u.)	(63 d.u.)	(126 d.u.)	(252 d.u.)	(504 d.u.)
const	-4.5624*** (0.0129)	-3.9818*** (0.0119)	-3.3683*** (0.0109)	-2.8320*** (0.0096)	-2.4674*** (0.0088)
dummy_ata	0.2488*** (0.0705)	0.2257*** (0.0646)	0.2426*** (0.0592)	0.1794*** (0.0523)	0.1469*** (0.0480)
ΔINC_0			-11.3789 (12.3270)		
ΔINC_4				-17.7916* (9.7479)	
ΔINC_5	26.6799 (19.4747)				
R ² Ajustado	0.0022	0.0020	0.0029	0.0024	0.0015
Observações	5524	5524	5524	5524	5524

*** p<0.01, ** p<0.05, * p<0.1.

Fonte: Elaborado pelo autor (2025).

O resultado mais surpreendente desta análise é a ausência de impacto significativo dos índices de incerteza ΔINC na volatilidade, apenas uma variável relevante apresentou significância estatística. Este resultado contrasta diretamente com a intuição teórica, que associa incerteza a prêmio de risco e volatilidade, e com os achados empíricos de Chague *et al.* (2015), que identificaram um efeito significativo usando um índice de sentimento agregado.

O único resultado encontrado foi em relação ao Tópico 4 (Índices de Preços e Produção Industrial) que apresentou um coeficiente, marginalmente significativo, negativo de -17,79 para o vértice de 252 d.u. Os demais tópicos selecionados ΔINC_0 para 63 d.u. e ΔINC_5 para 21 d.u. não foram estatisticamente significativos.

Notavelmente, a incerteza no Tópico 2 (Setor Externo e Fluxos de Crédito), que demonstrou um impacto forte e significativo no nível dos juros, não se mostrou relevante para explicar a volatilidade em nenhum vértice. Enquanto a incerteza externa leva o mercado a reajustar suas expectativas sobre a trajetória da Selic (nível), ela não parece, no dia da publicação da ata, gerar um aumento ou diminuição da volatilidade do mercado.

5 CONSIDERAÇÕES FINAIS

Este trabalho partiu da premissa de que a comunicação dos bancos centrais vai muito além das decisões de juros, configurando-se como um instrumento ativo de política monetária capaz de ancorar expectativas. Embora o impacto do tom agregado da comunicação já tenha sido explorado no Brasil por Chague *et al.* (2015), havia, ainda, espaço para explorar melhor a granularidade do efeito da comunicação sobre as taxas de juros. Inspirado pela abordagem

internacional de Hansen e McMahon (2016), que decompõe a comunicação usando LDA, este trabalho buscou preencher esta lacuna. O objetivo central foi mensurar o impacto desagregado do tom (otimismo e incerteza) de diferentes tópicos da comunicação do Copom sobre o nível e a volatilidade da ETTJ brasileira.

Para alcançar este objetivo, a abordagem de Modelagem de Tópicos se mostrou eficaz, identificando seis tópicos distintos que compõem as atas do Copom. Uma primeira validação metodológica relevante foi a capacidade do modelo de capturar mudanças estruturais na comunicação da autarquia, principalmente a que ocorreu a partir de meados de 2016, como ilustrado na figura 2.

A análise econométrica demonstrada na tabela 2 foi capaz de revelar que o mercado reage de forma diferente para cada tipo de otimismo. Um tom otimista sobre o mercado de trabalho e atividade real (Tópico 0) foi interpretado como um sinal inflacionário, associando-se a uma alta nos juros de curto prazo. Em contraste, o otimismo sobre o cenário e projeções de inflação (Tópico 1) foi recebido como um sinal de queda nos juros. Os resultados mais robustos, emergiram da análise de Incerteza (tabela 3). O principal achado deste trabalho é que o mercado decompõe a origem da incerteza e reage de formas opostas: a incerteza sobre índices de preços e produção industrial (Tópico 4) é lida como um prêmio de risco doméstico, elevando os juros, Ao passo que a incerteza sobre o setor externo e fluxos de crédito (Tópico 2) leva a uma queda nos juros, efeito que pode ser interpretado como o resultado dos agentes especificando uma futura reação expansionista do Banco Central para mitigar o choque externo.

Na análise de volatilidade, o achado mais claro (tabela 4) foi que o otimismo sobre a atividade econômica e crescimento (Tópico 3) reduz a volatilidade, sugerindo um efeito estabilizador no mercado. Em contrapartida, a incerteza (tabela 5) não se mostrou um vetor robusto para explicar a volatilidade. O único achado, relacionado ao Tópico 4, foi apenas marginalmente significativo ($p < 0.1$) e contraintuitivo (sinal negativo), carecendo de uma interpretação econômica clara e sendo possivelmente apenas ruído estatístico.

Este estudo oferece contribuições em múltiplas frentes. Em termos teóricos, avança na literatura sobre comunicação dos bancos centrais ao fornecer evidência empírica de que o efeito da comunicação não é uniforme, o mercado reage a tons específicos sobre tópicos específicos, sendo capaz de diferenciar sinais economicamente opostos, como os de incerteza externa *versus* interna. Do ponto de vista empírico, o trabalho introduz no contexto brasileiro uma metodologia híbrida (LDA + Dicionário Financeiro + LASSO) para analisar a ETTJ, avançando sobre a análise de Chague *et al.* (2015) e Moreno-Pérez e Minozzo (2024). As implicações práticas são relevantes: para o Banco Central, os resultados mostram que "como" se fala sobre "o que" importa.

Finalmente, o estudo reconhece limitações que abrem caminho para pesquisas futuras. A principal limitação é a suposição de parâmetros constantes, visto que a própria análise descritiva (figura 2) identificou uma forte quebra estrutural na comunicação do Copom a partir de 2016. A segunda limitação reside na linguagem, pois o uso de atas traduzidas para o inglês e de um dicionário focado no mercado financeiro americano, Loughran e McDonald (2011), pode negligenciar nuances da comunicação. Por fim, o baixo poder explicativo dos modelos, embora já esperado, indica que o tom da ata é um fator marginal no dia da sua divulgação, dado que a decisão principal sobre a taxa Selic já foi especificada dias antes. Sugere-se, para pesquisas futuras, a aplicação de modelos que incorporem a quebra estrutural de 2016, o desenvolvimento de um dicionário de sentimento focado para a comunicação da autarquia em português brasileiro, e a comparação da abordagem (LDA + Dicionário) com técnicas baseadas em arquitetura de *transformers*, como BERT ou ChatGPT.

REFERÊNCIAS

APEL, M.; BLIX-GRIMALDI, M. **The information content of Central Bank minutes.** Stockholm: Sveriges Riksbank, 2014. (Sveriges Riksbank Working Paper Series, n. 261). Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2092575. Acesso em: 19 jul. 2025.

BERNANKE, B. S. Inaugurating a new blog. **Brookings**, March 30, 2015. Disponível em: <https://www.brookings.edu/blog/ben-bernanke/2015/03/30/inaugurating-a-new-blog/>. Acesso em: 28 jun. 2025.

BHOLAT, D.; HANSEN, S.; SANTOS, P.; SCHONHARDT-BAILEY, C. **Text mining for central banks.** London: Bank of England, 2015. (Centre for Central Banking Studies, n. 33). Disponível em: <https://www.bankofengland.co.uk/-/media/boe/files/ccbs/resources/text-mining-for-central-banks.pdf>. Acesso em: 19 jul. 2025.

BLINDER, A. S.; EHRMANN, M.; FRATZSCHER, M.; DE HAAN, J.; JANSEN, D. J. Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence. **Journal of Economic Literature**, United States, v. 46, n. 4, p. 910–945, Dec. 2008. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/jel.46.4.910>. Acesso em: 23 jun. 2025.

CAMPBELL, J. R.; EVANS, C. L.; FISHER, J. D. M.; JUSTINIANO, A. **Macroeconomic Effects of Federal Reserve Forward Guidance.** Chicago: FRB of Chicago, 2012. (FRB of Chicago Working Paper No. 2012-3). Disponível em: <https://dx.doi.org/10.2139/ssrn.2166310>. Acesso em: 23 jun. 2025.

CHAGUE, F.; DE-LOSSO, R.; GIOVANNETTI, B; MANOEL, P. Central Bank Communication Affects the Term-Structure of Interest Rates. **Revista Brasileira de Economia**, Brasil, v. 69, n. 2, p. 147–162, jul. 2015. Disponível em: <https://doi.org/10.5935/0034-7140.20150007>. Acesso em: 16 abr. 2024.

De La HUERTA AVILA, C. A. The Predictive Power of Central Bank Communication: Evidence from Mexico. **Sobre Mexico Temas de Economia**, Mexico, v. 1, n. 9, p. 83-127, 2024. Disponível em: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4595144. Acesso em: 15 jun. 2025.

GENTZKOW, M.; KELLY, B.; TADDY, M. Text as Data. **Journal of Economic Literature**, United States, v. 57, n. 3, p. 535-574, Sep. 2019. Disponível em: <https://doi.org/10.1257/jel.20181020>. Acesso em: 01/07/2025.

GÜRKAYNAK, R. S.; SACK, B. P.; SWANSON, E. T. Do Actions Speak Louder than Words? The Response of Asset Prices to Monetary Policy Actions and Statements. **International Journal of Central Banking**, [S.I.], v. 1, n. 1, p. 55- 93, Jun. 2005. Disponível em: <https://mpra.ub.uni-muenchen.de/820/>. Acesso em: 14 jul. 2025.

GUTHRIE, G.; WRIGHT, J. Open mouth operations. **Journal of Monetary Economics**, Netherlands, v. 46, n. 2, p. 489-516, 2000. Disponível em: [https://doi.org/10.1016/S0304-3932\(00\)00035-0](https://doi.org/10.1016/S0304-3932(00)00035-0). Acesso em: 14 jul. 2025.

HANSEN, S.; MCMAHON, M. Shocking language: Understanding the macroeconomic effects of central bank communication. **Journal of International Economics**, Netherlands,

v. 99, p. 114-133, 2016. Disponível em: <https://doi.org/10.1016/j.jinteco.2015.12.008>. Acesso em: 02 abr. 2025.

LOUGHAN, T.; MCDONALD, B. When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks. **The Journal of Finance**, United States, v. 66, n. 1, p. 35-65, Jan. 2011. Disponível em: <https://doi.org/10.1111/j.1540-6261.2010.01625.x>. Acesso em: 02 jul. 2025.

MORENO-PÉREZ, C.; MINOZZO, M. ‘Making text talk’: The minutes of the Central Bank of Brazil and the real economy. **Journal of International Money and Finance**, Netherlands, v. 147, 2024. Disponível em: <https://doi.org/10.1016/j.jimonfin.2024.103133>. Acesso em: 14 jul. 2025.

SHAPIRO, A. H; SUDHOF, M.; WILSON, D. J. Measuring News Sentiment. **Journal of Econometrics**, Netherlands, v. 228, n. 2, p. 221-243, Jun. 2022. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0304407620303535>. Acesso em: 17 jul. 2025.

VASWANI, A.; SHAZER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention Is All You Need. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS 30. Proceding [...]. Long Beach: NIPS, 2017. Disponível em: <https://arxiv.org/abs/1706.03762>. Acesso em: 14 abr. 2025.

WOODFORD, M. **Monetary Policy in the Information Economy**. Cambridge: National Bureau of Economic Research, 2001. (NBER Working Paper No. w8674). Disponível em: <https://ssrn.com/abstract=294723>. Acesso em: 19 jul. 2025.

APÊNDICE A – FOLHA DE APROVAÇÃO

MATHEUS DOS SANTOS BREITENBACH

TÓPICOS, TOM E TAXA DE JUROS: DESAGREGANDO O IMPACTO DA COMUNICAÇÃO DO COPOM NA CURVA DE JUROS

Trabalho de conclusão submetido ao Curso de Graduação em Ciências Econômicas da Faculdade de Ciências Econômicas da UFRGS, como requisito parcial para obtenção do título Bacharel em Economia.

Aprovada em: Porto Alegre, _____ de _____ de 2025.

BANCA EXAMINADORA:

Prof. Dr. Carlos Eduardo Schonerwald da Silva – Orientador
UFRGS

Prof. Dr. Hudson da Silva Torrent
UFRGS

Prof. Dr. Sabino da Silva Porto Júnior
UFRGS