



COMUNICACIÓN DE LA POLÍTICA MONETARIA EN COLOMBIA: UN ANÁLISIS A TRAVÉS DE LA MINERÍA DE TEXTO*

Felipe Roldan Ferrin**

Recibido: 16 de enero de 2023 - Aprobado: 30 de junio de 2023

<https://doi.org/10.22395/seec.v26n60a4431>

RESUMEN

Este artículo examina el efecto de la COVID-19 en la estructura discursiva de las minutas de política monetaria del Banco de la República de Colombia entre 2018 y 2021. Para llevar a cabo esta investigación, se utilizó el algoritmo *Latent Dirichlet Allocation* (LDA) y un corpus de texto con 35 minutas. Los resultados indican que el modelo estimado con las minutas de 2020-2021 presentó un tópico adicional y palabras nuevas relacionadas con la crisis sanitaria en comparación con el modelo de 2018-2019. Además, el análisis de las redes mostró cambios en las conexiones entre las palabras y las temáticas principales subyacentes en el discurso plasmado en las minutas que se produjeron durante la crisis. Esto sugiere que la Junta Directiva del Banco se adaptó a la coyuntura económica y agregó información de vanguardia en sus comunicaciones, dando un giro en torno al uso de sintagmas en la estructura sintáctica de las minutas. En este sentido, se evidencia el compromiso del Banco con respecto a la inclusión de información de primera mano en sus comunicaciones en aras de una comunicación transparente de la política monetaria.

PALABRAS CLAVE:

Comunicación de la política monetaria; LDA; política económica; Banco de la República de Colombia.

CLASIFICACIÓN JEL:

C63, E52, E58

CONTENIDO:

Introducción; 1. Revisión de la literatura; 2. Metodología; 3. Resultados; 4. Conclusiones; Referencias.

* Este artículo es producto del trabajo de investigación (Comunicación de la política monetaria en tiempos de cólera. Un análisis exploratorio de la estructura sintáctica de las minutas del Banco de la República de Colombia, a través del modelamiento de tópicos durante la crisis de la COVID-19) presentado para optar el título de economista en la Universidad del Valle, reconocido por los evaluadores como trabajo meritorio y presentado como ponencia en el XI seminario de economía aplicada auspiciado por la universidad de Medellín y la red de investigadores del Banco de la República. Agradezco a Juan Acosta (director del trabajo de investigación y profesor titular de la Universidad del Valle) por sus valiosos aportes e importantes sugerencias.

** Economista, Universidad del Valle, Cali, Colombia. Estudiante de la maestría en economía aplicada, Universidad del Valle, Cali, Colombia. Correo electrónico: roldan.felipe@correounivalle.ude.co. Orchid: <https://orcid.org/0000-0001-7703-4864>

MONETARY POLICY COMMUNICATION IN COLOMBIA: AN ANALYSIS THROUGH TEXT MINING

ABSTRACT

This article examines the effect of COVID-19 on the discursive structure of the monetary policy minutes of the Banco de la República de Colombia between 2018 and 2021. To conduct this research, the Latent Dirichlet Allocation (LDA) algorithm and a text corpus with 35 minutes were used. The results indicate that the model estimated with the minutes from 2020-2021 presented an additional topic and new words related to the health crisis compared to the 2018-2019 model. In addition, the analysis of the networks showed changes in the connections between words and the underlying main themes in the discourse contained in the minutes produced during the crisis. This suggests that the Board of the Bank adapted to the economic situation and added cutting-edge information to their communications, shifting the use of syntagms in the syntactic structure of the minutes. In this sense, it demonstrates the Bank's commitment to including firsthand information in its communications for the sake of transparent monetary policy communication.

KEYWORDS:

Monetary policy communication; LDA; economic policy; Banco de la República de Colombia.

COMUNICAÇÃO DA POLÍTICA MONETÁRIA NA COLÔMBIA: UMA ANÁLISE ATRAVÉS DA MINERAÇÃO DE TEXTO

RESUMO

Este artigo examina o efeito da COVID-19 na estrutura discursiva das atas de política monetária do Banco de la República de Colômbia entre 2018 e 2021. Para realizar essa pesquisa, foi utilizado o algoritmo de Alocação de Dirichlet Latente (LDA) e um corpus de texto com 35 atas. Os resultados indicam que o modelo estimado com as atas de 2020-2021 apresentou um tópico adicional e novas palavras relacionadas à crise de saúde em comparação com o modelo de 2018-2019. Além disso, a análise das redes mostrou mudanças nas conexões entre palavras e nos principais temas subjacentes no discurso contido nas atas produzidas durante a crise. Isso sugere que o Conselho do Banco se adaptou à situação econômica e adicionou informações de ponta às suas comunicações, mudando o uso de sintagmas na estrutura sintática das atas. Nesse sentido, demonstra o compromisso do Banco em incluir informações em primeira mão em suas comunicações em prol de uma comunicação transparente da política monetária.

PALAVRAS-CHAVE:

Comunicação da política monetária; LDA; política econômica; Banco de la República de Colômbia.

INTRODUCCIÓN

La política monetaria es un instrumento fundamental para las economías y su estabilidad. En tiempos de crisis, la política monetaria ha desempeñado un papel esencial en la recuperación de las economías en estas situaciones. La comunicación de estos movimientos y acciones por parte de los bancos centrales desempeña un papel protagónico en el diseño y la efectividad de la política (Sims, 2009). Por lo tanto, es de gran importancia para los agentes económicos incorporar, en su proceso de toma de decisiones, no solo las acciones de los bancos centrales, sino también su justificación y la percepción de estos movimientos por parte del público, así como las expectativas sobre el rumbo de la economía de cara al futuro.

La comunicación de los bancos centrales se realiza a través de diversos canales, entre los cuales se incluyen las minutas de las reuniones de la junta directiva, los informes del equipo técnico, los informes, las conferencias de prensa, así como los discursos y entrevistas a los responsables de la toma de decisiones en el banco central (la Junta Directiva en el caso colombiano). Estos canales constituyen la principal vía que conecta a los banqueros centrales con los agentes económicos en su conjunto. A través de estos medios se proporciona información de naturaleza numérica, como proyecciones e indicadores, y de naturaleza narrativa, como consideraciones cualitativas sobre la coyuntura económica. En el caso del Banco de la República de Colombia, las minutas desempeñan un papel fundamental en la comunicación y la exposición de las discusiones en torno a la política monetaria por parte de la Junta Directiva (JDBR). La importancia de estas minutas, publicadas unos días después de la reunión de la Junta, radica en que contienen información de primera mano discutida por la Junta y resumen sus consideraciones. Como lo menciona Valencia (2021):

La Resolución Interna 1 de 2007 (24 agosto) en su artículo 2.º señala que las minutas contendrán: 1. Un resumen de la información económica analizada por la Junta para tomar la decisión de tasas de interés. 2. Un análisis de la situación de los precios y de las perspectivas de la economía. 3. Una evaluación de los riesgos para el cumplimiento de la meta de inflación. 4. La decisión en materia de tasas de interés de intervención, indicando si ésta se tomó de manera unánime o por mayoría. (p. 290)

Analizar estos documentos puede revelar qué tipo de temáticas están presentes en el discurso de la JDBR.

Durante la crisis sanitaria provocada por la COVID-19, los bancos centrales jugaron un papel central para contrarrestar sus efectos nocivos. Este fenómeno sin precedentes llevó a los banqueros centrales a entablar otro tipo de discusiones

más acordes a la situación. Ante los diversos problemas que la pandemia produjo en la economía, como choques de oferta, aislamientos, rupturas de las cadenas de abastecimiento, desempleo, entre otros, los bancos centrales llevaron a cabo políticas monetarias no convencionales, como la expansión cuantitativa (QE)¹ y el uso de los derechos especiales de giro (SDR)². Además, se implementaron políticas más comunes, como operaciones repo, flexibilización del crédito, reducción del encaje bancario, operaciones de mercado abierto, modificación de la tasa de intervención, entre otras, todo con el fin de impulsar la demanda. El Banco de la República no fue la excepción.

Este escenario motiva la presente investigación, que tiene como objetivo determinar si la pandemia de la COVID-19 provocó un cambio en los sintagmas (palabras o términos) y/o tópicos (temas) principales que aparecen en el discurso subyacente en las minutas de la Junta Directiva del Banco de la República de Colombia durante el periodo 2018 al 2021.

Con las herramientas computacionales que ofrece el minado de texto y la construcción de un corpus de texto basado en las 35 minutas producidas entre 2018 y 2021, se implementa la metodología LDA (Latent Dirichlet Allocation) para estimar dos modelos: uno para el periodo comprendido entre 2018 y 2019 y otro para el periodo comprendido entre 2020 y 2021. A través de estos modelos, una revisión individual de las minutas y una visualización de las palabras más frecuentes dentro de las mismas, se construyeron gráficos de redes para analizar la interacción entre los términos implícitos en las minutas y cambios en el protagonismo de los nodos (términos) ante la presencia de la pandemia. Además, se crearon gráficos para mostrar el peso relativo (presencia) de los tópicos en ambos modelos dentro de las minutas producidas antes y durante la pandemia.

Los principales resultados muestran que las minutas producidas en el periodo previo a la pandemia (2018-2019) se centran en torno a 3 tópicos, mientras que las producidas durante la pandemia (2020-2021) se enfocan en torno a 4 tópicos. Con la llegada de la pandemia, surgió un tópico adicional. Dentro de los tópicos del modelo 2020-2021, aparecen palabras relacionadas con el fenómeno de la crisis sanitaria, como liquidez, recuperación, pandemia, aislamiento, entre otras. En lo que respecta a la presencia de los tópicos dentro de las minutas, se evidenció que durante el periodo 2018-2019, la importancia de los tópicos varía, y ciertos tópicos

¹ La expansión cuantitativa (*Quantitative Easing*) consiste en aumentar la liquidez de la economía mediante la compra de activos financieros más allá de la deuda pública de corto plazo—a la que suelen limitarse los bancos centrales.

² Los Derechos Especiales de Giro (*Special Drawing Rights*, SDR) son un activo de reserva internacional creado en 1989 por el FMI para complementar las reservas oficiales de los países miembros.

son protagonistas en el discurso de la JDBR durante ciertos periodos. Dentro de las minutas producidas durante la pandemia (2020-2021), la presencia de los tópicos es mucho más balanceada; sin embargo, hay tópicos que adquieren mayor relevancia en algunas minutas, lo que demuestra la adaptación de las discusiones de la JDBR a la situación coyuntural de la economía colombiana.

Además, en el gráfico de redes que representa el periodo de la pandemia, se observa que los términos relacionados con la misma ganan protagonismo e interactúan con los nodos principales, como política monetaria y tasa de intervención. Esto pone de manifiesto el cambio en la estructura discursiva dentro de las minutas y en las discusiones de la JDBR sobre la política monetaria.

En adición a la introducción, este artículo se divide en cuatro secciones. La primera expone la revisión de la literatura existente sobre el tema de investigación. La segunda sección describe las herramientas y algoritmos utilizados para el análisis del texto subyacente en las minutas. La tercera sección presenta los resultados obtenidos mediante la aplicación de los algoritmos y el uso de redes lingüísticas. Finalmente, se presentan las conclusiones.

1. REVISIÓN DE LITERATURA

El uso de las herramientas que ofrece el análisis de texto ha ganado cierta popularidad en los últimos años en relación a las comunicaciones de los bancos centrales, ya que permite aproximarse, desde una perspectiva diferente, a las discusiones y decisiones que se toman en torno a la política monetaria. En los comunicados de los bancos centrales subyace información que respalda las acciones de política monetaria y, por lo tanto, recolectar, condensar y utilizar esa información puede ser de gran utilidad para los agentes económicos. Con el propósito de aumentar las posibilidades de reunir datos y analizar la estructura discursiva y narrativa presente en los comunicados de los bancos centrales, se ha empezado a implementar recientemente el uso de las herramientas que brinda el análisis de texto.

En el contexto internacional, se encuentran varios trabajos que incorporan las herramientas de la lingüística computacional en el análisis de las comunicaciones de los bancos centrales. Hendry y Madeley (2010) emplean el método LSA (Latent Semantic Analysis) para extraer información de las declaraciones en los comunicados del Banco de Canadá, con el propósito de evaluar qué tipo de información afecta los rendimientos y la volatilidad en los mercados de tasas de interés a corto y largo plazo durante el periodo 2002-2008. En términos generales, encuentran que temáticas en torno al riesgo geopolítico, shocks externos, principales shocks internos

(SARS y BSE), entre otros, afectan significativamente los rendimientos y la volatilidad de los mercados, especialmente a corto plazo.

Boukus y Rosemberg (2006) analizan, para el periodo comprendido entre 1987 y 2005, las minutas del Comité Federal de Mercado Abierto (FOMC³ por sus siglas en inglés). Aplicando la metodología LSA, descomponen cada acta en torno a sus temas centrales. Los investigadores demuestran que los temas están estrechamente correlacionados con el contexto macroeconómico (condiciones económicas actuales y futuras) y financiero. Además, Boukus y Rosemberg (2006) encuentran que la información sobre el nivel de incertidumbre de la política monetaria y las perspectivas económicas tienen un efecto en los rendimientos de los bonos del Tesoro estadounidense. Los investigadores sugieren que es posible extraer información valiosa y de gran utilidad para comprender mejor las acciones de política monetaria a partir de estas minutas.

Para Colombia, también existen algunos trabajos que han implementado los algoritmos proporcionados por el análisis de texto en las comunicaciones del banco central. Guío et al. (2020) analizaron dos tipos de documentos fundamentales mediante los cuales se comunica el Banco de la República: las minutas y los informes de política monetaria (IPM) durante el periodo de marzo de 2007 a diciembre de 2018. Los autores encontraron que dichos instrumentos de comunicación se centraban en ocho temas⁴, siendo el tema de sectores económicos y demanda interna el más frecuente en los textos. Además, descubrieron que después de 2012, el tema de inflación, meta y expectativas adquirió mayor relevancia en las minutas, en contraste con su menor importancia en los IPM durante el mismo período. El tema de inflación de alimentos y regulados tuvo una presencia baja y similar en ambas minutas e IPM, mientras que el tema de proyecciones macroeconómicas tuvo una mayor presencia en ambos documentos. Guío et al. (2020) también calcularon un indicador de ciclicidad que muestra la importancia de un tema en función del ciclo de la política monetaria. Se encontró que el tema relacionado con la demanda interna y sectores fue el más relevante tanto en períodos de aumento como de disminución de la tasa de interés.

Meneses y Ávila (2019), del grupo de investigaciones económicas de Bancolombia, emplearon diversos métodos de análisis de texto para caracterizar las minutas y comunicados del Banco de la República, desde julio de 2007 hasta diciembre de

³ El Comité Federal de Mercado Abierto corresponde al Sistema de la Reserva Federal que tiene la obligación de supervisar las operaciones de mercado abierto en los Estados Unidos.

⁴ La diferencia con respecto a la cantidad de tópicos del presente artículo y los resultados de Guío et al. (2020), radica en la cantidad de documentos que se usaron. Ellos usan aproximadamente 216 y en el presente artículo solamente se usan 35.

2019 (139 documentos), con el propósito de reducir los errores de pronóstico de la tasa de intervención. Realizaron un conteo de las palabras subyacentes en los documentos y utilizaron los métodos LDA y LSA. El primero se utilizó para realizar un análisis probabilístico de los tópicos en cada documento, mientras que el segundo se empleó para llevar a cabo una inspección semántica de los textos.

Basándose en el uso de estos métodos, entre otros resultados, Meneses y Ávila (2019) también encontraron que la relevancia de los tópicos variaba de un año a otro. El tópico de inflación predominó en la estructura sintáctica de los documentos entre 2014 y 2015, lo que coincidió con choques inflacionarios durante ese período. El tópico relacionado con el sector externo cobró relevancia en 2016, y el tópico relacionado con el crecimiento económico fue protagonista en 2017. Los autores pronosticaron que la tasa de intervención de política monetaria se mantendría estable en el corto plazo, pero con signos de entrar en un período alcista en 2020. También descubrieron que al agregar la información subyacente en los comunicados y minutas en forma de tópicos, les permitió pronosticar una tasa del 4.25% para el primer trimestre, un resultado que coincidió con lo que sucedió en Colombia, y un aumento de 25 p.b. para los períodos siguientes.

Arango et al. (2017) centraron su investigación en el análisis de los comunicados de prensa del Banco para el período comprendido entre septiembre de 2004 y marzo de 2016. A diferencia de Guío et al. (2020) y Meneses y Ávila (2019), los autores utilizaron exclusivamente la metodología LSA. Este algoritmo es muy similar al LDA, ya que también se utiliza para caracterizar y analizar un conjunto de temáticas subyacentes en los textos, pero difiere en que los tópicos no cuentan con la propiedad de ser probabilísticos. La metodología LSA no se basa en la lógica de la probabilidad bayesiana utilizada en el método LDA. En el algoritmo LSA, las palabras y los documentos se asignan a los tópicos con una probabilidad igual a 1 (cada palabra se ubica en un único tema). Implementando el método LSA, los autores encontraron los tópicos más relevantes subyacentes en los textos. Además, descubrieron que en las minutas, el tema más relevante dentro de las discusiones llevadas a cabo por la JDBR era la inflación, lo cual está en sintonía con el propósito constitucional del Banco. Luego, utilizando los tópicos identificados previamente, calcularon un modelo VAR de tipo estructural en el cual incluyeron un factor que modela las expectativas de inflación y un indicador que modela la situación económica. Al estimar este modelo, los autores encontraron que los comunicados de prensa del Banco sí tenían un efecto en la formación de las expectativas de los agentes económicos y, por lo tanto, podían ser un instrumento de política monetaria útil para intervenir en la economía.

Por último, Taborda (2015) amplía su análisis a otros países. En su investigación, examina las minutas de cinco bancos centrales del continente americano: Chile, Colombia, Brasil, Perú y México. El autor enmarca su investigación en lo que denomina "transparencia procedimental"⁵, centrándose en tres aspectos: extensión, complejidad del vocabulario empleado y legibilidad, y su evolución a lo largo del tiempo. Taborda (2015) analiza un total de 564 documentos, entre actas y comunicados (80 para el caso colombiano), abarcando el período comprendido entre 2007 y 2014. Para el caso de Colombia, el autor encuentra que las minutas presentaron una tendencia creciente en cuanto a su longitud. En relación al vocabulario empleado, la palabra "alimentos" (que hace referencia a la inflación de alimentos) desempeña un papel fundamental en la estructura sintáctica de las minutas. Según el autor, esto refleja la importancia que tiene este tema en las discusiones del Banco de la República de Colombia. Un resultado quizás menos sorprendente, dada la estrategia de inflación objetivo del país, es la asociación que el autor encuentra entre los términos inflación, meta, expectativas y precios para el caso colombiano.

En términos de legibilidad, las minutas se mantuvieron constantes y se ubicaron en una escala de complejidad intermedia, una característica similar a la de otros bancos centrales analizados en la investigación.

Aunque este artículo se enfoca en revelar los cambios en los sintagmas y temas centrales de discusión de la JDBR tras la pandemia, no se debe pasar por alto la amplia literatura existente sobre los efectos que tiene la comunicación de la política monetaria y la transparencia en las diversas variables macroeconómicas. Anzoátegui-Zapata y Galvis-Ciro (2019, 2022), Peteiro, Salcines y Orosa (2002), Se-laive y González (2015), Vega (2018) y Camora (2015) destacan la importancia de una comunicación precisa, una redacción clara y una estructura adecuada para reducir la volatilidad en los mercados y fomentar la confianza de los agentes, lo que pone de manifiesto el papel fundamental de la comunicación de la política monetaria por parte de los bancos centrales en la economía.

La transparencia también desempeña un papel fundamental en las comunicaciones de la política monetaria (Bascand, 2013; Dincer y Eichengreen, 2009). Como afirmó Poole (2001), los agentes pueden tomar decisiones más acertadas y realizar predicciones más precisas cuando pueden anticipar las acciones de los bancos centrales. Por lo tanto, la transparencia en las comunicaciones juega un papel crucial en las decisiones de los diferentes agentes. Trabajos como los de Eusepi y Preston (2007), Svensson (1998) y Neuenkirch (2012) sugieren que proporcionar más

⁵ Un aspecto fundamental dentro de los bancos centrales que poseen esquemas de inflación objetivo, que consiste en la relevancia y la transparencia con la que se hacen los comunicados de los banqueros centrales.

información sobre el proceso de toma de decisiones de política monetaria y el uso de parámetros y variables puede mejorar el bienestar de los agentes. Sin embargo, investigaciones como las de Van der Cruisen, Eijffinger y Hoogduin (2010) y Walsh (2006) señalan que la transparencia puede generar bienestar hasta cierto punto, lo que depende en gran medida de las características de cada mercado.

Junto con las investigaciones mencionadas anteriormente y mediante el uso del algoritmo LDA, este trabajo analiza los posibles efectos de un choque externo en el discurso subyacente de las minutas del Banco de la República de Colombia. Aunque la literatura anteriormente referenciada ha explorado los períodos de tiempo previos a 2019, este artículo llena el vacío existente al analizar todas las minutas comprendidas entre 2018 y 2021. Además, se aborda un aspecto poco explorado: los efectos de choques externos a las discusiones y comunicados del Banco de la República.

Además de la literatura mencionada, esta investigación se basa en el trabajo de Fontan et al. (2021), quienes se centran en un tema específico, la inequidad, y lo analizan mediante la implementación de herramientas de análisis de texto. Los autores examinan 2500 documentos de la Reserva Federal (Fed) y el Banco de Canadá (BoC) desde 2015 hasta finales de 2020, rastreando la presencia de palabras asociadas con el tema de la inequidad para determinar si estas palabras, o las relacionadas con el tema, comienzan a aparecer con mayor frecuencia después de la aparición de la COVID-19. Los autores concluyen que, efectivamente, el discurso de estos bancos centrales cambió debido a la situación coyuntural que atravesaba la economía durante la pandemia. Además, destacan los efectos redistributivos de las políticas monetarias implementadas durante la pandemia para abordar los impactos negativos de dicho fenómeno. En última instancia, los autores concluyen que el discurso de los banqueros centrales puede adaptarse a las situaciones coyunturales y abordar temas adicionales, como la redistribución y la inequidad.

El trabajo de Fontan et al. (2021) complementa las aproximaciones realizadas en los artículos mencionados anteriormente, ya que se enfoca en analizar la estructura discursiva de los comunicados, centrándose en un tema específico relacionado con la coyuntura de las políticas monetarias implementadas en todo el mundo durante la pandemia.

Hasta donde se sabe, no ha habido un enfoque desde la perspectiva discursiva que se centre en comprender los efectos de fenómenos externos en las discusiones del Banco de la República de Colombia. Por lo tanto, este trabajo busca llenar ese vacío al integrar las aproximaciones realizadas previamente por otros

investigadores en relación a las comunicaciones del banco y poner énfasis en los efectos que los choques exógenos y la situación coyuntural tienen en el discurso de los banqueros centrales.

2. METODOLOGÍA

Esta sección se divide en tres partes. En la primera, se expone el algoritmo utilizado para llevar a cabo el modelado de tópicos. En la segunda parte, se describe la forma en que se obtuvieron y prepararon los datos (corpus de texto). La tercera parte describe la estimación y la calibración de los modelos.

2.1 Latent Dirichlet Allocation (LDA)

El machine learning o aprendizaje automático es un campo de la ciencia de datos que se basa en programar computadoras para que puedan aprender a descifrar patrones complejos de manera autónoma en conjuntos de datos (Géron, 2017). Estos sistemas no se limitan exclusivamente al análisis de datos numéricos, sino que también permiten el análisis de texto de manera que pueda ser cuantificado. Para esta tarea, se han diseñado diversas técnicas computacionales y algoritmos estadísticos que facilitan el procesamiento del lenguaje natural con el propósito de investigar de forma cuantitativa datos en formato de texto.

Dentro de las técnicas en el área de la lingüística computacional o el análisis de texto, existen algoritmos que se basan en dos enfoques: aprendizaje supervisado y no supervisado. Dentro de la categoría de los algoritmos no supervisados se encuentran Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA), entre otros. LSA se basa en capturar la frecuencia con que aparecen los términos dentro de una matriz de documentos o palabras, explorando las posibles combinaciones lineales que puedan explicar tanto la varianza de los términos a través de los documentos como la varianza de los documentos a través de los términos (Bholat et al. 2015, p. 10-11). El algoritmo funciona de manera similar a los análisis de componentes principales. Como mencionan Bholat et al. (2015), el método LSA opera de manera determinista ya que asigna las palabras y los documentos únicamente a un tema, lo que limita el análisis y elimina la posibilidad de que la misma palabra pueda aparecer en diferentes documentos para su uso en varios contextos (p. 11).

La asignación de LDA resuelve este problema. El método LDA asigna probabilidades a las palabras y a los documentos, lo que permite que la misma palabra aparezca en diferentes temas y las clasifica en función de la probabilidad, lo que a su vez conserva el contexto en el que se utilizaron. Esto hace que el análisis de

texto sea más completo y profundo. Como lo describen Blei et al. (2003), LDA es un modelo probabilístico que permite identificar temas (tópicos) comunes latentes en un corpus de texto cualquiera. Cualquier texto podría considerarse una mezcla de un número específico de temas (tópicos), asignados a cada uno con una cierta probabilidad de ocurrencia, y cada tópico está compuesto por una distribución de probabilidad sobre las palabras que aparecen en él.

En términos formales, el algoritmo trabaja estimando una probabilidad condicional (bayesiana) basada en las palabras (el único aspecto observable) que aparecen en los documentos:

$$p(\theta, z; w, \alpha, \beta) = \frac{p(\theta, z, w; \alpha, \beta)}{p(w; \alpha, \beta)} \quad (1)$$

Donde:

z es el conjunto con n tópicos.

w es el conjunto de las palabras dentro de todos los textos.

α es la probabilidad de que un tópico z_i aparezca en un documento .

β es la probabilidad de que una palabra w_i aparezca en un tópico z_i .

θ es la probabilidad de α y β mezcladas.

Como enfatizan Blei et al. (2003), estimar esta distribución se torna difícil (p. 1003) y, a pesar de la existencia de métodos como las técnicas bayesianas variacionales, hay una restricción en términos computacionales para poner en marcha este método. No obstante, Hoffman et al. (2010) elaboraron un algoritmo capaz de analizar cantidades considerables de documentos y que permite realizar inferencia entrenando al algoritmo.

Existen técnicas implementadas en paquetes de software que permiten aproximarse de forma más sencilla a los parámetros dentro de la distribución. Siguiendo las instrucciones de Single y Robinson (2017), en el presente artículo se hará uso del software R y los paquetes `tm`, `topicmodels`, y `quanteda`, que permitirán implementar el algoritmo LDA y facilitarán la obtención de los parámetros.

2.2 Elaboración del corpus de texto y construcción de la base de datos

Para la aplicación del algoritmo LDA, es fundamental seleccionar los documentos o textos que servirán para armar el corpus de texto, que es el insumo principal

para llevar a cabo el entrenamiento del algoritmo y, posteriormente, la obtención de los resultados mediante el modelado de tópicos. Para el presente trabajo de investigación, se condensan todas las minutas producidas por la JDBR entre enero de 2018 y diciembre de 2021.

Como se mencionó en la introducción, existen varios tipos de documentos que funcionan como instrumentos de comunicación por parte del banco central. Sin embargo, para la presente investigación, solo se utilizan las minutas. La razón de esta elección es que las minutas condensan las discusiones de la JDBR, las razones y argumentos mediante los cuales se justifican los movimientos en la tasa de intervención. Documentos como los informes de política monetaria, informes de estabilidad fiscal y demás, pueden tener información valiosa, pero son producidos por equipos técnicos del banco y no son el resultado directo del debate entre los miembros de la JDBR.

Durante el periodo mencionado se produjeron 35 minutas en total, 16 entre enero de 2018 a diciembre de 2019 y 19 entre enero de 2020 a diciembre de 2021. Dado que la crisis sanitaria provocada por la pandemia requería tomar decisiones de forma constante, la JDBR se reunió tres veces más de lo habitual durante el 2020 (mayo, agosto y noviembre). La selección del periodo de análisis no es arbitraria, el propósito de tomar las minutas de 2018 a 2021 tiene dos justificaciones. Primero, captura el efecto de la pandemia en la estructura sintáctica de las discusiones de la JDBR y, segundo, permite tener una cantidad de texto considerable para que los algoritmos puedan trabajar.

Las minutas son uno de los varios instrumentos de comunicación que tiene el Banco de la República. Básicamente, son las reseñas escritas de lo que se discute y se decide en las reuniones de la JDBR. Estas minutas se publican días después de cada reunión de la Junta y están disponibles en el sitio web oficial del Banco. Generalmente, su extensión no supera las dos páginas. Como lo menciona Guío et al. (2020), a partir de 2018 las reuniones se redujeron a 8 por año (en los meses de enero, marzo, abril, junio, julio, septiembre, octubre y diciembre) y su publicación se redujo a una semana. Desde octubre del 2019, ese plazo se redujo. Las minutas se publican ahora el siguiente día hábil posterior a las reuniones de carácter normativo (ordinarias) de la JDBR (p. 20).

Para el uso de los documentos, se llevó a cabo un proceso de depuración para consolidar el corpus de texto. Primero se extrajeron todas las minutas en formato PDF⁶. Después de extraer todas las minutas de la página del Banco, se realizó un

⁶ Para esta tarea se usó el paquete pdftools.

proceso de concatenación donde se unieron las palabras que comparten la misma sintaxis y semántica para evitar perder el verdadero significado durante el proceso de tokenización. Palabras como "déficit en cuenta corriente", "economías emergentes", entre otras⁷, se reorganizaron de tal forma que quedaran como un solo elemento: "déficit _ cuenta _ corriente", "economías _ emergentes", etc. Este paso fue necesario porque no se encontró la forma de concatenarlo con algún paquete dentro del software R.

Luego de esto, se llevó a cabo el proceso de tokenización⁸. Este proceso consiste en fraccionar las frases dentro de los textos (minutas) en tokens, es decir, elementos textuales independientes (sintagmas) que contienen un significado sintáctico dentro de un texto. Estos tokens pueden definirse como n-gramas, donde n representa la cantidad de palabras que componen el token. En esta investigación se realizaron dos procesos de tokenización. En el primero, se tokenizó de la forma habitual, tal y como lo presentan Single y Robinson (2017), tomando solamente palabras o términos, es decir, n-gramas con n igual a 1. Con estos tokens se consiguió armar la bolsa de palabras con la que se estimaron los modelos. Y en el segundo, se tomaron los bigramas, es decir, n-gramas con n igual a 2. Esto con el propósito de realizar unos gráficos de redes (ver la sección de resultados) que permitieron dejar en evidencia el cambio de las interacciones entre las palabras/términos antes de 2018 y 2019 y durante la pandemia de 2020 y 2021. Esto, junto con las evidencias del modelamiento de tópicos, consolida los resultados de esta investigación.

La bolsa de palabras también pasó por un proceso antes de estar totalmente lista. Primero se eliminó la puntuación, números, símbolos y demás elementos que no pertenecen al espectro de sintagma. Seguido de ello, se filtraron todas las stopwords, es decir, las palabras que no aportan al sentido del texto: conectores, preposiciones, artículos, entre otros. Aunque en la mayoría de los casos estos filtros vienen estandarizados en ciertos paquetes estadísticos, se hace necesario personalizarlos ya que diferentes temas usan diferentes palabras y es altamente probable que no estén incorporadas en el filtro estandarizado⁹. Por último, se eliminaron los tokens que aparecían en la bolsa de palabras con una frecuencia menor a 3 con el propósito de optimizar los resultados, dejando de lado información poco relevante dentro de las minutas.

⁷ Condiciones_ financieras_ internacionales, expectativa_ inflación, tasa_ intervención entre otros fueron los términos concatenados.

⁸ Para esta tarea se usaron los paquetes dplyr, tidyr, widyr, tidytext.

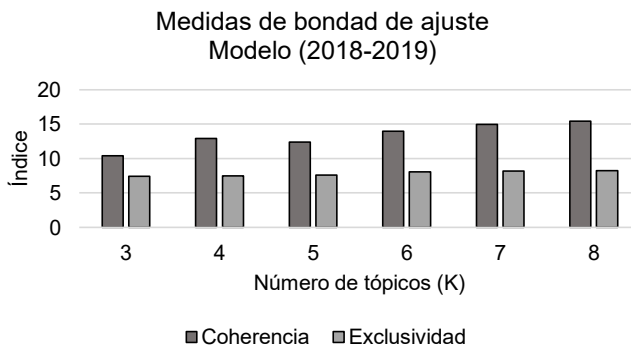
⁹ Palabras como: favorable, materia, esperados, términos entre otras fueron agregadas a las stopwords.

Luego de todo este largo proceso, la bolsa de palabras y el corpus de texto quedan listos para implementar la estrategia empírica a cabalidad.

2.3 Calibración¹⁰

Para la calibración del modelo, se debe elegir el parámetro k , el cual define a priori la cantidad de tópicos que los modelos contendrán. Elegir este parámetro es una parte muy importante y es algo de sumo cuidado, ya que un número considerablemente grande de k dará lugar a temas muy específicos, pero uno bajo dará lugar a tópicos demasiado generales. La elección de este parámetro no es arbitraria y debe realizarse con base en medidas de bondad de ajuste. Entre estas medidas se encuentran el índice de exclusividad y el índice de coherencia. El primero hace referencia a qué tan diferentes son los temas entre sí y, por lo tanto, los temas describen cosas diferentes. El segundo es la frecuencia con que coexisten las características que describen un tema y, por lo tanto, los temas parecen ser más coherentes. Siguiendo a Single y Robinson (2017) y a Röder et al. (2015), se calcularon los índices de exclusividad y coherencia tanto para el modelo con los datos comprendidos entre 2018 y 2019 como para el modelo con los datos comprendidos entre 2020 y 2021.

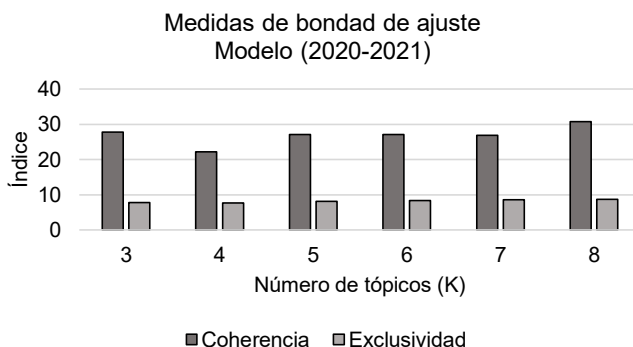
Figura 1. Índices de coherencia y exclusividad para el modelo 1.



Fuente: elaboración propia (2023).

¹⁰ Para esta tarea se usó el paquete stm.

Figura 2. Índices de coherencia y exclusividad para el modelo 2.



Fuente: elaboración propia (2023).

Los índices deben ser los más altos posibles. Sin embargo, dado que el índice de coherencia da valores negativos, se decidió analizarlo en términos de valor absoluto. En otras palabras, entre más pequeño (la barra más pequeña), mejor ajuste presenta el modelo. La exclusividad es muy similar para todos los niveles de k en las dos gráficas; por lo tanto, la coherencia será la que decida el número de tópicos "k" en los modelos.

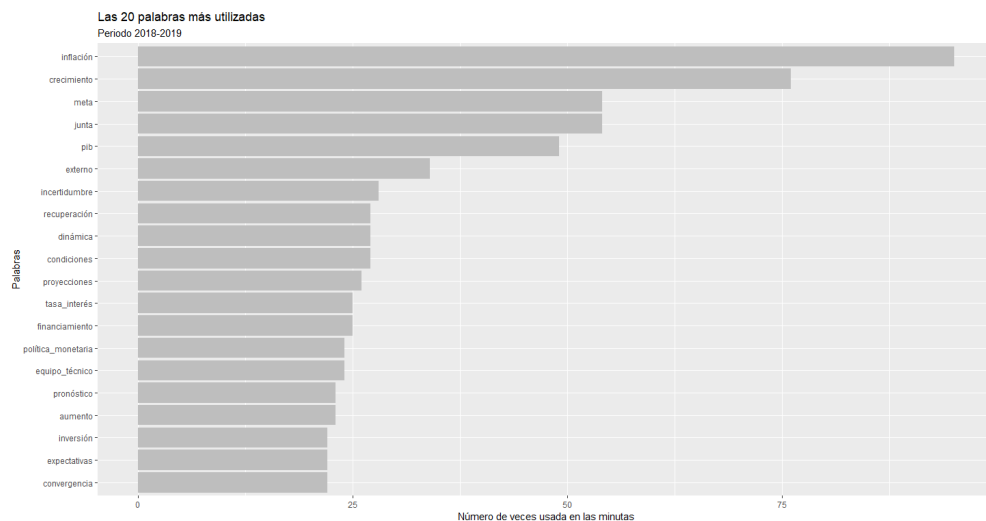
En la figura 1, se aprecia que el modelo #1 (2018-2019) tiene un buen ajuste en términos de coherencia y exclusividad con un k igual a 3 (tres tópicos), ya que la barra que representa el índice de coherencia es la menor dentro de todas las posibilidades. Bajo el mismo criterio, en la figura 2, se aprecia que el modelo #2 (2020-2021) presenta un buen ajuste con un k igual a 4 (cuatro tópicos). Cabe resaltar que el algoritmo fue entrenado con varios valores de k (3, 4, 5, 6, 7, 8).

3. RESULTADOS

3.1 Visualización de los términos más frecuentes

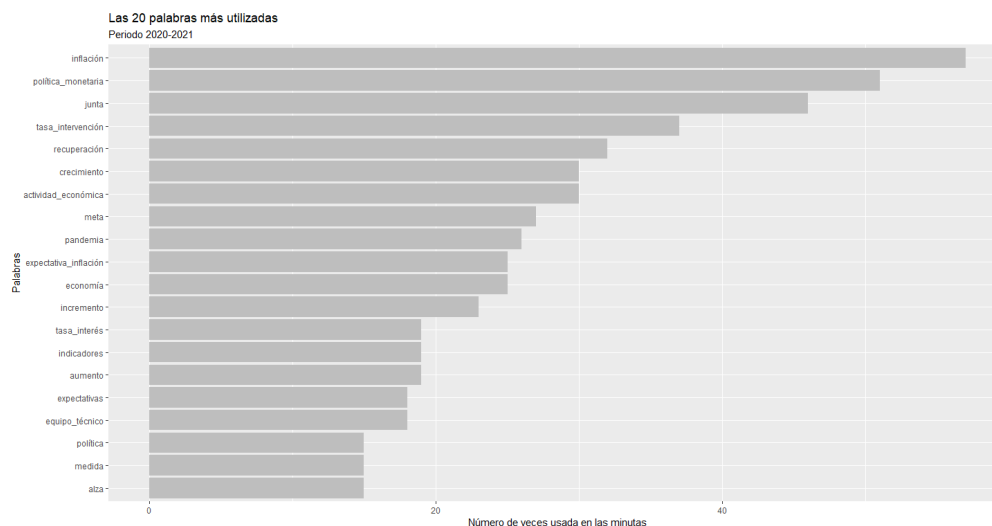
Previo a la designación del nombre de los tópicos, se realizó una visualización de las palabras más frecuentes dentro de los dos periodos de análisis: 2018-2019 y 2020-2021.

Figura 3. Frecuencia de las 20 palabras más usadas en los dos periodos de análisis: 2018-2019.



Fuente: Elaboración propia (2023).

Figura 4. Frecuencia de las 20 palabras más usadas en los dos periodos de análisis: 2020-2021.



Fuente: Elaboración propia (2023).

En la figura 3, se puede apreciar que entre los 20 términos más utilizados en el periodo 2018-2019, aparecen la inflación, crecimiento, junta, PIB, meta, entre otros.

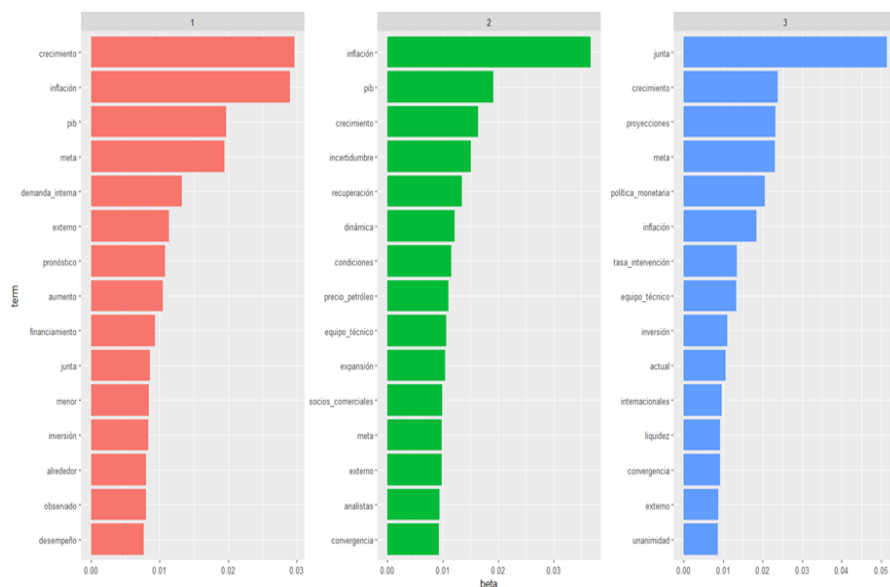
No es algo sorprendente, ya que son términos muy usuales dentro de las minutas analizadas y forman parte fundamental del léxico que utilizan los banqueros centrales.

Sin embargo, en la figura 4, correspondiente al periodo 2020-2021, aparecen nuevas palabras dentro de las 20 más utilizadas. Ahora las palabras pandemia e indicadores, así como los términos expectativa _ inflación y tasa _ intervención, ganan protagonismo dentro de las minutas. Esto refleja inmediatamente los cambios que se dieron durante la pandemia en torno a los términos utilizados dentro de las discusiones de la JDBR.

3.2 Inferencia de los tópicos

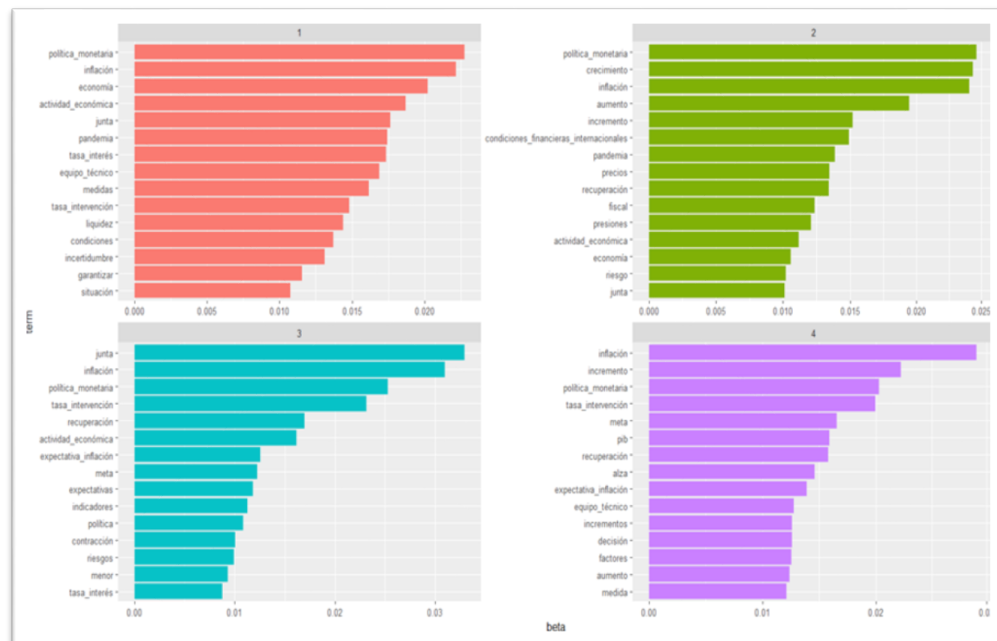
Es de suma importancia tener en cuenta que el algoritmo LDA no produce etiquetas específicas para los modelos que estima. El investigador debe realizar un proceso de inferencia basado en los parámetros α y β . Basado en el parámetro β , se construyeron las figuras 4 y 5 que ponen de manifiesto las 15 palabras más usuales dentro de cada tópico. La figura 4 representa el resultado de la estimación de un modelo con tres tópicos correspondiente al periodo (2018-2019), y la figura 5 representa el resultado de la estimación de un modelo con cuatro tópicos correspondiente al periodo 2020-2021.

Figura 5. Las 15 palabras más relevantes dentro de los tópicos del modelo #1.



Fuente: elaboración propia (2023).

Figura 6. Las 15 palabras más relevantes dentro de los tópicos del modelo #2.



Fuente: elaboración propia (2023).

En el primer modelo aparecen palabras como crecimiento, inflación, junta, proyecciones, convergencia, unanimidad, entre muchas otras. Todas son palabras muy usuales dentro de la jerga económica de los banqueros centrales y muy ligadas a su propósito constitucional, lo que indicaría que estos términos están basados en las acciones usuales que lleva a cabo el Banco. Los tópicos del modelo #2 contienen, adicionalmente, palabras estrechamente relacionadas con la coyuntura económica provocada por la pandemia. Palabras como pandemia, incertidumbre, liquidez, fiscal, presiones, riesgo, recuperación, contracción, entre otras. Estas palabras pasaron a ser importantes dentro de los temas centrales (los tópicos) que fueron tratados durante las reuniones de la JDBR. Como bien se aprecia en la figura 6, en todos los cuatro tópicos, aparecen palabras relacionadas con la crisis sanitaria. Esto deja en evidencia el giro discursivo dentro de las discusiones de la JDBR. Con base en las 15 palabras más usuales dentro de los tópicos, el parámetro Alpha¹¹ y la lectura previa de las 35 minutas, se logró etiquetar los tópicos de ambos modelos.

¹¹ En las figuras 6 y 7 se hace uso del parámetro Alpha de cada modelo.

Tabla 1. Etiquetas de los tópicos del modelo #1

Modelo #1 (2018-2019)	
Tópico #1	Contexto macroeconómico.
Tópico #2	Contexto externo.
Tópico #3	Política monetaria y acciones de tasa de interés.

Fuente: elaboración propia.

Tabla 2. Etiquetas de los tópicos del modelo #2

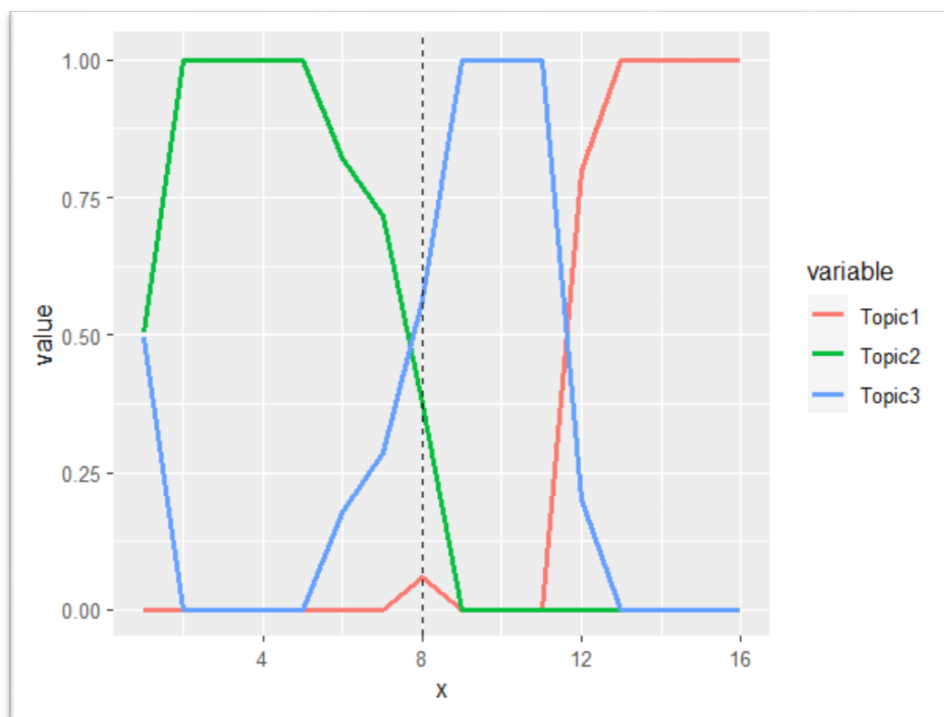
Modelo #2 (2020-2021)	
Tópico #1	Política monetaria y acciones de tasa de interés.
Tópico #2	Contexto macroeconómico externo e interno.
Tópico #3	Consideraciones de la junta directiva.
Tópico #4	Expectativas e inflación.

Fuente: elaboración propia.

El surgimiento del tópico #3 dentro del modelo #2 muestra cómo las discusiones y debates de la JDBR tomaron especial protagonismo durante la crisis. Si bien el banco tiene como principal objetivo controlar la inflación, este propósito pasó a un segundo plano en el periodo de la crisis. Durante la pandemia, aparecieron otros problemas y dificultades más relevantes que la inflación, lo que se evidencia en la aparición de términos como liquidez, actividad económica, condiciones financieras internacionales, entre otros.

El uso de los tópicos a lo largo de las minutas no fue constante. Tal y como se prevé en situaciones adversas, hay diferentes preocupaciones que se vuelven más o menos relevantes en relación con la coyuntura económica. En la figura 6 se aprecia que el tópico "contexto externo" fue, con mucho, el más relevante durante las primeras 5 minutas producidas en 2018. Sin embargo, en las primeras 4 minutas producidas durante 2019, el tópico "política monetaria y acciones de tasa de interés" tomó mayor relevancia en las discusiones de la Junta Directiva. Finalmente, las últimas 4 minutas del periodo 2019 estuvieron mucho más permeadas por el tópico "contexto macroeconómico". En este sentido, se evidencia que las discusiones de la Junta realmente se adaptan a las necesidades de la economía colombiana y no son decisiones aisladas y sesgadas por los acontecimientos o las dolencias que padece la economía.

Figura 7.¹² Presencia relativa de los tópicos del modelo#1 a lo largo de las minutas producidas durante el periodo 2018-2019.



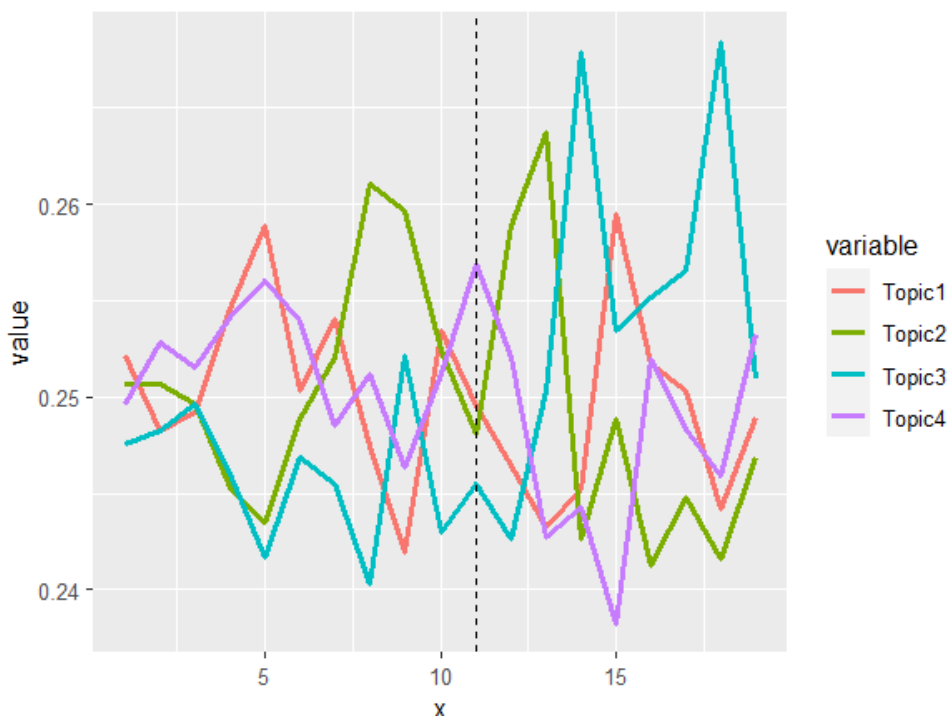
Fuente: elaboración propia (2023).

Ahora bien, con respecto a las minutas producidas durante el periodo de crisis (2020-2021), el uso de los tópicos es un poco más balanceado. Tal como se aprecia en la figura 8, las discusiones de la Junta Directiva registradas en las primeras 6 minutas dejan en evidencia una mayor presencia de los tópicos "política monetaria y acciones de tasa de interés" y "Expectativas e inflación". Durante los primeros meses de la pandemia en 2020, la política monetaria fue bastante agresiva y empezó a tomar un matiz supremamente expansivo para empezar a mitigar los efectos nocivos de la COVID-19 en el aparato económico. Explorando más a detalle la figura 7, se aprecia que los tópicos "contexto macroeconómico interno y externo" y "consideraciones de la junta directiva" empiezan a tomar mayor relevancia a partir de la duodécima minuta (producida a inicios del 2021). Aunque la diferencia entre los pesos relativos de la presencia de los tópicos en las minutas no es muy grande, es evidente que

¹² La línea negra punteada separa las minutas producidas durante 2018 (izquierda) y las producidas durante el 2019 (derecha).

durante la pandemia hubo temas más relevantes que otros y su presencia en las discusiones fue variando a medida que la crisis lo requería.

Figura 8.¹³ Presencia relativa de los tópicos del modelo#2 a lo largo de las minutas producidas durante el periodo 2020-2021.



Fuente: elaboración propia (2023).

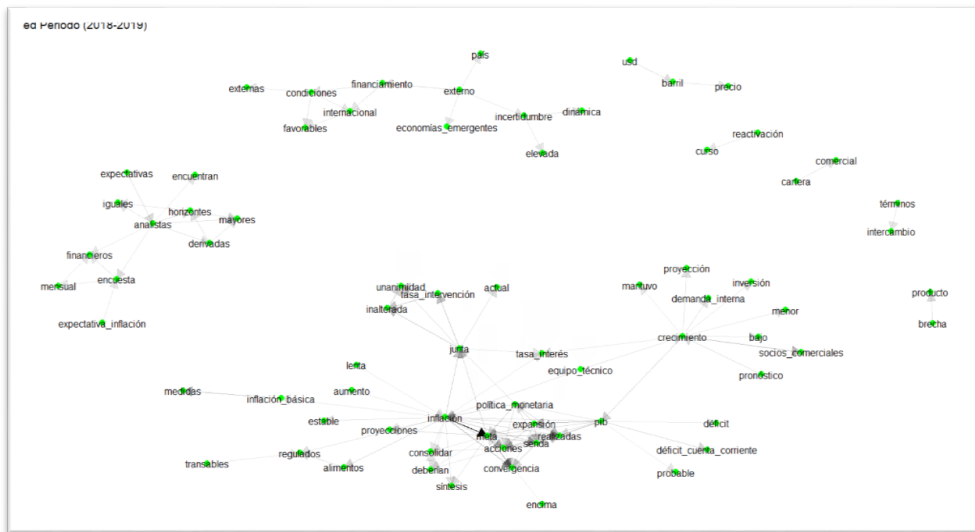
3.3 Redes¹⁴

Con el uso de la segunda tokenización, explicada en la sección anterior, se tomó la decisión de realizar un par de gráficos de redes con base en los bigramas con el propósito de evidenciar la interacción entre las palabras y términos dentro de las minutas y ver el comportamiento en cuanto al protagonismo de los nodos (en este caso, palabras o términos). En estos gráficos, la intensidad del color de la flecha refleja la frecuencia con que se conectan y relacionan los términos o palabras.

¹³ La línea negra punteada separa las minutas producidas durante 2020 (izquierda) y las producidas durante el 2021 (derecha).

¹⁴ Para construir las redes y las figuras 6 y 7, se usaron los paquetes `ggraph`, `igraph`, `ggplot2`, `scales`.

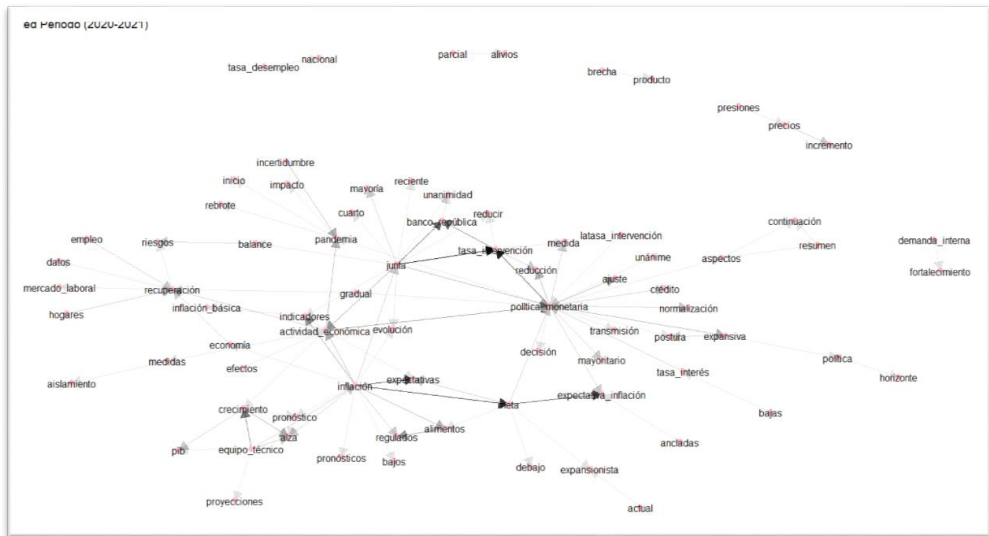
Figura 9. Red de los bigramas implícitos en las minutas elaboradas entre 2018 y 2019



Fuente: elaboración propia (2023).

En la figura 9, es evidente que los nodos "inflación" y "meta" tuvieron bastante protagonismo dentro del discurso de la JDBR. Además, estos términos también se conectan con palabras como "convergencia," "senda," "política monetaria" y "junta." A su vez, el nodo "junta" muestra una relación no despreciable con palabras como "unanimidad," "tasa de intervención," "inalterada," entre otras. Esta red muestra de forma clara que las discusiones de la JDBR estaban girando en torno a términos muy usuales en un contexto de tranquilidad económica, tasas de inflación bajas y estables, y condiciones macroeconómicas poco alarmantes. No obstante, también se puede apreciar que no solamente estos términos forman parte de las discusiones de la JDBR, ya que en la red se evidencian otros conjuntos de nodos (palabras) relacionándose entre sí. Palabras propias del contexto externo, demanda interna y otros temas aparecen en la red.

Figura 10. Red de los bigramas implícitos en las minutas elaboradas entre 2020 y 2021



Fuente: elaboración propia (2023).

En la figura 10, el nodo “inflación” pierde relevancia y el nodo “política monetaria” pasa a ser protagonista en las interacciones. Además, se puede apreciar la estrecha relación con la “tasa de intervención.” Esto muestra lo sucedido durante la pandemia en torno a los recurrentes movimientos como respuesta a la crisis. Adicional a esas particularidades, el nodo “pandemia” aparece como uno de los protagonistas dentro de la red y muestra relaciones con nodos como “incertidumbre,” “rebrote,” “actividad económica,” entre otros. Estos resultados refuerzan lo encontrado con el ejercicio de modelamiento de tópicos presentado en la primera parte de estos resultados. Cabe resaltar que dentro de la red aparecen palabras concernientes al mercado laboral y términos como la “tasa de desempleo,” “hogares,” “crédito,” “aislamiento,” entre muchos más. Estos términos también aparecen como una novedad ya que en la figura 8 no tienen presencia.

El objetivo principal de una correcta comunicación de la política monetaria radica en incidir de forma certera sobre las expectativas de los agentes y, de dicha forma, contribuir a los objetivos de la política monetaria (Sims, 2009). El hecho de encontrar un cambio en las temáticas y/o estructura discursiva de las discusiones de la JDBR muestra cómo el Banco enfatizó en políticas en pro de subsanar los efectos de la pandemia y deja en evidencia que optó por priorizar aspectos como la liquidez, la reactivación de la economía y el desempleo. Además,

empezó a incorporar en sus discusiones temas relacionados con los hogares y las empresas. Tratando con esto de comunicar de forma precisa todo lo que se estaba haciendo durante el periodo de crisis para generar confianza y credibilidad en los agentes, garantizando con esto el éxito de la política monetaria.

Estos hallazgos, junto con lo encontrado en el modelamiento de tópicos, demuestran empíricamente el efecto que tuvo la pandemia en los temas que hacen parte de las discusiones de la JDBR a la hora de tomar decisiones en torno a la política monetaria.

4. CONCLUSIONES

A partir de la aplicación del método LDA, herramientas de análisis de texto y gráficos de redes, se pudo evidenciar el efecto de la pandemia en la estructura discursiva implícita en las minutas producidas por el Banco de la República de Colombia en el período 2018-2021. Las palabras usadas dentro de los tópicos principales de las minutas cambiaron. La tasa de intervención y su relación con términos relacionados con la pandemia muestran que el Banco de la República incorporó y se interesó en temas como la actividad económica, la situación de las empresas y los hogares durante la pandemia, el desempleo, entre otros, a la hora de tomar las decisiones de política monetaria durante 2020 y 2021.

Dentro de los cuatro tópicos del modelo #2, se puede apreciar que los términos pertenecientes al tópico Política monetaria y acciones de tasa de interés están estrechamente relacionados con la pandemia. La política monetaria fue discutida incorporando perspectivas diferentes a las habituales. Se dio un cambio discursivo evidente dentro de las discusiones y temas tratados durante la pandemia. La rigidez no es una opción cuando la situación requiere acciones menos convencionales. La política monetaria puede también preocuparse por otros aspectos de la economía e incorporar en sus discusiones otros temas igual de relevantes a la inflación. Esta investigación deja ver cómo los bancos centrales incorporan información según el contexto para transmitir a los agentes confianza, credibilidad y tranquilidad en momentos de crisis.

Este artículo busca promover futuras investigaciones sobre la comunicación de los bancos centrales, que incorporen estos algoritmos al análisis de textos en economía y, sobre todo, a documentos de comunicación de la política monetaria colombiana y del mundo. Otro tipo de choques pueden ser analizados con estos métodos. Para futuras investigaciones se sugiere aplicar estos algoritmos extendiendo el período de análisis, procurando capturar en dichos periodos otro tipo

de fenómenos trascendentales para la economía colombiana. Adicionalmente, se recomienda explorar los efectos que estos cambios discursivos pueden tener en las variables macroeconómicas o en la formación de expectativas de los agentes.

Entender la política monetaria no solamente puede quedar en analizar cifras, predicciones y modelos, sino que debe trascender a entender el cómo, el por qué y el para qué se llevan a cabo diversas políticas.

REFERENCIAS

- Anzoátegui Zapata, J. C., & Galvis Ciro, J. C. (2019). Efectos de la comunicación del banco central sobre los títulos públicos: evidencia empírica para Colombia. *Cuadernos de Economía*, 38(77), 337-364. <https://doi.org/10.15446/cuad.econ.v38n77.64706>
- Anzoátegui Zapata, J. C., & Galvis Ciro, J. C. (2022). Efectos de la comunicación del banco central sobre el desacuerdo en las expectativas de la tasa de política monetaria: evidencias para Colombia. *Revista Finanzas Y Política Económica*, 14(2), 375-409. <https://doi.org/10.14718/revfinanzpolitecon.v14.n2.2022.4>
- Arango, L. E., Pantoja, J., & Velásquez, C. (2017). Effects of central bank's communications in Colombia. A content analysis. *Borradores de Economía*.
- Bascand, G. (2013). Communication, understanding, and credibility. Discurso en el Admirals Breakfast Club, Nueva Zelanda.
- Bholat, D., Hansen, S., Santos, P., & Schonhardt-Bailey, C. (2015). Text mining for central banks: handbook. *Centre for Central Banking Studies* (33), 1-19.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Boukus, E., & Rosenberg, J. (2006). The Information Content of FOMC Minutes. Available at SSRN: <https://dx.doi.org/10.2139/ssrn.922312>
- Camora, M. (2015). Impacto de la comunicación del Banco Central de Bolivia en la eficiencia de la política monetaria. Serie de Documentos de Trabajo del Banco Central de Bolivia, (4), 1-21. <https://www.bcb.gob.bo/webdocs/publicacionesbcb/2017/06/27/23%20-%20Impacto%20comunicaci%C3%B3n%20BCB.pdf>
- Dincer, N. N., & Eichengreen, B. (2009). Central Bank Transparency: Causes, Consequences and Updates. *National Bureau of Economic Research, Working Paper Series*. Estados Unidos.
- Eusepi, S. and Preston, B., 2007, Central Bank Communication and Expectations Stabilization. *NBER Working Paper No. 13259*, Issued in July 2007.
- Fontan, C., Dietsch, P., Claveau, F., & Dion, J. (2021, October 21). "Central Banking and Inequalities: Old Tropes and New Practices". <https://doi.org/10.31235/osf.io/wvtjx>
- Géron, A. (2017). *Hands-On Machine Learning with Scikit-Learn & TensorFlow*. O'Reilly.
- Guío, D., Ospina, J., Muñoz, J., & Parra, J. (2020). Descripción de las Minutas e Informes de Política monetaria a partir de herramientas de lingüística computacional. *Borradores de Economía*, 1108.

- Hendry, S., & Madeley, A. (2010). Text mining and the information content of Bank of Canada communications (No. 2010, 31). *Bank of Canada Working Paper*. <https://doi.org/10.34989/swp-2010-31>
- Hoffman, M. D., Blei, D., & Bach, F. (2010). Online learning for Latent Dirichlet Allocation. NIPS'10 Proceedings of the 23rd International Conference on Neural Information Processing Systems, 856-864.
- Meneses, J., & Ávila, C. (2019). ¿Permiten los comunicados del Banco de la República anticipar el futuro de las tasas de interés? Dirección de investigaciones económicas. Grupo Bancolombia. <https://www.bancolombia.com/empresas/capital-inteligente/actualidad-economica-sectorial/permiten-comunicados-banco-republica-anticipar-tasas-interes>
- Neuenkirch, M., 2012, Managing financial market expectations: The role of central bank transparency and central bank communication. *European Journal of Political Economy*, 28,1-13.
- Peteiro, C., Salcines, J., & Orosa, J. (2002). La estrategia de la política monetaria: efectividad y comunicación en el caso de BCE. *Momento económico*, (2), 2-32. <https://biblioteca.org.ar/libros/92172.pdf>
- Poole, W., 2001, Expectations. *Federal Reserve Bank of St. Louis Review*, 83, 1, 1-10.
- Röder, M., Both, A., & Hinneburg, A. (2015). Exploring the Space of Topic Coherence Measures. Proceedings of the Eighth ACM International Conference on Web Search and Data Mining, 399-408.
- Selaive, J., & González, H. (2015). Comunicación de la Política Monetaria del BCCCh: Evaluación y propuestas a la luz de la evidencia internacional. *Boletín BBVA Research*, 1-16. https://www.bbvarsearch.com/wp-content/uploads/2015/03/Chile_Comunicacion-de-PM_marzo_2015_vf.pdf
- Sims, C. (2009). Inflation expectations, uncertainty and monetary policy [BIS Working Papers No 275].
- Single, J., & Robinson, D. (2017). *Text Mining with R: A Tidy Approach* (1.ª ed.). O'Reilly Media, Inc. <https://www.tidytextmining.com/index.html>
- Svensson, L. E. (1998). Inflation Targeting as a Monetary Policy Rule.
- Taborda, R. (2015). Procedural transparency in Latin American central banks under inflation targeting schemes. A text analysis of the minutes of the Boards of Directors. *Ensayos sobre Política Económica*, 33(76), 76-92. <https://doi.org/10.1016/j.espe.2015.01.002>
- Valencia, G. (2021). Junta Directiva del Banco de la República: grandes episodios en 30 años de historia. Libros Banco de la República.
- Van der Cruysen, C., Eijffinger, S., and Hoogduin, L., 2010, Optimal central bank transparency. *Journal of International Money and Finance*, 29(8), 1482-1507.
- Vega, M. (2018). La comunicación de la política monetaria: una primera evaluación. *Revista Moneda*, (175), 4-10. <https://www.bcrp.gob.pe/docs/Publicaciones/Revista-Moneda/moneda-175/moneda-175.pdf>
- Walsh, C. E. (2006). Optimal Economic Transparency. In Workshop, May (Vol. 8, p. 10).