# METHODOLOGY OF THE ECONOMIC POLICY UNCERTAINTY INDEX IN COLOMBIA (IPEC)

## I. OVERVIEW

The Foundation for Higher Education and Development (Fedesarrollo) is a private, non-profit organization established in 1970. It is dedicated to researching economic and social policy issues, with the mission of contributing to the design, monitoring, and improvement of public policies. Through its studies, publications, and policy debates, Fedesarrollo promotes Colombia's economic and social development.

Recognized as Colombia's leading economic think tank, Fedesarrollo has pioneered the development of key opinion indicators that have become essential references for corporate and public policy decision-making. Its reputation for independence and analytical rigor has been solidified over decades of impactful work.

As part of its commitment to providing technical, rigorous, and timely analyses of Colombia's economic conditions, Fedesarrollo introduces the methodology for the **Index of Economic Policy Uncertainty in Colombia (IPEC)**. This index aims to measure monthly economic policy uncertainty in Colombia by analyzing word frequencies in media coverage related to economic conditions. The IPEC also enables sector-specific uncertainty analysis, offering valuable insights across various economic domains[1].

## II. INTRODUCTION

Uncertainty[2] is a critical variable influencing the behavior of economic agents. It can stem from factors such as regulatory instability, unpredictable policy decisions, and income volatility. Regardless of its origin, increased economic uncertainty typically delays investment and consumption decisions by firms and households. This is due to perceptions of lower returns on investments (Eberly, 1994) and higher costs of accessing credit (Pastor and Veronesi, 2012). As a result, reduced investment and consumption can decelerate economic growth (Balcilar, Gupta, and Segnon, 2016) and amplify business cycles (Bloom, 2014).

Quantifying economic uncertainty, however, has posed a persistent challenge due to its inherently unobservable nature. Traditional proxies such as the Chicago Board Options Exchange Market Volatility Index (VIX), Credit Default Swaps (CDS), and the

[1] This project was supervised by Luis Fernando Mejía (Executive Director of Fedesarrollo) and Carlos Manuel Parra (Professor at Florida International University), with the support of Sara Ramírez (Director of Macroeconomic and Sectoral Analysis at Fedesarrollo), Luis Felipe González and José Julián Parra (analysts of the Direction of Macroeconomic and Sectoral Analysis at Fedesarrollo). In its initial stage, the project had the contributions of Martha Elena Delgado, Diego Gutiérrez, César Pabón and Carolina Celis.

[2] According to Bloom (2024), uncertainty can be understood as the perception of a high probability of negative outcomes occurring (risk), or as the lack of knowledge about the scenarios that may unfold in the future and their associated probabilities (ambiguity).

Emerging Market Bond Index (EMBIG) have been widely used. Some authors have also employed the variance of interest rates, inflation, or fiscal indicators as indirect measures. In this context, Baker, Bloom, and Davis (2013) developed the **Economic Policy Uncertainty Index (EPU)**, which uses keyword analysis in media content to quantify uncertainty.

Fedesarrollo's IPEC builds upon the methodology proposed by Baker et al. (2016). By counting keywords related to economic policy uncertainty in Colombian press articles, the IPEC enables both aggregate and sector-specific analysis of uncertainty. This methodology complements traditional risk measures such as CDS, EMBIG, or exchange rate fluctuations, providing a narrative rooted in the frequency of sector-specific news.

To detail the construction of the IPEC and guide users in interpreting its results, this document is organized as follows:

1. **Background:** A review of the international and local development of uncertainty indices.
2. **Methodology:** Technical details for constructing both the aggregate index and sectoral estimates.
3. **Interpretation of IPEC results:** Explanation and guidance on interpreting the aggregate and sectoral results.
4. **Relationship with other indicators:** Analysis of the IPEC's correlation with traditional country risk indicators.
5. **Additional information:** Details on data periodicity and availability.

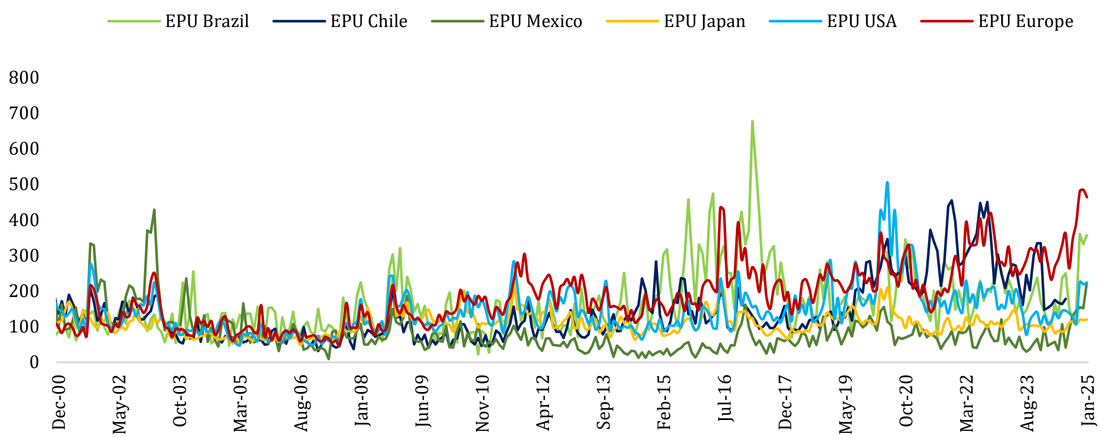
## III. BACKGROUND: A REVIEW OF THE INTERNATIONAL AND LOCAL DEVELOPMENT OF UNCERTAINTY INDICES

The first attempt to quantify economic policy uncertainty using keyword analysis dates back to Baker et al. (2016), who developed the **Economic Policy Uncertainty (EPU) Index** for the United States. This monthly index, covering the period from January 1985 to November 2011, was constructed by analyzing the frequency of terms related to economic policy uncertainty in press articles.

To explore the relationship between uncertainty and business cycles, Baker et al. employed a vector autoregressive (VAR) model. Their findings revealed that increased policy uncertainty during 2006–2009 was followed by a persistent decline in real industrial production and a sustained drop in aggregate employment. These results underscore the negative effects of economic policy uncertainty on economic activity.

This methodological breakthrough provided a quantitative approach to measuring uncertainty, and the authors extended their research in collaboration with Stanford University to include indices for countries such as the United States, France, Italy, the United Kingdom, and Japan (Figure 1). In Latin America, EPU indices have been developed for Brazil, Chile, and Mexico, offering regional perspectives on economic uncertainty[3].

**Figure 1. International comparison of the Economic Policy Uncertainty (EPU) Index**




*The index for Europe includes the countries of the United Kingdom, France, Italy, Spain, and Germany.
Source: Economic Policy Uncertainty.

In Colombia, two notable efforts have adapted the methodology proposed by Baker et al. (2016):

- **Perico (2018)** constructed an EPU index using data from the digital archive of the newspaper *El Tiempo* for the period 1994–2016. The study found high correlations between the index and periods of violence or economic crises.

- **Gil and Silva (2019)** developed an EPU index for 2000–2017, highlighting increased uncertainty during episodes of economic volatility.

Building on these foundations, Fedesarrollo introduces a monthly updated **Index of Economic Policy Uncertainty in Colombia (IPEC)**. The IPEC is updated within ten days after the end of each month, ensuring timely and relevant information for decision-making. Additionally, the IPEC enables sectoral analysis, providing insights into the impact of uncertainty on specific industries.

## IV. METHODOLOGY FOR THE CONSTRUCTION FEDESARROLLO'S IPEC

[3] The EPU indices for these countries are available on https://www.policyuncertainty.com/.

This section outlines the methodology used to construct the **Index of Economic Policy Uncertainty in Colombia (IPEC)**, focusing on its data sources, filtering process, standardization, and sectoral classification.

## Data collection and processing

The IPEC is based on the frequency of economic policy uncertainty keywords in news articles from *El Tiempo*, a Colombian newspaper with nationwide and international coverage and a digital archive dating back to 1990. To extract the relevant data, a web scraping algorithm gathers articles in text format, removes accent marks and special characters to minimize errors in the subsequent search for relevant words, and prepares the data for keyword filtering.

## Keyword filtering and article selection

A news article is included in the IPEC calculation if it satisfies three criteria, based on keywords from predefined categories (Table 1):

- Contains at least one term from the **Uncertainty** category.
- Contains at least one term from the **Economy** category.
- Contains at least one term from the **Policy** category.

The selected articles are expressed as a proportion of the total number of articles published during the reference month. To validate the filtering process, a manual review ensures the algorithm's accuracy and confirms that the articles focus on economic policy topics.

**Table 1. Keywords for news filtering**[4]

| Category | | Keywords in spanish | Keywords in english |
|---|---|---|---|
| I: News indicating **Uncertainty** | | Incertidumbre, Incierto, Incierta | Uncertainty, Uncertain |
| E: News about the **Economy** | Economy | Economía, Económico, Económica | Economy, Economic |
| P: News about **Policy** | Fiscal policy | Gobierno, Política fiscal, Presupuesto, Déficit fiscal, Deuda pública, Impuesto, Tributaria, Tributario, Ministerio de Hacienda, Gasto público | Government, Fiscal policy, Budget, Fiscal deficit, Public debt, Tax, Tax authorities, Ministry of Finance, Public spending |
| | Monetary | Política monetaria, Banco de la República, Emisor | Monetary policy, Central Bank of |

[4] These keywords are similar to those used in Perico (2018), where policy-related terms include fiscal policy, budget, monetary policy, tariffs, and broader public policy concepts. Likewise, the policy keywords identified in Gil and Silva (2018) encompass fiscal policy, deficit, the Central Bank of Colombia, among others. The keywords in the Economy and Uncertainty categories are mostly consistent across both papers and align closely those presented in Table 1.

| | | | |
|---|---|---|---|
| | policy | | Colombia, Issuer |
| | Trade policy | Arancel, Arancelaria, Aranceles, Política comercial | Tariff/Tariffs, Trade policy |

Source: Fedesarrollo.

## Standardization and normalization

To account for variations in article volume over time, the index is standardized and normalized[5].

- **Standardization**: The series is divided by its standard deviation for the period 2000–2019 to achieve unit standard deviation. This period excludes the disruptions caused by the COVID-19 pandemic[6].
- **Normalization**: The series is scaled to have a mean of 100 by multiplying by the factor $\frac{100}{\mu}$, where $\mu$ is the mean of the series for 2000-2019.

The resulting value represents the IPEC for each month.

## Sectorial classification

Each news in the IPEC is then categorized by the main economic sector it relates to. To achieve this, a method called SVC (Support Vector Classification) is used[7]. This method works like a smart sorting machine that helps manage large amounts of data and identify complex patterns. It is particularly effective for classifying news articles because it can analyze the content and group it into categories such as agriculture, transportation, or finance, among others.

The SVC model operates by converting news articles into numerical representations (called "feature vectors"), which essentially summarize the key characteristics of the text. These feature vectors are created using natural language processing (NLP) techniques, such as term frequency-inverse document frequency (TF-IDF) or word embeddings. These methods help the model understand not just the words but also the meaning and context of the text. For example, TF-IDF identifies the most important words in a document by considering how often they appear relative to other documents, while word embeddings capture relationships between words (e.g., identifying that "irrigation systems" and "crop yields" are closely related within the

[5] The standardization process takes into account a sample mean of 1.8 and a standard deviation of 0.0016.
[6] The authors who estimate uncertainty indices for different countries use different standardization periods. Although these windows vary in the estimation of the index for each country, they coincide in isolating periods of high volatility. For example, in Baker et al. (2016), it is observed that after 2009-2010 there are notable increases in the uncertainty index for the countries analyzed, so the standardization isolates the periods of highest volatility and disruptions in the series.
[7] Other classification algorithms tested for this task included Random Forest, Multinomial Naive Bayes, and Complement Naive Bayes. Among these, the SVC algorithm demonstrated the best performance.

agriculture sector).

To train the SVC model, a labeled dataset is used, which includes examples of news articles that have already been assigned to their respective sectors. This training process allows the model to learn patterns and associations, so it can accurately classify new articles in the future.

However, it is recognized that the language of news can be complicated and sometimes ambiguous, so the SVC model may make mistakes. These challenges arise due to factors like the use of metaphors, vague terminology, or sectoral overlaps in news content, which may not be straightforward for the algorithm to interpret. To address this, a manual review is conducted to ensure that the classifications are correct and to fix any errors made by the algorithm.

Finally, in this classification approach, each article is assigned to a single predominant economic sector. This decision ensures analytical clarity and avoids dilution of the impact of sector-specific trends. Allowing multiple sectoral classification could introduce ambiguity and complicate subsequent analyses, such as tracking sectoral sentiment.

**Table 2. IPEC sectoral classification**

| Classification | Categories |
|---|---|
| Sectoral | Economic activity, Agriculture, Construction, Communications, Education, Electricity, Gas and Water (EGA), Security, Poverty, Economic, Social, and Geopolitical Policy, Health, Transportation, and Financial variables. |

## V. INTERPRETATION OF IPEC RESULTS

### Aggregate IPEC results

The IPEC is a monthly index with a lower bound of 0 and a mean of 100. An IPEC result for month $m$ of year $t$ ($IPEC_{m,t}$) of 0 indicates that no newspaper articles meeting the keyword filter were found during the analyzed period.

Key interpretations of the IPEC include:

- If $IPEC_{m,t} = 100$: The proportion of articles that meet the keyword filter[8] in month $m$ of year $t$ equals the 2000–2019 average, reflecting average economic policy uncertainty.
- If $IPEC_{m,t} > 100$ ($< 100$), the proportion of articles that meet the keyword

[8] This proportion is calculated with respect to the total number of news articles analyzed during the period.

filter in month $m$ of year $t$, is above (below) the average between 2000 and 2019.

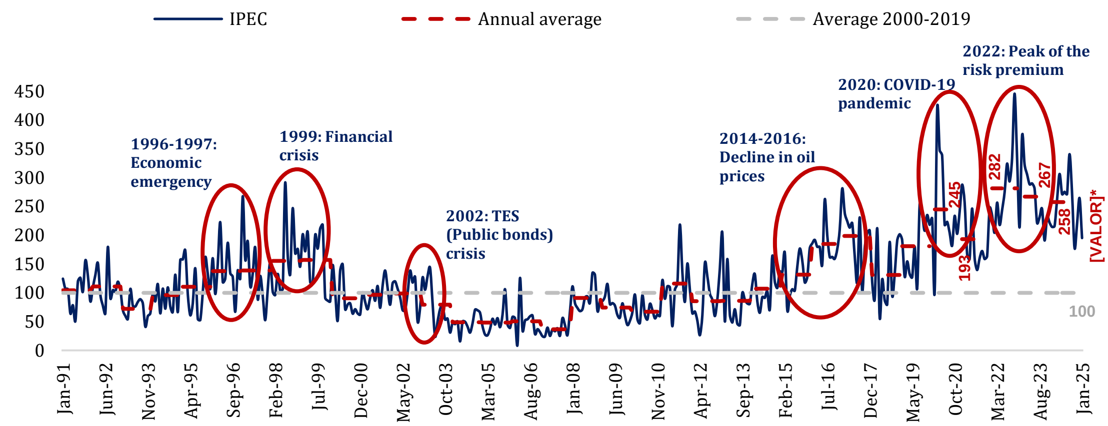
Trends over time are evaluated by comparing consecutive monthly values $(IPEC_{m+1,t})$- $(IPEC_{m,t})$. Positive differences indicate increasing uncertainty, while negative differences indicate decreasing uncertainty.

The results of the IPEC are presented in Figure 2. Between 2000 and 2019, the IPEC averaged 100. However, since the second half of 2021, the index has shown an upward trend, with a temporary correction in the first half of 2023, after which it resumed its upward trend. Since 2022, the IPEC yearly average has remained above 200, and in 2024, the average IPEC stood at 258. This implies that news related to economic policy uncertainty were 2,58 times more frequent in 2024 than the 2000-2019 average.

One of the main strengths of the IPEC is its ability to effectively capture periods of heightened uncertainty triggered by significant economic events, as illustrated in Figure 2. The index reflects sharp increases during critical moments, such as the economic emergency of 1996-1997, the financial crisis of 1999, and the TES crisis in 2002. Similarly, it highlights episodes like the 2014-2016 decline in oil prices, the uncertainty caused by the COVID-19 pandemic in 2020, and the peak of the risk premium in 2022. These spikes demonstrate the IPEC's robustness in tracking shifts in economic policy uncertainty, making it a valuable tool for analyzing the evolution and intensity of uncertainty over time. Notably, several of the periods of uncertainty captured by the IPEC have also been identified by Perico (2018) and Gil and Silva (2019). These studies highlighted episodes such as the economic emergency, TES crisis, and the decline in oil prices[9], which align with the periods identified by Fedesarrollo's IPEC.

**Figure 2. Historical evolution of the IPEC**

[9] The economic emergency period was identified by Perico (2018), the TES crisis was recognized by both authors, and the decline in oil prices period was highlighted by Gil and Silva (2019).




Source: Fedesarrollo. (*) Year-to-date average.

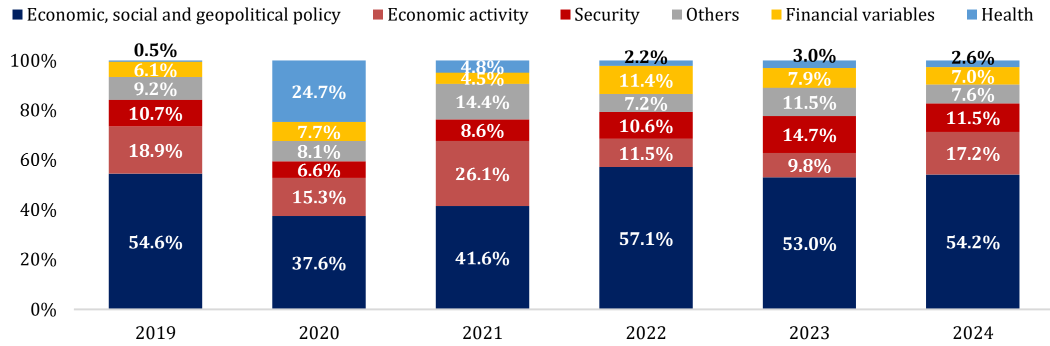
## Sectoral IPEC results

Sectoral classifications offer insights into the distribution of uncertainty across economic activities. Key interpretations include:

- A value of 0 for a sector indicates no articles related to that sector.
- A value of 100 for a sector indicates all relevant articles in that month pertain to that sector.
- The categories of economic sectors are mutually exclusive, as both the SVC algorithm and the subsequent manual review identify the predominant sector discussed in the news, classifying it into a single category.

As presented in Table 2, the categories used in the sectoral classification correspond to **economic activity, economic policy, social and geopolitical issues, security, financial variables, and other sectors,** which are grouped into the "other" category. **Financial variables** include news about financial and currency markets, credit ratings, monetary policy interest rates, and public debt, among others. **Economic activity** covers news on economic, sectoral, and business analysis, trade policy, and foreign trade, among others. **Economic, social, and geopolitical policy** includes news about economic and social reforms, electoral cycles, and recent geopolitical events, among others. **Security** includes news on armed conflict, peace negotiations, and international conflicts, among others. **Health** contains news about the healthcare system, reforms, and the persistent effects of the pandemic, among others. Finally, the **"other"** category encompasses the sectors of transportation, poverty, communications, Electricity, Gas and Water (EGA), agriculture, construction, education, and health.

**Figure 3. IPEC sectoral classification**




Source: Fedesarrollo.

Note: *Financial variables* include news about financial and currency markets, risk rating, monetary policy interest rate and public debt, among others. *Economic activity* includes news about the economic, sectoral and business situation, trade policy and foreign trade, among others. *Economic, social and geopolitical policy* includes news about economic and social reforms, electoral cycles and recent events in geopolitics, among others. *Security* includes news about armed conflict, peace negotiations and international conflicts, among others. *Health* contains news about the health system, reforms and persistent effects of the pandemic, among others. *Others* includes the sectors of transportation, poverty, communications, EGA, agriculture, construction, education, and health.

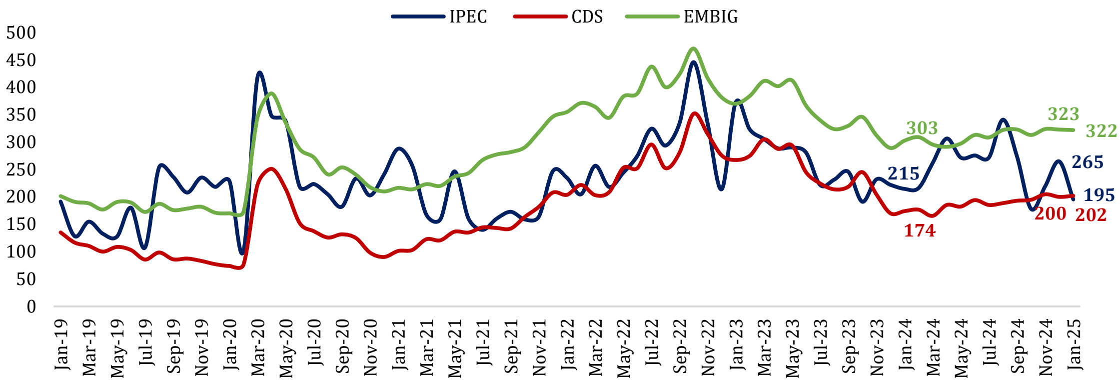
Figure 3 shows the percentage distribution of the sectoral breakdown of IPEC news between 2019 and 2024. During this period, on average, 49.7% of the news refers to issues of economic, social, and geopolitical policy, while a smaller percentage corresponds to economic activity (16.5%), security (10.4%), others (9.7%), financial variables (7.4%) and health (6.3%). In 2020 and 2021, the importance of the health sector stands out, with 24.7% and 4.8% of the news, respectively, primarily focused on this sector.

## VI. CORRELATION OF THE IPEC WITH OTHER COUNTRY RISK INDICATORS

The IPEC reflects changes in a country's economic policy uncertainty, which can also be captured by other indicators such as *Credit Default Swaps* (CDS) and the *Emerging Markets Bonds Index Global* (EMBIG). On the one hand, the CDS indicator measures the cost of insurance against the default of a country's public debt securities. On the other hand, the EMBIG captures the spread or additional cost that a country must pay the market for issuing public debt securities, compared to the cost of securities considered risk-free. An increase in both indicators is associated with higher country risk. Given the link between economic policy uncertainty and the perception of country risk, a similar trend can be observed between the IPEC and these variables, reflecting the indicator's accuracy in capturing economic policy uncertainty in the Colombia[10].

**Figure 4. Evolution of the IPEC, EMBIG, and CDS**

[10] For the period 2019-2024, the Pearson correlation coefficient is approximately 0.66 between the IPEC and CDS, and 0.65 between the IPEC and EMBIG, and it is statistically significant at the 1% confidence level in both cases.




Source: Fedesarrollo, Bloomberg.

## VII. IPEC PUBLICATION

The monthly result of the IPEC will be published in the Fedesarrollo repository during the month following the one to which the result corresponds. This publication will be made through a newsletter that will include the following sections:

- Evolution of the overall IPEC result.
- Evolution of the IPEC by sector.

Users can freely access the newsletters and historical results at the following link: [www.repository.fedesarrollo.org.co](www.repository.fedesarrollo.org.co). Access to the raw dataset and algorithms requires a formal request and is subject to approval by Fedesarrollo.

## VIII. REFERENCES


- Balcilar, Mehmet, Gupta, Rangan, & Segnon, Mawuli (2016). *The role of economic policy uncertainty in predicting US recessions: A mixed-frequency Markov-switching vector autoregressive approach*. Economics, 10(1), 20160027.
- Baker, Scott, Bloom, Nick & Davis, Steven (2016). *Measuring Economic Policy Uncertainty.* The Quarterly Journal of Economics, President and Fellows of Harvard College, 131(4), 1593-1636.
- Bloom, Nick (2014). *Fluctuations in uncertainty*. Journal of Economic Perspectives, 28(2), 153-176.
- Eberly, Janice (1994). *Adjustment of Consumers' Durables Stocks: Evidence from Automobile Purchases*. Journal of Political Economy, 102(3), 403-36.
- Gil, J. & Silva, D. (2019). *Índice de incertidumbre de política económica (EPU) para Colombia, 2000-2017*. Ensayos de Economía, Universidad Nacional de Colombia Sede Medellín
- Pastor, Lubos, & Veronesi, Pietro (2011). *Political Uncertainty and Risk Premia*. NBER Working Paper 17464.
- Perico, Daniel (2018). *Measuring Economic Policy Uncertainty in Colombia: a News Based Approach*. Available at SSRN 4193944.