

# Capturing international influences in U.S. monetary policy through a NLP approach

LAURENT FERRARA

Skema Business School - University Côte d'Azur  
[laurent.ferrara@skema.edu](mailto:laurent.ferrara@skema.edu)

NICOLAS DE ROUX \*

QuantCube Technology, EconomiX  
[n.deroux@quant-cube.com](mailto:n.deroux@quant-cube.com)

May 22, 2023

## Abstract

*The U.S. Federal Reserve has a statutory dual domestic mandate of price stability and full employment. In this paper, we question the role of the international environment in shaping Fed monetary policy decisions. In this respect, we apply natural language processing (NLP) techniques to minutes of the Federal Open Market Committee (FOMC) and construct indexes of the attention paid by U.S. monetary policymakers to the international economic and financial situation, as well as sentiment indexes. By integrating those text-based indicators into a Taylor rule, we derive various quantitative measures of the external influences on Fed decisions. Our results show that when there is a focus on international topics within the FOMC, the Fed's monetary policy generally tends to be more accommodative than expected by a standard Taylor rule. This result is robust to various alternatives that includes a time-varying neutral interest rate or a shadow central bank interest rate.*

**Keywords :** Monetary policy, Federal Reserve, FOMC minutes, International environment, Natural Language Processing, Machine Learning

---

\*We would like to thank Luca Barbaglia, Benoit Bellone, Oussama Belmejdoub, Menzie Chinn, Thomas Chuffart, Thomas Drechsel, Leif Anders Thorsrud and Francis Wolinski for useful comments, as well as seminar participants at EconomiX - Université Paris Nanterre and researchers at QuantCube Technology for useful discussions.

---

## 1. Introduction

To manage monetary policy in the United States, the Federal Reserve (the Fed hereafter) has a dual statutory mandate to fulfill of promoting maximum employment and stable prices, as well as moderate long-term interest rates. Clearly, those objectives are domestic ones and do not account, at least directly, for the international environment. If we add to this the fact that the U.S. economy is the archetype of a large, closed economy, the common wisdom is that the Fed is considered to be largely indifferent to what's going on in the rest of the world. In an influential paper, Barry Eichengreen was asking in 2013: *Does the Federal Reserve care about the Rest of the World?* (Eichengreen, 2013).

However, a first sight at the meeting minutes of the Federal Market Open Committee (FOMC, the committee in charge of the monetary policy within the Fed), shows that the international situation is often discussed among members during meetings, sometimes at lengths. Indeed, international factors play a role in the stability of the U.S. economy and financial system, through various mechanisms. The main factors are the global integration of financial markets which influences domestic market conditions, the role of international prices on domestic inflation (*imported inflation*) and the potential spillbacks on the U.S. economy of U.S. monetary policy's influence on foreign economies (Obstfeld, 2020). Adverse economic conditions that could affect the U.S. economy can therefore lead the FOMC to raise or to cut the Federal Funds Rates (FFR), the main conventional tool of monetary policy (Ihrig and Wolla, 2020).

The influence of the international environment on U.S. monetary policy is sometimes explicitly evoked by policymakers themselves. In the midst of the global financial turbulence in 1998, marked by the Russian Federation's default on its debt, Fed Chairman Alan Greenspan remarked in a September 4th speech in Berkeley that *"it is just not credible that the United States can remain an oasis of prosperity unaffected by a world that is experiencing greatly increased stress"* (Greenspan, 1998). This take was widely seen by market participants as an early warning of the FFR cut decided later that month. Another example of this supposed attention to the international environment took place in August 2015, at a time when the Fed was expected to start hiking rates from the zero lower bound, where they had been stuck at for several years. At the same time, concerns over the rebalancing of China's growth model and the possible manipulation of the renminbi made international headlines. Vice Chairman Stanley Fisher declared in a Jackson Hole speech that the FOMC had to *"consider [...] the influence of foreign economies on the U.S. economy as [they] reach [their] judgment on whether and how to change monetary policy"*, explicitly referring to the Chinese slowdown in the same speech (Fischer, 2015). Some interpreted this concern as the driving force behind the decision to shift the first rate hike since the global financial crisis towards December 2015.

Based on this anecdotal evidence, our research question is to quantitatively assess the possible effects of the global economy on U.S. monetary policy decisions. Our empirical strategy is to use the textual information contained in the minutes of the FOMC, published after each meeting of the committee. In this respect, we construct various indicators of attention to the international environment through Natural Language Processing methodologies. Then, we integrate those indicators into extended econometric equations building on the Taylor rule as initially put forward by John Taylor (Taylor, 1993, and Taylor, 1999).

Our work contributes to two major strands of the economic literature (see next section for more

---

details). The first one considers the role of the U.S. Fed within the global economy. Generally, authors look at the spillovers of Fed monetary policy decisions to the rest of the world (see among others [Dedola et al., 2017](#)), and to emerging countries more specifically. This literature has been highlighted by the famous words of Dilma Rousseff in 2012, then-President of Brazil, referring to the ultra-accommodative U.S. monetary policy as a driver of the *monetary tsunami* that was overwhelming emerging market with large capital flows. However, spillbacks from the rest of the world to the U.S. monetary policy are less a concern in the literature (see however [Breitenlechner et al., 2022](#)). The other strand of the literature is the quantitative analysis of central bank speeches and especially Federal Reserve communication. This literature starts with the influential work of [Romer and Romer, 2004](#), who dig into *Greenbook* documents prepared by Fed economists in order to estimate monetary policy shocks as the differences between FFR and forecasts of inflation, output and unemployment rate. With the development of Machine Learning tools and Natural Language Processing (*NLP*) techniques, authors have developed approaches to extract more efficiently textual information. For example, [Aruoba and Drechsel, 2022](#) have recently proposed new NLP methods to estimate U.S. monetary policy shocks.

Our work fits into both fields by using NLP approaches to extract from official Fed documents an index summarizing the attention of FOMC members to international economic conditions. In this respect, we extend and assess econometrically the idea put forward in the blog post by [Ferrara and Teuf, 2018](#), who are simply counting some international words contained in FOMC minutes.

The rest of this work is structured as follows. Section 2 briefly reviews the literature related to our work. Section 3 presents the data and the methods we consider to create our international environment indices and discusses our results in light of major international events. Section 4 presents the framework under which we will consider our index to assess its significance on the conduct of U.S. Monetary Policy. Section 5 concludes.

## 2. Literature review

### 2.1. U.S. monetary policy and the rest of the world

The situation of the U.S. central bank as a pivot of the world economy has been well established in the literature. Relationships between the Fed and the rest of the world go in both ways: from the U.S. monetary policy to the rest of the world (*spillovers*) and from the rest of the world to the U.S. monetary policy (*spillbacks*). However the literature appears somewhat asymmetric as the first direction has been by far the most studied. Indeed, most of the time, authors look at the spillovers of Fed monetary policy decisions to the rest of the world, and to emerging countries more specifically. Without being exhaustive, the works of [Bruno and Shin, 2015](#), [Dedola et al., 2017](#), or [Miranda-Agrippino and Rey, 2020](#), provide a good overview of the concept of global financial cycle and of the risk-taking channel of monetary policy, while [Degasperi et al., 2021](#), specifically study the effects of U.S. monetary policy tightening on advanced and emerging economies.

However, spillbacks from the rest of the world to the U.S. monetary policy are less a concern in the literature. [Eichengreen, 2013](#), offers a very thorough history of the attention paid by the Federal Reserve to international considerations in its century of existence. He widely describes periods of lower attention and renewed interest in the international environment by U.S. monetary policy-makers. In the first half of the 1960s, as worries mounted over a possible devaluation of the dollar and the amount of U.S. foreign liabilities, the Fed paid "considerable attention" to the balance of

---

payments, as argued by [Bordo and Eichengreen, 2013](#), resulting in a tighter monetary policy than what would have been expected from inflation and output gap considerations, something [Taylor, 1999](#), has called a *policy mistake*. [Eichengreen, 2013](#), underlines that this influence has drastically diminished after the late 1970s, when the conjunction of the collapse of the Bretton Woods system in 1973, the Federal Reserve Reform Act of 1977 (which carved into stone the dual mandate) and the arrival of Paul Volcker (who made fighting inflation his top priority), made international situation far less important for the Fed. The share of the U.S. in world GDP (33 percent in 1985, its peak), the moderate rise in the Trade/GDP ratio (compared to the following decades) and weaker international finance links than in the subsequent years meant that in this period and through the 1980s, the Fed was able to decide its policy under the rough approximation that the U.S. could be considered as a closed economy. This situation of low attention to international environment (relative to what was seen before) has dominated ever since, even though the decades since the 1980s have seen an increasing coordination of the Federal Reserve and other central banks (for meetings, coordinated interventions like what was seen in 2008, or through dollar swap lines opened during crises such as the global financial crisis). But Eichengreen argues that these occurrences show that "what happens abroad doesn't stay abroad" and in a prospective part, contends that international considerations should be more explicitly taken into account by the FOMC in the future.

In a more analytical work, [Obstfeld, 2020](#), explores the way the global economic situation can influence the U.S. economy, and in particular the economic variables monetary policymakers are interested in, especially inflation, employment, output and financial conditions. He notes that if global factors do not necessarily affect the ability of the central bank to exert a control on long-term price changes, they can have a strong influence in the short run. The author identifies the main channels of transmission through which these global factors can impact U.S. GDP. First of all, he highlights the role of international prices (such as commodities for example) but also of international competition in shaping U.S. inflation. The role of imported inflation is shown to be more and more important in setting domestic prices. Another channel originates from the integration of global financial markets, which affects the determination of asset returns (including  $r^*$ , the natural real rate of interest) and domestic financial stability. Overall, the author concludes that in a world made ever more complex by increasing cross-border interconnections, the U.S. economy is likely to be increasingly affected by global savings and investment trends, therefore requiring attention from U.S. monetary policy-makers. Those spillback effects have been highlighted in multiple studies, such as [Sachs, 1985](#), who argues that higher interest rates mean a strong dollar which pushes import prices lower and favors disinflation, and extensively studied in [Breitenlechner et al., 2022](#).

An overall result from empirical and theoretical works tends to acknowledge evidence of significant spillbacks, but of small magnitude. For example, [Caldara et al., 2022](#), use the Sigma model in its dollar-dominance version, developed at the Fed Board, and show that a monetary tightening of 100 basis point in the foreign economies results in a drop of U.S. GDP of 0.15 percent at the trough, that is about half of the impact of U.S. monetary policy on foreign GDP.

## 2.2. Textual analysis of monetary policy material

Our work also fits into the literature on textual analysis of documents produced by monetary policymakers, and especially FOMC material. Within that common theme, researchers have used

---

a large range of techniques to study widely different issues.

This literature starts with the influential work of [Romer and Romer, 2004](#), who dig into *Greenbook* documents prepared by Fed economists in order to estimate monetary policy shocks as the differences between FFR and forecasts of inflation, output and unemployment rate done by the Fed staff. Those monetary policy shocks have been then extensively used in the macroeconometric field to assess their effects on macroeconomic and financial variables (see for example [Coibion et al., 2017](#)).

With the development of Machine Learning tools and Natural Language Processing (*NLP*) techniques, authors have developed approaches to extract more efficiently and more accurately textual information from monetary policy documents. Perhaps the first example of computational textual analysis of FOMC content is the paper by [Boukous and Rosenberg, 2006](#) who are using *Latent Semantic Analysis* (LSA), a form of principal component analysis applied to text. This approach allows recognition of textual patterns in documents to decompose minutes into a set of themes, which are in turn linked they link to market moves. They conclude that market participants can gain "multifaceted signal" from the minutes, paving the way for more computational analysis of minutes text data. The same LSA methodology has been also applied by [Mazis and Tsekrekos, 2017](#) who identify themes in FOMC statements, among them credit spread, short or long end of the yield curve, with significant effects on the moves of the bond markets. [Fligstein et al., 2014](#) apply another topic modeling technique, *Latent Dirichlet Allocation* (LDA, see [Blei et al., 2003](#)), to analyze how FOMC members recognize financial crises. [Hansen and McMahon, 2016](#) also use a *LDA* to carry out an investigation of the effects of Fed communication on macro variables. Both LSA and LDA approaches are unsupervised topic modeling techniques which sets them apart from the methods we use in section 3.2.2. These topic modeling techniques have in recent years become a *de facto* default for textual data analysis in this field, but several others have been used in the related literature or in other economics-related natural language processing papers. [Aruoba and Drechsel, 2022](#) use NLP techniques to develop a novel method to identify monetary policy shocks by analyzing Fed *Greenbook* documents. In this respect, they construct multiple time series of the sentiment surrounding various economic concepts they identify in the documents ('aspect-based' sentiments) using the dictionary first put forward by [Loughran and McDonald, 2011](#). This approach through dictionaries is close to what we study in section 3.2.1 and in line with the initial work by [Ferrara and Teuf, 2018](#).

Some research works have also been conducted in a *supervised* learning framework. For example, [Handlan, 2022](#), applies a Neural Network that predicts change in Fed funds futures using FOMC statements to create a series of monetary policy shock. [Rohlfes et al., 2016](#), use a classification method (Max-Entropy Discrimination LDA or *MedLDA*) to classify whether a given topic leads to changes in interest rates. Then, they infer from this classification whether interest rates are likely to move in a given direction after the publication of the minutes of a meeting. Many other studies focus on the interactions between Fed decisions and financial markets. [Cieslak and Vissing-Jorgensen, 2021](#), count market-related words in Fed minutes to show that Fed monetary policy pays attention to the stock market. [Sharpe et al., 2022](#), use *Greenbook* documents to estimate a sentiment index which is in turn put into quantile regressions to show that sentiment can predict monetary policy surprises and stock returns. [Shapiro and Wilson, 2021](#), estimate FOMC's preferences, including the implicit inflation target, from the tone of the language used in transcripts, minutes, and members' speeches using predefined dictionaries of words associated with sentiments ("Bag of Words" or "lexical" approach). [Ochs, 2021](#), uses a measure of sentiment

---

in FOMC documents to disentangle effects of a monetary policy shock between the actual effects of monetary policy and the reaction of private agents to the newly acquired information (measured by a text-extracted variable). [Arseneau et al., 2022](#), analyze central banks speeches with NLP approaches to understand how central banks communicate about climate change.

Note that some other works have been conducted on other countries and central banks. Among others, we can quote the work by [Shibamoto, 2016](#) who uses a text-mining approach to show that asset prices react to Bank of Japan communication, while [Hubert and Labondance, 2021](#) finds that the ECB statements' tone explains monetary surprises. Also [TerEllen et al., 2022](#), use NLP techniques to identify narrative monetary policy surprises in Norway by looking at the difference between official central bank communication and economic media coverage.

### 3. Computing International Attention Indexes

In this section, we describe data and methods used to compute our International Attention Indexes (IAI).

#### 3.1. Data

Published three weeks after each FOMC meeting, FOMC minutes are a more complete but less timely overview of the mechanisms behind monetary policy decisions than the statement, published immediately after each meeting. As such, they are a major tool of transparency and accountability policy of the Federal Reserve. Despite this delayed publication, they can have a sizeable impact on financial markets. [Nechio and Wilson, 2016](#), show that information contained in FOMC minutes can impact Treasury bond yields at the time of their release, and that these impacts are unsurprisingly larger when the tone of the minutes differs from the tone of the statement.

We choose to work on FOMC minutes (and not on speeches made by its members or press conferences by its chair) for three main reasons. First, our aim is to study whether discussions of international matters *within* the FOMC are likely to influence monetary policy decisions, and we therefore limit ourselves to materials reflecting communication within the Federal Reserve, not from the FOMC towards outsiders<sup>1</sup>. Second, among internal materials, minutes offer the best compromise between completeness and delay, since the more complete *Meeting Transcripts* are published with a longer lag<sup>2</sup> therefore preventing us from studying recent periods. Finally, minutes feature a general structure which is stable throughout the period under study (1993-2022), making comparisons between values at different times both easier and more relevant. Overall, we have access to two sets of FOMC minutes per quarter.

We retrieve text files of the minutes of FOMC meetings from the dedicated section on the [Federal Reserve Board of Governors website](#). We include the minutes of conference calls held by the FOMC in case of emergency meetings (mainly during economic or financial crises) that are included at the end of the minutes of the following meetings. We also remove the list of all attendants of the

---

<sup>1</sup>Additionally, while we recognize that all central bank communication is political communication, we assume that since these documents, including the minutes, reflect internal discussion and are released with a delay, they are less prone to suffer from the biases usually highlighted in central bank communication analysis, such as carefully chosen words.

<sup>2</sup>As of January 2023, the latest transcripts published date back to mid-2016.

---

meeting (included at the start of every document). Our dataset comprises 238 meetings, around 30,000 paragraphs and 60,000 sentences. We transform each html document into a text file, which enables us to read the raw textual content and process it with **Python**.

All economic data time series were retrieved from the Atlanta Fed's [Taylor Rule Utility website](#), which includes data from the Federal Reserve, Bureau of Economic Analysis and Congressional Budget Office, among others.

### 3.1.1 Data Preprocessing

To efficiently use text data in computational textual analysis models, it is necessary to pre-process our raw text to "normalize" it. Our preprocessing<sup>3</sup> follows standard procedures for Natural Language Process projects, applied in much of the literature (see [Thorsrud, 2020](#), or [Handlan, 2022](#), for example).

#### Cleaning and stemming

First, stopwords (such as *the, a, ...*) words which do not bring any meaning to the sentence, are removed. We also remove numbers from the text. Words are then lowercased and tokenized. Finally, we stem all words, to reduce inflected words to their root form. For example, stemming transforms *dollars* into *dollar* or *fairly* into *fair*. To stem our text, we use the Snowball Stemmer ([Porter, 2001](#)), a stemmer which builds on the Porter stemmer and is widely accepted to be more efficient, for example at stemming adverbs. To clarify how it works, we provide examples of this 2-step process in [Table 1](#) below.

"A meeting of the Federal Open Market Committee was held."	meeting	Federal	Open	meet	feder	open	market
	Market	Committee	held.	committe	held		

**Table 1:** Examples of the 2-step preprocessing of a sentence

#### Vectorization

Since machine learning models are not capable of processing raw text but only numbers, we have to transform our sentences into vectors. For that, we select two methods whose results will be compared in our empirical analysis: (i) Term Frequency - Inverse Document Frequency (*TF-IDF*) and (ii) Word2Vec.

TF-IDF vectorization, based on a given corpus  $D$ , transforms a given sentence  $d$  into a vector of size  $n$  (the number of different words in the corpus) by including for each word  $t$  in the corpus a measure of its importance in the sentence, denoted  $tfidf(t, d, D)$ . TF-IDF is the combination of two measures: Term Frequency, put forward by [Luhn, 1958](#), and Inverse Document Frequency, first devised by [Spurck Jones, 1972](#). For a term  $t$  in a sentence  $d$  in a corpus  $D$ , we define:

$$tfidf(t, d, D) = tf(t, d) \times idf(t, D) \tag{1}$$

where:

---

<sup>3</sup>A diagram of our preprocessing is featured in [Figure 6](#), [Appendix A](#).

- $tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$  is the relative frequency of term  $t$  within sentence  $d$ , where  $f_{t,d}$  is the raw count of term  $t$  in sentence  $d$ .
- $idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$  is the inverse document frequency of term  $t$  in corpus  $D$ , where  $|D|$  is the total number of documents in the corpus<sup>4</sup>.

Word2Vec, published in Mikolov et al., 2013, is a word embedding method: it transforms words into vectors by using a neural networks to infer word meanings from a corpus of text. The method generates a vector space, where, ideally, words are positioned to be close in the space (in terms of cosine similarity) if they are close in meaning. The vector space means that the representation can subtract or add vectors to each other. For example, we have:  $king - man + woman \approx queen$ .

Both vectorization methods need to be trained on a given corpus of textual data: TF-IDF derives frequency from a corpus and Word2Vec needs to learn word association. In our case, both trainings are performed on the corpus of cleaned, stemmed minutes from 1993 to 2018. Details on those vectorization methods are featured in Appendix C.

## 3.2. NLP Methods

We describe here the two NLP methods that we carry out after the vectorization step of the FOMC minutes, namely the *dictionary* approach and the *classification with supervised models*.

### 3.2.1 Dictionary count

In this first method, we track simple counts of international-related words contained in a given *dictionary*<sup>5</sup>. For each FOMC document, we count the number of occurrences of the dictionary’s words contained in the document and divide it by the total number of words. This ratio is our first indicator of attention to foreign matters by the FOMC, shown in Figure 3 and defined by:

$$IAI_M = \frac{\sum_{w \in M} \mathbb{1}_{w \in D}}{|M|} \quad (2)$$

where  $M$  is the set of all words in stemmed minutes and  $|\cdot|$  denotes the cardinality of a set. To obtain a quarterly time series, we perform this count on  $M_Q$ , the set of all words in stemmed minutes of a given quarter  $Q$  that includes two FOMC meetings.

Table 2 lists the words and groups of words we use in this dictionary count, organized in three different categories: *foreign-exchange related*, *trade-related*, *others*. Appendix E reports the number of times each search phrase appears in the corpus. Obviously, the choice of the dictionary is crucial in this exercise. In most of the literature on dictionary count, words are chosen according to economic intuition, in spite of recent efforts to develop new dictionaries (see for example Renault, 2017, Gardner et al., 2023, or Barbaglia et al., 2023). To justify our dictionary, we follow an algorithm inspired by Aruoba and Drechsel, 2022, detailed in Appendix B.

<sup>4</sup>The denominator (number of documents where the term  $t$  appears) is commonly adjusted to  $1 + |\{d \in D : t \in d\}|$  to avoid division by 0 if the word is not in the original corpus.

<sup>5</sup>While it is true that the index depends of the phrases selected, our methodology is robust to small changes in the dictionary. The effects of such changes are discussed in Appendix F.



Foreign Exchange	Trade	Others
depreciation	international trade	international
currency	imports	global
foreign currencies	exports	foreign economies
foreign exchange	external sector	emerging market
dollar		EME
euro		advanced economies
pound		china
sterling		russia
renminbi		europa
ruble		india
		japan
		asia
		world
		abroad

**Table 2:** Dictionary used in the study.

### 3.2.2 Classification with supervised models

In this second NLP method to construct international attention indicators from our textual data, we divide the text into sentences<sup>6</sup>, which we then classify using Machine Learning models as either relating to international matters or not<sup>7</sup>. Once sentences are classified, we compute the ratio of the number of words in a text belonging to international-classified sentences on the total number of words in the text. This ratio is defined as our international attention indicator for the minutes considered.

$$IAI_M = \frac{\sum_{s \in M} \mathbb{1}_{cl(s)=1} |s|}{\sum_{s \in M} |s|}, \quad (3)$$

where  $M$  is the set of all sentences in stemmed minutes,  $s$  is a sentence (set of words) and  $Cl$  is the sentence classifier that returns 1 if the sentence is classified as international-related. To obtain a quarterly series, we perform this count on  $M_Q$ , the set of all sentences in stemmed minutes of a given quarter  $Q$ .

Building classification-based indexes is essentially a four-steps process : (i) labeling data to train the algorithms, (ii) training algorithms on a *training set*, (iii) determining the most relevant ones using a *test set* and (iv) finally computing the classification of all sentences and the international environment index.

### Data Labelling

In a first step, we manually label a total of 4,300 sentences to classify them between those which relate to the international situation (class 1, or "Positive") and those which do not (class 2, or "Negative"). To avoid labeling only data from a specific time period or specific part of the

<sup>6</sup>Extracted from the text using the `nltk` library in Python (Bird et al., 2009)

<sup>7</sup>We performed tests before the manual labeling to decide whether to use sentences or paragraphs as our reference atom for the study, and decided for sentences. Tests for paragraphs are featured in Appendix H.

minutes structure, we shuffle the dataset and label sentences that have been randomly chosen. We also avoid inserting a forward-looking bias on recent events by only labeling data until 2018. Descriptive statistics of labeled data and total data are displayed in Table 3 below:

	Labeled data	Total data
Total Number	4314	66557
Average Length (in words)	11.55	12.66
Number of positives	401 (9.3%)	

**Table 3:** Descriptive statistics of sentences and labeled sentences. Positive means relating to international situation.

Since the dataset of labeled data is biased (there are 90.7% of non-international-related sentences and 9.3% of international-related sentences), we rebalance it by keeping only 1/9th of the negative (i.e. non-international) observations. Indeed, we find that even though the imbalance is not extreme<sup>8</sup>, the models trained on a dataset on which we do not apply this bias correction tend to exhibit a very low recall (percentage of positives well predicted by our model), i.e. many positive observations are likely to be classified as negative<sup>9</sup>.

### Classification algorithms training

We first divide the dataset of minute sentences between a training dataset (two thirds of the dataset) and a testing dataset (the remaining third of the sentences). The attribution of a sentence to one set or the other is done randomly. The training dataset is the one the algorithms will have access to estimate parameters, while the testing dataset will be used to assess their performance. For every algorithm, the parameters are determined via a *k-fold Cross Validation* approach, as it is common in Machine Learning. The training dataset is divided into five subsets, and the algorithm is trained on four of them. It is then tested on the subset that was left out, and its performance on this subset is assessed. We retain the model version that performs the best on its testing subset. We also perform a grid search of hyperparameters, where every combination of hyperparameters is tested by cross validation.

Here we test two vectorization methods, TF-IDF and Word2Vec, and three machine learning classifiers : Naive Bayes, Logistic Regression, Random Forest<sup>10</sup>. That makes six pipelines (combination of a preprocessing and an algorithm).

We add a seventh pipeline: Dictionary-based naive classification. In this pipeline, we classify sentences as positive or negative based on whether or not they contain one of the phrases in the international dictionary described in Table 2. Each sentence that contains at least one word or group of word belonging to the dictionary is classified as international-related.

<sup>8</sup>Extreme imbalance would be defined in Machine Learning literature as 1, 2 or 5% of class 1 observations.

<sup>9</sup>We could have used a resampling technique to augment the number of positive observations in our unbalanced labeled dataset. Since the models we use do not require huge volumes of labeled data and more manual labeling is possible for our minutes sentences dataset, we decided against this option, which never reaches the results obtained with a non-augmented unbiased dataset, as shown by Baesens et al., 2022

<sup>10</sup>More information on these algorithms can be found in Appendix D

---

## Results of classification pipelines on the testing set

We therefore compare the results of seven pipelines, using standard classification metrics: Precision, Recall, F1 Score. For every class, *Precision* is the proportion of true positives (i.e. rightly classified in this class) on all predicted positives, *Recall* is the proportion of actual positives that have been predicted as such, and *F1 Score* is the harmonic mean of precision and recall. All metrics are reported in Table 4 and have to be maximized. We can also compute a macro average of every metric for the whole dataset (not just one class), detailed results of the seven pipelines on the *testing dataset* are featured in Appendix G.

Vectorization	NLP Algorithm	Precision	Recall	F1-Score
Word2Vec	Logistic Regression	0.5770	0.5154	0.3617
	Naive Bayes	0.2841	0.5000	0.3624
	Random Forest	0.7852	0.5043	0.3717
TF-IDF	Logistic Regression	0.8987	0.8843	0.8892
	Naive Bayes	0.8427	0.8378	0.8191
	Random Forest	0.9077	0.8971	0.9011
	Dictionary Classification	0.8692	0.8532	0.8582

**Table 4:** Results of the seven pipelines

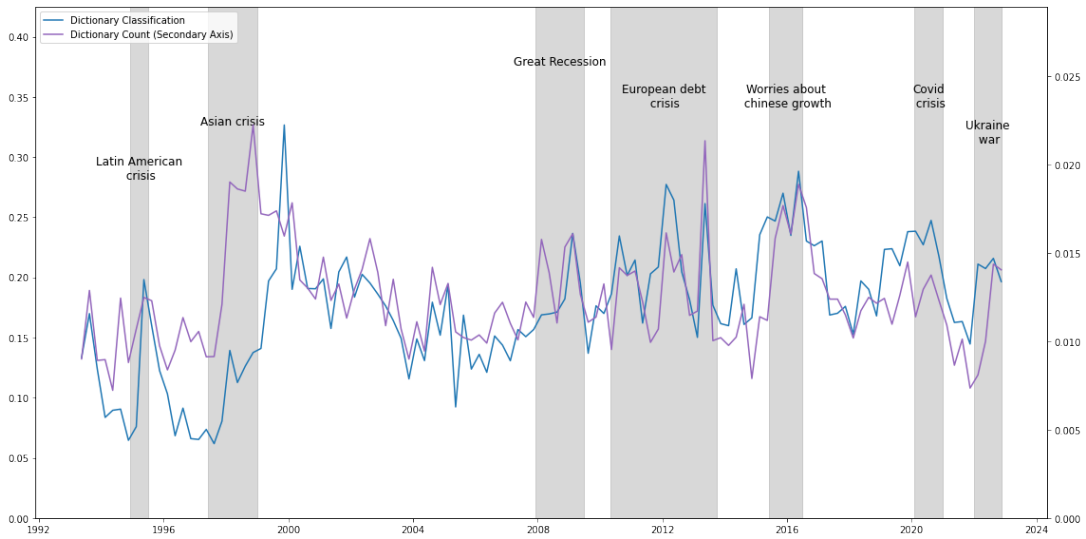
Results show that the models using Word2Vec for vectorization perform poorly on our test dataset. One explanation could be that the *training dataset* of the Word2Vec algorithm, that is our corpus, is not large enough for the neural network to learn words association. This is why we will only focus on the TF-IDF vectorization approach in our results. As regards the TF-IDF-based pipelines, performance for Naive Bayes seem to be lower than other algorithms, as it performs worse in Precision and Recall criteria, and Random Forest seems to perform best, boasting the best results in terms of both Precision and Recall criteria. Our Naive Dictionary classifier has performance on par with the best models in terms of precision, which validates this dictionary-based approach: the phrases selected in our dictionary are relevant to detect the occurrences of attention to the international situation in the FOMC minutes.

We therefore choose to built indexes with four pipelines: the Naive Dictionary classifier and all classifiers based on TF-IDF.

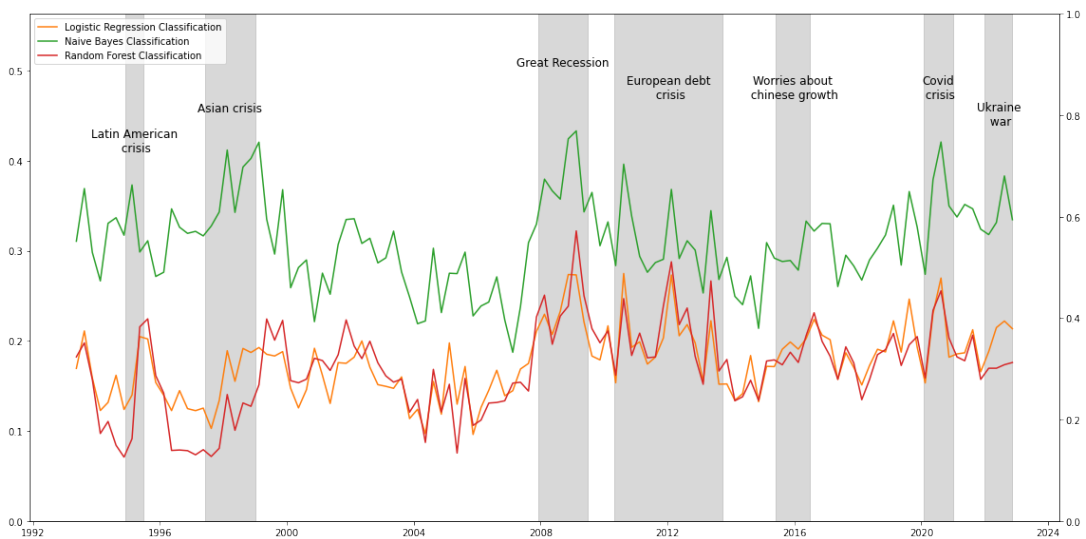
### 3.3. International Attention Indexes

In this subsection, we present and discuss the various versions of our International Attention Index (IAI). Indexes stemming from our dictionary count approach described in section 3.2.1 and the four classification methods described in section 3.2.2 are shown in Figures 1 and 2.

By eye-balling those graphs, we note that spikes in the indexes, that is periods of accrued international attention in FOMC minutes, match the major international events of the past three decades. In particular, there is a good fit with crises of any type outside the U.S.: we point out spikes at the time of the Latin-American crisis of the mid 1990s, the Asian crisis, the 2008-09 Great Recession,



**Figure 1:** International Attention Indexes obtained with dictionary count and dictionary classifier.

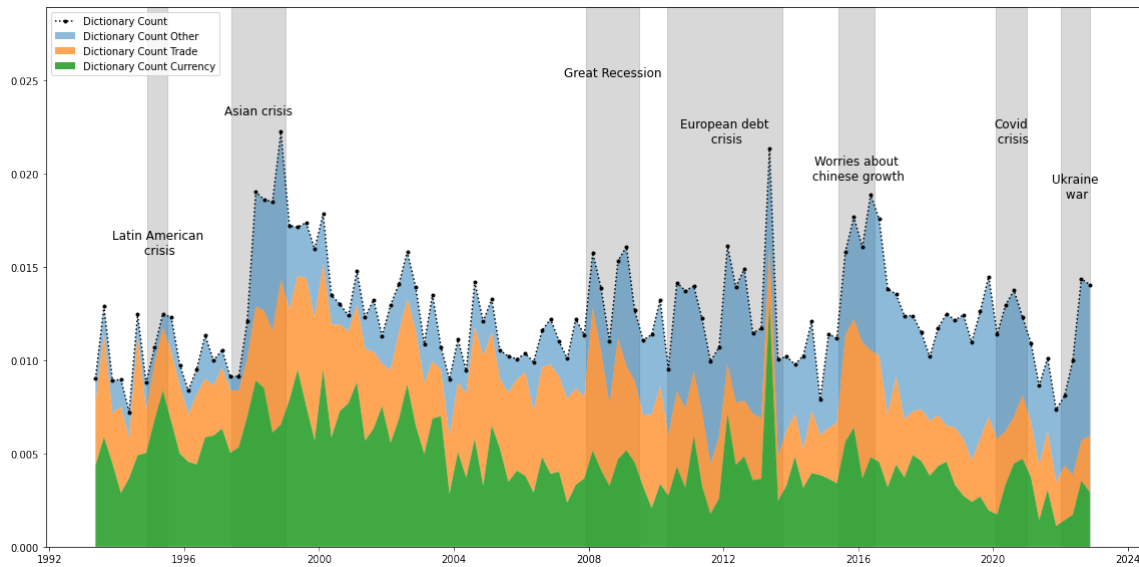


**Figure 2:** International Attention Indexes obtained with sentences classification through Machine Learning methods.

the European debt crisis in the 2011-13 period, the concerns about the Chinese economy in 2015-2016, as well as both recent crises, namely the Covid pandemic and the Russian invasion of Ukraine.

The size of these spikes differs depending on the version of the index. The relatively small rise seen during the Great Recession in dictionary-based models (relative to the Machine Learning based ones) might come from a different type of attention in these years (through swap lines with foreign central banks for example) that is not as well captured in our dictionary. Similarly, Machine Learning based models are able to catch a significant spike during the Covid crisis (a period for which they have seen no training data, since we only labeled sentences from 1993 to 2018) while dictionary count does not move.

All indexes tend to show an asymmetric behaviour in the sense that discussions about the international environment seem to take place mainly when there are negative news on foreign economic activity or markets<sup>11</sup>.



**Figure 3:** Value of the International Attention Index for dictionary count method and sub-dictionaries counts.

One of the advantages of the dictionary approach is that we can split international words into categories that sum up to the global index. Thus we decompose the dictionary-based IAI into three main categories, namely trade, currency and others (see Table 2). Figure 3 shows this decomposition of the IAI into specific channels through which foreign events might affect the Fed monetary policy discussions. We first note a long-run decline in the component related to currencies. This was the topic of discussions before the 2000s, maybe related to the introduction of the euro as competitor for the dollar, but since then it does seem to be an issue among FOMC members. Let's just note the spike in 2013q1 at the end of the European debt crisis, a long period of dollar over-valuation with respect to the euro. In comparison, the trade component tends to show more spikes, the last one being between mid-2015 and beginning 2016 amidst concerns on the Chinese economic growth and the renminbi. Interestingly, the trade war between the U.S. and China/Europe started by the Trump's administration in 2018 didn't affect discussions among FOMC members. Increase in tariffs from imports of those countries was not perceived as a concern for monetary policy by Fed's officials.

The last component, named *others* is driving the headline value of the international attention index since the mid 2000s. This is likely due to the fact that most of the mentions of foreign crises link the name of the involved countries with broad economic concepts (growth, inflation, etc.), which we could not include in our dictionary because these concepts are also associated with separate domestic news in other parts of the documents that are of not of direct interest for our research work.

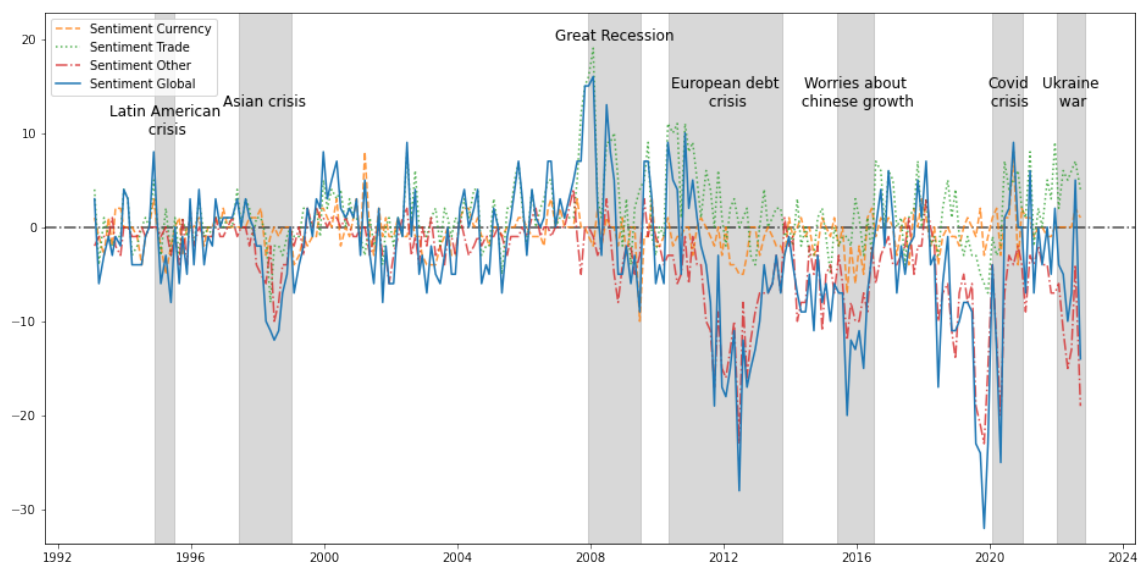
<sup>11</sup>Section 3.4 investigates this point by studying the positive or negative sentiments of international-related words, but we do not conclude on the systematic negativity of attention to international matters.

### 3.4. Sentiment around international-related words

Against our background, it is also tempting to extract a sentiment around some international topics evoked during FOMC meetings. Sentiment indexes have proved extremely useful for economists to assess the tone of large textual databases (see [Algaba et al., 2020](#)), including for economic forecasting ([Barbaglia et al., 2022](#)). Here, we investigate evidence of a negative bias of Fed attention to the international economic situation, in the sense that FOMC members are likely to discuss more about international situation during crises. In this respect, we follow a methodology used by [Aruoba and Drechsel, 2022](#): for a given set of phrases (our international dictionary in [Table 2](#)), we isolate the 10 words<sup>12</sup> surrounding each occurrence of any of the phrases and compute the "sentiment" (positivity or negativity) of these words using the [Loughran and McDonald, 2011](#), dictionary. The sentiment around a phrase is estimated as the number of positively connoted words around its vicinity minus the number of negatively connoted ones. *In fine*, for a given FOMC meeting, a final international sentiment score is obtained by summing up the sentiment scores of all phrases contained in the related document. Given the structure of our dictionary in [Table 2](#), it is also possible to compute a sentiment index for the three categories, namely *Currency*, *Trade* and *Others*. Our estimated sentiment indexes are presented in [Figure 4](#).

By eyeballing [Figure 4](#), we note that the global international sentiment estimated from FOMC minutes seems to be relatively well balanced until the Great Recession. However, starting from 2009, the sentiment index is clearly tilted to the downside. We observe large negative spikes during various phases such as the European sovereign crisis, the period of rising concerns about the Chinese economy or the Covid pandemic. The sequence of international crises seems to have introduced a negative bias in the attention of FOMC members.

<sup>12</sup>As a robustness check, we test the size of the "halo", i.e. the number of surrounding words taken into account. Results remain extremely similar.



**Figure 4:** Sentiment around international-related words, for the global dictionary and each of our 3 channels (10 words surrounding each international-related word).

Breaking down the global sentiment into the three categorical sentiments, also featured on Figure 4, shows that the sentiment related to category *Others* is driving the global dynamics. The currency-related international sentiment exhibits much less volatility and seems overall balanced. However, we note negative values during the period of concern about China economic growth in line with rising uncertainties at that time on the composition of the panel of currencies against which the renminbi was managed by Chinese authorities. The trade-related international sentiment does show a clear pattern and appears well balanced overtime. Interestingly, the trade war launched in 2018 by the Trump's administration seems to positively affect the sentiment index.

## 4. Effects of International Attention Indexes on monetary policy

This section empirically investigates if the attention to international environment by the FOMC (as measured by the different indexes detailed in section 3.3) can have an influence on monetary policy decisions. In this respect, we follow a framework in the spirit of [Cecchetti et al., 2000](#), by estimating a Taylor rule augmented with a given International Attention Index (IAI). This framework enables to assess whether our indexes provide significant information on Fed decisions. We assume in this section that despite the wider range of tools policymakers have now at their disposal (forward guidance, asset purchases ...), the Fed Funds Rate remains the main monetary policy tool during the period under study. We will also consider the U.S. shadow rate put forward by [Wu and Xia, 2016](#) in order to account for unconventional monetary policy tools launched since 2009. The estimation sample goes from 1994q1 to 2022q3.

### 4.1. Framework

The so-called Taylor rule, initially put forward by [Taylor, 1993](#), estimates the response of Fed interest rate to macroeconomic conditions. In its basic form, it is the endogenous response of monetary policy to domestic macroeconomic conditions. The basic version of the Taylor rule (Equation (4) below) linearly relates the nominal Fed Funds Rate  $r_t$ , to core PCE inflation rate  $\pi_t$ , the inflation target  $\pi_t^*$ , the output gap  $g_t$  and a to Gaussian white noise  $\varepsilon_t$  in the following way:

$$r_t = \alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \varepsilon_t \quad (4)$$

where  $\alpha$  can be interpreted as the nominal neutral interest rate that prevails when inflation is at the target and output gap is closed (i.e.  $\alpha = \pi_t + r_t^*$ , where  $r_t^*$  is the real neutral interest rate). In this work, we use an extended version of the Taylor Rule given by Equation (5) below, which includes  $\rho$ , a smoothing term for the FFR and  $i_t$ , our International Attention Index (IAI)<sup>13</sup>.

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t) + \varepsilon_t \quad (5)$$

We chose to work on data provided by the Atlanta Fed on its [Taylor Rule Utility website](#) and not on Greenbook forecasts as for example [Istrefi et al., 2021](#), since those latter forecasts are no longer available on the Fed website from 2010 onward, and we would also like to study the influence of our index in the 2010s and early 2020s. We choose to use as  $\pi_t^*$  the four-quarter core PCE inflation, constructed by the U.S. Bureau of Economic Analysis (BEA), and as  $g_t$  the output gap derived from the Congressional Budget Office's (CBO) estimate of potential real GDP. All variables are expressed in percent, and our index  $i_t$  is standardized before entering the equation. We estimate this Taylor Rule by Non-Linear Least Squares<sup>14</sup> and get the results featured in the following

<sup>13</sup>Another version of the Taylor Rule, with a time-varying  $r^*$ , is discussed in Appendix J.

<sup>14</sup>Non-Linear Least Squares approach is implemented in R with the `nls` function.

---

section.

## 4.2. Empirical results

### 4.2.1 Main results

Table 5 presents the results of our various regressions. Estimated parameters of the standard Taylor rule (column 0) are quite close for the usual calibration based on initial results of Taylor. The smoothing parameter is highly significant underlying the strong persistence of FFR. The intercept is close to 4%, meaning that the real neutral interest rate  $r^*$  can be proxied by a constant around 2% over the 1994-2022 period. We will later relax this assumption in line with recent research showing time-variation in  $r^*$  (see [Laubach and Williams, 2003](#)). The estimated weight of the inflation gap is above 1 (1.28) but doesn't appear significantly different from zero. In opposition the estimated coefficient in front of the output gap is clearly significant. We now integrate our various IAIs into the Taylor equation (5). First, we note that the integration of the IAIs doesn't change much the other coefficients, highlighting the stability of the equation. Then, we get that the estimated IAI coefficients are always negative, and generally significant, except for models (1) and (5) (dictionary and Naive Bayes, respectively). The version of the IAI that yields the best results in terms of significativity, AIC and MSE is model (4), that is the one based on Random Forest classification of sentences. This is also consistent with the fact that it is the method that performs best in terms of out-of-sample classification (Table 4). In this version, the (*standardized*) IAI has a coefficient of -8.15, which means that when the index is equal to 0.1, the marginal effect of the international environment on FFR is about 81 basis points on average. So for example, at the heart of the Global Financial Crisis, the IAI went to 0.3, meaning that the contribution of the international outlook to the zero lower bound was about 240 basis point at that time.

### 4.2.2 Robustness checks

Let's focus on Model (4), which can be considered as the best model according to all criteria, and assess some robustness checks.

Let's first introduce some additional control variables in equation (5) to ensure that our IAI provides information beyond that of a simple measure of international financial instability. In this respect, we estimate the following augmented equation (6):

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t + \phi s_t) + \varepsilon_t. \quad (6)$$

where  $s_t$  is an index of global conditions. We consider two indexes: the MSCI World (equity) index and the Geopolitical Risk (GPR) index of [Caldara and Iacoviello, 2022](#) (both taken in differences). Results are reported in Table 6. The coefficient for the global equity index does not appear statistically significant when we control by our IAI, but becomes significantly positive when it is integrated alone in the basic Taylor rule. This may reflect the fact that the IAI and the MSCI are somewhat negatively correlated, suggesting that a drop in the MSCI is associated to higher international attention by the FOMC. In opposition, the geopolitical risk as estimated by the GPR index appears always negatively significant in the Taylor rule, independently of the inclusion of the IAI. This indicates that the two indicators are not necessarily correlated, suggesting that the IAI doesn't capture geopolitical tensions. From a policy perspective, this result suggests that the international attention of the FOMC is more focused on international financial markets than on



Dependent variable Model	FFR						
	(Cal)	(0)	(1)	(2)	(3)	(4)	(5)
$r_{t-1}$	0.85	0.915*** (0.024)	0.913*** (0.024)	0.894*** (0.026)	0.91*** (0.023)	0.903*** (0.023)	0.92*** (0.024)
$\alpha$	4.0	4.17*** (0.557)	4.191*** (0.548)	3.853*** (0.44)	3.801*** (0.508)	3.529*** (0.459)	4.121*** (0.59)
$PCEGap_t$	1.5	1.285 (0.794)	0.889 (0.777)	0.858 (0.598)	1.236* (0.725)	0.922 (0.615)	1.364 (0.863)
$g_t$	1.0	1.209*** (0.244)	1.251*** (0.246)	1.005*** (0.195)	0.96*** (0.225)	0.79*** (0.206)	1.178*** (0.257)
$i_t$			-3.06 (2.506)	-4.084** (1.824)	-5.185** (2.322)	-8.149*** (2.639)	-3.32 (2.693)
Observations		115	115	115	115	115	115
Period		1994- 2022	1994- 2022	1994- 2022	1994- 2022	1994- 2022	1994- 2022
AIC		115.637	115.95	113.197	110.812	102.444	115.546
MSE		16.872	16.626	16.233	15.9	14.784	16.568
Taylor Version		9	12	12	12	12	12

**Table 5:** Taylor rule regression results — Equation 5 specification

*Note:* The table shows Taylor Rule (Equation 5) results for Non-Linear Least Squares Regression, with different inputs as the international attention index  $i_t$ : (0) is no international environment index, (1) the index derived from the dictionary count, (2) the index from dictionary classification of sentences, (3) the index from logistic regression classification of sentences, (4) the index from random forest classification of sentences, (5) the index from Naive Bayes classification of sentences. The (Cal) column features the calibrated parameters (from the Atlanta Fed)

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

geopolitical tensions between countries. This is in line with the absence of movement in the IAIs during the trade war between U.S. and China.

Second, we study the variation over time of  $\delta$ , the estimated coefficient of our IAI, in search for possible shifts during some specific periods. We carry out a recursive analysis starting from the period 1994-2005, then we progressively extend the window by adding a new data point, and we re-estimate the augmented Taylor rule. For each point in time, we report the estimated value of  $\delta$ , the coefficient of  $i_t$ , and its 68% confidence interval. Figure 5 shows the evolution overtime of  $\hat{\delta}$ . We observe that the estimate is always largely significantly different from zero, underlining the robustness overtime of the model. The international environment was playing negatively on central bank interest rate before the Global Financial Crisis (GFC), but there is a clear downward shift in 2008 due to the cross-border propagation of the crisis and its spill-back effect on U.S. monetary policy as the financial crisis was really global. In the wake of the GFC, the estimated coefficient stayed at a low level, reflecting an always strong sensitivity of policy-makers to external developments. Additional results involving Model (4) are reported in the Appendix, for example when taking unemployment gap instead of output gap.

Dependent variable Model	FFR					
	(Cal)	(4)	(4.4.1)	(4.4.2)	(4.5.1)	(4.5.2)
$r_{t-1}$	0.85	0.903*** (0.023)	0.912*** (0.023)	0.926*** (0.024)	0.909*** (0.022)	0.92*** (0.023)
$\alpha$	4.0	3.529*** (0.459)	3.397*** (0.507)	3.927*** (0.618)	3.657*** (0.476)	4.329*** (0.594)
$PCEGap_t$	1.5	0.922 (0.615)	1.154 (0.722)	1.698* (1.004)	1.161* (0.672)	1.586* (0.882)
$g_t$	1.0	0.79*** (0.206)	0.769*** (0.224)	1.167*** (0.27)	0.774*** (0.211)	1.186*** (0.248)
$i_t$		-8.149*** (2.639)	-8.111*** (2.862)		-8.115*** (2.707)	
$MSCI_t$			7.538 (5.689)	13.566* (8.068)		
$Risk_t$					-1.619** (0.733)	-2.062** (0.959)
Observations		115	115	115	115	115
Period		1994-2022	1994-2022	1994-2022	1994-2022	1994-2022
AIC		102.444	101.967	112.32	96.663	109.004
MSE		14.784	14.469	16.109	13.816	15.652
Taylor Version		12	17	18	19	20

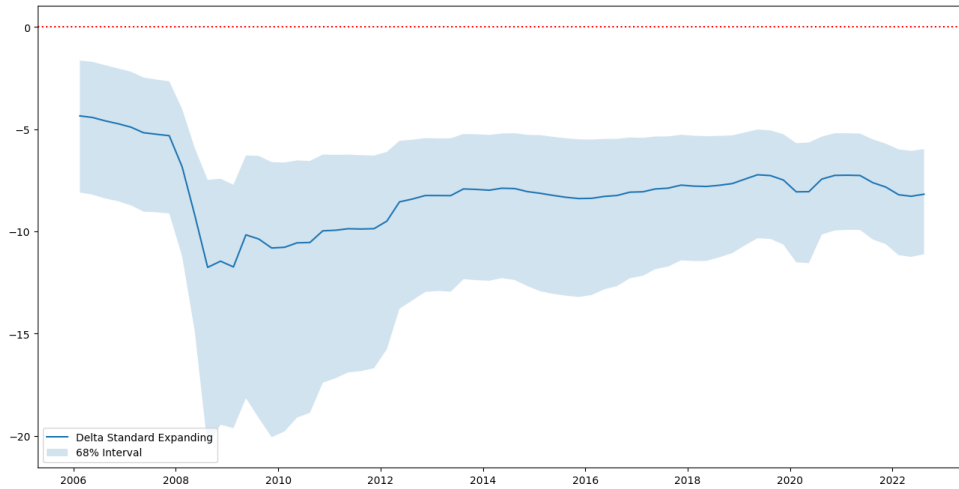
**Table 6:** Taylor rule regression results - with financial situation control variables

Note: The table shows Taylor Rule regression with additional variables to control for the global financial environment. All models use Non-Linear Least Squares Regression, with the same international attention index. Model (4) features no control, (4.4.1) and (4.4.2) the Quarter over Quarter evolution of the MSCI World Equity index, respectively with and without the IAI index included (Equations 17 and 18), (4.5.1) and (4.5.2) the Quarter over Quarter evolution of the GPR index, respectively with and without the IAI index included (Equations 19 and 20). The (Cal) column features the calibrated parameters (from the Atlanta Fed).

\* p<0.1; \*\* p<0.05; \*\*\* p<0.01.

Third, let's consider some robustness checks on interest rates. Monetary policy tools have changed since the GFC and now unconventional tools, such as quantitative easing or forward guidance, have joined the toolbox of policy-makers. To account for this, some researchers have proposed a modified version of central bank interest rates that reflects additional accommodation in the monetary policy stance. In this respect, Wu and Xia, 2016, provide a *shadow* interest rate for the U.S. economy that we integrate into equation (5). This shadow Fed rate has the advantage of taking into account unconventional monetary measures put in place after 2009, leading to some negative values for the FFR. Appendix I present the results obtained when the Fed Funds Rate is replaced as the dependent variable by the Wu-Xia *shadow rate*. Results are stronger in the sense that the estimated coefficient of the IAI is highly significant and increases to -10.48 (compared to -8.15 when using the standard FFR).

Fourth, another stylized fact from the monetary policy literature is that there is clear evidence of long-run decline in  $r_t^*$ , the real natural interest rates. This is mainly due to structural issues related



**Figure 5:** Time-varying estimated value  $\hat{\delta}$ , the IAI coefficient, and its 68% confidence interval in an expanding window.

to population ageing or secular stagnation hypothesis (see for example [Laubach and Williams, 2003](#)). Appendix J explores other versions of the augmented Taylor rule that incorporate the time-varying  $r_t^*$  estimated by [Laubach and Williams, 2003](#) and updated by the Federal Reserve Bank of New-York. When  $r_t^*$  is integrated in a standard Taylor rule, both coefficients of inflation gap and output gap sharply decline, but the one related to the output gap stays significant. There is no change as regards the IAI estimated coefficient that stays at -8.03 (*vs* -8.15 in the standard version without  $r_t^*$ ). So results appear extremely robust to the inclusion of time-varying neutral interest rate.

## 5. Conclusions

The U.S. Federal Reserve has officially domestic objectives of price stability and full employment and is often considered to pay little attention to the rest of the world. Yet economic analysis and anecdotal evidence suggest that the international situation is likely to play a role in U.S. monetary policy, or at least is considered to some extent by policy-makers. To test this hypothesis, we develop a quantitative approach to assess the attention paid by members of the Federal Open Market Committee (FOMC) to foreign economic conditions and its potential effects on monetary policy decisions from 1993 to 2022.

In this respect, we deploy various Natural Language Processing (NLP) methods of computer-based textual analysis to construct International Attention Indexes (IAIs) from the minutes of FOMC meetings. Those indexes are proxies of the attention paid by FOMC members to economic situation abroad. We focus on the FOMC minutes as they reflect internal discussions within the committee and offer a good trade-off between publication delays, on the one hand, and completeness in terms of policy discussions, on the other hand. After pre-processing this large textual dataset, we include it into both dictionary-based models and Machine Learning classification models. As an output, we get various IAIs reflecting the main economic events of the last decades.

In a second step, we use our IAIs to augment standard Taylor rules in order to assess Fed monetary

---

policy reaction to both U.S. domestic and international conditions. Overall, we get that those indexes have a significantly negative effect on the level of Fed Funds Rates. In other words, those indexes reinforce a dovish pattern with respect to a standard Taylor rule. We interpret those results as evidence that FOMC members tend to take global conditions into account when reaching a decision. Those results turn out to be robust to various changes in modelling, such as the integration of shadow central bank interest rate or the use of time-varying real neutral interest rates.

One possible continuation of this work could be to conduct the same type of study on a longer time frame to try to identify the longer-term shifts put forward by [Eichengreen, 2013](#). While very appealing, this potential continuation runs into data-linked obstacles, namely the availability of comparable documents. *Minutes* of FOMC meetings are available from 1993 onward, but only *Minutes of Actions* exist from 1967 to 1992, and *Historical Minutes* from 1936 to 1967. If historical minutes are pretty comparable to current-format minutes, the *Minutes of Actions* are much shorter in length and feature significantly less discussion.

---

## References

- Algaba, A., Ardia, D., Bluteau, K., Borms, S., & Boudt, K. (2020). Econometrics Meets Sentiment: An Overview Of Methodology And Applications. *Journal of Economic Surveys*, 34(3), 512–547.
- Arseneau, D. M., Drexler, A., & Osada, M. (2022). Central bank communication about climate change. *Finance and Economics Discussion Series, Board of Governors of the Federal Reserve System*, (No. 2022-031).
- Aruoba, B., & Drechsel, T. (2022). Identifying monetary policy shocks: A natural language approach. *CEPR Discussion Papers*, 17133.
- Baesens, B., Banulescu-Radu, D., Hurlin, C., Koughblenou, Y., & Verdonck, T. (2022). Benchmarking state-of-the-art resampling techniques for classification models. *mimeo*.
- Barbaglia, L., Consoli, S., & Manzan, S. (2022). Forecasting with economic news. *Journal of Business and Economic Statistics*, forthcoming.
- Barbaglia, L., Consoli, S., Manzan, S., Pezzoli, L. T., & Tosetti, E. (2023). Sentiment analysis of economic text: A lexicon-based approach. *SSRN*.
- Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with python: Analyzing text with the natural language toolkit*. O'Reilly Media, Inc.
- Blei, D., Ng, A., & Jordan, M. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 993–1022.
- Bordo, M., & Eichengreen, B. (2013). Bretton Woods and the Great Inflation. In *The Great Inflation: The Rebirth of Modern Central Banking* (pp. 449–489). National Bureau of Economic Research, Inc.
- Boukous, E., & Rosenberg, J. V. (2006). The information content of FOMC minutes. *Federal Reserve Bank of New York*, (Working Paper).
- Breitenlechner, M., Georgiadis, G., & Schumann, B. (2022). *What goes around comes around how large are spillbacks from US monetary policy?* European Central Bank.
- Bruno, V., & Shin, H. S. (2015). Capital flows and the risk-taking channel of monetary policy. *Journal of Monetary Economics*, 71(100), 119–132.
- Caldara, D., Ferrante, F., & Queralto, A. (2022). *International spillovers of tighter monetary policy* (FEDS Note No. 26609). Board of Governors of the Federal Reserve System.
- Caldara, D., & Iacoviello, M. (2022). Measuring geopolitical risk. *American Economic Review*, 112(4), 1194–1225.
- Cecchetti, S., Genberg, H., Lipsky, J., & Wadhvani, S. (2000). Asset prices and central bank policy. *The Geneva Report on the World Economy*.
- Chakraborty, C., & Joseph, A. (2017). Machine learning at central banks. *SSRN Electronic Journal*.
- Cieslak, A., & Vissing-Jorgensen, A. (2021). The economics of the fed put. *The Review of Financial Studies*, 34(9), 4045–4089.
- Coibion, O., Goronidchenko, Y., Kueng, L., & Silvia, J. (2017). Innocent bystanders? monetary policy and inequality. *Journal of Monetary Economics*, 88, 70–89.

- 
- Dedola, L., Rivolta, G., & Stracca, L. (2017). If the Fed sneezes, who catches a cold? *Journal of International Economics*, 108(S1), 23–41.
- Degasperi, R., Hong, S. S., & Ricco, G. (2021). The global transmission of u.s. monetary policy. *Sciences Po. OFCE Working Paper*, (9), 49.
- Eichengreen, B. (2013). *Does the federal reserve care about the rest of the world?* (Working Paper No. 19405) [Series: Working Paper Series]. National Bureau of Economic Research.
- Ferrara, L., & Teuf, C.-E. (2018). *International environment and US monetary policy: A textual analysis* [Banque de France blog].
- Fischer, S. (2015). *Speech on u.s. inflation developments* [Board of governors of the federal reserve system].
- Fligstein, N., Brundage, J. S., & Schultz, M. (2014). *Why the federal reserve failed to see the financial crisis of 2008: The role of “macroeconomics” as a sense making and cultural frame* (Institute for Research on Labor and Employment, Working Paper Series). Institute of Industrial Relations, UC Berkeley.
- Gardner, B., Scotti, C., & Vega, C. (2023). Words Speak as Loudly as Actions: Central Bank Communication and the Response of Equity Prices to Macroeconomic Announcements. *Journal of Econometrics*, forthcoming.
- Greenspan, A. (1998). Question: Is there a new economy? *Remarks at Haas Annual Business Faculty Research Dialogue, University of California, Berkeley, September 4.*
- Handlan, A. (2022). Text shocks and monetary surprises: *mimeo*.
- Hansen, S., & McMahon, M. (2016). Shocking language: Understanding the macroeconomic effects of central bank communication. *Journal of International Economics*, 99, S114–S133.
- Hubert, P., & Labondance, F. (2021). The signaling effects of central bank tone. *European Economic Review*, 133, 103684.
- Ihrig, J., & Wolla, S. A. (2020). The feds new monetary policy tools. *Page One Economics*, 11.
- Istrefi, K., Odendahl, F., & Sestieri, G. (2021). *Fed communication on financial stability concerns and monetary policy decisions: Revelations from speeches.* (Working Paper No. 2110). Banco de España.
- Laubach, T., & Williams, J. C. (2003). Measuring the natural rate of interest. *The Review of Economics and Statistics*, 85(4), 1063–1070.
- Loughran, T., & McDonald, B. (2011). When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The Journal of Finance*, 66(1), 35–65.
- Luhn, H. P. (1958). The automatic creation of literature abstracts. *IBM Journal of Research and Development*, 2(2), 159–165.
- Mazis, P., & Tsekrekos, A. (2017). Latent semantic analysis of the fomc statements. *Review of Accounting and Finance*, 16(2), 179–217.
- Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient estimation of word representations in vector space. *International Conference on Learning Representations*.

- 
- Miranda-Agrippino, S., & Rey, H. (2020). U.s. monetary policy and the global financial cycle. *The Review of Economic Studies*, 87(6), 2754–2776.
- Nechio, F., & Wilson, D. J. (2016). How Important Is Information from FOMC Minutes? *FRBSF Economic Letter*.
- Obstfeld, M. (2020). Global dimensions of u.s. monetary policy. *International Journal of Central Banking*, 16(1), 73–132.
- Ochs, A. C. R. (2021). *A new monetary policy shock with text analysis* (Working Paper). Faculty of Economics, University of Cambridge.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(85), 2825–2830.
- Porter, M.-F. (2001). *Snowball: A language for stemming algorithms*.
- Rehurek, R., & Sojka, P. (2011). Gensim–python framework for vector space modelling. *NLP Centre, Faculty of Informatics, Masaryk University, Brno, Czech Republic*, 3(2).
- Renault, T. (2017). Intraday online investor sentiment and return patterns in the U.S. stock market. *Journal of Banking & Finance*, 84(100), 25–40.
- Rohlf, C., Chakraborty, S., & Subramanian, L. (2016). The effects of the content of FOMC communications on US treasury rates. *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, 2096–2102.
- Romer, C., & Romer, D. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4), 1055–1084.
- Sachs, J. D. (1985). The dollar and the policy mix: 1985. *Brookings Papers on Economic Activity*, (1), 117–197.
- Shapiro, A., & Wilson, D. (2021). Taking the fed at its word: A new approach to estimating central bank objectives using text analytics. *The Review of Economic Studies*, 89(5), 2768–2805.
- Sharpe, S. A., Sinha, N. R., & Hollrah, C. A. (2022). The power of narrative sentiment in economic forecasts. *International Journal of Forecasting*.
- Shibamoto, M. (2016). *Empirical Assessment of the Impact of Monetary Policy Communication on the Financial Market* (Discussion Paper Series DP2016-19). Research Institute for Economics & Business Administration, Kobe University.
- Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application in retrieval. *Journal of Documentation*, 28(1), 11–21.
- Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.
- Taylor, J. B. (1999). A historical analysis of monetary policy rules. In *Monetary policy rules* (pp. 319–348). University of Chicago Press.
- TerEllen, S., Larsen, V. H., & Thorsrud, L. A. (2022). Narrative Monetary Policy Surprises and the Media. *Journal of Money, Credit and Banking*, 54(5), 1525–1549.

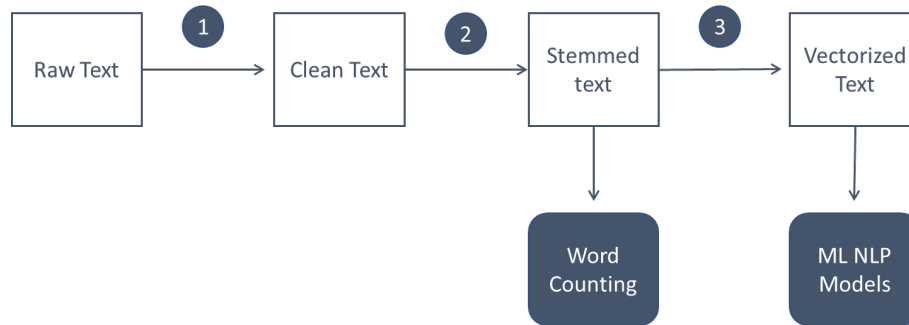
- 
- Thorsrud, L. A. (2020). Words are the new numbers: A newsy coincident index of the business cycle. *Journal of Business & Economic Statistics*, 38(2), 393–409.
- Wu, J. C., & Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2), 253–291.



---

# Appendices

## A. Preprocessing



**Figure 6:** Preprocessing of our text data

- (1) *Cleaning.* Stopwords and numbers removal
- (2) *Stemming.* The text is ready to be used as input for Word Counting models.
- (3) *Vectorization.* The text is ready to be used as input for Machine Learning models.

## B. Selecting international-related phrases

1. We sort all words, doubles (sets of two words) and triples (three words) in the documents by their number of occurrences.
2. We only keep those related to international economic matters.
3. We limit ourselves to phrases above a frequency threshold (above 460 occurrences for words, 230 for doubles, 115 for triples).
4. For phrases contained in or containing other phrases, we select either the containing phrase or the contained: *foreign economies*, for example, is contained in *advanced foreign economies* and contains *foreign*, and we settle on including *foreign economies*.

## C. Vectorization Methods

In this section, we give more information on the workings of the two vectorization methods we use, TF-IDF and Word2Vec. Both methods are implemented in standard python libraries <sup>15</sup>.

### C.1. TF-IDF

TF-IDF vectorization based on corpus  $D$  transforms a sentence  $d$  into a vector of size  $n$  (the number of different words in the corpus) by including for each word  $t$  in the corpus a measure of its importance in the sentence: its  $tfidf(t, d, D)$ . An example of this process can be found in Table ??.

$tfidf(t, d, D)$  is the combination of two measures: Term Frequency, put forward by Luhn, 1958 and Inverse Document Frequency, first devised by Sparck Jones, 1972. For a term  $t$  in a document  $d$ , in a corpus  $D$ ,  $tfidf(t, d, D) = tf(t, d) \times idf(t, D)$ .

- $tf(t, d) = \frac{f_{t,d}}{\sum_{t' \in d} f_{t',d}}$  is the relative frequency of term  $t$  within document  $d$ , where  $f_{t,d}$  is the raw count of term  $t$  in document  $d$ .
- $idf(t, D) = \log \frac{|D|}{|\{d \in D : t \in d\}|}$  is the inverse document frequency of term  $t$  in corpus  $D$ , where  $|D|$  is the total number of documents in the corpus<sup>16</sup>.

As a consequence, for a given word, high weight in TF-IDF is reached by a high term frequency (in the given document) and a low document frequency of the term in the whole collection of documents; TF-IDF vectorization tends to filter out common terms.

Sentence 1: Inflation today is very high and lasting.

Sentence 2: Inflation today is not high and is temporary.

Sentence 3: Inflation today is entrenched and dangerous.

Term	Count			TF			IDF			TF-IDF		
	1	2	3	1	2	3	1	2	3	1	2	3
Inflation	1	1	1	1/7	1/8	1/6	0	0	0	0	0	0
today	1	1	1	1/7	1/8	1/6	0	0	0	0	0	0
is	1	2	1	1/7	1/4	1/6	0	0	0	0	0	0
very	1	0	0	1/7	0	0	0.477	0.477	0.477	<b>0.068</b>	0	0
high	1	1	0	1/7	1/8	0	0.176	0.176	0.176	<b>0.025</b>	<b>0.022</b>	0
and	1	1	1	1/7	1/8	1/6	0	0	0	0	0	0
lasting	1	0	0	1/7	0	0	0.477	0.477	0.477	<b>0.068</b>	0	0
not	0	1	0	0	1/8	0	0.477	0.477	0.477	0	<b>0.060</b>	0
temporary	0	1	0	0	1/8	0	0.477	0.477	0.477	0	<b>0.060</b>	0
entrenched	0	0	1	0	0	1/6	0.477	0.477	0.477	0	0	<b>0.080</b>
dangerous	0	0	1	0	0	1/6	0.477	0.477	0.477	0	0	<b>0.080</b>

Table 7: TF-IDF Vectorization example

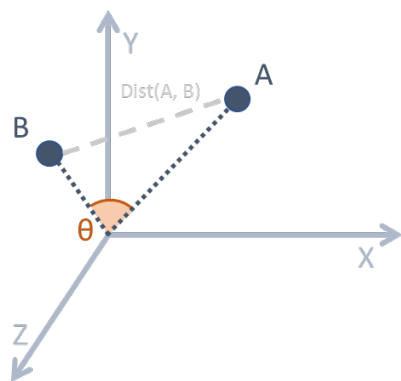
<sup>15</sup>We use the sklearn package of Python, developed by Pedregosa et al., 2011 for TF-IDF and the gensim package of Python, developed by Rehkurek and Sojka, 2011 for Word2Vec.

<sup>16</sup>The denominator (number of documents where the term  $t$  appears) is commonly adjusted to  $1 + |\{d \in D : t \in d\}|$  to avoid division by 0 if the word is not in the original corpus.

## C.2. Word2Vec

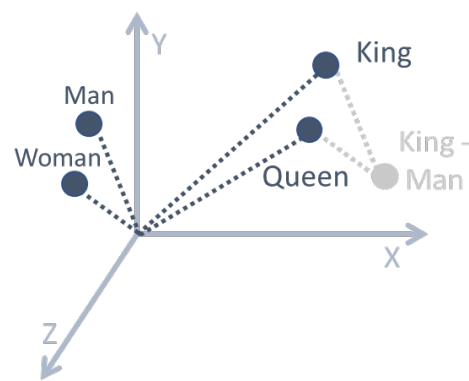
Word2Vec, published in [Mikolov et al., 2013](#) is a word embedding method: it transforms words into vectors by using a neural network to infer word meanings from a corpus of text. The method generates a vector space (generally of several hundreds dimensions), where, ideally, words representations (*embeddings*) are positioned to be close in the space (in terms of cosine similarity) if they are close in meaning.

In the vector space, it is possible to infer the semantic proximity of two words from the distance between their representations in the vector space, usually determined with cosine similarity, i.e. the cosine of the angle between the two vectors. The vector space also means that the representation can subtract or add vectors to each other: in that representation,  $king - man + woman = queen$ .



**Figure 7:** Distance between two vectors in Word2Vec representation: an example

Here,  $\cos(\theta)$  is the distance between A and B.



**Figure 8:** Vector composition in Word2Vec representation: an example

The Word2Vec technique can use one either continuous bag-of-words (CBOW) or continuous skip-gram to produce the representation of a word in the vector space. Both methods use a 3-layer neural network, and consider both individual words and the context words surrounding them. They differ in that in the CBOW method the model predicts the current word from its surrounding words (without taking into account the order of the context words, hence *bag-of-words*) while the continuous skip-gram uses the current word to predict its surrounding context words (and weights context words differently based on their place in the text).

Words which are semantically similar (and should be used in similar contexts), will usually share a large part of their context words and have similar influences on the relative word probabilities in the model, meaning that they will therefore be positioned in similar embeddings. That explains that semantically similar words are represented by similar vectors.

---

## D. Supervised Machine Learning algorithms

In this section, we give an overview <sup>17</sup> of the theory behind the machine learning models we used in Section 3.2.2, namely Naive Bayes Classifier and Random Forest Classifier.

Both models are supervised learning, which means that the target  $Y$  is given. If we represent the model by  $h(\cdot)$ , and assume that the objective function to minimize is the Mean Squared Error (MSE), training the model consists in determining the parameters  $\beta$  minimizing the error.

$$\begin{aligned} \min_{\beta} \quad & \text{ERR}(X, Y, \beta) \\ \text{(MSE)} \quad & = \frac{1}{m} \sum_{i=1}^m (h(x_i, \beta) - y_i)^2 \end{aligned}$$

In the case of a classification algorithm,  $Y$  is a vector of 0s and 1s, and the function  $h(\cdot)$  returns 0 or 1.

### D.1. Naive Bayes

The Naive Bayes Classifier, often taken as a baseline to evaluate more complex methods (Chakraborty and Joseph, 2017), is based on applying the Bayes' rule of conditional probability to determine the probability that observation  $c_i$  is a part of class  $c_i$ . In computing the probability, the assumption that all features are independent from each other is made, hence the *Naive* denomination. In a situation with  $C$  classes (in our problem,  $C = 2$ ),  $c_1, \dots, c_C$ , the classifier is based on the maximising-a-posteriori (MAP) decision rule :

$$\begin{aligned} h(x_i) &= \underset{c}{\operatorname{argmax}} P(c|x_i) = \underset{c}{\operatorname{argmax}} \left( \frac{P(x_i|c)P(c)}{P(x_i)} \right) \\ \text{(Independence)} &= \underset{c}{\operatorname{argmax}} P(c) \prod_{j=1}^n P(x_{i,j}|c) \end{aligned}$$

$P(c_i)$  is the probability of being assigned to a class (and therefore is constant).

The  $P(x_{i,j}|c)$  are the class-conditional likelihood factors for  $x_i$  being labeled  $c$  according to feature  $j$  <sup>18</sup>.

### D.2. Random Forest

The random forest classifier operates by constructing multiple decision trees and aggregating their output.

The idea behind the working of a decision tree is to divide the dataset  $X$  based on the  $x_{i,j}$  features of the  $x_i$  to minimize the entropy  $H(Y|X)$  within areas of  $Y$  defined by the tree, which acts as the objective function in classification problems. This is done through successive identifications

---

<sup>17</sup>Among multiple others, a more complete explanation of the working of these models can be found in Chakraborty and Joseph, 2017, Section 3.

of the feature  $x$  of  $X$  which, when used as the split criterion, leads to the highest information gain  $I(Y|x)$ . The decision tree model consists of the sequence of splitting rules and the final value assignment defined.

$$I(Y|x) = H(Y|X) - \sum_v \frac{|X_v|}{|m|} H(Y|X_v)$$

$$H(Y|X) = - \sum_{c=1}^C p(Y = c|X) \log(p(Y = c|X))$$

$p(Y = c|X)$  is the relative frequency of class  $c$  observations in  $X$ .

While they can be very powerful and can handle complex relations within data, on top of having the advantage of being relatively interpretable, decision trees can be prone to severe over-fitting. Indeed, the learning method can lead the model to learn very specific patterns of the training dataset, which leads to high variance in out-of-sample testing, and therefore large errors.

Random forest models, developed to alleviate this issue, are based on separately training a number of roughly independent tree models from randomly selected subsamples of  $X$ , and make the classification decision based on majority voting of all different trees. Figure 9 illustrates this process.

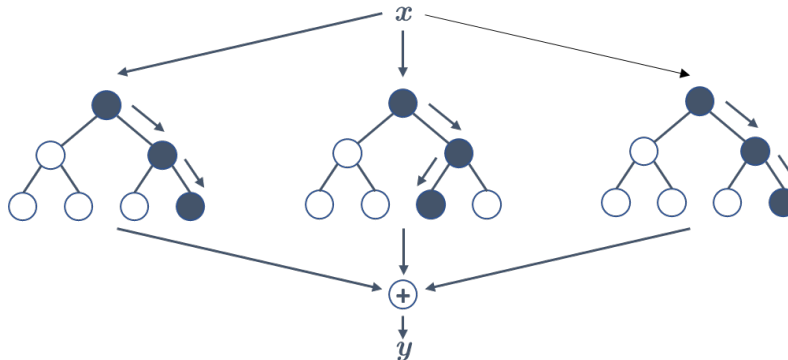


Figure 9: Random Forest

Even though trees in the forest are not independent from each other (as they are constructed from subsets of the same data), the variance of the final prediction is expected to decline compared to a single decision tree, therefore making random forest a reliable choice for classification. While they are less interpretable than decision trees<sup>19</sup>, as a consequence of the randomness inherent to the construction of the model, random forests can be seen as a manifestation of the *wisdom of the crowd* phenomenon, putting crowd intelligence into artificial intelligence.

<sup>19</sup>It can be possible, by studying which branch splittings are performed by the trees, to extract information about the most important features, and even economic rationale (Chakraborty and Joseph, 2017).

---

## E. Counts by phrase in the international dictionary

	Number of Occurrences
dollar	1415
currency	932
foreign exchange	488
foreign currencies	355
euro	304
depreciation	206
pound	36
renminbi	28
sterling	17
ruble	3
imports	1782
exports	1207
international trade	179
external sector	43
global	693
international	459
abroad	377
japan	363
foreign economies	314
china	278
emerging market	262
europa	201
EME	188
asia	147
world	111
russia	54
advanced economies	43
india	17

**Table 8:** *Counts by phrase.*

## F. Dictionary count index with variations in the dictionary

To ensure that our results are robust to small changes in the dictionary, we study whether the random exclusion of 3 of the 28 phrases in the dictionary yields different indexes.

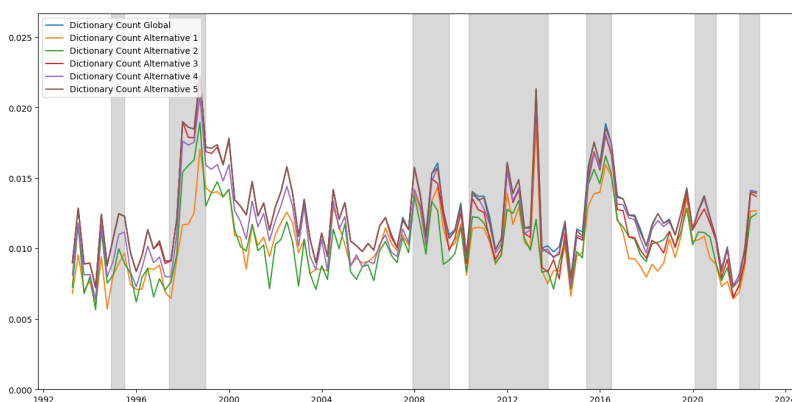
Missing Phrases	
Alternative 1	dollar, asia, EME
Alternative 2	abroad, foreign currencies, currency
Alternative 3	sterling, emerging market, EME
Alternative 4	sterling, foreign exchange, ruble
Alternative 5	yuan, india, advanced economies

**Table 9:** Missing phrases for each of our alternative dictionaries

Table 10 shows the Pearson correlation between indexes obtained with the reference dictionary and the 5 alternatives. What is shown is that all correlations are high (in the 90s), and even extremely high (above 98 %) for alternatives 3 to 5. Alternative 1 has the lowest correlation to other indexes, but this stems from the fact that *dollar*, the phrase with the second-highest count (Table 8), has been excluded, as well as *Asia*, which explains why this index is not able to catch the start of the late 90s Asian crisis while the other alternatives manage to do it, as shown in Figure 10.

	Basis	1	2	3	4	5
Basis	1.000000	0.922090	0.934825	0.982972	0.986841	0.999289
1	0.922090	1.000000	0.860623	0.911276	0.926918	0.919711
2	0.934825	0.860623	1.000000	0.904087	0.938426	0.931466
3	0.982972	0.911276	0.904087	1.000000	0.952997	0.986321
4	0.986841	0.926918	0.938426	0.952997	1.000000	0.983821
5	0.999289	0.919711	0.931466	0.986321	0.983821	1.000000

**Table 10:** Pearson correlation of headline dictionary count (Basis) index and its 5 alternatives



**Figure 10:** Headline dictionary count index and its alternatives

We conclude from Table 10 and Figure 10 that our dictionary count index is robust to small changes in the dictionary, even though more significant changes (exclusion of words with higher counts) can lead to slightly more different results.

## G. Detailed results of classification models

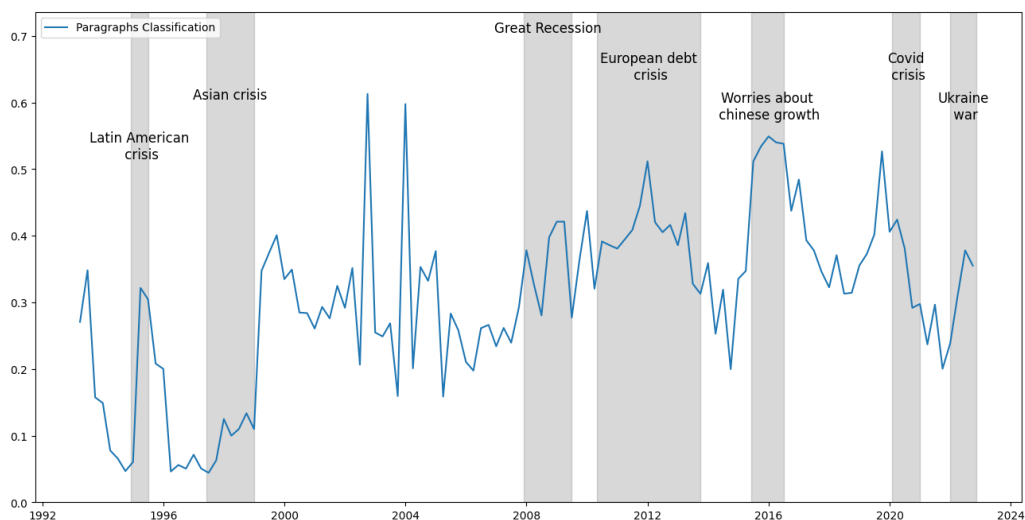
Vectorization	Algorithm	Class	Precision	Recall	F1-Score
TF-IDF	Dictionary	0.0	0.821622	0.987013	0.896755
		1.0	0.976744	0.717949	0.827586
		Macro Average	0.899183	0.852481	0.862171
	Logistic Regression	0.0	0.874251	0.948052	0.909657
		1.0	0.923077	0.820513	0.868778
		Macro Average	0.898664	0.884282	0.889218
	Naive Bayes	0.0	0.972973	0.701299	0.815094
		1.0	0.712500	0.974359	0.823105
		Macro Average	0.842736	0.837829	0.819100
Random Forest	0.0	0.886905	0.967532	0.925466	
	1.0	0.951456	0.837607	0.890909	
	Macro Average	0.919181	0.902570	0.908187	
Word2Vec	Logistic Regression	0.0	0.736842	0.636364	0.682927
		1.0	0.594203	0.700855	0.643137
		Macro Average	0.665523	0.668609	0.663032
	Naive Bayes	0.0	0.764706	0.253247	0.380488
		1.0	0.477273	0.897436	0.623145
		Macro Average	0.620989	0.575341	0.501817
	Random Forest	0.0	0.661654	0.571429	0.613240
		1.0	0.521739	0.615385	0.564706
		Macro Average	0.591697	0.593407	0.588973

**Table 11:** Classification results: detailed results



## H. Classification of paragraphs

We test the naive dictionary classification method on paragraphs, i.e. we classify paragraphs as positive or negative based on whether or not they contain one of the phrases in the international dictionary described in Table 2. We then compute the ratio of words in paragraphs classified as positive on the total number of words, and obtain the index featured in Figure 11.



**Figure 11:** *Index of Fed attention to international, computed via paragraphs classification.*

A quick comparison with the naive dictionary classification on sentences shows that this classification option is less promising. Indeed, the highs of the index are close to 60%, with a mean value around 30%, which would mean that about half of the minutes conversation is directly related to international matters (something that contradicts basic assumptions about the role of the Fed). This might be due to the size of minutes paragraphs, which can start with discussion of domestic matters and end with an international situation briefing. Therefore, classifying the whole paragraph as related or not to international matters does not make economic sense. We also observe shocks in the value of the index, for example at the start of 1999. These jumps do not seem to be related to international events, but rather to structural breaks in minutes structure and paragraph size (since we use the paragraphs delimitations provided on the Fed website).

## I. Results of Taylor Rule Regression with Shadow Rates

Since the monetary policy of the late 2000s and early 2010s has been marked by rates that stayed for a long time at the Zero Lower Bound, we want to investigate in this section whether it affected the potential significance of the Taylor rule. To offset the potential hurdles posed by the monetary policy of this period and account for the Zero Lower Bound, we test another dependent variable in the Taylor rule by replacing the Fed Funds Rates by the Shadow Rates put forward by [Wu and Xia, 2016](#).

They compute an alternative to the Fed Funds Rates to reflect the stimulus provided by the Fed through quantitative easing in the period when additional stimulus through the conventional channel of monetary policy (lowering interest rates) was not available because of the Zero Lower Bound. The interest rate derived from [Wu and Xia, 2016](#) summarizes monetary policy in the Zero Lower Bound years and enables us to link monetary policy (with all its tools) to the economy.

Dependent variable Model	FFR		Shadow Rates		
	(Cal)	(0)	(4)	(0)	(4)
$r_{t-1}$	0.85	0.915*** (0.024)	0.903*** (0.023)	0.921*** (0.026)	0.916*** (0.025)
$\alpha$	4.0	4.17*** (0.557)	3.529*** (0.459)	4.204*** (0.856)	3.421*** (0.76)
$PCEGap_t$	1.5	1.285 (0.794)	0.922 (0.615)	1.979 (1.286)	1.615 (1.1)
$g_t$	1.0	1.209*** (0.244)	0.79*** (0.206)	1.459*** (0.374)	0.945*** (0.34)
$i_t$			-8.149*** (2.639)		-10.483** (4.449)
Observations		115	115	115	115
Period		1994-2022	1994-2022	1994-2022	1994-2022
AIC		115.637	102.444	193.755	186.188
MSE		16.872	14.784	33.279	30.622
Taylor Version		9	12		

**Table 12:** Results of Taylor Rule Regression with Shadow Rates as the dependent variable

*Note:* The table shows Taylor Rule (Equations 4 & 5) results for Non-Linear Least Squares Regression, with two possible dependent variables (Fed Funds Rates, Shadow Rates) and different inputs as the international attention index  $i_t$ : (0) is no international environment index, (4) the index from random forest classification of sentences. The (Cal) column features the calibrated parameters (from the Atlanta Fed).

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

We conclude from Table 12 that when used to explain variations in the Shadow Rates of [Wu and Xia, 2016](#), our Fed IAI retains significance and the conclusions of 4.2 still hold : attention to the international situation has a negative effect on Shadow Rates. A rise in the Fed's attention to the global economy translates into a more accommodating monetary policy (conventional and unconventional).

## J. Results of Taylor Rule Regression with time-varying $r^*$

In this section, we estimate a different version of the Taylor Rule which differs from Equation 5 in that  $\alpha$  is replaced by a time varying neutral interest rate.

$$r_t = \rho r_{t-1} + (1 - \rho)((r_t^* + \pi_t) + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t) \quad (7)$$

We take as  $r_t^*$  the Laubach-Williams model 1-sided estimate, which is based on [Laubach and Williams, 2003](#) and updated by the Federal Reserve Bank of New York.

Using the same non-linear least square method as in Section 4, we obtain the following results.

Dependent variable Model	FFR	FFR (TVR)	
	(Cal)	(0)	(4)
$r_{t-1}$	0.85	0.913*** (0.027)	0.919*** (0.025)
$\alpha$	4.0		
$PCEGap_t$	1.5	0.567 (0.808)	0.537 (0.82)
$g_t$	1.0	0.828*** (0.191)	0.604*** (0.197)
$i_t$			-8.034** (3.569)
Observations		115	115
Period		1994-2022	1994-2022
AIC		117.029	107.8
MSE		17.377	15.76
Taylor Version		10	13

**Table 13:** Results of Taylor Rule Regression with time-varying  $R^*$

*Note:* The table shows time-varying  $R^*$  Taylor Rule (Equation 7) results for Non-Linear Least Squares Regression, with different inputs as the international attention index  $i_t$ : (0) is no international environment index, (4) the index from random forest classification of sentences. The (Cal) column features the calibrated parameters (from the Atlanta Fed).

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

The results are globally on par with those in Table 5: the coefficient our IAI is negative and significant (although less than in the standard version of the Taylor rule) and the conclusions of Section 4 hold in terms of the effect of international attention. Yet, Table 13 shows that the PCE gap is never statistically significant.

## K. Results of Taylor Rule Regression with unemployment gap

Dependent variable Model	FFR		
	(Cal)	(4)	(4.3)
$r_{t-1}$	0.85	0.899*** (0.026)	0.935*** (0.024)
$\alpha$	4.0	3.592*** (0.478)	2.94*** (0.715)
$PCEGap_t$	1.5	1.142 (1.05)	2.244 (1.724)
$g_t$	1.0	0.75*** (0.231)	
$i_t$		-8.067*** (2.771)	-13.955** (5.656)
$u_t$			0.092 (0.209)
Observations		104	104
Period		1994-2019	1994-2019
AIC		86.072	91.334
MSE		12.413	13.057
Taylor Version		<b>12</b>	<b>15</b>

**Table 14:** Results of Taylor Rule Regression with unemployment gap

*Note:* The table shows Taylor Rule regression (Equations 12 and 15) results for Non-Linear Least Squares Regression, with the same international attention index and two options for the economic slack variable : (4) uses the output gap, (4.3) the unemployment gap (from the Congressional Budget Office). The (Cal) column features the calibrated parameters (from the Atlanta Fed).

\* p<0.1; \*\* p<0.05; \*\*\* p<0.01.

Results in Table 14 show that including the unemployment gap as the slack measure in an augmented Taylor Rule does not lead to improve performances in terms of AIC and MSE. In such a model, the unemployment gap does not seem to be statistically significant.

---

## L. Different versions of Taylor Rule Regression Used

In the equations below:

- $r_t$  is the Fed Funds Rate (or the Shadow Rate if relevant)
- $\pi_t$  is the (core PCE) inflation rate
- $\pi_t^*$  is the inflation target
- $g_t$  is the output gap
- $\rho$  is a smoothing term for the FFR
- $\alpha$  is a constant playing the role of the neutral interest rate
- $R_t^*$  is the time-varying neutral interest rate
- $i_t$  is our International Attention Index (*IAI*)
- $u_t$  is the unemployment gap
- $s_t$  is a variable derived from an equity index
- $v_t$  is a variable derived from a risk index

### L.1. Non-augmented versions

$$r_t = \pi_t + R_t^* + \gamma(\pi_t - \pi_t^*) + \beta g_t \quad (8)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t) \quad (9)$$

$$r_t = \rho r_{t-1} + (1 - \rho)((R_t^* + \pi_t) + \gamma(\pi_t - \pi_t^*) + \beta g_t) \quad (10)$$

### L.2. Augmented versions

$$r_t = \pi_t + R_t^* + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t \quad (11)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t) \quad (12)$$

$$r_t = \rho r_{t-1} + (1 - \rho)((R_t^* + \pi_t) + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t) \quad (13)$$

### L.3. With Unemployment Gap

$$r_t = \pi_t + R_t^* + \gamma(\pi_t - \pi_t^*) + \beta u_t + \delta i_t \quad (14)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta u_t + \delta i_t) \quad (15)$$

$$r_t = \rho r_{t-1} + (1 - \rho)((R_t^* + \pi_t) + \gamma(\pi_t - \pi_t^*) + \beta u_t + \delta i_t) \quad (16)$$

---

#### L.4. With additional control variables

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t + \sigma s_t) \quad (17)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \sigma s_t) \quad (18)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \delta i_t + \mu v_t) \quad (19)$$

$$r_t = \rho r_{t-1} + (1 - \rho)(\alpha + \gamma(\pi_t - \pi_t^*) + \beta g_t + \mu v_t) \quad (20)$$