A. Y. 2023 - 2024

Monetary Words, Market Echoes

The power of ECB wording in shaping financial landscapes

Index

Executive Summary

Monetary policy decisions, while often perceived as detached from daily life, profoundly influence the everyday experiences of individuals, echoing the foundational theories of economists like John Maynard Keynes and Milton Friedman. These policies, primarily executed through adjustments in interest rates and money supply by Central Banks, have significant financial impacts both regionally and globally.

Inspired by a desire to explore the shady relationship between language and monetary policy in modern society, this study focuses on the strategic use of words in central banking communication. The choice of one word over another is far from trivial, carrying significant implications as demonstrated in Blinder et al. (2008) and further explored by Gorodnichenko et al. (2023) in their analysis of non-verbal cues.

Carrying on in their footsteps, the aim of this work is to analyze the correlation between the thematic content of speeches by the European Central Bank Chairpersons (Mr. Draghi and Ms. Lagarde) and their impact on financial markets, using 10Y ECB swap rates, EUR/USD exchange rate, Italian and German 10Y treasury yields as measures of market response in the period from November 2011 to June 2023.

In the first part, the study encompasses a detailed examination of the four financial variables in the specific timeframes analyzed. These choices aimed to capture the immediate market reactions to the ECB's communications, drawing on insights from a window of 2 hours and 30 minutes, starting 20 minutes before the press release (hold at 13:45 CET) and ending 20 minutes after the official press conference (at 15:50 CET).

In the second, third and fourth sections, the study presents the application models, starting from text preprocessing through tokenization and lemmatization till obtaining the recurrent topics through LDA technique. Each topic was then fitted within four linear regression models, in which the inclusion of the governor's dummy expands the scope of the analysis even further.

The principal results indicate that certain topics such as "*Asset Convergence*", "*Financial Conditions*", "*Structural Stability*", and "*Inflation Expectations*" significantly influence the aforementioned financial metrics. For instance, it was found that discussions around financial conditions negatively impacted the 10Y OIS rate but increased the yields on Italian bonds, thus suggesting divergent market interpretations.

The dedicated final chapter will then attempt to provide a more financial interpretation of these phenomena. Similarly, the study revealed that the topic of inflation expectations was linked to a decrease in long-term interest rates but led to an increase in yields in the Italian bond market, denoting, once again, a divergent response from the Italian treasury market.

For each of the models, the inclusion of a dummy variable pertaining to the incumbent governor, either Mr. Mario Draghi or Ms. Christine Lagarde, expanded the analytical possibilities allowing the reader to compare the styles of two leaders. The results here are remarkable. They suggest that the Chairwoman's tenure did not have a pronounced effect on the financial variables, particularly on Italian bond yields, probably due to the established policies of the ECB and broader economic challenges, such as COVID-19, faced during her term, which necessarily required a commitment much more widespread across the entire Eurozone.

Conversely, Mr. Draghi's communication style is associated to lower long-term interest, a slight depreciation of the euro and a decrease in the yields of both Italian and German government bonds. Notably, it is interesting to observe how the mentioned dummy had a 5x more pronounced effect, on average and ceteris paribus, on Italian State bonds compared to their German counterparts, a topic that we will endeavor to explain in the closing chapter.

The paper then concludes with a discussion of the project’s limitations, which are fundamental for further research.

Literature review

The scholarly landscape of monetary policy and central bank communication, while not widely discussed, has seen the significant contribution of a handful of important works. Over the past decade, this niche has increasingly intersected with the wider realm of natural language processing, a field that has gained momentum since the 1980s with the advent of machine learning algorithms, thanks to pioneers such as Prof. Geoffrey Hinton and Prof. Marvin Minsky. This intersection has elevated the importance of the subject, especially in light of its integration with speech recognition and generation technologies and its broader implications in the field of artificial intelligence.

The application of NLP to banking research thus represents a novelty, at least for three reasons.

Firstly, as pointed out by Blinder et al. (2008), the enigmatic nature of central banks prevalent until 1990s filled their operations in mystery limiting scholars' exploration of its verbal communication.

Secondly, the feasibility of correlating linguistic patterns with financial indicators has only recently expanded, fuelled by advancements in machine learning and deep learning techniques (Gorodnichenko et al., 2023). These technological novelties have then allowed for deeper examinations of CBs communication patterns.

Finally, the third point pertains to data availability. To provide a case, the European Central Bank did not produce nor release full transcripts of its conferences (offering only brief summaries) until 2005. This lack of comprehensive data accounts for why the most significant contributions to the literature emerged primarily post-2008, with key publications dating to 2008, 2010, and 2022.

For the purpose of this paper, the three articles by Blinder et al. (2008), Gorodnichenko et al. (2023) and Jansen & de Haan (2010) set the way, each from a single point of view.

Blinder et al. study the change in FED communication over two decades, highlighting the shift from secrecy to transparency (as anticipated above) and pointing out the absence of a consensus on optimal communication strategies. Their seminal work suggests that the final solution is ambiguous: there is not an optimal communication style thus different CBs adopt different approaches, without following any best practice. This, back in 2008, was the first work to highlight the complexity of CBs wordings. Jansen & de Haan's research hence builds on Blinder et al. by developing a "*Wordscores*" methodology on ECB conferences' transcripts. In particular, what they did was to focus on the consistency of the institution's wording, using previous ECB statements as reference texts and comparing these with later statements (2002-2009) to assess any observable pattern in ECB's policy communication. This approach represented a huge methodological advancement, providing a detailed and quantitative analysis of how language in official communications reflects and influences monetary policy over time, an aspect not specifically addressed in Blinder et al.

Finally, the study by Gorodnichenko et al. (2023) introduces a new perspective. Indeed, while previous work had highlighted the role of verbal communication, Gorodnichenko and colleagues have ventured into a relatively new territory, that of non-verbal expressions. As a matter of fact, their research concerns the subtle but impactful issue being the tone of voice analyzed on FED presidents from April 2011 to June 2019. Leveraging deep learning, the authors isolated the emotional nuances of the Fed presidents' responses, and classified these segments as positive, negative or neutral. Then, they assessed the influence on key variables as stock market performance, inflation expectations, and exchange rates[1], like this paper does. One of their notable findings was that a positive shift in tone by 1 standard deviation could lead, on average and ceteris paribus, to a 75-basis point increase in S&P 500 returns (+0.75%). Needless to say, these findings filled a literature gap and demonstrated that not only what is said, but also how it is said, matters.

It is precisely in this variegated but relatively new context that the paper inserts.

[1] The final specification they obtained is the following: Outcomet,t = $\beta 0 + \beta 1 VoiceTone\ t + \beta 2 TextSentiment\ t + \beta 3 FFRShock\ t + \beta 4 FGShock\ t + \beta 5 APShock\ t + \beta 6 ShadowRate\ t + \beta 7\ 1\{NoPressConferencet\} + error\ t$

Variables description

Text data retrieval

This study focuses on the European Central Bank (ECB) and, for the purpose of simplifying analysis and enhancing data collection, specifically examines the mandates of two governors: Mr. Mario Draghi and the current, Ms. Christine Lagarde.

The dataset covers the period from November 2011 to October 2023, with the aim of understanding both periods of relative economic stability and unpredictable events that have permanently marked recent history. These events include the sovereign debt crisis with the risk of default for several European countries, the COVID-19 pandemic and the recent conflicts in Ukraine and Palestine.

The two governors' approaches diverge significantly in response to the challenges they face, as we will see in the final chapters. Mr. Draghi has tried to reassure Europeans in an almost paternalistic manner, taking important decisions such as massive debt purchases of some member states.

Ms. Lagarde, on the other hand, was more interested in expanding the role of the ECB, emphasising sustainability-oriented policies and transparent communication, mostly as a consequence of the macroevents she witnessed.

These shocks, unfolding from 2019 onward, led the Ms. Lagarde to undergo an important shift in approach. Initially characterized by an expansive monetary policy featuring low interest rates, increased Quantitative Easing (QE), and lending programs to support businesses, the policy shifted towards a highly restrictive stance due to wars and rising inflation, resulting in a subsequent increase in interest rates.

For what regards textual data, all transcripts addressing Monetary Policy Decisions from November 2011 onward were retrieved from the ECB Official Database[2].

For the sake of completeness, it is hereby clarified that each monetary policy shift includes 3 steps:

- Decisions on Monetary Policy: Deliberations of the ECB's Governing Council on interest rates and monetary policy instruments during its most recent meeting.
- Monetary Policy Statement and Q&A: An explanation of the rationale behind the latest monetary policy decisions.
- Review of Monetary Policy: An examination of recent developments and policy alternatives, including a summary of the discussions and decisions made.

[2] https://www.ecb.europa.eu/press/govcdec/mopo/html/index.en.html

For our analysis, we focused on the Monetary Policy Statement, excluding the Q&A.
This decision comes from two reasons: first, the statements provide more comprehensive explanations of the reasons behind policy decisions, and second, excluding Q&A helps maintain focus on the real motivations guiding the ECB's choices. As a matter of fact, Q&A sessions often flow into the realm of public interest, diverting attention from the core messages.
To extrapolate the data, we downloaded the transcripts as *txt file*st, so to not present any type of formatting issues. After that, the files were named with acronyms that differentiated by governor, month and year in order to allow R to distinguish and classifying.

Example:
`DRG_Rel_11_11` *for* Draghi – November 2011
`LGD_Rel_19_12` *for* Lagarde – December 2019

The first observation from *Table 1* shows that Mr. Draghi's speeches are shorter and straighter while Ms. Lagarde's are longer and full-bodied. This may confirm what has been expressed earlier in terms of mandates' periods.
If Mr. Draghi tried to be as direct as possible, avoiding long speeches and preferring short and direct phrases (famous is in fact the phrase said on 26 July 2012 "Whatever it takes" in the context of the European sovereign debt crisis), Ms. Lagarde instead relies more on a discursive approach, made up of long and discursive sentences. Her approach is likely intended to provide a broader picture of the shock and as mentioned above, to expand the role of the ECB within a wider European financial and banking system.

GOV	N° SPEECHS	AVG TYPES	AVG TOKENS
DRG	76	515	1463
LGD	32	600	1822

Table 1. Draghi vs Lagarde: speech properties

To provide a clear picture, we believe that these differences mostly depend on two factors.
In the first place, their communication style is quite divergent: Mr. Mario Draghi is known for his concise and straight language, whereas Ms. Christine Lagarde is more inclined to use more articulate language and provide technical details, also due to her purely macroeconomics background at IMF presidency.
The second factor regards the historical period: when Mr. Draghi became the ECB president in 2011, Europe was in the midst of a financial crisis and monetary policy had to be very expansionary to support the economy. In this context, he needed to clearly communicate his strategy and clear up the

road for the markets and the public. His speeches were therefore short and concise, but at the same time clear and understandable.

Ms. Lagarde, on the other hand, took over the ECB presidency in a more stable but entangled context. The covid19 pandemic then brought economic activity to a halt and the Russian invasion of Ukraine created further uncertainties. In this context, Ms. Lagarde needed to provide far more detailed information on ECB decisions, with a focus on macroeconomic backlashes. Her speeches are therefore longer and deeper, although still clear and understandable, as data will show in the final chapter.

Financial Data retrieval

Temporal Window choice

Before identifying the reasons and origin of the included financial data, the reader must be clear about the timeline that ECB has set for its public relations, since it includes several critical stages.

These events traditionally occur on the second Thursday of each month and the sequence begins with the Policy Decision, announced at 13:45 CET in the so-called “Pre-Release window”. Following this moment, the “Post-Release window” occurs at 14:00 CET. At 14:30 CET, the ECB President delivers the Introductory Statement in a “Pre-Conference window”, explaining the reasoning behind the monetary policy decision. This is a key moment for clarifying the ECB's stance and intentions. Afterward, a press conference takes place, including a Question & Answer (Q&A) session starting at around 15:30 CET, offering further elaboration on the ECB's policy trajectory.

These communications are pivotal, as they can significantly influence market conditions, marked by changes due to Policy Decisions and changes due to the Press-Conference. The timeline concludes with the Post-conference Window at 15:50 CET.

In conclusion, our group had three distinct time windows to investigate rate fluctuations: the pre-release window (13:25-13:35), the post-release window (14:00-14:10), and the post-conference one (15:40-15:50 CET), as the reader can observe in the illustration below. After careful consideration, we decided to focus on the rate representing the gap between the pre-release window and the post-conference window. Algebrically, the rate we chose to analyze is the change in the median quote from the window 13:25-13:35 before the press release to the median quote in the window 15:40-15:50 CET after the press conference.

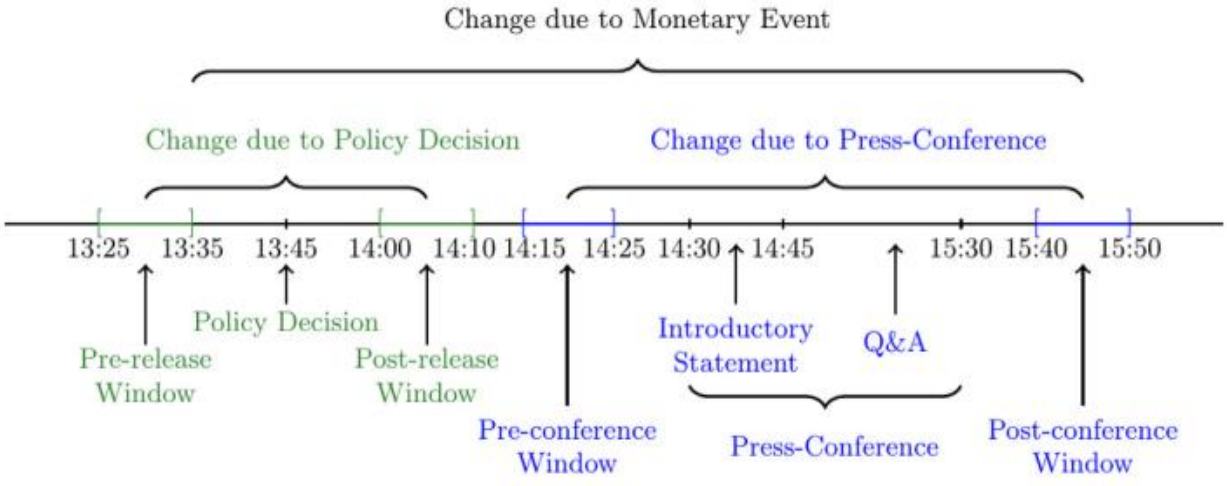


Illustration. Press release and Press conference - Temporal windows

It's important to note that as of 21 July 2022, following the change in the timing of the press release and the press conference to 14:15 and 14:45, respectively, the corresponding windows for our analysis shifted to 13:55-14:05 before the press release and 15:55-16:05 after the press conference. However, this change does not impact our raw financial data observations.
This specific period captures the market's immediate reaction to the policy decision and the subsequent press conference, which are likely to have a pronounced impact on the 10-year IOS rate, EUR/USD and Italian & German 10-year treasury yields.

Focus on IOS 10Y ECB Rate

Our first empirical analysis focuses on the 10Y IOS rate, a choice inspired by its strong correlation with the ECB's long-term monetary objectives. Indeed, as noted in Gürkaynak, Sack, and Swanson (2005), long-term rates change like the 10Y IOS are instrumental in understanding the impact of monetary policy. As the ECB's policies are aimed at maintaining price stability, the 10Y IOS rate (where IOS stands for *Overnight Indexed Swap*, namely the interest rate charged on overnight loans between banks) thus serves as an essential indicator of the market's response to ECB communications. We retrieved data from November 2011 to June 2023 (for homogeneity reasons the same time period was chosen for the other financial variables).
The rate selected, based upon the aforementioned intervals, is algebraically represented by the variation in the median quote from the pre-release window of 13:25 - 13:35 CET to the median quote in the post-conference window of 15:40 - 15:50 CET. Intriguingly, this rate inherently accounts for the change before and after the press release, thereby obviating the need for any preliminary deductions of pre-release rates that the initial model may have anticipated.

These rate adjustments by the ECB bear profound implications, as explored in the broader context of economic literature. Higher rates can lead to increased borrowing costs, affecting loans and mortgages, while possibly bolstering savings at the expense of certain investments. Conversely, lower rates might stimulate spending and investment but adversely impact savings returns, as noted in studies examining the multifaceted effects of monetary policy on economies.
Additionally, these rate changes significantly influence the value of domestic currency against foreign currencies (as we will see for EUR/USD change variable) and profoundly impact the national debt costs of European states (as we will discuss for 10 years German and Italian bond yields rate change), a theme extensively discussed in the context of economic policy impacts. For the reader's better understanding, especially in view of the inferential analyses in the following chapters, the trend of the rate over the mentioned period is shown in *Chart 1*.

Chart 1. IOS 10Y trend, Nov 2011 - Current

Focus on EUR/USD

An upward adjustment in interest rates by the European Central Bank (ECB) can significantly influence the EUR/USD exchange rate. This is why we could not exclude such an analysis for the exchange rate as well. Indeed, it would be reductive and biased to discuss the impact of European financial markets without considering the delicate relationship with the strongest currency, the Dollar. For instance, an increase in exchange rates may attract investments into the Eurozone, strengthening the Euro against the USD, while, for a European citizen, this could mean greater affordability when purchasing goods or services priced in dollars. The same upward movement may lead an appreciated Euro to negatively affect exports, as EU goods become less competitive in the US market.

These dynamics are crucial in analyzing the impact of ECB announcements on financial markets, as they alter the cost structure of international transactions and the expectations of economic actors. We will see more in the section related to the results' interpretations. To facilitate the reader's comprehension, especially in preparation for the inferential analyses in the subsequent chapters, *Chart 2* illustrates the trend of the rate over the mentioned period.

Chart 2. EUR-USD trend, Nov 2011 - Current

Focus on 10 years IT & DE Treasury Bond

As exchange rates, yields on sovereign debt could not miss out. These rates differ from the IOS 10Y as those are set by the ECB, whereas sovereign rates are freely determined by market actors based on the underlying national systems' solidity and perceived stability.

Generally, the relationship between the two is quite clear, and it is positive: when the ECB sets higher interest rates, it often leads to increased confidence in the economy, boosting the perceived stability, and improved economic outlook.

The reason for this choice can be found both in the fact that the relationship between the two sovereign debts is easily measurable through the so-called spread, which reflects the additional risk perception on Italian titles over German ones, and because, in light of the historical period under consideration, it is particularly interesting to observe how Europe's first and third largest economies have fought back.

One factor that heavily impacted spread, for instance, is Governor Mario Draghi itself, who had a far more significant effect on Italian public debt than on German homologue instruments. His renowned *"whatever it takes*" on July 26, 2012, is certainly a clear example of how his actions have been associated with a narrowing of the spread, indicating an improvement in investor confidence towards the *Belpaese*.

In the following pages, we will empirically note that there is indeed a differing impact. For the reader's better understanding, especially in view of the inferential analyses in the following chapters, the trend of the rate over the mentioned period is shown in *Charts 3,4.*

Chart 3.Germany 10Y Bund yields' trend, Nov 2011 - Current

Chart 4. Italy 10Y BTP yields' trend, Nov 2011 - Current

Model Building

Preprocessing

The initial phase of our statistical analysis involved a detailed preprocessing of data, a crucial step in ensuring the data's suitability for comprehensive analysis. Our selection process focused on 108 speeches and press conferences delivered by Mr. Mario Draghi and Ms. Christine Lagarde, covering a period from November 2011 to October 2023. This timeframe is significant, encompassing several key historical events such as the Eurozone debt crisis and the economic upheavals of the covid19 pandemic, which our model is adept at identifying and examining.

Starting from the beginning, in the data preparation phase, we used the readtext() function to import the data. Next, we exploited the file names to create the essential document variables, including the Chairperson, year and month of the conference. This was achieved by strategically sectioning the file names at the underscores, a methodical approach that aligns with standard textual analysis practices. Such methodical division will be highly useful for the following comparisons by time and by governor.

```
ECBspeeches <- readtext("Speeches", docvarsfrom = "filenames",
                        docvarnames = c("gov", "type", "year", "month"),
                        dvsep = "_",
                        encoding = "UTF-8")
```

After data have been imported, we proceeded by creating the corpus which is based on the document variables created earlier.

```
ECBcorpus <- corpus(ECBspeeches)
summary(ECBcorpus)
```

At this point in our research, we embarked on the primary phase of preprocessing: tokenization. This step involves splitting paragraphs and sentences into smaller, more analyzable units, that allow us to go way deeper in the analysis.

In the following code, we have meticulously prepared the text by removing extraneous elements such as punctuation, numbers, symbols and URLs, which are clearly unnecessary and misleading for our semantic analysis. Furthermore, to resolve the complexity of spelled words, we split them into separate tokens, trying to improve the granularity of the analysis. Finally, the tokenization was not

limited to dissecting the text into individual words, but also included the incorporation of critical document variables, such as dates and speakers' names.

```
tokens = quanteda::tokens(ECBcorpus,
                          remove_punct = TRUE,
                          remove_numbers = TRUE,
                          remove_symbols = TRUE,
                          remove_url = TRUE,
                          split_hyphens = TRUE,
                          include_docvars = TRUE)
```

Tokenization, while being an essential component of text analysis, is subject to certain limitations mainly associated with the nature of the language and its writing system, as outlined by Kapur (2020). Some of these, prevalent in other languages, such as Chinese or Japanese, are fortunately not present in our model. The inherent structure of English, characterized by a clear and spatial separation of words, facilitates a much more direct and simple tokenization process.

Pre-processing then continued with the conversion of all tokens to lower case, performed primarily to reduce potential discrepancies.

Subsequently, we proceeded to remove the stop-words. Common words such as determinative and indefinite articles, although frequently used in speech, do not contribute significantly to our analytical goals. Following Blinder et al. (2008), we paid attention that too much removal did not obscure the meaning of the text.

We therefore decided to selectively remove terms irrelevant to our search purpose, including mentions of months and references to percentages such as 'for' and 'cent'. This selective approach to token removal was carefully designed to focus the analysis on the most relevant and impactful terms, thus facilitating a more targeted and in-depth exploration of the ECB's communication patterns. The implicit code was as follows:

```
toks_clean = tokens_remove(toks_clean, stopwords(), valuetype = "fixed")
toks_clean = tokens_remove(toks_clean, pattern = c("january", "february", "march",
                                                   "april", "may", "june", "july", "august",
                                                   "september", "october","november",
                                                   "december","monetary", "policy",
                                                   "per", "cent"),
                           valuetype = "fixed")
```

At this point, we were faced with a critical decision: to choose between stemming or lemmatisation for further text refinement. We opted for lemmatisation, as it is a more comprehensive approach to text processing. Unlike stemming, lemmatisation involves an in-depth lexical and morphological

analysis of words. This process aims to remove inflections from words, returning them to their basic or dictionary form - the lemma - as described by Kapur (2020). This choice was dictated by the need for precision in our analysis; lemmatisation ensures that the various inflected forms of a word are treated uniformly, thus simplifying textual analysis and improving the accuracy of our models.

```
lemmas = tokens_replace(toks_clean,
                        pattern = lexicon::hash_lemmas$token,
                        replacement = lexicon::hash_lemmas$lemma)
```

We then proceeded with the trimming step, performed to refine the dataset by removing words that were too common and too uncommon. Using the dfm.trim() function, we proceeded with "document frequency trimming," in which we defined the frequency of occurrence of words as the basis for exclusion. However, we approached this step with discernment, recognizing that some frequent terms, such as *"inflation"*, had substantial relevance to our analysis.

Rather than omitting these terms because of their high frequency in discourse, thus treating them as outliers, we opted for a different approach, preserving and recontextualizing these words with bigrams. Through this, the word *"inflation"* was, for example, linked to related terms such as *"rate"*, *"term"*, *"expectation"*, *and "outlook"*, which naturally gave greater meaning to the word inflation itself. This strategy, as illustrated in our analytical graphs, thus maximized the significance associated to high-frequency words.

```
bigrams = tokens_ngrams(lemmas, n = 2)
```

Our approach to processing, particularly the use of bigrams, was instrumental in ensuring that frequently used words and terms in the speeches, like *"inflation"* and *"financial*", were not lost in the trimming process. By bundling these words with others, we effectively managed their "absolute frequency," as clearly demonstrated in the comparison between *Appendix chart 1* (the as-is situation) and *Appendix chart 2* (the post-bigrams situation), that, for graphic reasons, have been reported in the appendix.

We then chose to exclude words that appeared less than 10% and more than 90%. This specific threshold was determined after several trials, during which we observed that a 10-80% range, while simplifying the texts, led to unintentional loss of key terms. For example, narrower selection parameters would have omitted significant words such as "finance" from the analysis, causing a serious error. Therefore, the 10-90% trimming proportion was optimal to balance the complexity of the data with the preservation of essential terms.

```
ecb_dfm = dfm(bigrams)
ECB = dfm_trim(ecb_dfm,
         min_docfreq = 0.10,max_docfreq = 0.90,
         docfreq_type = "prop")
```

At this point we used the DFM, the Document-Feature Matrix, to create a quantitative object to be used to proceed with the analysis. The results of this action, with the graph of frequencies by year and by governor, are presented in Chart 5 and Chart 3 in the Appendix.

As a result of the cleaning actions implemented on the dataset, the graph in Appendix 3 provides a lucid representation of the predominant themes in ECB speeches over time. This illustration clearly shows how the most frequently used terms in the speeches align with the main economic concerns of the respective periods. For example, the term "*financial markets*" featured prominently in the discourse following the 2011 financial market crisis, and, similarly, the term *"asset purchases*" emerged as a significant theme in 2017-2018, while the term *"energy prices*" gained prominence in the years 2022-2023.

In the following chart, we present a comparative chart of the speeches by Mr. Mario Draghi and Ms. Christine Lagarde. This clearly sheds light on the evolving priorities and concerns within the ECB's

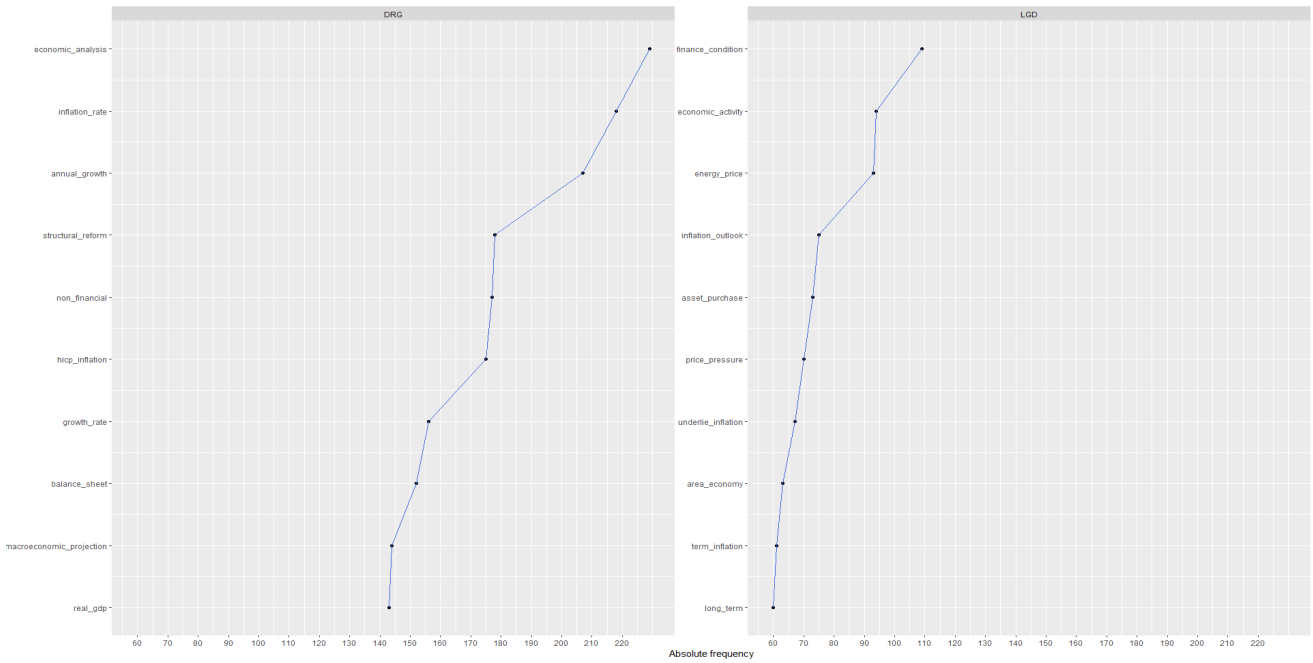


Chart 5. Dfm distribution, by Governor

communication under different leaderships, confirming a blurred understanding of how central bank communication strategies and focus areas have adapted over time to address varying economic challenges and contexts.

Text classification

We proceed to examine the results obtained. In particular, we are interested in gaining a broad, non-detailed understanding of whether and how differences in communication exist both between the two presidents and between different years, indicative of profoundly distinct historical periods. Using a dictionary-based approach, we classify (and categorize) speeches according to the frequency of the words they contain. As dictionary, we choose the "*LexiCoder Policy Agenda*" dictionary provided by the Comparative Agendas Project (CAP), a research initiative that focuses on the systematic analysis of government agendas across time and countries. The decision to use this dictionary resides in the fact that this one allowed us to reliably compare the issue horizontally, across different periods and, in our case, across different ECB governors. The Dictionary was aimed at capturing topics in news content, legislative debates, and policy documents[3].

To carry out this specific analysis, a new *dfm* was created to include the same results obtained in *ECB* but as unigrams rather than bigrams, to work better with this dictionary. However, to properly analyze the textual data, tokens with a frequency > 80% were excluded (not 90% as for the bigram cases), thereby eliminating all those elements that are redundant and not useful for the analysis. The analysis then focused only on a subset of topics from the entire dictionary, selecting those that are most closely related thematically to speeches belonging to the realm of economic policy discourse.

```
topics = c ("year", "macroeconomics", "civil_rights", "healthcare",
            "labour", "immigration", "education", "environment", "energy",
            "transportation", "social_welfare", "finance", "defence","
             government_ops", "culture","intergovernmental", "religion")
```

The output displayed in *Chart 6* proves to be highly interesting and representative of the connection between the speeches delivered by the presidents of the ECB and global geopolitical events.

Attempting a comprehensive interpretation, it can be asserted, firstly, that the theme of *"macroeconomics"* has progressively assumed a marginal role in economic policy releases over the years, particularly during Ms. Lagarde's presidency. Just as the Covid-19 pandemic has had a profound and devastating impact on our lives, the voice of economic policy has followed this trend, with the *"Healthcare"* topic almost monopolizing the discourse in the biennium 2020-2021.

Following the Russo-Ukrainian conflict and later the Israeli-Palestinian conflict, the ECB President dedicated a significant portion of her speeches to the topic of *"Defense*", a theme that had occupied

[3] https://www.snsoroka.com/data-lexicoder

an extremely limited space in the preceding 10 years. *"Labour"* is the only topic that somehow manages to carve out a space despite the various vicissitudes.

On the other hand, the increasing frequency of the *"Macroeconomics"* topic between 2011 and 2013 in European Central Bank communication reflects this period of intense institutional effort in preserving the stability of the currency throughout the European sovereign debt crisis (2010-2013). As a matter of fact, this tumultuous period saw Mario Draghi's famous *"whatever it takes"* moment in July 2012.

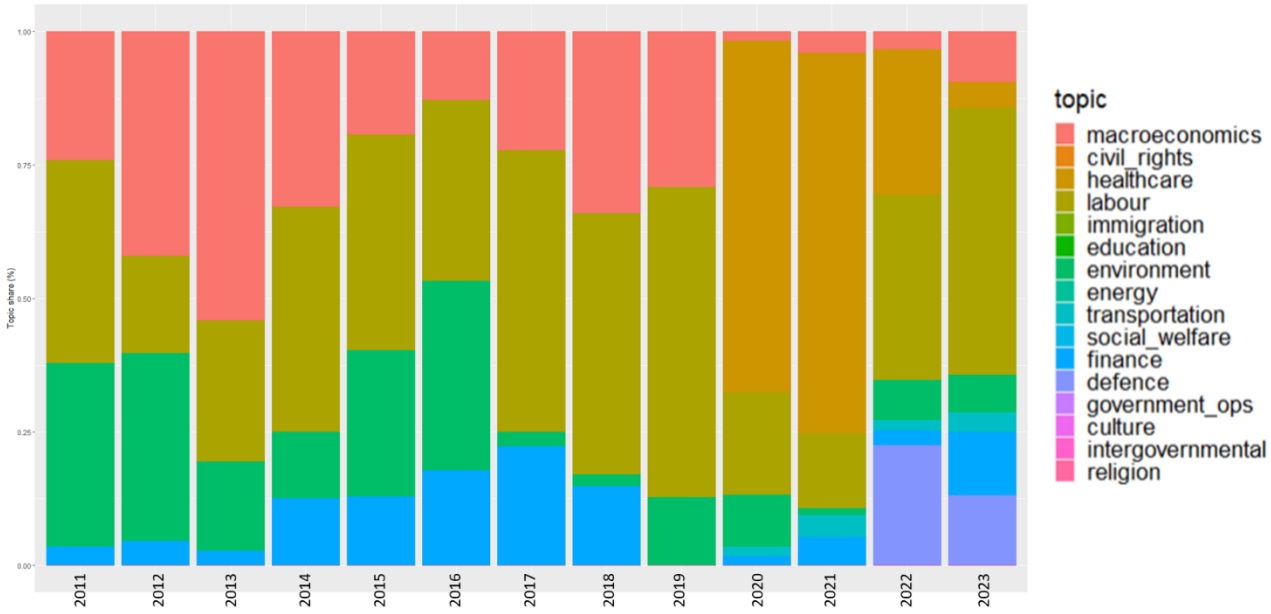


Chart 6. Topic distribution, by year

Chart 7, reported below, illustrates the comparison among the two governors. In this case, it is important to focus attention on two main aspects. Mr. Draghi appears to be more balanced in the presence of different topics, with a strong emphasis on macroeconomic issues. This, in addition to what was learned earlier about his concise communication style, may represent a spotlight of the Italian banker's approach to the role.

Ms. Lagarde, also due to significant global events under her mandate, dedicates very little space to pure macroeconomics topics, as well as to the environmental theme, an issue highly valued by the European Union as an institution.

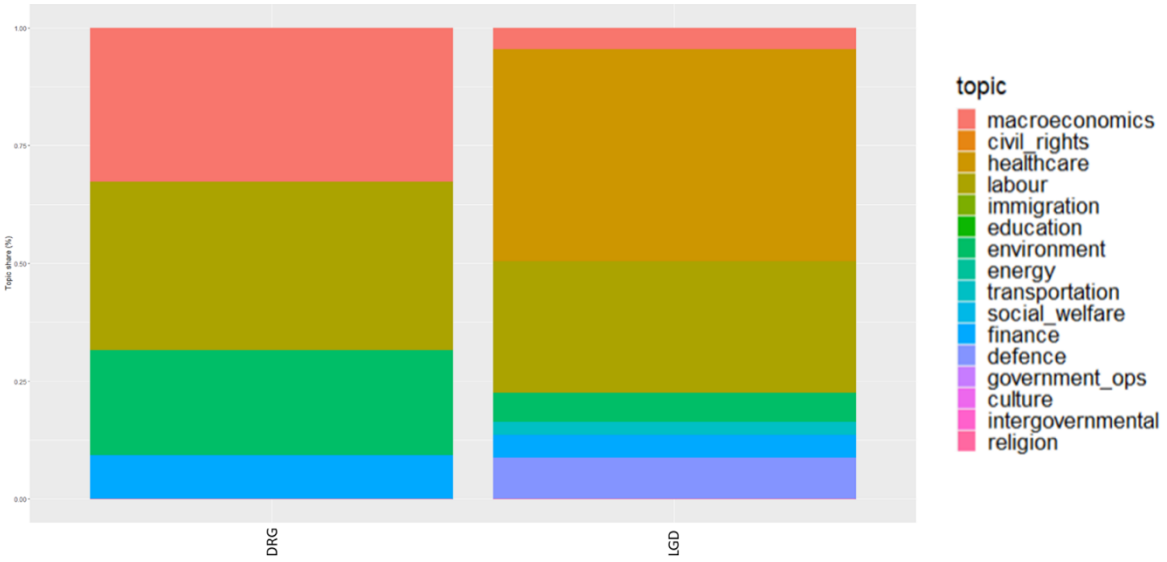


Chart 7. Topic distribution, by Governor

Topic Modeling

The analysis now delves into the database obtained after the entire pre-processing phase. The goal is to obtain a quantitative measure that can contribute to our regression analysis in the next chapter. We chose to use Latent Dirichlet Allocation (LDA), a popular topic modeling technique, to extract topics from the corpus. The model identifies patterns of co-occurring words and groups them into topics, giving as output to their composition and the probability that a specific topic appears in the single discourse under consideration.

Before running the topic modeling model, however, it is necessary to decide on the number of topics to be obtained. This choice of number of topics is a critical decision in topic modeling because it directly affects the granularity, quality, interpretability, and usefulness of the results. The risks are that if the number of topics is too low, the model may fail to capture the inherent diversity of the corpus, while with too many topics there may be overfitting. In our model, the values considered range from 2 to 200.

This step involves running the Latent Dirichlet Allocation (LDA) algorithm for a different number of topics. The process is executed through a for loop, iterating over each value of k and computing an LDA model for that number of topics.

```
k = c(2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150,
200)
lda_list = vector("list", length(k))
names(lda_list) = str_c("LDA_", kvec)
for (i in seq_along(k)) {
  cat("Computing LDA with", k[i], "topics\n")
  lda_list[[i]] = LDA(dfmLDA, k[i])}
```

To determine the ideal number of topics, we utilized a perplexity index by number of topics.

Graph 8 helps to select an optimal number of topics: it represents the perplexity index as a function of the number of topics in our model. The curve shows an inflection point between a number of topics between 40 and 60. In our model, for simplicity of analysis, we chose to proceed with 30 topics.

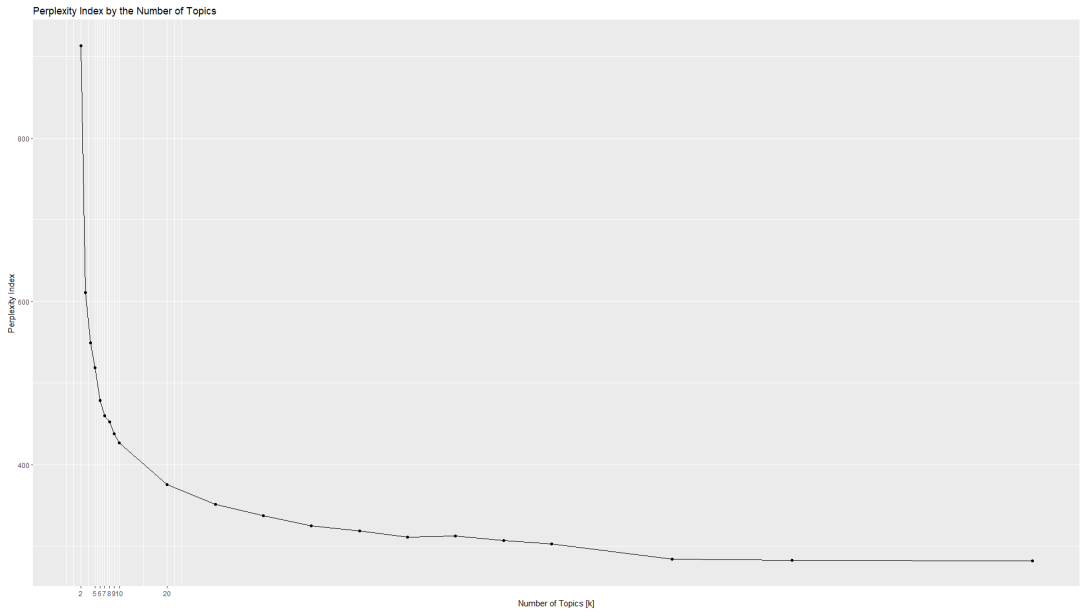


Chart 8. Perplexity index, by number of topics

Once the model with 30 topics was launched, each topic was analysed to understand what bigrams it was composed by. Based on the composition of the topics, evaluated on the first 15 bigrams (have a look at the appendix for a complete overview), a title was assigned to each topic. The *Table 2* details all the topics with their respective assigned titles.

Topic	Topic Label	Topic	Topic Label	Topic	Topic Label
V1	Asset_Strategies	V11	Financial_Sustainability	V21	Staff_Analysis
V2	Non_Financial_Impact	V12	Banks_and_Outlook	V22	Price_Pressures
V3	Risk_Analysis	V13	Inflation_Projections	V23	Inflation_Targets
V4	Refinancing_Trends	V14	Macroeconomic_Outlook	V24	Annual_Growth
V5	Inflation_Dynamics	V15	Energy_Prices	V25	Decision_Making
V6	Structural_Impact	V16	Inflation_Expectations	V26	Sustainable_Adjustments
V7	GDP_Projections	V17	Oil_and_Development	V27	Structural_Stability
V8	LTRO_III_Impact	V18	Financial_Analysis	V28	Market_Dynamics
V9	Containment_Measures	V19	Financial_Conditions	V29	Market_Trends
V10	Asset_Convergence	V20	Economic_Measures	V30	Reform_Impac

Table 2. Topics' assigned titles

This juncture, indeed, constitutes the pivotal linkage between the qualitative textual analysis and the quantitative aspects of regression modeling, as it facilitates understanding, enabling a robust statistical interpretation of the underlying patterns within the data set.

As a matter of fact, we report below the normalized weight proportions with color-coded labels for each topic, providing an overview of the thematic distribution in ECB presidents' speeches over the considered timespan. In *Chart 9,* we observe that Mr. Draghi exhibits a greater variety of topics,

despite being more concise in his speeches and focusing on macroeconomic themes. In contrast, Ms. Lagarde almost monopolizes her later years of speeches with fewer topics. This is influenced not only by personal characteristics but also by the highly complex historical period, both geopolitically and macroeconomically, marked by unprecedented inflation levels in recent years.

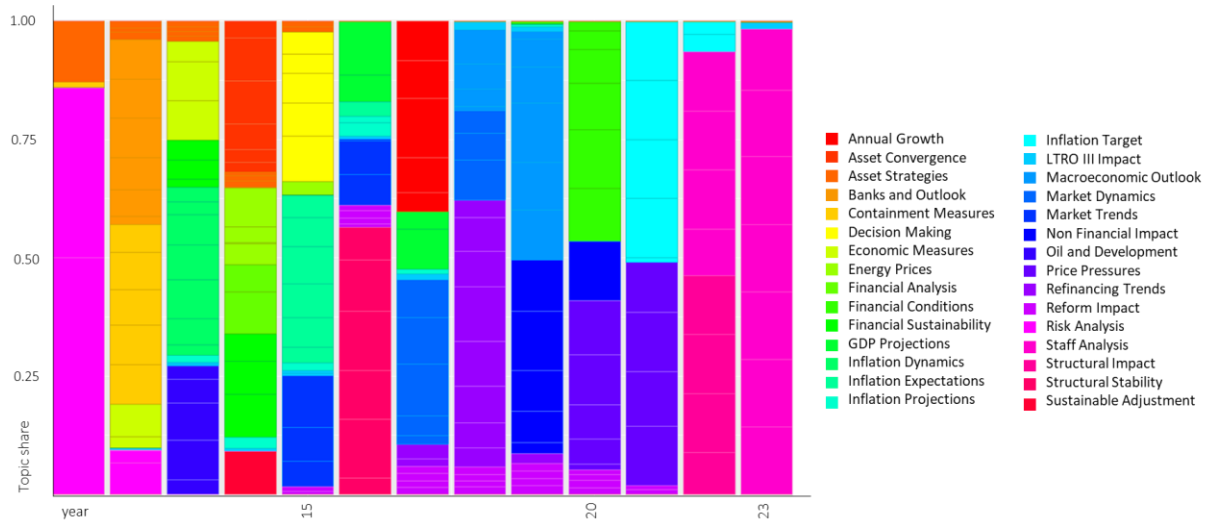


Chart 9. Topic Modelling - Distribution of topics, by year

Linear Regression model

The use of linear regression analysis in our study aligns with established academic precedents examining the same issues. For example, the aforementioned work by Blinder et al. (2008), which highlighted the importance of such communications in anchoring inflation expectations, was based upon an OLS linear regression model.

Accordingly, we used these models to investigate the relationship between the discourse on economic issues and several financial measures, including the change in the 10-year IOS rate, the change in the EUR/USD exchange rate, and the change in Italian and German government bond yields. Each of these variables was observed in the precise time windows highlighted in the Financial Data Retrieval section. This approach not only echoes methods applied in the literature, but also provides a transparent (and most importantly, statistically significant!) framework for assessing the effects of economic communication on financial variables.

$$\Delta OIS\ 10Y = \beta_0 + \beta_n Topics_n + \beta_{GOV} Draghi + \varepsilon$$

$$\Delta EURUSD = \beta_0 + \beta_n Topics_n + \beta_{GOV} Draghi + \varepsilon$$

$$\Delta IT\ 10Y = \beta_0 + \beta_n Topics_n + \beta_{GOV} Draghi + \varepsilon$$

$$\Delta DE\ 10Y = \beta_0 + \beta_n Topics_n + \beta_{GOV} Draghi + \varepsilon$$

To accomplish this, we created a database that merged the Topic Modeling model constructed with the *EA-MPD* dataset, which contains the impacts of various release windows on rate changes. The data were filtered by selecting only indices starting from November 11, 2011, which represents the first speech analyzed in our model. For the sake of completeness in the analysis, it is important to emphasize that our regression model concludes on June 15, 2023, as this is the most recent data available in the updated dataset.

A crucial aspect of our model revolved around determining how to manage and interpret the element related to the topic to be included. The LDA model has allowed us to come up with the probability of each topic being present in each individual speech. An initial regression model was constructed, analyzing each topic for every speech. Our analysis was therefore based on a 105x30 matrix. However, the results obtained were not satisfactory in terms of the statistical significance of the model. To achieve statistically relevant results, only the major topic for each speech was selected and isolated.

In other words, we focused on the topic that, according to our model, had the highest probability of being present in an individual speech. By doing so and using *doc_id*, *year*, and *month* as keys, all the data were combined into a single database: "*Regression.csv*".

In addition to what was said before, linear regression proves to be the optimal choice for several reasons. First, it provides a clear and interpretable framework for assessing the independent effect of each topic on the dependent variable, which as we shall see will be the focus of our discussion. Second, the simplicity of the model is a great advantage. Finally, despite the simplicity of the model, it can be efficiently extended to control for time fixed effects, as in our case through months and years. However, do bear in mind that such a model may have limitations, so we will devote an entire section later to define firstly the limitations and secondly the achievable next steps and reasonably attainable results of a deeper analysis.

In addition, a dummy independent variable named "*draghi*", is incorporated to capture the effect of the ECB in-charge presidency. This dummy variable takes the unitary value if the ECB president at the time was Mr. Mario Draghi, and 0 if the president was Ms. Christine Lagarde.

Aware of the variables within the model we then proceeded to interpret the output. For a thorough comprehension the reader can find the complete outputs in the Appendix.

Before proceeding with the interpretation of the results let us consider the values of R-square and adjusted R-square, which reflect the ability of the models to explain the variance of our financial indicators. The R-square values, which are high for such an empirical model and range from about

18.65% to 24.90%, suggest that a reasonable fraction of the variance is captured by our independent variables. On the other hand, the adjusted R-square provides a more sophisticated indicator of fit, penalizing for the number of predictors used. In this case, negative values such as -0.105107 to -0.039806 suggest that the inclusion of an additional variable only marginally worsens the explanatory power of the model compared to a simpler model without the same variable.

In addition, the RMSE (Root Mean Square Error), which ranges from 3.08924 to 7.95393, indicates a satisfactory level of prediction accuracy. The high significance levels of the β coefficients, indicated by their respective p-values, further ensure the validity of the independent variables in the models. Therefore, while the adjusted R-square values might initially suggest a suboptimal model fit, the combination of low RMSE values and statistically significant predictors confirms the significance of our four models.

Another very important metric to evaluate is the plot of residuals on fitted values. As shown in *graphs 4, 5, 6 and 7* in the *Appendix*, it can be stated that in all four regressions the residuals appear to be randomly distributed around zero, showing no specific patterns, thus rejecting the hypothesis of heteroskedasticity. Furthermore, it is evident from the graphs that there are no significant outliers that would affect the pattern.

The absence of patterns in the residuals suggests that the linear regression model captures the underlying relationships in the data well. This is critical to the model's predictions and the validity of subsequent inferences drawn from the analysis.

Results' Interpretation

In order to proceed with the interpretation of the data, we have decided to focus on four topics that we consider particularly telling, and to leave the remaining independent variables in the appendix for the reader to read. However, we point out that most of these variables are statistically significant (p-value to reject the null hypothesis of coefficient = 0), with the exception of a very few non-significant variables.

One of these is the case of the LTRO III topic for Italian and German sovereign bonds. In particular, it is noticeable that this variable does not appear to have a significant impact on changes in sovereign bond yields, but the reason is quickly explained: the European Central Bank's long-term refinancing operations (LTRO III) have the main objective of providing liquidity to banks, not to directly influence countries' sovereign debt.

In our analysis, we chose to study the cases of the topics *Asset Convergence, Financial Conditions, Inflation Expectations* and *Structural Stability*.

Starting from the first point, it is interesting to note that it shows a positive β in relation to IOS 10Y (16.663), meaning that, on average and all other things being equal, a unitary increase in the probability of the topic in question corresponds to an increase of 16.66 basis points, or 0.1666%, in the OIS 10Y rate. Here, one must not make the mistake of thinking that a change of 16 basis points in the ECB's interest rates cannot substantially influence the cost of credit for banks. In fact, a change of this magnitude is sufficient to alter loan conditions for businesses and consumers.

Conversely, the same topic has a much smaller impact on EUR/USD, showing a lower β (1.730) and suggesting a still positive but less pronounced effect. This indication will be repeated several times: the impacts on the exchange rate are much smaller, all other things being equal, probably an empirical explanation of the fact that exchange rates are truly impacted by many metrics, not only European (in our consideration) but also, obviously, by American homologues.

To better understand the topic's impact, the reader is provided with an overview of the bigrams contained in V10, *Asset Convergence (Table 3*).

Topic 10			
"asset_purchase"	"favourable_finance"	"previous_month"	"finance_condition"
"underlie_inflation"	"bank_lend"	"towards_level"	"area_economy"
"3_quarter"	"rate_towards"	"economic_analysis"	"secure_sustain"
"annual_growth"	"look_ahead"	"economic_expansion"	

Table 3. First 15 bigrams contained in topic 10.

Taking a look at table 3, it's possible to observe that the topic "*Asset Convergence*" is characterized by terms such as *"asset_purchase", "favourable_finance", "finance_condition",* and "*economic_expansion*". At this point, we can refine our previous interpretation, also in more financial terms.

For example, the highly significant positive β coefficient for "*Asset Convergence*" in the OIS 10Y regression suggests that when topics related to expansionary monetary policy, such as the quantitative easing of 2015-2016, are discussed, the market tends to react extremely well, especially in terms of national debt (β on IT BTP 10Y = 20.7633, and for DE Bund 10Y = 16.633).

In other words, when the debate focuses on the purchase of public debt, the entire eurozone seems to breathe a sigh of relief and receive a kind of bonus from the markets. On the contrary, the common currency will experience this effect to a lesser extent, being, as already anticipated, defined by many other occurrencies (trade balances, foreign interest rates, and geopolitics).

Regarding the second topic we selected, it is interesting to see how this, *V19 "Financial Conditions"*, emphasizes a dichotomy in market reactions. This dichotomy is expressed in a rather evident contrast: a β of -1.400 in the OIS 10Y regression model and one of 11.260 associated with the yields of Italian BTP 10Y bonds. The first indicator suggests that the market may interpret discussions on financial conditions as an implicit need for monetary easing, which as seen in 2022 would in turn lead to a reduction in the rates set by the ECB (by 1.4 basis points, on average and ceteris paribus).

Conversely, the reaction in sovereign bonds reflects a different market sentiment, this time negative, raising concerns about the Italian system and thus leading investors to demand a higher rate, indicative of greater risk (by 11.26 basis points, on average and ceteris paribus). Indeed, a negative β is actually an unfavorable indicator, as an increase in rates is not beneficial to the economy, symbolizing greater risk and volatility.

A similar behavior, positive at the IOS 10Y level and negative domestically, is also observed in the topic V16 *"Inflation Expectations"*. While it shows a negative β of -2.870 in the OIS 10Y regression, indicating a market's clear expectation of lowering rates to counter potential incoming inflation (Castelnuovo et al., 2003), the regression instead produces a positive β of 8.160 on IT titles, implying that, as expected, talking about inflation does not do well for domestic systems, to which inflation actually benefits (a state always benefits from inflation, just think of the maximization of VAT revenue).

This impact, it is repeated, is anything but unpredictable, but is not observed in Germany (β = -4.3), whose bonds are considered somewhat safer. From our point of view, this has been explained as

Germany being seen as the engine of Europe and in case of higher inflation, the average investor would turn to German Bunds, which are undoubtedly the safest asset of the continent.
If the reader wants to delve deeper into the topic, to better understand the logic, she has been provided with an overview of the bigrams contained in V16, "*Inflation Expectations"*.

Topic 16			
"energy_price"	"energy_food"	"inflation_expectation"	"high_energy"
"supply_bottleneck"	"price_pressure"	"high_inflation"	"term_inflation"
"economic_activity"	"price_inflation"	"bank_lend"	"basis_point"
"near_term"	"inflation_outlook"	"food_price"	

Table 4. First 15 bigrams contained in topic 16.

Regarding the fourth and final variable investigated, V26 *"Structural Stability"*, the discussion around it is quantified by a large negative β of -11.20 in the OIS 10Y regression, which clearly suggests that market participants associate structural stability with reduced bond volatility. In the government bond yield model, the positive β of 15.1458 for German bonds and the positive β of 18.0583 for Italian bonds indicate a significant confidence in the stability of the domestic market, similar to the reasons mentioned for the V19 *"financial conditions"* topic.

As we have seen, a pattern often emerges in which domestic bond yields and European financing rates move in opposite directions in response to the same phenomenon. However, it is evident that this dichotomy is driven by purely macroeconomic and validated reasons. Something that has no econometric or macroeconomic properties is the role of the dummy variable concerning the governor who gave the speech. This variable, as mentioned earlier, has a value of 1 if the speaker is Mr. Mario Draghi (in office from November 1, 2011, to October 31, 2019), and 0 if it's Ms. Christine Lagarde (in office since November 1, 2019). The first thing the reader can notice is that the variable consistently has a negative β, indicating that on average and ceteris paribus, a speech given by Mr. Draghi lowered the OIS 10Y interest rates, slightly depreciated the euro currency, and lowered the yields of Italian and German government bonds (consequently the spread). This is undoubtedly a product of the turbulent times in which Mr. Draghi operated successfully.
Let's now correctly quantify the distinct effects that the communications of the former ECB President Mario Draghi had on various financial parameters in contrast to the communications during Ms. Christine Lagarde's tenure. In particular, a negative coefficient of -2.350 for the OIS 10Y rate implies that the market associates Mr. Draghi's speeches with a reduction in long-term interest rates. A similarly negative but smaller coefficient of -0.035 for the EUR/USD exchange rate suggests a slight depreciation of the Euro against the Dollar during Draghi's term (as mentioned before, this exchange

rate is not particularly impacted in our model since it depends on both European and American factors).

In general, however, it can be stated that the negative β indicates that the market viewed Mr. Draghi's tenure as a period of stabilization for both interest rates and government bond yields. This interpretation is supported by Mr. Draghi's decisive measures, such as OMT and PSPP, which not only directly contributed to financial restructuring but also instilled a great deal of confidence in the role of the ECB itself.

Worthy of attention is the relationship with Italian government bonds: in particular, a significant negative coefficient of -11.700 for Italian bond yields indicates that Mr. Draghi's communication is linked to a marked decline in these yields and, consequently, a lower perceived risk. This could be due to various factors, such as his Italian background, his esteemed reputation, and his in-depth knowledge of the Italian fiscal system (Mr. Draghi served as the Governor of the Bank of Italy from 2005 to 2011). Furthermore, the credibility of his assurances regarding the ECB's support, epitomized by "whatever it takes", certainly had a greater impact on countries like Italy, Spain, or Portugal, which were already in extreme difficulty, compared to those considered stable. It's worth noting that such a strong effect is not observed on German government bond yields, whose impact amounts to only -2.350 basis points, almost a fifth of the effect on equivalent Italian instruments and of the same duration.

Regarding the impact of Ms. Christine Lagarde, as we did not provide a specific coefficient for her mandate (this will be a point of discussion on project limitations), we can only hypothesize. For example, it is highly likely that her mandate had a limited impact on individual country systems, as she took office at a time when the ECB's policy orientation was already well-established. Additionally, given the turbulent period of her mandate (in some ways even more turbulent than that experienced by Mr. Draghi), it required the challenging management of the consequences of the pandemic, where the ECB's response had to be more extensive and uniformly applied across the Eurozone. This could be a reason why the perceived benefit to Italian government bonds was, in fact, diluted. As mentioned earlier, her speeches are much more discursive, inclusive, and, so to speak, widespread.

Conclusions

Our work featuring ECB speech data and financial metrics led to the creation of a statistically robust model that highlights the influence that monetary policy orations have on financial markets. We realized that the use of one word rather than another has a tangible effect on financial actors, explicating itself either by a different attitude toward sovereign bonds or on monetary exchange rates. Relevant here was the intentional choice of short time intervals, and this selection proved acute: lengthening such windows to longer periods, such as one or three days (instead of just over two hours), would in fact have weakened the significance of the observed impacts.

The model likewise showed a considerable influence of the topics discussed by ECB leaders on variables such as the 10-year OIS rate, the EUR/USD exchange rate, and yields on Italian and German government bonds. The significant β coefficients, found for topics such as Asset Convergence, Financial Conditions, Inflationary Expectations, and Structural Stability, align with what has already been discussed in the existing literature, e.g., in the work of Blinder et al. (2008) and Gorodnichenko et al. (2023), which emphasize the crucial role of central bank communications in modeling market perceptions. But not only that, such impacts align with financial theory in an intrinsic blend of theorems and econometric principles.

From a statistical point of view, considering the high R-square values, our model demonstrates strong explanatory power, which is further strengthened by significant β coefficients and satisfactory RMSE values. Moreover, the residuals appear to be normally distributed around zero, without any defined pattern, which reassures us about the homoskedasticity of the regression.

Extremely interesting trends concern the dummy variable *'draghi'*. The consistently negative coefficients of this variable in the different models indicate that Mr. Mario Draghi's presidency was perceived as a stabilizing influence on the markets, especially when contrasted with Ms. Christine Lagarde's current mandate (It should be noted, however, that the lack of a dedicated dummy variable is a limitation of the model). This effect was particularly pronounced in Italian bond yields, suggesting an extremely sensitive market response to Mr. Mario Draghi's leadership, for a variety of reasons. In conclusion, the next chapter will discuss the limits of the model, including the scope of the speeches analyzed and the selection of topics for the LDA analysis, which provide fertile ground for future research in this area.

Project limitations

While providing relevant insights and tangible results on the relationship between central banks and financial indicators, ours is subject to some limitations that require attention for future research efforts.

The first of the limitations of our study is the limited scope of the speeches analyzed, particularly in the context of the ECB presidency. Given that Mr. Mario Draghi's presidential term has ended while Ms. Christine Lagarde's is still ongoing, it creates a potential bias in favor of Mr. Draghi. This disparity in the time periods taken into account can distort the results, as it does not consider the full impact of Ms. Christine Lagarde's presidency, which with great certainty will be best explored until 2027 (end of term). Therefore, the results for the dummy variable "Draghi" should be interpreted with caution, taking into account this time imbalance.

Another aspect concerns the decision to include in the regression model only the most dominant topic in the LDA analysis for each discourse. While this choice simplifies the model and promotes clarity, it also limits our analysis to a single topic perspective for each discourse. Yet, as mentioned earlier, if we had incorporated all the topics identified by the LDA, the model could have captured more but with no (less) significance.

A third point is the absence of a specific *"lagarde"* dummy variable. The current model primarily contrasts the periods of Mr. Draghi's presidency with a baseline that also includes Ms. Lagarde's tenure, but without explicit differentiation. In all likelihood, the inclusion of a specific variable could offer insights into its relationship with financial markets, distinguishing itself from both Mr. Mario Draghi's tenure and the broader ECB narrative.

A fourth limitation is the potential divergence between the thematic content of ECB speeches and the actual implementation of monetary policies. It must be remembered that financial markets are indeed heavily influenced by the tangible actions of the central bank, as well as by expectations and sentiments. The risk is that our analysis attributes market movements to speech themes, potentially overlooking the actual but tangible confusing movement of markets.

From a purely statistical point of view, there are all the limitations of linear regression models. Although our models appear robust from this perspective, continuous vigilance is required in future analyses.

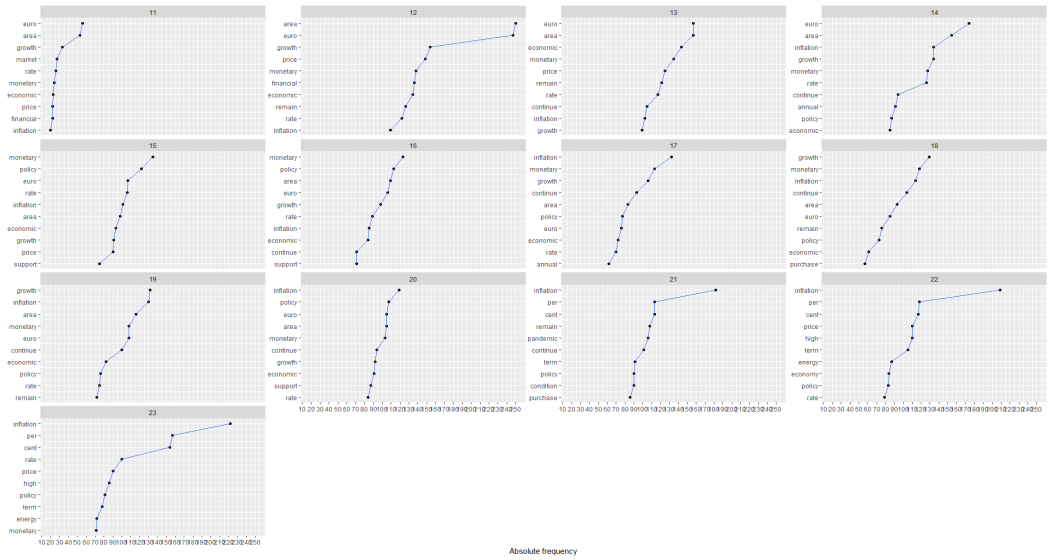
Lastly, potential discrepancies in the initial databases represent an idiosyncratic risk. As with any empirical research, the reliability of our conclusions is inherently linked to the accuracy of the underlying data. In the case of this work, the data were obtained from the ECB database, the same used by Gorodnichenko et al. (2023), thus this risk should be minimized. Yet the timing of the conferences has recently changed, so there may have been discrepancies in observing the deltas.

In conclusion, while on one hand our study advances the understanding of the impact of ECB communications on financial markets, on the other hand these limits highlight areas for improvement and further research. Addressing them in future studies will enhance the inferential capacity of the model and contribute to a more solid and thorough understanding of the role of central bank communications on the financial metrics.

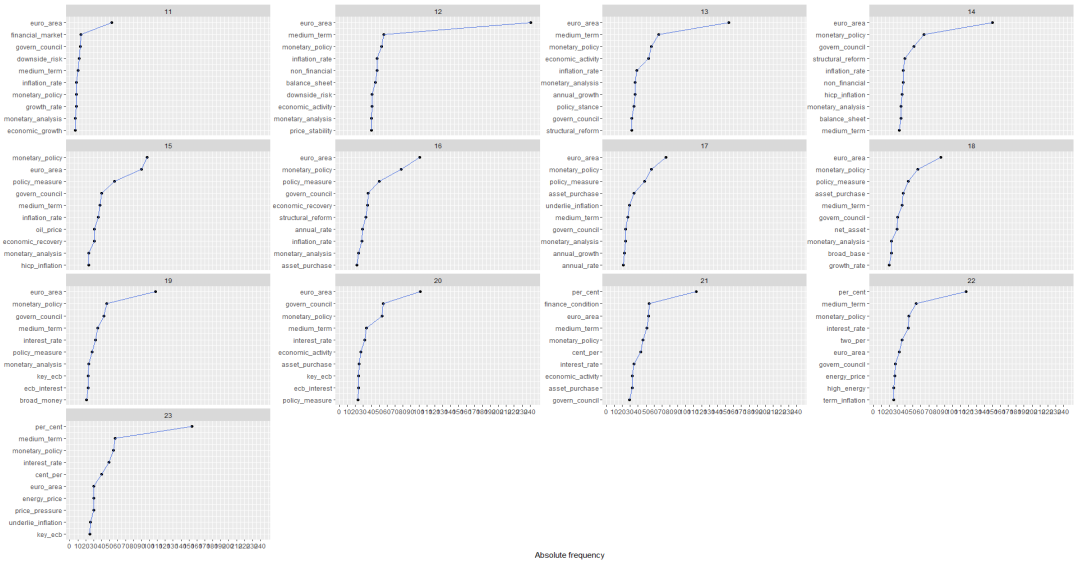
Bibliography

- Blinder, A. S., Ehrmann, M., Fratzscher, M., de Haan, J., & Jansen, D.-J. (2008). Central bank communication and monetary policy: A survey of theory and evidence. Journal of Economic Literature, 46(4), 910-945.
- Castelnuovo, E., Nicoletti-Altimari, S., & Rodriguez-Palenzuela, D. (2003). Definition of price stability, range and point inflation targets: The anchoring of long-term inflation expectations. SSRN Electronic Journal. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1949682
- Gorodnichenko, Y., Sockin, M., & Talavera, O. (2023). The voice of monetary policy. The American Economic Review, 113(4), 700-736.
- Gürkaynak, R. S., Sack, B., & Swanson, E. T. (2005). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review*, *95*(1), 425-436
- Jansen, D.-J., & Haan, J. de. (2010). An Assessment of the Consistency of ECB Communication Using Wordscores. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.1949682
- Kapur, N. (2020, July 18). Text preprocessing — NLP basics. Medium. https://medium.com/analytics-vidhya/text-preprocessing-nlp-basics-430d54016048

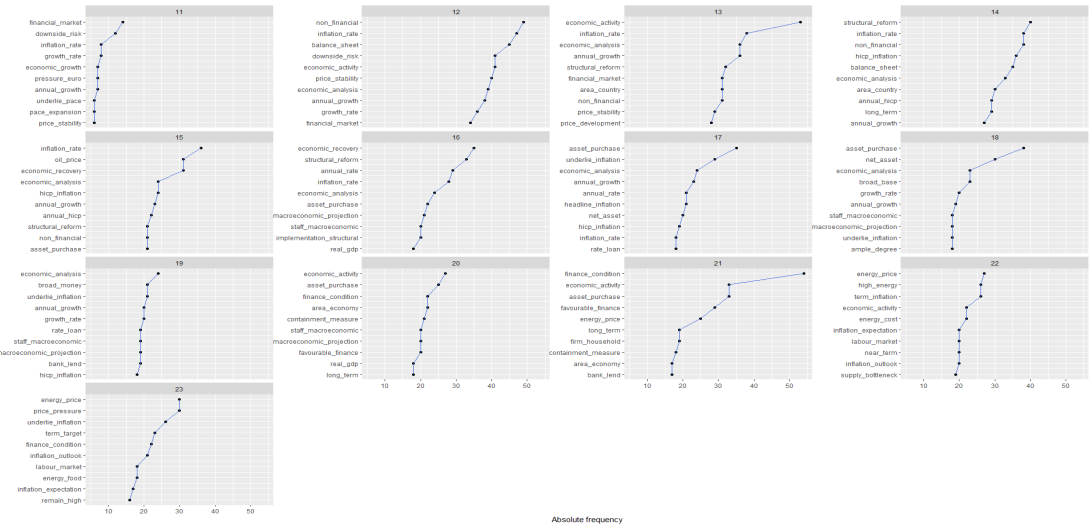
Appendix



Appendix Chart 1. Token Frequency distribution



Appendix Chart 2. Bigram Frequency distribution



Appendix Chart 3. DFM Distribution, by year

Compositions of the 30 topics identified (first 15 bigrams)

	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7
[1,]	"asset_purchase"	"non_financial"	"downside_risk"	"refinance_operation"	"inflation_rate"	"balance_sheet"	"real_gdp"
[2,]	"structural_reform"	"economic_activity"	"financial_market"	"inflation_rate"	"economic_recovery"	"structural_reform"	"area_economy"
[3,]	"underlie_inflation"	"inflation_rate"	"growth_rate"	"fix_rate"	"oil_price"	"euro_area's"	"staff_macroeconomic"
[4,]	"annual_rate"	"downside_risk"	"inflation_rate"	"rate_close"	"growth_euro"	"growth_momentum"	"containment_measure"
[5,]	"purchase_programme"	"balance_sheet"	"annual_growth"	"tender_procedure"	"annual_growth"	"non_financial"	"raise_key"
[6,]	"headline_inflation"	"remain_subdue"	"fiscal_consolidation"	"long_term"	"hicp_inflation"	"economic_analysis"	"macroeconomic_projection"
[7,]	"rate_loan"	"area_economy"	"underlie_pace"	"term_refinance"	"asset_purchase"	"annual_growth"	"economic_activity"
[8,]	"annual_growth"	"price_stability"	"pace_expansion"	"full_allotment"	"economic_analysis"	"energy_price"	"start_raise"
[9,]	"economic_analysis"	"non_standard"	"economic_activity"	"continue_conduct"	"annual_hicp"	"inflation_rate"	"asset_purchase"
[10,]	"monthly_pace"	"standard_measure"	"growth_euro"	"procedure_full"	"close_medium"	"area's_growth"	"price_stability"
[11,]	"pace_billion"	"economic_analysis"	"pressure_euro"	"rate_tender"	"firm_household"	"fiscal_policy"	"eurosystem_staff"
[12,]	"hicp_inflation"	"look_ahead"	"economic_growth"	"less_end"	"exchange_rate"	"private_sector"	"degree_accommodation"
[13,]	"level_close"	"financial_corporation"	"high_uncertainty"	"real_economy"	"non_financial"	"credit_supply"	"security_purchase"
[14,]	"inflation_rate"	"growth_rate"	"price_stability"	"three_month"	"oil_future"	"annual_hicp"	"long_term"
[15,]	"net_asset"	"loan_non"	"market_tension"	"basis_point"	"expect_economic"	"hicp_inflation"	"long_necessary"

	Topic 8	Topic 9	Topic 10	Topic 11	Topic 12	Topic 13
[1,]	"tltro_iii"	"tltro_iii"	"asset_purchase"	"convergence_inflation"	"underlie_inflation"	"inflation_rate"
[2,]	"annual_growth"	"refinance_operation"	"favourable_finance"	"sustain_convergence"	"broad_money"	"economic_recovery"
[3,]	"growth_rate"	"economic_activity"	"previous_month"	"asset_purchase"	"economic_analysis"	"macroeconomic_projection"
[4,]	"economic_analysis"	"containment_measure"	"finance_condition"	"economic_analysis"	"bank_lend"	"staff_macroeconomic"
[5,]	"rate_loan"	"hicp_inflation"	"underlie_inflation"	"continue_sustain"	"inflation_outlook"	"hicp_inflation"
[6,]	"refinance_operation"	"private_sector"	"bank_lend"	"degree_accommodation"	"accommodative_stance"	"price_stability"
[7,]	"level_close"	"asset_purchase"	"towards_level"	"close_medium"	"area_economy"	"real_gdp"
[8,]	"sustain_convergence"	"basis_point"	"area_economy"	"m3_growth"	"low_level"	"annual_rate"
[9,]	"close_medium"	"measure_take"	"3_quarter"	"level_close"	"period_time"	"structural_reform"
[10,]	"broad_money"	"long_term"	"rate_towards"	"ample_degree"	"annual_growth"	"economic_analysis"
[11,]	"bank_lend"	"staff_macroeconomic"	"economic_analysis"	"inflation_level"	"money_growth"	"return_inflation"
[12,]	"uncertainty_relate"	"macroeconomic_projection"	"secure_sustain"	"rise_wage"	"area_fiscal"	"level_close"
[13,]	"ample_degree"	"economic_analysis"	"annual_growth"	"net_asset"	"fiscal_stance"	"refinance_operation"
[14,]	"degree_accommodation"	"finance_condition"	"look_ahead"	"long_necessary"	"growth_rate"	"asset_purchase"
[15,]	"long_necessary"	"downside_risk"	"economic_expansion"	"case_long"	"area_growth"	"annual_hicp"

	Topic 14	Topic 15	Topic 16	Topic 17	Topic 18	Topic 19
[1,]	"staff_macroeconomic"	"energy_price"	"energy_price"	"oil_price"	"economic_activity"	"finance_condition"
[2,]	"macroeconomic_projection"	"finance_condition"	"energy_food"	"price_development"	"financial_market"	"asset_purchase"
[3,]	"ecb_staff"	"term_inflation"	"inflation_expectation"	"economic_recovery"	"inflation_rate"	"economic_activity"
[4,]	"projection_euro"	"asset_purchase"	"high_energy"	"inflation_rate"	"economic_analysis"	"favourable_finance"
[5,]	"foresee_annual"	"economic_activity"	"supply_bottleneck"	"balance_sheet"	"annual_growth"	"bank_lend"
[6,]	"real_gdp"	"near_term"	"price_pressure"	"economic_analysis"	"balance_sheet"	"impact_pandemic"
[7,]	"broadly_reflect"	"long_term"	"high_inflation"	"increase_investment"	"accommodative_stance"	"structural_policy"
[8,]	"projection_outlook"	"area_economy"	"term_inflation"	"non_financial"	"non_financial"	"weak_demand"
[9,]	"eurosystem_staff"	"labour_market"	"economic_activity"	"annual_growth"	"area_country"	"fiscal_policy"
[10,]	"hicp_inflation"	"high_energy"	"price_inflation"	"structural_reform"	"financial_sector"	"containment_measure"
[11,]	"area_foresee"	"financial_condition"	"bank_lend"	"fall_oil"	"continue_subdue"	"continue_support"
[12,]	"reflect_ecb"	"staff_projection"	"basis_point"	"support_economic"	"expansion_continue"	"annual_growth"
[13,]	"gdp_growth"	"supply_bottleneck"	"near_term"	"annual_rate"	"pace_expansion"	"growth_rate"
[14,]	"annual_real"	"inflation_outlook"	"inflation_outlook"	"outlook_price"	"underlie_pace"	"pandemic_emergency"
[15,]	"growth_revise"	"energy_cost"	"food_price"	"exchange_rate"	"pressure_remain"	"purchase_pepp"

	Topic 20	Topic 21	Topic 22	Topic 23	Topic 24	Topic 25
[1,]	"finance_condition"	"macroeconomic_projection"	"price_pressure"	"economic_recovery"	"underlie_inflation"	"asset_purchase"
[2,]	"economic_activity"	"staff_macroeconomic"	"underlie_inflation"	"structural_reform"	"asset_purchase"	"today's_decision"
[3,]	"asset_purchase"	"ecb_staff"	"finance_condition"	"annual_rate"	"annual_growth"	"recent_month"
[4,]	"containment_measure"	"eurosystem_staff"	"term_target"	"inflation_rate"	"economic_analysis"	"sustain_adjustment"
[5,]	"favourable_finance"	"projection_euro"	"energy_price"	"economic_analysis"	"headline_inflation"	"degree_accommodation"
[6,]	"energy_price"	"area_foresee"	"inflation_outlook"	"implementation_structural"	"economic_expansion"	"mainly_reflect"
[7,]	"annual_growth"	"foresee_annual"	"remain_high"	"look_ahead"	"inflation_rate"	"adjustment_path"
[8,]	"global_demand"	"real_gdp"	"labour_market"	"support_economic"	"path_inflation"	"purchase_programme"
[9,]	"inflation_outlook"	"projection_projection"	"energy_food"	"expect_economic"	"annual_rate"	"high_level"
[10,]	"price_inflation"	"hicp_inflation"	"assessment_inflation"	"growth_euro"	"degree_accommodation"	"oil_price"
[11,]	"growth_rate"	"annual_hicp"	"two_medium"	"fiscal_policy"	"term_outlook"	"european_institution"
[12,]	"remain_subdue"	"assessment_also"	"ensure_inflation"	"asset_purchase"	"sustain_adjustment"	"remain_weak"
[13,]	"fiscal_policy"	"annual_real"	"inflation_return"	"balance_sheet"	"adjustment_path"	"council_decide"
[14,]	"staff_macroeconomic"	"also_reflect"	"rate_increase"	"rate_change"	"growth_rate"	"dynamic_continue"
[15,]	"macroeconomic_projection"	"reflect_eurosystem"	"today_decide"	"annual_growth"	"purchase_programme"	"area_also"

	Topic 26	Topic 27	Topic 28	Topic 29	Topic 30
[1,]	"economic_activity"	"structural_reform"	"asset_purchase"	"financial_market"	"inflation_rate"
[2,]	"remain_subdue"	"economic_activity"	"net_asset"	"annual_growth"	"hicp_inflation"
[3,]	"annual_growth"	"area_country"	"broad_base"	"balance_sheet"	"structural_reform"
[4,]	"economic_analysis"	"price_stability"	"underlie_inflation"	"non_financial"	"non_financial"
[5,]	"low_level"	"economic_analysis"	"annual_rate"	"economic_analysis"	"long_term"
[6,]	"price_pressure"	"annual_growth"	"economic_analysis"	"price_stability"	"area_country"
[7,]	"area_country"	"price_development"	"annual_growth"	"financial_sector"	"balance_sheet"
[8,]	"underlie_price"	"inflation_rate"	"growth_rate"	"area_country"	"economic_analysis"
[9,]	"expect_remain"	"weak_economic"	"growth_potential"	"growth_rate"	"medium_long"
[10,]	"outlook_price"	"growth_rate"	"increase_resilience"	"growth_euro"	"annual_hicp"
[11,]	"price_stability"	"non_financial"	"continue_support"	"inflation_rate"	"fiscal_consolidation"
[12,]	"market_condition"	"progress_make"	"inflation_aim"	"next_year"	"inflation_expectation"
[13,]	"price_development"	"line_price"	"economic_union"	"increase_indirect"	"private_sector"
[14,]	"inflation_rate"	"since_last"	"period_time"	"fiscal_consolidation"	"subdue_underlie"
[15,]	"non_financial"	"domestic_demand"	"headline_inflation"	"indirect_tax"	"period_low"

Linear Regression Results - OIS 10 Y

```
OLS estimation, Dep. Var.: OIS_10Y
Observations: 105
Fixed-effects: year: 13
Standard-errors: Clustered (year)
                                        Estimate   Std. Error       t value   Pr(>|t|)
topic_label_facAsset_Convergence        12.030000 6.837600e-15  1.759398e+15  < 2.2e-16 ***
topic_label_facAsset_Strategies          1.040000 3.629900e-15  2.865106e+14  < 2.2e-16 ***
topic_label_facBanks_and_Outlook         0.890000 1.178190e-14  7.553985e+13  < 2.2e-16 ***
topic_label_facContainment_Measures     -1.240370 2.581210e-01 -4.805384e+00 4.2958e-04 ***
topic_label_facDecision_Making           0.655833 1.473900e-15  4.449737e+14  < 2.2e-16 ***
topic_label_facEconomic_Measures        -0.588333 1.891000e-14 -3.111226e+13  < 2.2e-16 ***
topic_label_facEnergy_Prices            -2.347037 2.581210e-01 -9.092779e+00 9.9063e-07 ***
topic_label_facFinancial_Analysis       -2.505000 9.546700e-15 -2.623930e+14  < 2.2e-16 ***
topic_label_facFinancial_Conditions      0.396667 2.229720e-14  1.778995e+13  < 2.2e-16 ***
topic_label_facFinancial_Sustainability  0.900000 3.175800e-15  2.833945e+14  < 2.2e-16 ***
topic_label_facGDP_Projections           2.603333 9.050000e-17  2.875369e+16  < 2.2e-16 ***
topic_label_facInflation_Dynamics       -2.080370 2.581210e-01 -8.059672e+00 3.4842e-06 ***
topic_label_facInflation_Expectations   -2.870000 1.075820e-14 -2.667740e+14  < 2.2e-16 ***
topic_label_facInflation_Targets         2.664167 2.092210e-14  1.273373e+14  < 2.2e-16 ***
topic_label_facLTRO_III_Impact           1.139630 2.581210e-01  4.415099e+00 8.4279e-04 ***
topic_label_facMacroeconomic_Outlook     3.036667 1.211530e-14  2.506473e+14  < 2.2e-16 ***
topic_label_facMarket_Dynamics          -4.150000 1.142380e-14 -3.632769e+14  < 2.2e-16 ***
topic_label_facNon_Financial_Impact     -0.650000 8.742800e-15 -7.434705e+13  < 2.2e-16 ***
topic_label_facPrice_Pressures           2.800000 5.747900e-15  4.871326e+14  < 2.2e-16 ***
topic_label_facRefinancing_Trends        4.100000 1.104800e-15  3.711117e+15  < 2.2e-16 ***
topic_label_facReform_Impact            -0.276296 9.034234e-01 -3.058330e-01 7.6497e-01
topic_label_facStructural_Impact         4.450000 1.351700e-15  3.292198e+15  < 2.2e-16 ***
topic_label_facStructural_Stability     11.230000 7.442700e-15  1.508868e+15  < 2.2e-16 ***
draghi                                  -1.100000 5.638000e-16 -1.950994e+15  < 2.2e-16 ***
... 2 variables were removed because of collinearity (topic_label_facStaff_Analysis and topic_label_facSustainable_Adjustments)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 3.08924     Adj. R2: -0.168683
                Within R2:  0.18111
```

Liner Regression Results – EURUSD

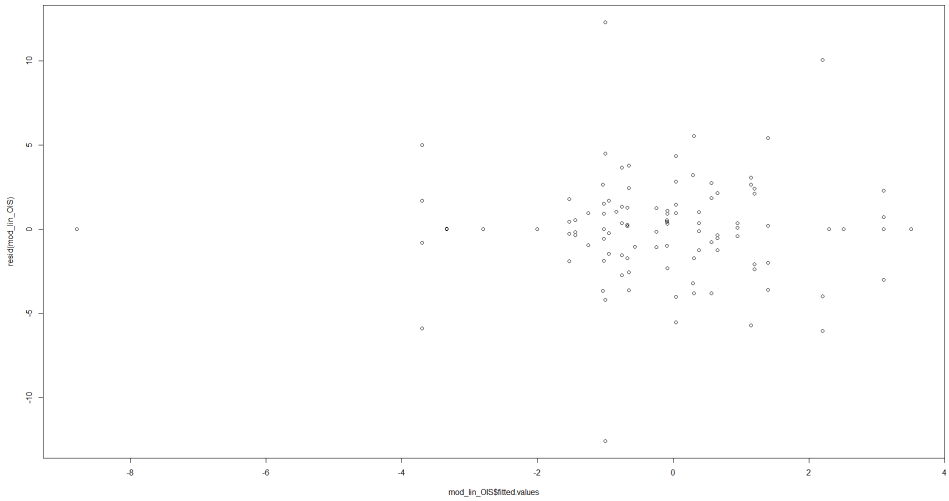
```
OLS estimation, Dep. Var.: EURUSD
Observations: 105
Fixed-effects: year: 13
Standard-errors: Clustered (year)
                                        Estimate   Std. Error       t value   Pr(>|t|)
topic_label_facAsset_Convergence         1.730000 1.206000e-15  1.434189e+15  < 2.2e-16 ***
topic_label_facAsset_Strategies          0.433333 5.010000e-16  8.640865e+14  < 2.2e-16 ***
topic_label_facBanks_and_Outlook        -0.421667 1.848000e-15 -2.282236e+14  < 2.2e-16 ***
topic_label_facContainment_Measures     -0.168519 6.031724e-02 -2.793870e+00 1.6228e-02 *
topic_label_facDecision_Making           0.615000 2.920000e-16  2.107726e+15  < 2.2e-16 ***
topic_label_facEconomic_Measures         0.683333 3.824000e-15  1.786855e+14  < 2.2e-16 ***
topic_label_facEnergy_Prices            -0.481852 6.031724e-02 -7.988630e+00 3.8151e-06 ***
topic_label_facFinancial_Analysis       -1.217500 2.365000e-15 -5.146921e+14  < 2.2e-16 ***
topic_label_facFinancial_Conditions     -0.115000 3.244000e-15 -3.544871e+13  < 2.2e-16 ***
topic_label_facFinancial_Sustainability -1.320000 1.387000e-15 -9.514135e+14  < 2.2e-16 ***
topic_label_facGDP_Projections           0.190000 1.770000e-16  1.073211e+15  < 2.2e-16 ***
topic_label_facInflation_Dynamics       -0.285185 6.031724e-02 -4.728090e+00 4.9014e-04 ***
topic_label_facInflation_Expectations   -0.895000 1.657000e-15 -5.399760e+14  < 2.2e-16 ***
topic_label_facInflation_Targets         0.635833 1.953000e-15  3.255618e+14  < 2.2e-16 ***
topic_label_facLTRO_III_Impact           0.211481 6.031724e-02  3.506150e+00 4.3325e-03 **
topic_label_facMacroeconomic_Outlook    -0.221667 1.943000e-15 -1.140842e+14  < 2.2e-16 ***
topic_label_facMarket_Dynamics          -1.289167 2.271000e-15 -5.676516e+14  < 2.2e-16 ***
topic_label_facNon_Financial_Impact     -1.100000 1.364000e-15 -8.063829e+14  < 2.2e-16 ***
topic_label_facPrice_Pressures          -0.607500 9.920000e-16 -6.126755e+14  < 2.2e-16 ***
topic_label_facRefinancing_Trends       -0.177500 3.690000e-16 -4.811345e+14  < 2.2e-16 ***
topic_label_facReform_Impact             0.415185 2.111104e-01  1.966670e+00 7.2779e-02 .
topic_label_facStructural_Impact        -0.480000 5.480000e-16 -8.758260e+14  < 2.2e-16 ***
topic_label_facStructural_Stability      1.883333 8.440000e-16  2.230466e+15  < 2.2e-16 ***
draghi                                  -0.035000 3.060000e-16 -1.143329e+14  < 2.2e-16 ***
... 2 variables were removed because of collinearity (topic_label_facStaff_Analysis and topic_label_facSustainable_Adjustments)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 0.498881     Adj. R2: -0.206546
                 Within R2:  0.19354
```

Linear Regression Results – IT BTP 10 Y

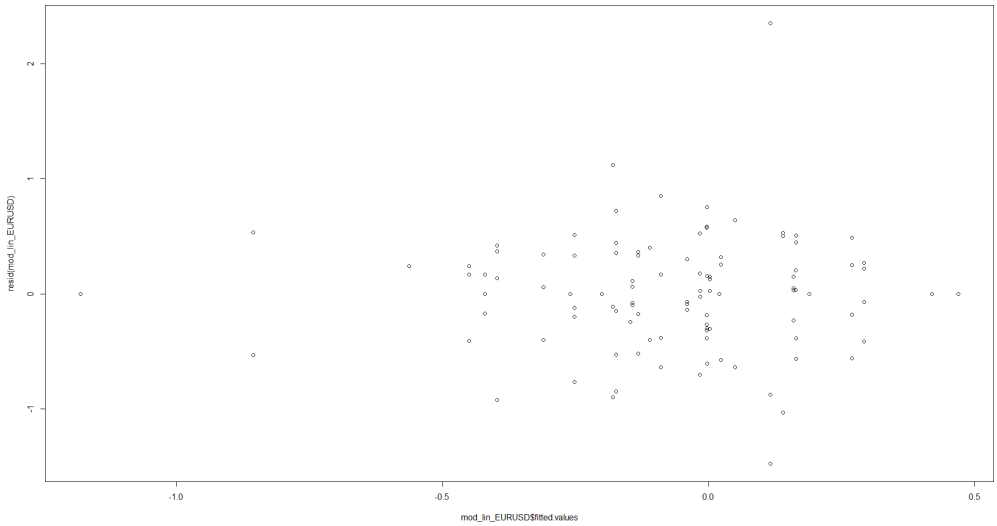
```
OLS estimation, Dep. Var.: IT10Y
Observations: 105
Fixed-effects: year: 13
Standard-errors: Clustered (year)
                                        Estimate   Std. Error       t value  Pr(>|t|)
topic_label_facAsset_Convergence        20.763333 1.732000e-14  1.198667e+15 < 2.2e-16 ***
topic_label_facAsset_Strategies          2.583333 7.810000e-15  3.305760e+14 < 2.2e-16 ***
topic_label_facBanks_and_Outlook        13.260000 5.899000e-14  2.247926e+14 < 2.2e-16 ***
topic_label_facContainment_Measures     -5.798148 2.563253e+00 -2.262028e+00  0.043053 *
topic_label_facDecision_Making           0.650000 3.820000e-15  1.701188e+14 < 2.2e-16 ***
topic_label_facEconomic_Measures        11.675000 6.687000e-14  1.745933e+14 < 2.2e-16 ***
topic_label_facEnergy_Prices            -7.148148 2.563253e+00 -2.788702e+00  0.016384 *
topic_label_facFinancial_Analysis        6.660000 6.361000e-14  1.047085e+14 < 2.2e-16 ***
topic_label_facFinancial_Conditions     11.260000 5.996000e-14  1.877989e+14 < 2.2e-16 ***
topic_label_facFinancial_Sustainability 23.410000 6.561000e-14  3.568209e+14 < 2.2e-16 ***
topic_label_facGDP_Projections          -3.700000 7.380000e-15 -5.014767e+14 < 2.2e-16 ***
topic_label_facInflation_Dynamics       -0.931481 2.563253e+00 -3.633980e-01  0.722629
topic_label_facInflation_Expectations    8.160000 5.899000e-14  1.383362e+14 < 2.2e-16 ***
topic_label_facInflation_Targets        18.902500 6.828000e-14  2.768202e+14 < 2.2e-16 ***
topic_label_facLTRO_III_Impact          -2.435648 2.563253e+00 -9.502180e-01  0.360748
topic_label_facMacroeconomic_Outlook    17.660000 5.897000e-14  2.994514e+14 < 2.2e-16 ***
topic_label_facMarket_Dynamics           2.910000 7.549000e-14  3.854740e+13 < 2.2e-16 ***
topic_label_facNon_Financial_Impact     -2.340000 1.391000e-14 -1.682563e+14 < 2.2e-16 ***
topic_label_facPrice_Pressures          -7.190000 4.540000e-14 -1.583860e+14 < 2.2e-16 ***
topic_label_facRefinancing_Trends        7.937500 7.860000e-15  1.009247e+15 < 2.2e-16 ***
topic_label_facReform_Impact            -1.318519 8.971384e+00 -1.469690e-01  0.885597
topic_label_facStructural_Impact         4.210000 1.387000e-14  3.034811e+14 < 2.2e-16 ***
topic_label_facStructural_Stability     18.058333 1.577000e-14  1.144805e+15 < 2.2e-16 ***
draghi                                 -11.700000 3.960000e-15 -2.956375e+15 < 2.2e-16 ***
... 2 variables were removed because of collinearity (topic_label_facStaff_Analysis and topic_label_facSustainable_Adjustments)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 7.95393     Adj. R2: -0.039806
                Within R2:  0.249016
```

Linear Regression Results – DE 10 Y

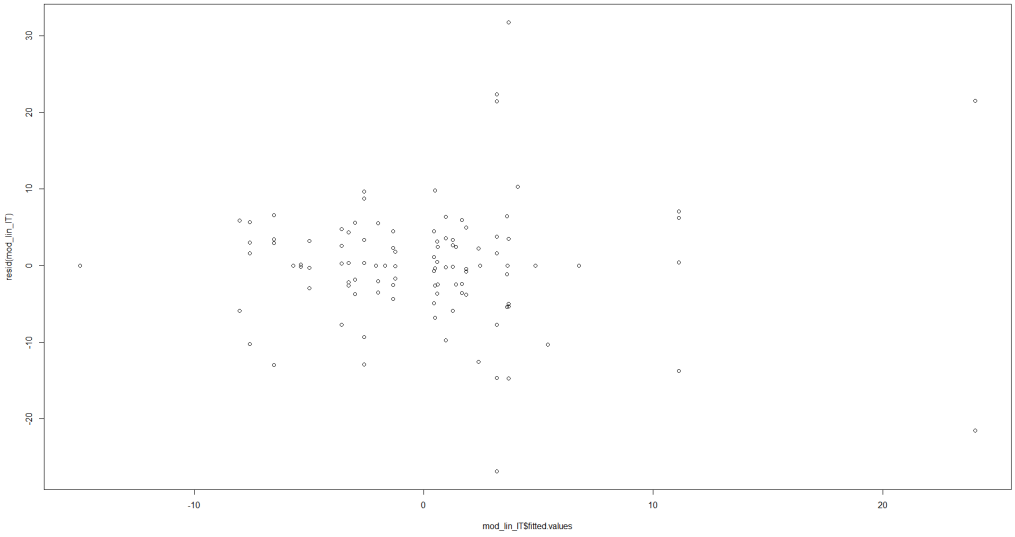
```
OLS estimation, Dep. Var.: DE10Y
Observations: 105
Fixed-effects: year: 13
Standard-errors: Clustered (year)
                                        Estimate   Std. Error       t value   Pr(>|t|)
topic_label_facAsset_Convergence        16.663333 7.361000e-15  2.263870e+15  < 2.2e-16 ***
topic_label_facAsset_Strategies          1.383333 3.831000e-15  3.610934e+14  < 2.2e-16 ***
topic_label_facBanks_and_Outlook        -0.200000 1.711100e-14 -1.168872e+13  < 2.2e-16 ***
topic_label_facContainment_Measures     -0.666667 4.185372e-01 -1.592849e+00 1.3718e-01
topic_label_facDecision_Making           0.937500 1.279000e-15  7.330467e+14  < 2.2e-16 ***
topic_label_facEconomic_Measures        -1.158333 1.751800e-14 -6.612290e+13  < 2.2e-16 ***
topic_label_facEnergy_Prices            -2.000000 4.185372e-01 -4.778548e+00 4.4967e-04 ***
topic_label_facFinancial_Analysis       -3.650000 1.373400e-14 -2.657717e+14  < 2.2e-16 ***
topic_label_facFinancial_Conditions     -1.400000 2.776100e-14 -5.043037e+13  < 2.2e-16 ***
topic_label_facFinancial_Sustainability -1.200000 4.906000e-15 -2.446231e+14  < 2.2e-16 ***
topic_label_facGDP_Projections           3.900000 6.700000e-16  5.817615e+15  < 2.2e-16 ***
topic_label_facInflation_Dynamics       -2.466667 4.185372e-01 -5.893542e+00 7.3257e-05 ***
topic_label_facInflation_Expectations   -4.300000 8.721000e-15 -4.930641e+14  < 2.2e-16 ***
topic_label_facInflation_Targets         1.729167 2.212600e-14  7.814927e+13  < 2.2e-16 ***
topic_label_facLTRO_III_Impact           0.870833 4.185372e-01  2.080659e+00 5.9553e-02 .
topic_label_facMacroeconomic_Outlook     1.566667 1.734100e-14  9.034673e+13  < 2.2e-16 ***
topic_label_facMarket_Dynamics          -6.200000 9.265000e-15 -6.691704e+14  < 2.2e-16 ***
topic_label_facNon_Financial_Impact     -2.750000 5.473000e-15 -5.024361e+14  < 2.2e-16 ***
topic_label_facPrice_Pressures           1.137500 4.978000e-15  2.285225e+14  < 2.2e-16 ***
topic_label_facRefinancing_Trends        3.325000 3.706000e-15  8.970931e+14  < 2.2e-16 ***
topic_label_facReform_Impact             0.066667 1.464880e+00  4.551000e-02 9.6445e-01
topic_label_facStructural_Impact         3.100000 4.628000e-15  6.698617e+14  < 2.2e-16 ***
topic_label_facStructural_Stability     15.145833 4.498000e-15  3.367310e+15  < 2.2e-16 ***
draghi                                  -2.350000 1.355000e-15 -1.733949e+15  < 2.2e-16 ***
... 2 variables were removed because of collinearity (topic_label_facStaff_Analysis and topic_label_facSustainable_Adjustments)
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
RMSE: 3.61626     Adj. R2: -0.105107
                Within R2:  0.186554
```



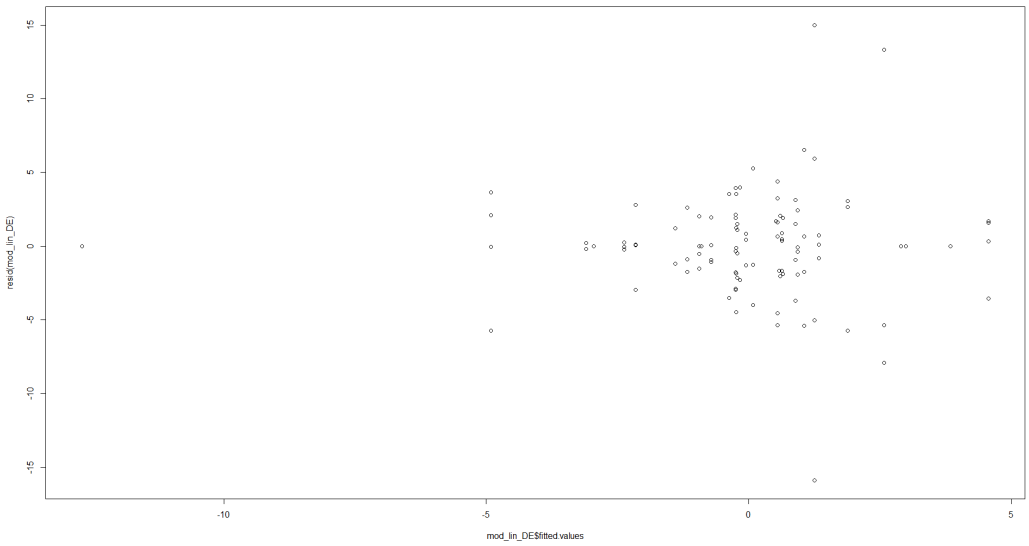
Appendix Chart. 4 Residual error, OIS 10 Y



Appendix Chart. 5 Residual error, EUR – USD



Appendix Chart. 6 Residual error, IT 10 Y



Appendix Chart. 7 Residual error, DE 10 Y