DISCUSSION PAPER SERIES

DP17931

MEASURING THE TEMPORAL DIMENSION OF TEXT: AN APPLICATION TO POLICYMAKER SPEECHES

David Byrne, Robert Goodhead, Michael McMahon and Conor Parle

MONETARY ECONOMICS AND FLUCTUATIONS

CEPR

ISSN 0265-8003

MEASURING THE TEMPORAL DIMENSION OF TEXT: AN APPLICATION TO POLICYMAKER SPEECHES

David Byrne, Robert Goodhead, Michael McMahon and Conor Parle

Discussion Paper DP17931
Published 21 February 2023
Submitted 21 February 2023

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

MEASURING THE TEMPORAL DIMENSION OF TEXT: AN APPLICATION TO POLICYMAKER SPEECHES

Abstract

Discussions of time are central to many questions in the social sciences and to official announcements of policy. Despite the growing popularity of applying Natural Language Processing (NLP) techniques to social science research questions, before now there have been few attempts to measure expressions of time. This paper provides a methodology to measure the "third T of Text": the Time dimension. We also survey the techniques used to measure the other Ts, namely Topic and Tone. We document key stylised facts relating to temporal information in a corpus of policymaker speeches.

JEL Classification: C55, C80, E58

Keywords: Textual analysis, Communication

David Byrne - david.byrne@centralbank.ie
Central Bank Of Ireland

Robert Goodhead - robert.goodhead@centralbank.ie
Central Bank Of Ireland and CEPR

Michael McMahon - michael.mcmahon@economics.ox.ac.uk
Department of Economics, University of Oxford and CEPR

Conor Parle - conor.parle@ecb.europa.eu
European Central Bank

Acknowledgements
McMahon gratefully acknowledges financial support from the European Research Council (Consolidator Grant Agreement 819131). We thank Mengqi Qiu and Delia Macaluso for excellent research assistance. We have made code available to replicate our approach to temporal tagging. Please see author webpages. The views presented in this paper are those of the authors alone and do not represent the official views of the Central Bank of Ireland, ECB, or the European System of Central Banks. Any errors remain ours alone.

Measuring the Temporal Dimension of Text: An Application to Policymaker Speeches*

David Byrne[†], Robert Goodhead[‡], Michael McMahon[§] and Conor Parle[¶]

February 17, 2023

Abstract

Discussions of time are central to many questions in the social sciences and to official announcements of policy. Despite the growing popularity of applying Natural Language Processing (NLP) techniques to social science research questions, before now there have been few attempts to measure expressions of time. This paper provides a methodology to measure the "third T of Text": the Time dimension. We also survey the techniques used to measure the other Ts, namely Topic and Tone. We document key stylised facts relating to temporal information in a corpus of policymaker speeches.

Keywords: Textual analysis, Machine Learning, Communication

JEL Codes: C55, C80, E58

*McMahon gratefully acknowledges financial support from the European Research Council (Consolidator Grant Agreement 819131). We thank Mengqi Qiu and Delia Macaluso for excellent research assistance. We have made code available to replicate our approach to temporal tagging. Please see author webpages. The views presented in this paper are those of the authors alone and do not represent the official views of the Central Bank of Ireland, ECB, or the European System of Central Banks. Any errors remain ours alone.

†Central Bank of Ireland. Email: david.byrne@centralbank.ie

‡Central Bank of Ireland. Email: robert.goodhead@centralbank.ie

§University of Oxford, CEPR, CfM (LSE). Email: michael.mcmahon@economics.ox.ac.uk

¶European Central Bank and Trinity College Dublin. Email: conor.parle@ecb.europa.eu

1 Introduction

In recent years, the use of unstructured data derived from human speech has grown substantially in the social sciences, particularly in economics. Researchers have leaned on innovations in the Natural Language Processing (NLP) literature to quantify dimensions of speech and use these to explain economically important phenomena, such as responses to announced policy changes. To date, these studies have largely focused on what we shall call the first two "Ts" of text: **Topic** and **Tone**. That is, a measurement of what is being discussed and a measurement of some sentiment being expressed. For example, NLP tools have been used to quantify the tone of text in corpora relating to firm statements (Tetlock 2007), to isolate references to uncertainty (Baker et al. 2016), and to study how both the topic and tone of central bank communication impacts the macroeconomy (Hansen and McMahon 2016). The present paper focuses on methods to measure the third "T": **Time**. This dimension of text is central to the understanding of a wide range of economic mechanisms, and yet its measurement has been under-studied before now due to practical and conceptual limitations. Our study fills that gap by synthesising a number of existing NLP methodologies into an algorithm applicable to economic questions.

The ability to programmatically extract temporal references is a very useful one for economists to possess, since discussions of economic problems abound with references to time. Questions of intertemporal allocation of resources can be traced back to the birth of the discipline of economics (Rae 1834). References to time will permeate any discussion of risk, asset pricing, economic or financial cycles, growth, social discount rates, and any evaluation of policy change. There are few questions in the social sciences more broadly that do not relate to time, in some fashion. Moreover, when policymakers undertake public communications as part of their response to economic and social questions, the information conveyed will have particular temporalities.

In this study, we quantify the use of temporal references in speeches by policymakers in multiple economic and political institutions. Since the sources of our data are derived from human language, our work forms part of an increasing body of research seeking to bring unstructured data to bear on the questions of social science. Unstructured data can be understood as any dataset that does not conform to conventional standards of organisation. Handling large volumes of unstructured data is becoming increasingly tractable, thanks to advances in the NLP literature, and to large increases in computational capacity in recent decades; the impact of unstructured data on economic research is likely only just beginning to manifest itself.

Of course, when dealing with unstructured data, researchers face an important data extraction problem. They must use specialised tools or techniques to filter desired information from text. This data extraction problem is particularly challenging for temporal references. Policymakers may invoke temporality in various linguistically and conceptually distinct ways, and varying degrees of context are necessary to successfully parse the date to which a speaker is referring. We can generally characterise temporal references in speech as being either categorical, numerical, or grammatical. Policymakers frequently make categorical references to time, which do not allow identification of a specific date, only a general orientation. Examples of categorical references would be "in the future", or "presently". Numerical references to time can be mapped to a clear calendar date, and trivial examples would be a reference to "June 2008" in a given piece of text. However, numerical references can come in forms that are harder to parse, such as "in February", which requires a mapping from text to a reference date (which could be ambiguous). Finally, grammatical references to time involve the use of verbal complexes that are interpreted as past, present, and future tenses.[1] Grammatical references to time can be rich with ambiguity, for example a statement such as "next year, we are enacting legislation" employs the present tense, and yet is a reference to the future. In this case, we learn the temporal orientation of the sentence from the numerical reference, not the grammatical reference.

In our study, we synthesise several existing tools from the NLP literature to quantify categorical, numerical, and grammatical temporal references in our corpus. We employ the SUTime algorithm of Chang and Manning (2012) to extract categorical and numerical references, and the TMV tool of Ramm et al. (2017) to tag cases of the future tense.[2] We also make a number of bespoke adjustments to the tools, to reflect the particularities of the communications by policymakers in our corpora. We argue that the temporal orientation of policymaker speeches is best summarised by a combination of the three measures, and we develop statistics to assess overall temporal orientation with respect to the past, present and future.

In the study of Byrne et al. (2023) we applied our measurement approach to textual datasets relating to central bank communication. We used such measures to quantify the evaluation and projection phases of central bank assessments of the macroeconomy, and found that both phases are an important source of news to financial market participants. The present study is designed to be a companion paper to that of Byrne et al. (2023), where we elaborate on the structure of the algorithms used, and any adjustments made,

[1]We focus on the English language in this study. Although English does not have an inflectional future tense as is found in other languages, such as French, we follow Ramm et al. (2017) in parsing certain verbal complexes into a future tense. We shall use the phrase "future tense" to refer to this parsed output throughout and we trust that the meaning is clear.

[2]The English language does not have a future tense. The "future tense" is an output of the TMV rules-based approach, designed to attribute temporal orientation to grammatical constructions used in English to convey future orientation.

in greater detail. In this study, we are also able to document a much broader array of descriptive statistics and stylised facts, to convey more information on the essential nature of how policymakers use temporal references in their discussions. In Byrne et al. (2023), we examined Federal Reserve (Fed) and ECB policy statements, press conferences, and speeches. In this study we will confine ourselves to the analysis of Fed and ECB speeches. We also broaden our corpus to include speeches by other types of policymaker. Specifically we use data from the EUSpeech dataset of Schumacher et al. (2016). These authors collect data for policymaker speeches from five institutions: the European Commission, the European Parliament, the European Council, the IMF, and the ECB.[3]

Our approach is flexible and would be straightforward to adapt to other corpora. We hope that part of the usefulness of this document will be to popularise the quantification of temporal information in the economics literature. We aim to make clear the decision-making process we followed when considering how best to use NLP methods to extract temporal references from text. It is important, however, to emphasise that in certain cases we have adjusted our algorithms to target specific features of how central banks communicate about time. For example, we capture key historical dates of particular interest to economists, such as the "Great Depression", as will be discussed. These features are also likely to be useful for measuring temporal references in the speeches of other economic policymakers. Other researchers may find it useful to adapt our approach, or even to overhaul it wholesale, depending on their context. We have aimed to make the discussions in this document of sufficient technical detail to represent a useful point of reference for social scientists in the quantification of the three Ts, and particularly time, in text.

The rest of the paper is structured as follows. In Section 2, we survey the existing literature on the three Ts of textual analysis. We then outline our NLP methodology to measure the temporal dimension of text through temporal tagging (Section 3) and tense tagging (Section 4). In Section 5 we present the the text corpora in our dataset. Section 6 presents our results on how policymakers express temporality in their communications. Section 7 concludes.

2 Literature Review

Our study employs techniques derived from the NLP literature, and applies them to policymaker communication. Research into NLP has expanded greatly in recent years,

[3]We prefer to use our dataset regarding ECB speeches (extracted from the corpus made available on the website of the ECB), rather than that contained within EUSpeech, for consistency with the treatment in Byrne et al. (2023). EUSpeech data also include data for national politician speeches that we do not study in this paper.

facilitated by continued expansion of computational power, as well as improved databases of textual content. For introductory treatments to NLP, one can consult Manning et al. (2008) and Jurafsky and Martin (2009). NLP methods are now used widely across industries and academic disciplines, including the social sciences. One can chart many influential applications of NLP in sociology and political science (see Evans and Aceves 2016 and Grimmer and Stewart 2013 for respective literature reviews). For a general overview of the use of NLP in economics, Gentzkow et al. (2019) provide a recent guide.

Within economics in particular, applied NLP studies have focused to date chiefly on topic and tone. The purpose of our study is to introduce the third T, time, to the economic literature. In this section we summarise the NLP literature across these three categories to date, and discuss relevant applications of these tools within economics.

2.1 Tone

Within the NLP literature, tone quantification can be broadly classified according to two methods: lexicon/dictionary based approaches and those that utilise machine learning (ML). Lexicon approaches generally involve the use of a dictionary to classify sentences into positive/negative sentiment (or indeed broader concepts such as uncertainty, or the hawkishness/dovishness of statements within the monetary policy literature). The lexicon based approach has been used for some time in fields such as psychology or sociology, with an early dictionary being that of Stone et al. (1966). Many sentiment dictionaries are now available, including those of Nielsen (2011), Hu and Liu (2004), Mohammad and Turney (2013), Mohammad (2018), Young and Soroka (2012). More sophisticated lexicon based methods make adjustments for the context of sentences, such as negation or punctuation, such as the SoCAL method (Taboada et al. 2011) or VADER (Hutto and Gilbert 2015).

The other stem of the tone literature involves the adaptation of ML methods for the classification of text into different categories of sentiment. These methods use approaches such as support vector machines, random forests and naïve bayes to classify bodies of text. These methods can be seen in the works of Lagrari et al. (2019), Crawford et al. (2015) and Al Amrani et al. (2018). Other machine learning influenced approaches include the BERT algorithm (Devlin et al. 2019) and the use of word embedding approaches such as Word2Vec (Mikolov et al. 2013) and GloVe (Pennington et al. 2014) for sentiment analysis.[4]

Tone extraction methods have been applied to great success within both finance and economics. Some early approaches involved the manual coding of text to measure tone,

[4]Further methods along these lines can be seen in the treatments of Pang et al. (2002), Wang and Manning (2012), Wilson et al. (2005a), Wilson et al. (2005b), Bradley and Lang (1999).

such as those of Ehrmann and Fratzscher (2009), Picault and Renault (2017) and Rosa and Verga (2007). Within finance, an influential initial attempt to create a specialist dictionary can be seen in the work of Loughran and McDonald (2011). A broad review of the literature can be found in Loughran and McDonald (2020). While it is possible to apply pre-existing dictionaries such as that of Loughran and McDonald (2011) to economic corpora (see for example Schmeling and Wagner 2019), many researchers have moved further and created economics or central banking specific dictionaries. A common approach has been to count hawkish and dovish keywords to measure monetary policy tilt, or the expansionary/contractionary nature of economic language (Parle 2022, Shapiro et al. 2022, Hubert and Labondance 2021, Hansen and McMahon 2016, Bennani and Neuenkirch 2017, Apel et al. 2022). Other applications have used dictionaries to measure uncertainty within texts (Baker et al. 2016, Hassan et al. 2019, Caldara and Iacoviello 2022). Non-dictionary approaches have also been increasingly seen, with early work applying pre-trained classification algorithms to measure the sentiment of internet stock market message boards seen in both Tetlock (2007) and Antweiler and Frank (2004). Similar non-dictionary approaches were seen in the work of Lucca and Trebbi (2009) and Tobback et al. (2017), while recent deep learning methods were seen in the neural network approach of Gorodnichenko et al. (2021).

2.2 Topic

The measurement of the topic of a body of text has become increasingly advanced in the machine learning literature. An early approach to the problem was the Latent Semantic Analysis (LSA) method of Deerwester et al. (1990), which grouped documents together based on their latent semantic structure. A key advancement in this field was the Latent Dirichlet Allocation (LDA) methodology (Blei et al. 2003), a three level hierarchical Bayesian model, in which documents are classified into an underlying set of topics (groups of words or n-grams). Other methods have extended this algorithm, such as the hierarchical topic model (hLDA, Blei et al. 2004) and dynamic topic models (DTM, Blei and Lafferty 2006) approaches. In more recent times, in line with the tone literature, advanced machine learning methods have allowed the further development of detailed topic modelling approaches. The BERT algorithm has, for example been adapted by Grootendorst (2022) to create the BERTopic approach. Other methods have leveraged word embedding algorithms to create clusters of topics with a notable example being the LDA2vec algorithm of Moody (2016). Some recent methods have combined machine learning and econometrics for text selection, such as that of Kelly et al. (2021)

The use of topic modelling within the analysis of central bank communication and the broader economic literature has continued to grow. Parle (2022) uses the dynamic topic modelling approach to create a dynamic measure of tone for each ECB president, while Cross and Greene (2020) use a non-negative matrix factorization approach is used to analyse ECB communication. Similarly, Hansen and McMahon (2016), Hendry and Madeley (2010) and Bybee et al. (2020) use standard LDA models to examine Bank of England inflation reports, Bank of Canada communications and Wall Street Journal articles respectively. Similarly, Istrefi et al. (2021) use an LDA approach to create a speech based measure of financial stability, while Aguilar and Pérez-Cervantes (2022) use a GloVe based approach to classify documents issues by the Mexican Central Bank.

2.3 Time

As already emphasised, within economics and finance, the time dimension is the least explored thus far of the three Ts of textual data. The extraction of references to time from language has been an active area of research in the NLP literature. Developments in the extraction of temporal information from language benefited from the creation of a schema for the annotation of temporal references, TimeML (Pustejovsky et al. 2003a). The TIMEBANK corpus of Pustejovsky et al. (2003b) provided an early annotated corpus according to the TimeML scheme. TimeML was the annotation scheme used for SemEval workshops involving multiple research teams foccussed on evaluating temporal expressions (Verhagen et al. 2007, 2009, 2010, UzZaman et al. 2013).

One can broadly divide studies focussed on extracting measures of time into those that use "rules-based" methods, and those that follow a statistical (ML) approach. Approaches can be designed to extract numerical references to time (e.g. "December 2008"), categorical references to time ("in the near future"), and use of the future tense. Rules-based approaches utilise lengthy lists of rules, that summarise how one typically discusses time in a given language. The TempoWordNet approach of Dias et al. (2014) extends the WordNet approach to create a temporal classifier based on three categories: past, present and future while other approaches can be seen in the work of Hafez et al. (2017) and Strötgen and Gertz (2010). An important example of a rules-based approach to the extraction of numerical and categorical temporal references is the SUTime algorithm of Chang and Manning (2012), while rules based classifier that identifies tense can be seen in the Tense, Mood, Voice (TMV) algorithm of Ramm et al. (2017). Our paper concentrates on the application of SUTime and TMV, but alternatives for measuring tense can be seen in the work of Palmer et al. (2005), Loáiciga et al. (2019) and Myers and Palmer (2019).

While temporal taggers have been applied in other fields in recent years, for example the study of social media data (Yan et al. 2011, Li and Cardie 2014, Tabassum et al. 2016, Kamila et al. 2019), applications in the field of economics are rare to date. We are aware of a small number of studies such as Galardo and Guerrieri (2017) and Coenen et al. (2017) that use simple counts of instances of "may" or "might" to measure the temporal orientation of text.

3 Temporal Tagging with SUTime

In order to quantify explicit references to time, we need a way to process our textual data and accurately isolate such references within a large corpus of documents. To achieve this, we employ the SUTime temporal tagger developed in Chang and Manning (2012). SUTime is a rules-based approach, and therefore does not employ a trained statistical model. Chang and Manning (2012) show that SUTime performs well in comparison to other temporal taggers in the NLP across a number of criteria. This evidence supports the rules-based approach for capturing time in text. SUTime is available as a part of the Stanford CoreNLP pipeline for NLP.

SUTime is built on regular expression patterns. That is, the algorithm searches strings of text for occurrences of certain sequences of characters of interest (so called "regular expressions"). For instance, regular expressions could be "June 2020" or "2020-06".[5] Rules can then be defined such that some output is returned whenever a particular regular expression is found. For the examples just given, the temporal tagger could return the same output showing that the speaker was referring to the sixth month of 2020.

A key benefit of SUTime is that it can tag a wide range of representations of time, allowing greater accuracy and fealty to the typical ways in which central bankers communicate. SUTime is not limited to absolute date formats such as YYYY/MM/DD, "June 2020" or similar simple temporal references. SUTime is also able to resolve relative date formats, such as "last Friday" or "two months from now", since the processor takes a reference date as an input. In cases of ambiguity as to the relative date, SUTime can also use grammatical tense to help resolution (Chang and Manning 2012).

The outputs of SUTime are temporal tags in the TIMEX3 format (Pustejovsky et al. 2003a). For an example of the output of the SUTime processor, see Table 4, which takes sentences from a representative introductory statement, and shows the TIMEX3 tags generated. Text that can be resolved to a specific date will result in a numerical time tag. For example, "June 2020" or "next June" are numeric dates. For more abstract date formats that do not resolve to a specific date, SUTime also produces categorical tags, covering the three general categories: past, present, or future. Examples of this include expressions such as "in the future" or "the current situation". SUTime will also identify

[5]SUTime builds upon the TokensRegex framework for analysing regular expressions in tokenized text as outlined in Chang and Manning (2014).

if the text is referring to a range of dates from one point to another, or a duration such as "for three months". We do not incorporate information from ranges in our approach, since it is unclear how to resolve such expressions into a single value (one could use the middle of the range, though it is not obvious whether such expressions should be treated in the same manner as dates, so we prefer to omit these cases).

Although the library of rules in SUTime is large, it is not tailored to the language of monetary policy or central banking. In our study make a number of additions to the standard SUTime rules, in order to best reflect the context in which we apply the tool, namely central bank communication. We implemented these amendments by: (1) pre-parsing the textual data entering the SUTime algorithm in a certain way; (2) by editing the rules applied by SUTime; (3) post-processing the output from the SUTime tagger.

The first addition to the standard SUTime routines was designed to handle frequently occurring references to events from economic history. Central bank communication often refers to dates, times and eras by commonly understood shorthand names. The audience hearing "Great Depression", "Bretton Woods era" or "Global Financial Crisis", for instance, is likely to know well to which point in time the speaker is referring. A temporal tagger, on the other hand, would not. As a result, it would not produce a time tag for these references, reducing the overall accuracy of our exercise, and biasing measures of time orientation toward the future if these shorthand references disproportionately refer to the past. We developed a list of relevant textual date expressions in economics and map them to numerical dates, allowing SUTime to process them (see Table 3). We chose to replace phrases identified as historical references with numeric dates *before* we applied the SUTime algorithm, for convenience.[6] We only performed this operation on the textual data prior to the application of SUTime, i.e. no such replacement was conducted prior to the application of TMV, or topic model estimation.

The second addition we made to the standard SUTime rules was to broaden the set of categorical temporal references understood by the tagger. This decision was in response to the existence of categorical temporal references that are regularly used by central bankers and easily understood by their audiences, but which are not covered by the standard SUTime rules. For instance, central bankers frequently refer to the "short-term", "long-run", and similar constructions. These expressions are however not included in the SUTime library. We thus expand SUTime's library of rules to capture better the ways in which central bankers speak about monetary policy or economics-related topics. These additional rules are listed in Table 2.

Though the addition of these central bank-specific rules to SUTime appears to capture a typical form of communication about the future, that is not included in the standard

[6]Obviously one could code additional rules for the SUTime algorithm to follow in these cases.

version of SUTime, our rules do lead to a certain complication. While it is common for a central bank to refer to "long-run growth", or a similar expression, they may also refer to concepts such as "long-run debt" (or "long-run yields", "long-run bonds", etc.). A reference to "long-run debt" clearly does contain a form of temporal reference, however we view these references as operating more akin to a form of proper noun ("long-run debt") as opposed to an indication of how a speaker views events transpiring in the future. We choose to remove such expressions from the output of our (amended) SUTime rules post processing.[7]

Another issue raised by our approach to identifying numerical temporal references in examples of central bank communication is the question of academic references, which are prevalent in the speeches. For example, many speeches contain references of the form "J. Doe (1999)". SUTime would record this as a past numerical date reference, "1999", however we do not wish to include such cases, which do not relate to any particular temporal signal the central banker might wish to send. For this reason we cut out all such date references (i.e. those that are academic references) from our output.[8]

Given that we have processed our textual data with SUTime, we are in a position to assign past, present, and future tags to identified temporal references in our corpora. For the case of numerical temporal references, to assign these categorical tags and reduce variation offered by a continuous measure into three broad categories. We take a given numerical time reference, and subtract the day of this speech from this value, to get a relative date. We code positive relative numerical date references as future, negative relative numerical date categorical references as past, and exact date references to the day of the speech as present. For example, if a speaker makes reference to the "1st of January 2009", and the speech was made on the 1st of January 2008, then the SUTime numerical reference will be 364 days. Note that we undertake this operation at daily frequency, so

[7]We do this by looping over categorical temporal references identified by SUTime, locating their position in the original corpus, and then ascertaining (using regular expressions) whether a reference to the "long-run" was followed by the token "bond". Specifically, we excerpt the cases that follow: References to "short run/long run" followed by the any of the tokens "rate/interest rate/bond/debt"; References to "long run" followed by "unemployed/unemployment"; References to "short run" followed by "rate/interest rate/bond/debt/money market/paper/corporate"; References to "medium run" followed by "rate/interest rate/debt". We cut references to "overnight" (which is a SUTime reference to the present) followed by any of the tokens: "rate/interest rate/repo/repurchase agreements/market rate/market". In these cases we account for basic semantic variation such as the use of "long term" or "longer run" in place of "long run", etc. We found that SUTime would identify "current account" as a reference to the present, on account of the word "current", and excerpted these cases. We also cut references to "longer term refinancing operations", which became particularly frequent in the post-crisis era, since they comprised part of the suite of unconventional monetary policies deployed by the ECB.

[8]We achieved this by looping over all numerical date references tagged by SUTime of the form YYYY, and ascertaining whether this date reference was enclosed within parentheses in the corpora to which we applied SUTime (we also account for lists of references within parentheses). Thankfully, academic references to publication dates are almost always made by enclosing the date in parentheses in our Fed and ECB documents, making their extraction straightforward.

there is no distinction between "this afternoon" and "this evening", both are references to the present. Note that if a numerical reference is to a year, and no month or day is specified (such as "2008"), we convert this to an exact day by assigning it to the first day of the first month (so "2008" becomes the 1st of January 2008). If a numerical reference is to a given month and year (for example, "March 2008"), then we assign this value to the first day of a the specified month (so "March 2008" becomes the 1st of March 2008).

It would be of interest to incorporate a more refined measure of numerical future orientation, that accounts the differences in the horizon of past and future references, but we leave this for future work. We have also experimented with more granular measures of the future orientation using the categorical references. Taking our collection of categorical future tags, we apply rules to map these tags to categories of short, medium, and long-run future. We also produce a final measure of the "ambiguous future" as a residual future category, which includes simple references to the "future". This split of categorical future references allows us to examine the broad future horizon about which the central banker is communicating, and whether this changes over time.

4 Tense Tagging with TMV

Applying the methods of computational linguistics to assess whether phrases within a given sentence are in the past, present, or future tenses is a non-trivial task. Standard computational tools allow one to assign "part of speech" (POS) tags to words from corpora of textual data. The tags themselves come from a list of potential word classes for the English language (nouns, verbs, determiners, etc.). The widely-used Penn TreeBank Tagset is an example of such a list of word classes. Given a catalogue of potential word classes, a POS tagger will assign words (tokens) to its appropriate class. The Stanford POS tagger is an example of a POS tagger that uses the Penn TreeBank Tagset (Toutanova et al. 2003). For example, this algorithm will tag the token "cat" with the class "noun", the token "eat" will be tagged as the class "verb base form". A similar POS tagger is included with the Natural Language Toolkit for the Python programming language. POS taggers are therefore widely available, and straightforward to implement.

However, because POS taggers are applied to tokens, and not verb phrases, their ability to detect tense is necessarily limited. A POS tagger using the TreeBank Tagset would be sufficient to identify verb phrases in the past tense, such as "I ate", since it would tag "I" with "PP" (personal pronoun) and "ate" as "VBD" (verb past tense). However, a standard POS tagger cannot determine whether sentences are in the future tense. For example, such a tagger would separate the sentence "I will go" into its three tokens, then the token "I" would be tagged as "PP" (personal pronoun), the token "will" as “MD” (modal auxiliary), and “go” as “VB” (verb base form). The tagger is not able to determine that the sentence “I will go” is in the simple future tense, since it takes as arguments only individual tokens. Therefore, POS taggers are necessary, but not sufficient for the identification of the future tense.

To identify the future tense, we therefore need to incorporate additional tools from the computational linguistics literature. To identify tense, this study applies the Tense-Mood-Voice tool, introduced by Ramm et al. (2017). This tool is designed to automatically classify verbal complexes (sequences of verbal tokens within a verbal phrase) into their tense, according to a rules-based method. Ramm et al. (2017) distinguish semantic tense from morphosyntactic tense. These authors give the example of the English sentence “He is leaving at noon”, which is semantically in the future tense, but has the morphosyntactic tense of present progressive. The TMV tool can only provide information about the morphosyntactic tense. The system takes as its argument individual sentences. It then identifies verbal complexes from these sentences, before applying a sequence of around 32 rules to these verbal complexes. For example, the system understands that, for the simple future tense, the model auxiliary “will” (or “shall”) precedes the infinitive form of the verb, so “I will go” is correctly identified as the future tense.

The system assigns verbal complexes to four forms of the present tense (present, present progressive, present perfect, present perfect progressive), four forms of the past tense (past, past progressive, past perfect, past perfect progressive), and four forms of the future tense (two respective forms of the future and future progressive tenses are identified). As well as tense, the tool also differentiates between the indicative and subjective moods, and the active and passive voices, though these aspects of verbal complexes are not studied in our paper. A distinction is also made between finite verbs, and non-finite verbs (which includes infinitives and gerunds). For an explanation of how these tenses differ in the English language, and how the TMV allocates the verb complexes according to its schema of tenses, see Table 1, which is a replication of Table 1 from Ramm et al. (2017).

The TMV tool takes as its input sentences that have already been assigned POS tags, and the format must be in the CoNLL form. Therefore, before running TMV, we apply the same POS parser used by Ramm et al. (2017), namely the MATE parser of Björkelund et al. (2010), which is implemented in Java language. The TMV tool itself is implemented in Python. To see an example of the output from the TMV tool, Table 5 shows the output when TMV is applied to the famous “whatever it takes” speech of ECB President Mario Draghi on July 27th 2012.

When applying the TMV tool, we do not distinguish between different forms of present, future or past tenses. We assign the four possible future tenses to a general future tense category, and likewise for the present and past tenses. We assign the two conditional tenses that are about the past to the past tense category, and we assign the two conditional future tenses to the future category.[9] We do not consider non-finite verbal complexes.

TMV classifies sentences according to their tense, however there are certain expressions in the present tense that in fact refer to the future, for example "we expect". Central bankers frequently make statements using expressions such as "we forecast", "we predict" or "we project" for example. We therefore re-assign present tense verbal complexes that contain these types of verbs to the future tense. We do this on an *ad hoc* basis, after the application of the TMV tool, and do not claim to altered the TMV system in a way that could handle semantic tense generally. The idea is to ensure we capture certain turns of phrase that a policymaker may use regularly that indicate discussion of the future, and we created an initial list based on our knowledge of the sorts of phrases used in our corpora. Of course, a legitimate concern when making such amendments is that our re-assignments introduce some arbitrariness to our tense measure. To discipline our choices of such verbs, we use the TempoWordNet dataset of Dias et al. (2014) as a guide. The TempoWordNet dataset maps each word in the English dictionary to a probability distribution as to whether it is about the past, present, or future, where the probability distribution is computed according to the predictions of a trained model. In Table 6 we report the list of 28 verb forms we additionally assign to the future tense (when they are in the present tense), and we cross-reference these cases against the future probabilities reported in Dias et al. (2014). In each case our chosen verb complexes are associated with the future with high probability. Table 15 shows that the proportion of future tense reference tags that are generated as a result of our amendment to the TMV tool is low.

5 Textual Data

5.1 Introduction to the Data

We examine textual data based on central bank policymaker speeches from two sources: the Fed and the ECB. For the case of the ECB data come from the ECB Speeches Dataset, which was created by ECB staff and made available on its website.[10] These

[9]Explicitly, according to the schema in Table 1, we tag verbal complexes assigned by TMV to the tenses present, presProg, presPerf, and presPerfProg as "present tense". We tag verbal complexes assigned as past, pastProg, pastPefr, pastPerfProg, condII, and condIIProg as "past tense". We tag verbal complexes assigned as futureI, futureIProg, futureII, futureIIProg, condI, and condIProg as "future".

[10]The dataset can be found here: https://www.ecb.europa.eu/press/key/html/downloads.en.html.

data represent an archive of all speeches by ECB Executive Board members, dating back to February 1997.[11] These data are continually updated by the ECB, we used a version of the dataset ending with a speech on the 15th September 2020. Data from the Fed are manually downloaded from the website of the Federal Reserve Board, as well as those of the regional Reserve Banks. The Reserve Board speeches include both speeches by the President, and those of other members of the board. The speeches from the regional Reserve Banks are exclusively those from Governors.

We also include textual data from other forms of policymaker, in order to perform a broad evaluation of the use of temporal references in communication content from related institutions. Our data include speeches by the European Commission, European Council, European Parliament, and the IMF. These data are taken from the EUSpeech dataset of Schumacher et al. (2016).[12] In the next two sub-sections we give more detail on the cleaning procedures applied to our corpora.

5.2 Construction of the Corpora

ECB Speeches There are 2,412 speeches in our raw sample from the ECB Speeches Dataset. There are 16 speeches in the ECB Speeches Dataset for which there is no textual data available at all. The transcripts of these speeches do not appear on the ECB website, they often take the form of lecture slides (many of these took place during the Covid-19 period). These speeches are removed from the corpus. There are 34 unusually short speeches for which a small amount of textual data is provided, however the available text merely summarises the title of the speech and the name of the speaker, and provides a hyperlink to lecture slides – these are also discarded.

We remove 159 speeches that are not in English from our sample.[13] In principle, our methods to extract temporal information could be applied to textual data in languages other than English (the TMV tool we use will also work for German and French text). However we wish to maintain a level of linguistic similarity between speeches for the purposes of successful comparison, so we do not consider these speeches. We wish to avoid complications that could arise if different languages have structurally different properties with respect to the way in which users of these languages express time. These operations leave us with a corpus of 2,203 English speeches.

[11] The data include speeches delivered by senior officials prior to the formal creation of the ECB in June 1998.

[12] Note that the Schumacher et al. (2016) also includes speeches by ECB policymakers. We prefer to source ECB speech data from the ECB Speeches Dataset, since the sample-period is longer.

[13] All of the ECB press conferences are held in English.

Fed Speeches In our raw sample of Fed speeches we have access to a raw sample of 4,715 documents. We delete 147 speeches that are merely references to conferences, or lecture slides, and have no content, leaving a final corpus of 4,668 speeches.

EUSpeech Data We analyse speech data from the European Parliament, European Commission, European Council, and the IMF, extracted from the EUSpeech dataset, which was created by Schumacher et al. (2016). EUSpeech contains all publicly available speeches from these institutions for the period between 2007 and 2015, which were scraped from the websites of these institutions.[14] For those speeches not in English, Schumacher et al. (2016) offers versions translated with *Google Translate*, however we only use speeches originally in English in our sample.

5.3 General Cleaning Procedures

For consistency, we apply cleaning routines to each of our corpora in an identical fashion.[15] All of the textual data is broken into sentences using the default sentence tokenizer of the Natural Language Toolkit (NLTK) package for Python.[16] Tracking the sentence number in given documents is essential for matching output from the various computational tools we apply in this study. We further apply a word tokenizer to separate words from punctuation.[17]

In Table 10 we display basic summary statistics regarding the number of sentences per document. One can observe that the central bank, IMF, and European Council speeches are of similar lengths overall. The speeches by members of the European Commission, and particularly the European Parliament, are shorter.

[14] EUSpeech also contains speeches of prime ministers (or presidents) for 10 EU countries, however, to avoid dealing with the complication of language differences, we do not examine these speeches in this study.

[15] The only difference relates to annexes and table notes in the ECB speeches, which we clean out manually in some cases. Many of the ECB speeches contain sections at the end of the text devoted to academic references, charts, tables, and other annexes. Many of these annexes contain large numbers of temporal references, of the sort we do not wish to include in our study. We do this by searching for lines of text that begin with "REFERENCES:", "Table:", or "Figure:" (these are merely examples, we use a range of indicators and match based on regular expressions).

[16] See: https://www.nltk.org/. The default algorithm used by NLTK is the Punkt sentence tokenizer, which follows Kiss and Strunk (2006).

[17] We use the default work tokenizer from the NLTK toolbox, which is the Treebank tokenizer and uses regular expressions to tokenize text in the same manner as the Penn Treebank.

6 Temporal References and the Speech Corpora

In this section, we examine the output of our temporal algorithm applied across the corpora. We are interested first in the frequency with which policymakers make temporal references in their communications. Then we examine the orientation of the temporal references: are they about the past, present or future? We then produce temporality statistics at the document level within the corpora, i.e., for individual speeches. The former approach allows us to examine distributions of overall temporality in our corpora. The latter allows us to better examine the distributions of our temporality measures as they would be used jointly, i.e., per communication event.

6.1 The Frequency of References to Time in Policymaker Speeches

To begin, we compute summary statistics on the number of temporal expressions identified per document across corpora, as shown in Table 11. We observe that the two central bank speech corpora typically contain a greater average number of temporal expressions per document, relative to the other policymaker speeches. This is true for categorical and tense temporal references, though European Council speeches contain a comparably large number of numerical temporal references relative to those of the ECB. The standard deviation of the number of temporal references per document is also greater for the central bank speeches.

Of course, the greater number of temporal references may simply reflect that the central bank speeches tend to be slightly longer than those of other policymakers (see Table 10). It is of interest therefore to examine how frequently given sentences contain temporal references. Table 12 displays information on the fraction of sentences in given corpora that receive at least one temporal tag. When we examine temporal expressions identified by SUTime, we see that the central bank speeches contain a greater fraction of sentences containing either categorical or numerical expressions, relative to other policy makers. However, the European Commission speeches remain generally comparable to the central bank speeches, and the fraction of sentences containing categorical references specifically exceeds that of the Fed speeches. When we examine the union measure, it can be seen that a large proportion of sentences contain at least one temporal tag (of any form), and that this is true for all of the corpora. This high coverage is driven by the presence of the tense measure of temporal orientation, since this measure merely requires that the TMV algorithm identifies a verbal complex in the past, present, or future tense in given sentences.

The previous discussion only examined the rate at which sentences contain at least one temporal reference, according to different parsers. A separate question is the quantity of references for given sentences. Statistics on the number of temporal references per sentence are displayed in Table 11. The central bank, IMF, and European Council speeches contain more numerous temporal references, though this is largely a function of the greater length of these speeches, as previously noted. We report these statistics since, when considering documents as communication events, it is not a priori obvious whether researchers should be interested in the rate of occurrences of temporal references (the *density* of temporal references), or the total number of temporal references during the event (the *volume* of temporal references).

The fraction of sentences tagged by each parser are displayed in Table 13. The central bank speeches contain a greater average number of categorical temporal references per sentence, relative to the other policymaker speeches. The greatest number of numerical references per sentence can be found in the Fed corpus, and that of the European Commission. Given that central bankers frequently discuss numerical forecasts pertaining to given dates in the future, one may have thought that the central bank corpora would contain more frequent numerical temporal references. The fact that the European Commission and European Council speeches contain a greater density of numerical references than those of the ECB shows this is not always the case. The central bank corpora, and that of the European Parliament, have the greatest frequency of tense temporal expressions per sentence.

To conclude, when we abstract from the nature of the temporal references, and study only the *frequency* of such references, we do observe a greater number of temporal expressions in central bank speeches. This result is broadly robust when we study the fraction of sentences containing temporal expressions, and the number of time expressions per sentence, where the central bank speeches are always at the top end of the scores across corpora. However, European Council speeches also contain a high number of numerical temporal references in given documents, while European Commission and Council speeches contain high frequencies of numerical references in given sentences.

6.2 Temporal Orientation

We then compute measures of overall temporal orientation across the corpora. These indicators are displayed in Table 14. It is clear that the distribution of temporal references between past, present, and future differs depending on the nature of the reference. We observe general differences in the allocations of the different forms of temporal reference across corpora, implying that these patterns are likely structural features of the way such expressions are used in the English language. Tense temporal references are mostly to the present, across corpora, with the share of present tense references typically taking a value of 70%. The remaining verbal complexes are assigned evenly between past and future, though there is some variation across corpora. For example, the ECB speeches contain future and past tense expressions to an exactly equal degree (15% each), while the Fed speeches contain greater use of the past tense (20%) relative to the future tense (13%).

Turning to those temporal expressions identified by SUTime, we observe that categorical temporal references are more typically about either the present, or the future, than they are about the past. The central bank and IMF corpora contain a greater fraction of categorical references to the future, relative to the other corpora.

When we examine numerical references to time, we show a clear preponderance of references to the past. We also observe that the central bank and IMF speeches contain a much greater frequency of the use of past numerical references, relative to the other corpora, while they also contain more limited use of *future* temporal references.

6.3 Document Level Measures of Temporal Orientation

The previous discussions focussed on measures of temporal orientation derived from aggregating all temporal references across corpora together. It is also of interest to study document-level measures of temporal orientation. That is, how a policymaker would express temporality per individual speech, which we might also describe as per communication event.

We compute disaggregated measures of document-level temporal orientation, according to categorical, numerical, and tense tags. Consider a document d that contains N_d^{num} numerical temporal references, N_d^{cat} categorical temporal references, and N_d^{vc} verbal complexes, each of which has been tagged as either past, present, or future. For a given numerical temporal reference $n \in \{1, \ldots, N_d^{num}\}$ in document d, let $T_{dn}^{fut,num} = 1$ represent an indicator variable taking the value 1 if this numerical temporal reference is to the future. For a categorical reference $c \in \{1, \ldots, N_d^{cat}\}$, define $T_{dc}^{fut,cat} = 1$ if this categorical reference is to the future. For a verbal complex $v \in \{1, \ldots, N_d^{tmv}\}$, let $T_{dv}^{fut,cat} = 1$ represent an indicator variable taking the value 1 if this verbal complex is assigned to the future tense. Let the indicator variables $T_{dn}^{pst,num}$, $T_{dc}^{pst,cat}$, and $T_{dv}^{pst,vc}$ be defined analogously as indicator variables recording references to the past, for the respective numerical, categorical, and verbal complexes n, c, and v in document d. Let indicator variables $T_{dn}^{prs,num}$, $T_{dc}^{prs,cat}$, and $T_{dv}^{prs,vc}$ indicate comparable indicators for present references. In these cases we create separate measures as follows:

$$p_d^{j,type} = \frac{1}{N_d^{type}} \sum_{i=1}^{i=N_d^{type}} T_{di}^{j,type}, \ j \in pst, prs, fut,$$

for $type \in \{num, cat, vc\}$.

We also create "union" measures of past, present, and future orientation across documents, denoted by p_d^{pst}, p_d^{prs}, and p_d^{fut} respectively. These are constructed on the basis that we assign a sentence to the future case if it contains at least one future reference, according to any of our parsers (with comparable definitions for past and present). Consider a given document d, with N_d sentences. Let T_{ds}^j represent an indicator variable taking a value of 1 if there is a temporal reference of form j in sentence s of document d, where $j \in \{past, present, future\}$. We compute document-level measures according to

$$p_d^j = \frac{1}{N^d} \sum_{s=1}^{s=N_d} T_{ds}^j, \quad j \in \{pst,\, prs,\, fut\}.$$

These measures therefore compute the fraction of sentences in given documents that are associated with at least one past, present, or future tag. Note that the overall temporal orientation measures p_d^{pst}, p_d^{prs}, and p_d^{fut} are constructed as averages over sentences. The measures $p_d^{pst,type}$, $p_d^{prs,type}$ and $p_d^{fut,type}$ for $type \in \{num, cat, vc\}$ are constructed as averages over total numerical temporal references, categorical temporal references, or verbal complexes, depending on *type*. Since there may be several forms of any given references in a given sentence, these averages are constructed over differing numbers of observations.

To understand the nature of the distribution of temporal orientation across documents in our corpora, we report statistics from the distribution across documents in Table 21. The median measure of future orientation across documents is broadly consistent for the corpora we study, when we consider the union measure. We values taken are typically just below 0.30. When we examine the median past orientation across documents, we observe comparable proportions across corpora.

We show fractions of sentences tagged with at least one reference to the past, present, or future in Figure 4. Note that for this exercise the fractions do not necessarily sum to one given that sentences can contain more than one type of temporal reference (e.g., both past and future). We can see that the distribution across past, present, and future more closely resembles that of the tense measure of temporal orientation arising from tags allocated by TMV. We observe a large share of sentences attributed to the present, with broadly equal shares allocated to the past and future. There is no clear difference between the shares across the corpora from economist speeches (ECB, Fed, and IMF), relative to those of other policymakers.

For a graphical representation of differences in document-level measures of temporal orientation, we show kernel density estimates for the whole distribution in Figure 5. These charts exhibit much the same story, in the sense that there is little evidence for a strong deviation between the future orientation of central bank speeches, when compared to those of other policy makers, across our union measure, or the categorical and tense measures. The exception is the numerical measure of future preponderance of central bank speeches, which has lower allocations of numerical temporal expressions to the future. A reflection of this fact is that we observe a greater fraction of central bank speeches with a large number of temporal expressions relating to the past, when compared to speeches of other policymakers.

6.4 The Relation Between Different Measures of Temporal Orientation

While the previous discussions uncovered heterogeneous patterns in the usage of categorical, numerical, and tense temporal expressions, as compared to measures that incorporate the union of all three measures, we have yet to discuss exactly how these measures relate to one another.

To gain a broad perspective on the ways in which policymakers interrelate different forms of temporal reference, we start by examining the frequency with which sentences contain multiple references to the time of different types. To do this we first aggregate the Fed and ECB corpora together to create an overall central bank corpus, while aggregating those of the European Commission, Parliament, Council and IMF to create a "policymaker" corpus. We then chart Venn diagrams, displaying the levels of overlap between the three measures. In panels (a) and (b) of Figure 2 we represent the extent to which sentences contain past, present, and future references (at least one) in combination. We observe that the majority of sentences contain only references to the present, with no additional references to future and past. Sentences containing references to the future and/or past mostly contain also a reference to the present. However, a small but not insignificant proportion of sentences containing either future or past references do not additionally include references to the present. There is also a fraction of sentences containing references to all three measures, though this proportion is small relative to the overall corpus, as is the fraction of sentences containing only past and future sentences. This implies that most sentences containing future or past references contain only such references (in addition to present), and only a smaller fraction contain both. We conclude that on average sentences are about *either* the past or present, though this is not always the case.

Panels (c)-(f) of Figure 2 consider the future tagged sentences and past tagged sentences separately, and sub-divide these tags into cases that arise from tense tags (TMV) or categorical and/or numerical (SUTime). We observe that the greatest proportion of tags are allocated on the basis of TMV, as has been previously observed. However, a non-negligible proportion of future tags are assigned to future on the basis that they contain SUTime tags relating to the future, and these sentences do *not* contain future TMV tags. This demonstrates that the information contained within the SUTime tags is not reducible to that of TMV. In an extreme case, one could imagine that every sentence containing a categorical or numerical reference to the future was framed in the future tense, making TMV sufficient to capture future orientation in given sentence. Panels (c) and (d) of Figure 2 demonstrate that this is not the case. Similar conclusions are show to hold for sentences containing past tags in panels (e) and (f). One difference between the past and future cases is that a greater fraction of the sentences containing categorical references to the past are also framed in the past tense, relative to the future case. However, again we observe a sizeable body of sentences that contain only past categorical or numerical references to the past. However, it remains of course true that the majority of past temporal references are allocated on the basis of the fact they contain usage of the past tense, though they may also contain other forms of past reference.

A related means to assess the relationships between the different measures of temporal orientation is to look at conditional probabilities. Here we compute, conditional on observing a tense temporal reference of a given form in a given sentence, the probability of observing at least one numerical or categorical reference of a given form in that same sentence. These conditional probabilities are displayed in Table 19. We also work in the opposite direction, computing the probability of observing at least one tense temporal reference of a given form in a sentence, conditional on observing a categorical or numerical reference of a given form. These estimates are displayed in Table 20. Note that we are treating categorical and numerical references as substitutes in this exercise, so we are essentially studying the conditional probability of observing SUTime tags in a sentence containing a TMV tag, and vice-versa.

From Table 19 we can observe some level of relation between SUTime tags and TMV. For each of the six corpora, for each TMV verbal complex in the past tense, the probability of observing a SUTime temporal reference in the past tense is greater than for the present. For each of the six corpora, the probability of observing at least one SUTime expression in the future tense, conditional on given TMV verbal complexes in the future tense, is always greater than the probability of observing SUTime expressions in the past or present tense. However, the relationship is less pronounced than was the case for the past, since the percentage point differences in probabilities across cases can be small. For example, the conditional probability of observing a SUTime future tag for a TMV future tag is 0.116 for the Fed speeches, while the probability of observing a SUTime past tag is 0.1.

From Table 20 we can see that, conditional on observing SUTime references in the past tense, we are much more likely to observe verbal complexes in the past tense than we are to observe those in the future tense. For the ECB speeches, for example, conditional on a past SUTime tag, the probability of observing a past TMV tag in that same sentence is 0.556, while the probability of observing a future tag is 0.127. This pattern is true across all six corpora. Conversely, conditional on a SUTime tag indicating a future temporal reference, we are always more likely to observe a TMV tag in the future than we are in the past. For the ECB speeches, the conditional probability of observing a future TMV tag in the case of a future SUTime tag is 0.311, while the probability of observing a past SUTime tag is 0.162. However, what is generally notable from Tables 19 and 20 is that the conditional probabilities for observing past tags from one parser conditional on past tags from the other parser are typically low, and this is especially the case for the probability of observing SUTime tags conditional on TMV tags. This provides further evidence that the parsers are not reducible to each other when summarising temporal information.

Another means to examine the relation between the different measures is to examine the correlation structure of the document-level indicators, which are displayed in Table 18. Unsurprisingly, the union measure of temporal orientation is positively correlated with each of the disaggregated measures across all corpora, and is most strongly correlated with the tense measure. Interestingly, the categorical and numerical measures are not always strongly correlated with each other, and are even negatively correlated for the case of the ECB and (weakly) for the Fed. The relationship of the tense measure of future orientation also relates to the categorical and numerical measures in a heterogeneous manner across corpora. For the case of the ECB, for example, speeches with a greater proportion of verbal complexes in the future tense actually have fewer categorical references in the future tense, though they have more numerical references to the future tense. Correlation measures for the past indicators are displayed in Table 17.

6.5 Additional Dimensions of Temporal Orientation

In the course of our previous discussions we have used measures of temporal orientation, attributed to the broad categories of past, present, and future. We have tended to apply the SUTime and TMV parsers in ways that made simplifying aggregations over different indicators, in order to reach such broad summary measures of temporal orientation. In this section we briefly discuss several forms of additional information generated by our parsers, that could be brought to bear on future research regarding policymaker speeches.

In our application of the SUTime algorithm we attributed the numerical references to time to past, present, and future by first computing the difference between the date of the reference and the date of the speech, before determining whether this difference was negative, zero, or positive. We therefore did not examine information on the *horizon* of the numerical references. In Figure 1 we display histograms demonstrating the distribution of the numerical temporal references across our corpora. We can see clearly that the distributions have a negative skew, with a greater tendency to make references to past dates relative to future dates. We also observe that the central bank speeches have a greater mass of observations across dates in the fairly recent past, which is also true of the IMF. The speeches of politicians tend to congregate numerical references either to the present, or to past dates into the longer-term past, or longer-term future. This is suggestive of the typical reference to recent macroeconomic data in central bank speeches, as well as the course of recent business cycles. As discussed in Byrne et al. (2023), this reflects the importance that central banks place in communicating their assessment of the state of the economy through interpreting recent economic data.

We can also attribute the categorical references to the future generated by SUTime to four broad categories: short-run, medium-run, long-run, or ambiguous. We display the breakdown of our future categorical measure in Table 16. We observe that the policy-maker speeches tend to make more ambiguous references to the future, relative to those of the central banks or the IMF. The European political speeches also tend to make a greater proportion of long-run references, relative to short- or medium-run, with the exception of the European Council that more frequently makes reference to the short-run future relative to the others. The ECB makes reference to the short-, medium-, and long-run in a monotonically increasing manner, while the Fed makes reference to the short- and long-run more infrequently mentions "medium-run". An explanation for this result is that ECB policymakers frequently make reference to their inflation target, which is formulated explicitly as a "medium-run" objective. The IMF future categorical references are broadly balanced across categories.

Finally, we examine the consequences of the amendment we made to the TMV classification of verbal complexes to the future tense. Recall that, based on a bespoke list of words associated with communication regarding monetary policy, we assigned present tense phrases such as "we expect" to the future tense. Table 15 displays the impact of the amendment we made to the TMV algorithm. We can observe that the impact of this change is minor, with the amendment typically increasing the fraction of verbal complexes assigned to the future by one percentage point.

7 Conclusion

In this study we have discussed how two existing tools from the Natural Language Processing literature, namely SUTime and TMV, can be applied in tandem to examine interesting features regarding the nature of policymaker communications. When necessary, we apply refinements and modifications of the two algorithms to best reflect the nature of our corpora.

We have examined the relations between the two tools in our representative corpora, and shown that each tool has the capacity to elucidate interesting features regarding policymakers' discussions of time. We examine also measures of future and past orientation based on the union of information provided by SUTime and TMV. One key finding from our analysis of the data is that these tools are not reducible to each other. That is, each captures distinct elements of the ways in which policymakers talk about temporality in their communications with the public. This justifies our synthesis of these tools into a single algorithm for measuring temporal references in text. A contribution of this paper is a guide for social science researchers to use NLP technology to measure temporality in text. We believe that this dimension has been under-studied and is an area for growth, particularly given the importance of understanding what policymakers are saying about the timing of their decisions, to which future horizons they are looking, and from which past data they are gleaning insights that inform their decisions.

A great deal of recent research in the expanding literature on central bank communication has focussed on extracting measures of topic, via LDA modelling, and tone, via dictionary methods. This paper surveys the existing literature that focus on topic and tone, and the methods that are used. It is our hope that this paper will be a useful starting point for researchers examining topic and tone, as well as new developments in textual analysis of temporality.

Our methodological contribution is to propose a means to evaluate the temporal dimension of policymaker communications: how they refer to time. Future research into the efficacy of central bank communication could therefore profitably focus on the interaction between the three Ts of topic, tone, and time.

References

Aguilar, Ana and Fernando Pérez-Cervantes (2022) "Communication, Monetary Policy, and Financial Markets in Mexico," BIS Working Papers 1025, Bank for International Settlements.

Al Amrani, Yassine, Mohamed Lazaar, and Kamal Eddine El Kadiri (2018) "Random Forest and Support Vector Machine based Hybrid Approach to Sentiment Analysis," *Procedia Computer Science*, 127, 511–520, Proceedings of the First International Conference on Intelligent Computing in Data Sciences.

Antweiler, Werner and Murray Z. Frank (2004) "Is All That Talk Just Noise? The Information Content of Internet Stock Message Boards," *The Journal of Finance*, 59 (3), 1259–1294.

Apel, Mikael, Marianna Blix Grimaldi, and Isaiah Hull (2022) "How Much Information Do Monetary Policy Committees Disclose? Evidence from the FOMC's Minutes and Transcripts," *Journal of Money, Credit and Banking*, 54 (5), 1459–1490.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis (2016) "Measuring Economic Policy Uncertainty," *The Quarterly Journal of Economics*, 131 (4), 1593–1636.

Bennani, Hamza and Matthias Neuenkirch (2017) "The (home) bias of European central bankers: new evidence based on speeches," *Applied Economics*, 49 (11), 1114–1131.

Björkelund, Anders, Bernd Bohnet, Love Hafdell, and Pierre Nuges (2010) "A High-Performance Syntactic and Semantic Dependency Parser,"Technical report, Coling 2010: Demonstration Volume.

Blei, David M., Thomas L. Griffiths, Michael I. Jordan, and Joshua B. Tenenbaum (2004) "Hierarchical Topic Models and the Nested Chinese Restaurant Process," in *Advances in Neural Information Processing Systems 16*: MIT Press.

Blei, David M. and John D. Lafferty (2006) "Dynamic Topic Models," in *Proceedings of the 23rd International Conference on Machine Learning*, ICML '06, 113–120, New York, NY, USA: Association for Computing Machinery.

Blei, David M., Andrew Y. Ng, and Michael I. Jordan (2003) "Latent dirichlet allocation," *J. Mach. Learn. Res.*, 3, 993–1022.

Bradley, Margaret M. and Peter J. Lang (1999) "Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings,"Technical report, The Center for Research in Psychophysiology, University of Florida.

Bybee, Leland, Bryan Kelly, Asaf Manela, and Dacheng Xiu (2020) "The Structure of Economic News," January, *mimeo*.

Byrne, David, Robert Goodhead, Michael McMahon, and Conor Parle (2023) "The Central Bank Crystal Ball: Temporal information in monetary policy communication," Research Technical Paper 2023/01, Central Bank of Ireland.

Caldara, Dario and Matteo Iacoviello (2022) "Measuring Geopolitical Risk," *American Economic Review*, 112 (4), 1194–1225.

Chang, Angel X. and Christopher D. Manning (2012) "SUTIME: A Library for Recognizing and Normalizing Time Expressions,"Technical report, 8th International Conference on Language Resources and Evaluation.

——— (2014) "TokensRegex: Defining Cascaded Regular Expressions Over Tokens," technical report, Stanford University.

Coenen, Günter, Michael Ehrmann, Gaetano Gaballo, Peter Hoffmann, Anton Nakov, Stefano Nardelli, Eric Persson, and Georg Strasser (2017) "Communication of monetary policy in unconventional times," Working Paper 2080, European Central Bank.

Crawford, Michael, Taghi M. Khoshgoftaar, Joseph D. Prusa, Aaron N. Richter, and Hamzah Al Najada (2015) "Survey of review spam detection using machine learning techniques," *Journal of Big Data*, 2 (23).

Cross, James P. and Derek Greene (2020) "Talk is not cheap: Policy agendas, information processing, and the unusually proportional nature of European Central Bank communications policy responses," *Governance*, 33, 425–444.

Deerwester, Scott, Susan T. Dumais, George W. Furnas, Thomas K. Landauer, and Richard Harshman (1990) "Indexing by Latent Semantic Analysis," *Journal of the American Society for Information Science*, 41 (6), 391–407.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova (2019) "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, 4171–4186, Minneapolis, Minnesota: Association for Computational Linguistics, June.

Dias, Gaël, Mohammed Hasanuzzaman, Stéphane Ferrari, and Yann Mathet (2014) "TempoWordNet for Sentence Time Tagging,"Technical report, 23rd International Conference on World Wide Web Companion.

Ehrmann, Michael and Marcel Fratzscher (2009) "Explaining Monetary Policy in Press Conferences," *International Journal of Central Banking*, 5 (2), 42–84.

Evans, James A. and Pedro Aceves (2016) "Machine Translation: Mining Text for Social Theory," *Annual Review of Sociology*, 42, 21–50.

Galardo, Maddalena and Cinzia Guerrieri (2017) "The effects of central bank's verbal guidance: evidence from the ECB," Temi di discussione (working papers) 1129, Banca d'Italia.

Gentzkow, Matthew, Brian Kelly, and Matt Taddy (2019) "Text as Data," *Journal of Economic Literature*, 57 (3), 535–574.

Gorodnichenko, Yuriy, Tho Pham, and Oleksandr Talavera (2021) "The Voice of Monetary Policy," NBER Working Papers 28592, National Bureau of Economic Research, Inc.

Grimmer, Justin and Brandon M. Stewart (2013) "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts," *Political Analysis*, 21, 267–297.

Grootendorst, Maarten (2022) "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," *arXiv preprint arXiv:2203.05794*.

Hafez, Ghada, Rasha Ismail, and Omar Karam (2017) "Temporal sentiment analysis and time tags for opinions," in *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)*, 373–378.

Hansen, Stephen and Michael McMahon (2016) "Shocking language: Understanding the macroeconomic effects of central bank communication," *Journal of International Economics*, 99, 114–133.

Hassan, Tarek A., Stephan Hollander, Laurence van Lent, and Ahmed Tahoun (2019) "Firm-Level Political Risk: Measurement and Effects," *The Quarterly Journal of Economics*, 134 (4), 2135–2202.

Hendry, Scott and Alison Madeley (2010) "Text Mining and the Information Content of Bank of Canada Communications,"Technical Report 31, Bank of Canada.

Hu, Minqing and Bing Liu (2004) "Mining and summarizing customer reviews," *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.*

Hubert, Paul and Fabien Labondance (2021) "The signaling effects of central bank tone," *European Economic Review*, 133 (C).

Hutto, C.J. and Eric Gilbert (2015) "VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text," 01.

Istrefi, Klodiana, Florens Odendahl, and Giulia Sestieri (2021) "Fed communication on financial stability concerns and monetary policy decisions: revelations from speeches," Working Papers 2110, Banco de España.

Jurafsky, Daniel and James J. Martin (2009) *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, 2nd ed. London: Pearson.

Kamila, Sabyasachi, Mohammed Hasanuzzaman, Asif Ekbal, and Pushpak Bhattacharyya (2019) "Resolution of grammatical tense into actual time, and its application in Time Perspective study in the tweet space," *PLOS ONE*, 1–20.

Kelly, Bryan, Asaf Manela, and Alan Moreira (2021) "Text Selection," *Journal of Business & Economic Statistics*, 39 (4), 859–879.

Kiss, Tibor and Jan Strunk (2006) "Unsupervised Multilingual Sentence Boundary Detection," *Computational Linguistics*, 32, 485–525.

Lagrari, Fatima-Ezzahra, Houssaine Ziyati, and Youssfi El Kettani (2019) "An Efficient Model of Text Categorization Based on Feature Selection and Random Forests: Case for Business Documents," in Ezziyyani, Mostafa ed. *Advanced Intelligent Systems for Sustainable Development (AI2SD'2018)*, 465–476, Cham: Springer International Publishing.

Li, Jiwei and Claire Cardie (2014) "Timeline generation: tracking individuals on Twitter,"Technical report, World Wide Web Conference.

Loáiciga, Sharid, Thomas Meyer, and Andrei Popescu-Belis (2019) "English-French Verb Phrase Alignment in Europarl for Tense Translation Modeling,"Technical report, European Language Resources Association (ELRA).

Loughran, Tim and Bill McDonald (2011) "When Is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *Journal of Finance*, 66 (1), 35–65.

——— (2020) "Textual Analysis in Finance," *Annual Review of Financial Economics*, 12, 357–75.

Lucca, David O. and Francesco Trebbi (2009) "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements,"Technical Report 15367, NBER.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze (2008) *Introduction to Information Retrieval*, Cambridge: Cambridge University Press.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeff Dean (2013) "Distributed Representations of Words and Phrases and their Compositionality," in Burges, C.J., L. Bottou, M. Welling, Z. Ghahramani, and K.Q. Weinberger eds. *Advances in Neural Information Processing Systems*, 26: Curran Associates, Inc.

Mohammad, Saif M. (2018) "Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words,"Technical report, Proceedings of the 56th Annual Meeting for the Association of Computational Linguistics (Long Papers).

Mohammad, Saif M. and Peter D. Turney (2013) "Crowdsourcing a Word-Emotion Association Lexicon," *Computational Intelligence*, 29 (3), 436–465.

Moody, Christopher E. (2016) "Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec."

Myers, Skatje and Martha Palmer (2019) "ClearTAC: Verb Tense, Aspect, and Form Classification Using Neural Nets,"Technical report, ClearTAC: Verb Tense, Aspect, and Form Classification Using Neural Nets.

Nielsen, Finn (2011) "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," *CoRR*.

Palmer, Martha, Daniel Gildea, and Paul Kingsbury (2005) "The Proposition Bank: An Annotated Corpus of Semantic Roles," *Computational Linguistics*, 31 (1), 71–106.

Pang, Bo, Lillian Lee, and Shivakumar Vaithyanathan (2002) "Thumbs up? Sentiment Classification using Machine Learning Techniques,"Technical report, Proceedsings of the 2002 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Parle, Conor (2022) "The financial market impact of ECB monetary policy press conferences — A text based approach," *European Journal of Political Economy*, 74, 102230.

Pennington, Jeffrey, Richard Socher, and Christopher D. Manning (2014) "GloVe: Global Vectors for Word Representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1532–1543: Association for Computational Linguistics, October.

Picault, Matthieu and Thomas Renault (2017) "Words are not all created equal: A new measure of ECB communication," *Journal of International Money and Finance*, 79, 136–156.

Pustejovsky, James, José Castaño, Robert Ingria, and Roser Saurí (2003a) "TimeML: Robust Specification of Event and Temporal Expressions in Text,"Technical report, American Association for Artificial Intelligence.

Pustejovsky, James, Patrick Hanks, Roser Saurí et al. (2003b) "The TIMEBANK Corpus,"Technical report, Lancaster.

Rae, John (1834) *The Sociological Theory of Capital*, London: Macmillan.

Ramm, Anita, Sharid Loáiciga, Annemarie Friedrich, and Alexander Fraser (2017) "Annotating Tense, Mood and Voice for English, French and German,"Technical report, Proceedings of ACL 2017, System Demonstrations.

Rosa, Carlo and Giovanni Verga (2007) "On the consistency and effectiveness of central bank communication: Evidence from the ECB," *European Journal of Political Economy*, 23 (1), 146–175, Central bank transparency and central bank communication.

Schmeling, Maik and Christian Wagner (2019) "Does Central Bank Tone Move Asset Prices?" October, *mimeo*.

Schumacher, Gijs, Martijn Schoonvelde, Denise Traber, Tanushree Dahiya, and Erik de Vries (2016) "EUSpeech: a New Dataset of EU Elite Speeches,"Technical report.

Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson (2022) "Measuring News Sentiment," *Journal of Econometrics*, 228 (2), 221–243.

Stone, Philip J., Dexter C. Dunphy, Marshall S. Smith, and Daniel M. Ogilvie (1966) *The General Inquirer: A Computer Approach to Content Analysis*, Cambridge, MA: MIT Press.

Strötgen, Jannik and Michael Gertz (2010) "HeidelTime: High Quality Rule-Based Extraction and Normalization of Temporal Expressions," in *Proceedings of the 5th International Workshop on Semantic Evaluation, SemEval@ACL 2010, Uppsala University, Uppsala, Sweden, July 15-16, 2010*, 321–324: The Association for Computer Linguistics.

Tabassum, Jeniya, Alan Ritter, and Wei Xu (2016) "TweeTime: A Minimally Supervised Method for Recognizing and Normalizing Time Expressions in Twitter,"Technical report, Association for Computational Linguistics.

Taboada, Maite, Julian Brooke, Milan Tofiloski, Kimberly Voll, and Manfred Stede (2011) "Lexicon-Based Methods for Sentiment Analysis," *Computational Linguistics*, 37, 267–307.

Tetlock, Paul C. (2007) "Giving Content to Investor Sentiment: The Role of Media in the Stock Market," *The Journal of Finance*, LXII (3), 1139–1168.

Tobback, Ellen, Stefano Nardelli, and David Martens (2017) "Between hawks and doves: measuring central bank communication," Working Paper Series 2085, European Central Bank.

Toutanova, Kristina, Dan Klein, Christopher D. Manning, and Yoram Singer (2003) "Feature-Rich Part-of-Speech Tagging With a Cyclical Dependency Network,"Technical report, Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL).

UzZaman, Naushad, Hector Llorens, Leon Derczynski, Marc Verhagen, James Allen, and James Pustejovsky (2013) "SemEval-2013 Task 1: TempEval-3: Evaluating Time Expressions, Events, and Temporal Relations,"Technical report, Association for Computational Linguistics.

Verhagen, Marc, Robert Gaizauskas, Frank Schilder, Mark Hepple, Graham Katz, and James Pustejovsky (2007) "SemEval-2007 Task 15: TempEval Temporal Relation Identification,"Technical report, Association for Computational Linguistics.

Verhagen, Marc, Robert Gaizauskas, Frank Schilder, Mark Hepple, Jessica Moszkowicz, and James Pustejovsky (2009) "The TempEval Challenge: Identifying Temporal Relations in Text," *Language Resources and Evaluation*, 43 (2), 161–179.

Verhagen, Marc, Roser Saurí, Tommaso Caselli, and James Pustejovsky (2010) "SemEval-2010 Task 13: TempEval-2,"Technical report, Association for Computational Linguistics.

Wang, Sida and Christopher D. Manning (2012) "Baselines and Bigrams: Simple, Good Sentiment and Topic Classification,"Technical report, Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics.

Wilson, Theresa, Paul Hoffmann, Swapna Somasundaran et al. (2005a) "OpinionFinder: A system for subjectivity analysis,"Technical report, Proceedings of HLT/EMNLP 2005 Demonstration Abstracts.

Wilson, Theresa, Janyce Wiebe, and Paul Hoffmann (2005b) "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis,"Technical report, Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing (HLT/EMNLP).

Yan, Rui, Liang Kong, Congrui Huang, Xiaojun Wan, Xiaoming Li, and Yan Zhang (2011) "Timeline Generation through Evolutionary Trans-Temporal Summarization,"Technical report, Association for Computational Linguistics.

Young, Lori and Stuart Soroka (2012) "Affective News: The Automated Coding of Sentiment in Political Texts," *Political Communication*, 29 (2), 205–231.

8 Tables

Table 1: Table 1 from Ramm et al. (2017)

finite	mood	tense	voice	example (active voice)
yes	ind	present	act pass	(I) work
		presProg		(I) am working
		presPerf		(I) have worked
		presPerfProg		(I) have been working
		past		(I) worked
		pastProg		(I) was working
		pastPerf		(I) had worked
		pastPerfProg		(I) have been working
		futureI		(I) will work
		futureIProg		(I) will be working
		futureII		(I) will have worked
		futureIIProg		(I) will have been working
	subj	condI		(I) would work
		condIProg		(I) would be working
		condII		(I) would have worked
		condIIProg		(I) would have been working
no	-	-	-	to work

Table 2: List of Categorical SUTime Rules

Time Orientation	Baseline SUTime	Additional
Past	past	looking back
	recently	
	at the time	
	once	
	medieval	
	previously	
Present	current	on-going / ongoing
	currently	at present / the present
	now	presently
		at the moment
Future	future	looking forward
		short/shorter term/run
		near term
		medium term/run
		long/longer term/run
		horizon

Notes: This table describes the phrases identified as categorical references to time in the baseline version of SUTime (Chang and Manning 2012). We also document some additional rules we added to the baseline algorithm, in order to capture certain commonly used phrases from our corpora, and typical of the discourse of central banking.

Table 3: Historical Dates

Phrase	Assigned Date	Phrase	Assigned Date
euro area crisis	2012	Bretton Woods system	from 1944 to 1971
sovereign debt crisis		Bretton Woods era	
global financial crisis	2008	post war	1946
the financial crisis		Second World War	1939
GFC		World War 2	
Great Recession		WWII	
pre crisis	2007	WW2	
dot com bubble	2000	interwar period	1930
collapse of the Soviet Union	1991	Great Depression	1929
dissolution of the Soviet Union		First World War	1914
German reunification	1990	Great War	
Gulf War		World War 1	
Iraq War		WWI / WW1	
fall of the Berlin Wall	1989	Gold Standard	1870
Greenspan era	from 1987 to 2006	French Revolution	1789
Volcker era	from 1979 to 1987	Industrial Revolution	1760
Volcker disinflation	1981		
Vietnam War	1965		
Schuman declaration	1950		
Korean War	1950		

Notes: This table describes the phrases identified as historical references according to our additional rules, which were added to the baseline version of SUTime (Chang and Manning 2012).

Table 4: Example of SUTime Output

27	With todays comprehensive package of monetary policy decisions , we are providing substantial monetary stimulus to ensure that financial conditions remain very favourable and support the euro area expansion , the ongoing build-up of domestic price pressures and , thus , the sustained convergence of inflation to our medium-term inflation aim . < TIMEX3 alt_value = "THISP1D" tid = "t1" type = "SET" > todays < /TIMEX3 > < TIMEX3 range = "(PRESENT_REF, PRESENT_REF,)" tid = "t2" type = "DATE" value = "PRESENT_REF" > ongoing < /TIMEX3 > < TIMEX3 range = "(PRESENT_REF, UNKNOWN,)" tid = "t3" type = "DATE" value = "FUTURE_REF" > mediumterm < /TIMEX3 >
28	Let me now explain our assessment in greater detail , starting with the economic analysis . < TIMEX3 range = "(PRESENT_REF, PRESENT_REF,)" tid = "t1" type = "DATE" value = "PRESENT_REF" > now < /TIMEX3 >
29	Euro area real GDP increased by 0.2 % , quarter on quarter , in the second quarter of 2019 , following a rise of 0.4 % in the previous quarter . < TIMEX3 range = "(2019 − Q2 − 04 − 01, 2019 − Q2 − 06 − 30, P3M)" tid = "t1" type = "DATE" value = "2019 − Q2" > thesecondquarterof2019 < /TIMEX3 > < TIMEX3 range = "(2019 − Q3 − 07 − 01, 2019 − Q3 − 09 − 30, P3M)" tid = "t2" type = "DATE" value = "2019 − Q3" > thepreviousquarter < /TIMEX3 >
30	Incoming economic data and survey information continue to point to moderate but positive growth in the third quarter of this year . < TIMEX3 range = "(2019 − Q3 − 07 − 01, 2019 − Q3 − 09 − 30, P3M)" tid = "t1" type = "DATE" value = "2019 − Q3" > thethirdquarterofthisyear < /TIMEX3 >
32	At the same time , the services and construction sectors show ongoing resilience and the euro area expansion is also supported by favourable financing conditions , further employment gains and rising wages , the mildly expansionary euro area fiscal stance and the ongoing albeit somewhat slower growth in global activity . < TIMEX3 range = "(PRESENT_REF, PRESENT_REF,)" tid = "t1" type = "DATE" value = "PRESENT_REF" > ongoing < /TIMEX3 > < TIMEX3 range = "(PRESENT_REF, PRESENT_REF,)" tid = "t2" type = "DATE" value = "PRESENT_REF" > ongoing < /TIMEX3 >
33	This assessment is broadly reflected in the September 2019 ECB staff macroeconomic projections for the euro area . < TIMEX3 range = "(2019 − 09 − 01, 2019 − 09 − 30, P1M)" tid = "t1" type = "DATE" value = "2019 − 09" > September2019 < /TIMEX3 >
34	These projections foresee annual real GDP increasing by 1.1 % in 2019 , 1.2 % in 2020 and 1.4 % in 2021 . < TIMEX3 range = "(2019 − 01 − 01, 2019 − 12 − 31, P1Y)" tid = "t1" type = "DATE" value = "2019" > 2019 < /TIMEX3 > < TIMEX3 range = "(2020 − 01 − 01, 2020 − 12 − 31, P1Y)" tid = "t2" type = "DATE" value = "2020" > 2020 < /TIMEX3 > < TIMEX3 range = "(2021 − 01 − 01, 2021 − 12 − 31, P1Y)" tid = "t3" type = "DATE" value = "2021" > 2021 < /TIMEX3 >

Notes: This table shows an example of the output of SUTime, applied to the introductory statement of the Governing Council press conference of September 12, 2019. The statement was delivered by ECB President Mario Draghi. We have parsed sentences 27, 28, 29, 30, 32, 33, and 34 (sentence 31 contained no numerical/categorical temporal references).

Table 5: Example of TMV Output

Parsed Text Data	Verbal Complex	Finite	Tense	Mood	Voice	Negation
`Within our mandate ,`	takes	yes	present	indicative	active	no
`the ECB is ready to do`	is	yes	present	indicative	active	no
`whatever it takes`	to preserve	no	-	-	-	-
`to preserve the euro .`	to do	no	-	-	-	-
`And believe me ,`	will be	yes	futureI	indicative	active	no
`it will be enough .`						

Notes: This table shows an example of output when the TMV tool of Ramm et al. (2017) is applied to two sentences from the speech of **?**. Raw text data is parsed according to MATE parser of Björkelund et al. (2010) before the TMV tool is applied. Tenses are assigned to verbal complexes according to the schema in Table 1.

Table 6: List of Future Verbs

Verb	Future Probability	Verb	Future Probability
anticipate	0.984	undertake	0.992
augur	0.992	warn	0.957
bode	0.998	wonder	0.998
forecast	0.998	pre-empt	0.998
forewarn	0.997	project	0.998
guesstimate	0.957	fear	0.998
lapse	0.997	envision	0.998
pledge	0.992	envisage	0.998
predict	0.998	foresee	0.998
prophesy	0.992	forestall	0.998
propose	0.967	look forward	0.998
retain	0.998	await	0.998
speculate	0.992	swear	0.998

Notes: Table displays list of verbal complexes we re-assign to the future tense, when found in the present tense. We also display the associated probability of being a future verb, where estimates are taken from the TempoWordNet dataset of Dias et al. (2014).

Table 7: List of Manual Stopwords

also	gentlemen	may	said
answer	go	mention	say
answers	here	month	second
april	january	november	see
can	july	now	september
clear	june	noyer	since
december	just	october	statement
discuss	know	one	take
discussion	ladies	particular	think
disposal	ladies	point	time
draghi	lagarde	press	today
duisenberg	let	question	trichet
fact	look	questions	want
february	m	regard	way
first	march	report	welcome

Table 8: List of N-Grams (Part 1 of 2)

accommodative monetary	commodity prices	employment growth	firm anchoring
adjustment path inflation	consumer confidence	euro area	firmly anchored
anchor inflation expectations	consumer prices	european commission	fiscal consolidation
annual growth	core inflation	european council	fiscal framework
annual growth rates	corporate sector purchase programme	european economy	fiscal imbalances
annual hicp inflation	covered bond purchase programme	european parliament	fiscal policies
annual inflation rates	credit growth	european union	fiscal policy
appreciation euro	credit risk	eurosystem staff macroecnomic projections	fixed rate full allotment
asset purchase porgramme	credit standards	eurosystem staff projections	fixed rate tender
asset purchases	current account	excess liquidity	foreign demand
assetbacked securities purchase programme	deposit facility	excess reserves	forward guidance
automatic stabilisers	deposit facility rate	excessive deficit	full allotment
balance sheet	developments remain	exchange rate	futures prices
balance sheet adjustment	disposable income	exchange rate policy	gdp growth
bank lending	domestic price pressures	executive board	global demand
bank lending survey	downside risks	expanded asset purchase programme	global economy
banking system	ecb interest rates	extended asset purchase programme	global imbalances
banking union	ecb staff macroeconomic projections	federal reserve	governing council
basel i	economic analysis	federal reserve system	gradual recovery
basel ii	economic conditions	financial conditions	growth prospects
basel iii	economic data	financial environment	growth rate loans
basel iv	economic developments	financial market	headline inflation
business cycle	economic expansion	financial markets	hicp inflation
capital markets	economic growth	financial stability	hicp inflation rates
central bank	economic recovery	financial system	inflation expectations
central banks	emerging markets	financing conditions	inflation rates

Table 9: List of N-Grams (Part 2 of 2)

inflationary expectations	marginal lending facility	outright monetary transactions	staff macroeconomic projections
inflationary pressure	market expectations	pandemic emergency purchase program	stress tests
inflationary pressures	medium term	policy measures	strong growth
inflationary risks	monetary accommodation	policy rate	structural policies
interest rate	monetary aggregate	policy rates	structural reform
interest rates	monetary aggregates	potential growth	structural reforms
interest rates unchanged	monetary analysis	potential output growth	structural unemployment
job creation	monetary credit	press conference	support economic recovery
labour costs	monetary developments	price developments	sustainable economic growth
labour market	monetary growth	price inflation	sustainable growth
labour market reforms	monetary policy	price rises	target ii
labour markets	monetary policy measures	price stability	targeted long term refinancing operations
labour product	monetary policy strategy	private consumption	targeted longerterm refinancing operations
labour productivity	monetary union	private sector	united states
lending rates	monetary union	productivity frowth	upside risks
liquidity conditions	money market	public finances	upward risks
liquidity situation	money markets	public sector purchase programme	wage developments
lognerterm inflation expectations	national central banks	purchasing power	wage growth
long term refinancing operations	negative rates	real economy	wage moderation
longer term refinancing operations	net purchases	real gdp	world economy
longerterm inflation	nonfinancial corporations	real gdp growth	yield curve
longerterm refinancing operations	nonstandard measures	remain broadly balanced	
longterm interest	nonstandard monetary	remain firmly anchored	
low inflation	nonstandard monetary policy	risk premia	
macroeconomic projections	oil prices	single currency	
main refinancing operation	outlook price	social partners	

Table 10: Corpora Characteristics

	Documents	Start	End	Median	Std	Min	Max
ECB Speeches	2203	1997/02/07	2020/09/15	112	69.38	4	793
Fed Speeches	4568	1993/04/02	2022/05/11	118	70.51	2	1276
Fed Speeches: President	438	1996/06/13	2022/03/21	99	47.73	8	311
Fed Speeches: Board	1109	1996/06/18	2022/04/05	113	75.38	7	1276
Fed Speeches: Governors	3021	1993/04/02	2022/05/11	123	70.82	2	758
IMF Speeches	509	2007/01/09	2015/11/30	106	47.86	4	335
European Commission	5987	2007/01/04	2015/12/18	69	40.10	1	535
European Council	220	2009/12/01	2015/09/07	110	58.34	12	357
European Parliament	2547	2007/01/15	2015/12/16	12	15.58	2	167

Notes: This table shows key summary statistics regarding the number of documents in each corpora and sub-corpora used in this study. The table also displays information on the number of sentences per document.

Table 11: Number of Temporal Expressions per Document

		Temporal Expressions per Document				
		Mean	Median	Std.	Min	Max
Type	Dataset					
Categorical	ECB Speeches	17.86	15	13.97	0	164
	Fed Speeches	16.11	14	12.78	0	146
	IMF	11.85	10	7.99	0	45
	European Commission	7.71	6	5.95	0	54
	European Council	10.30	9	6.70	0	33
	European Parliament	1.75	1	2.32	0	21
Numerical	ECB Speeches	18.83	16	13.95	1	142
	Fed Speeches	23.18	18	20.04	1	315
	IMF	14.46	13	9.12	0	54
	European Commission	12.63	11	8.38	0	102
	European Council	18.86	17	11.27	1	72
	European Parliament	1.91	1	2.72	0	29
Tense	ECB Speeches	204.03	183	135.00	6	1786
	Fed Speeches	204.99	202	114.79	1	1310
	IMF	151.39	152	71.29	7	463
	European Commission	116.00	107	64.76	4	874
	European Council	155.94	150	84.60	6	603
	European Parliament	33.88	23	29.89	1	246

Notes: This table shows summary statistics regarding the number of temporal expressions identified per document, according to the two parsers used in this study. Categorical and numerical temporal expressions are identified using SUTime. Tense temporal expressions are identified using TMV. We also display the number of temporal expressions per document identified as categorical or numerical time references according to SUTime. For the case of SUTime, a "temporal expression" is a TIMEX tag. For the case of TMV, a "temporal expression" is a verbal complex. These measures are computed after the application of our cleaning and refinement routines.

Table 12: Fraction of Sentences Containing Temporal Expressions

Parser	Categorical	Numerical	Tense	Union
ECB Speeches	0.12	0.13	0.94	0.95
Fed Speeches	0.16	0.12	0.88	0.90
IMF Speeches	0.12	0.10	0.88	0.89
European Commission	0.14	0.10	0.90	0.92
European Council	0.14	0.08	0.82	0.84
European Parliament	0.10	0.09	0.87	0.88

Notes: This table shows the fraction of sentences in each corpora that are associated with at least one temporal expression, according to the two parsers used in this study (SUTime and TMV). We also display the fraction of sentences tagged with categorical or numerical time references according to SUTime. The Union measure refers to sentences tagged with at least one temporal tag according to either parser.

Table 13: Number of Time Expressions per Sentence

Parser	Dataset	Mean	Median	Std	Max
Categorical	ECB Speeches	0.15	0.00	0.40	5.00
	Fed Speeches	0.13	0.00	0.39	7.00
	IMF Speeches	0.11	0.00	0.35	4.00
	European Commission	0.10	0.00	0.33	5.00
	European Council	0.09	0.00	0.31	3.00
	European Parliament	0.10	0.00	0.33	4.00
Numerical	ECB Speeches	0.15	0.00	0.46	8.00
	Fed Speeches	0.19	0.00	0.50	16.00
	IMF Speeches	0.14	0.00	0.41	7.00
	European Commission	0.17	0.00	0.45	14.00
	European Council	0.16	0.00	0.44	7.00
	European Parliament	0.11	0.00	0.36	5.00
Tense	ECB Speeches	1.66	1.00	1.03	16.00
	Fed Speeches	1.70	2.00	1.18	21.00
	IMF Speeches	1.46	1.00	0.99	14.00
	European Commission	1.57	1.00	1.05	36.00
	European Council	1.34	1.00	1.01	9.00
	European Parliament	1.93	2.00	1.48	17.00

Notes: This table shows summary statistics regarding the number of temporal references identified per sentence across the corpora, and sub-divided according to the parser used. Categorical and numerical temporal expressions are identified using SUTime. Tense temporal expressions are identified using TMV.

Table 14: Overall Temporal Orientation Across Corpora

Parser	Categorical			Numerical			Tense		
Dataset	Past	Present	Future	Past	Present	Future	Past	Present	Future
ECB Speeches	0.16	0.40	0.45	0.71	0.17	0.12	0.15	0.70	0.15
Fed Speeches	0.17	0.41	0.42	0.76	0.14	0.10	0.20	0.67	0.13
IMF Speeches	0.17	0.42	0.41	0.67	0.18	0.15	0.14	0.70	0.16
European Commission	0.14	0.46	0.40	0.49	0.26	0.25	0.11	0.74	0.15
European Council	0.20	0.43	0.37	0.55	0.23	0.22	0.19	0.70	0.11
European Parliament	0.19	0.59	0.22	0.49	0.30	0.21	0.14	0.73	0.13

Notes: This table shows overall temporal orientation across corpora. Categorical and numerical temporal expressions are identified using SUTime. Tense temporal expressions are identified using TMV. For the case of temporal orientation by tense, the fractions displayed indicate the proportions of verbal complexes identified by TMV that are assigned to past, present, or future. For the case of categorical and numerical temporal orientation, the fractions indicate the proportions of TIMEX tags attributed to past, present, or future (respectively computed for categorical and numerical tags).

Table 15: The Impact of the Amendment of the TMV Algorithm

	TMV	TMV*
ECB Speeches	0.14	0.15
Fed Speeches	0.12	0.13
IMF Speeches	0.14	0.16
European Commission	0.14	0.15
European Council	0.11	0.11
European Parliament	0.12	0.13

Notes: Figure shows the share of verbal complexes tagged as future using the baseline TMV algorithm of Ramm et al. (2017). The share of future references under the adjusted version of TMV employed in this study is indicated by the asterisk.

Table 16: Breakdown of Categorical Time-References by Horizon

Institution	Short-Run	Medium-Run	Long-Run	Ambiguous
ECB Speeches	0.14	0.19	0.26	0.42
Fed Speeches	0.19	0.03	0.36	0.42
IMF	0.18	0.19	0.22	0.41
European Commission	0.06	0.02	0.20	0.71
European Council	0.09	0.02	0.22	0.66
European Parliament	0.07	0.01	0.15	0.77

Notes: Table shows the decomposition of categorical temporal references about the future identified by SUTime. We assign the temporal expressions to four broad categories, which we class as short-run, medium-run, long-run, or ambiguous.

Table 17: The Correlation Structure of Temporal Indicators Across Corpora – Past Indicators

Institution		Categorical	Numerical	Tense	Union
ECB Speeches	Categorical		0.01	0.08	0.14
	Numerical	0.01		0.36	0.45
	Tense	0.08	0.36		0.91
	Union	0.14	0.45	0.91	
Fed Speeches	Categorical		0.04	0.11	0.10
	Numerical	0.04		0.05	0.17
	Tense	0.11	0.05		0.87
	Union	0.10	0.17	0.87	
IMF	Categorical		0.02	0.14	0.21
	Numerical	0.02		0.10	0.24
	Tense	0.14	0.10		0.89
	Union	0.21	0.24	0.89	
European Commission	Categorical		0.07	0.13	0.20
	Numerical	0.07		0.28	0.39
	Tense	0.13	0.28		0.87
	Union	0.20	0.39	0.87	
European Council	Categorical		-0.05	0.19	0.19
	Numerical	-0.05		0.39	0.51
	Tense	0.19	0.39		0.91
	Union	0.19	0.51	0.91	
European Parliament	Categorical		-0.07	0.02	0.17
	Numerical	-0.07		0.25	0.31
	Tense	0.02	0.25		0.83
	Union	0.17	0.31	0.83	

Notes: Table displays the correlation structure of the document-level measures of temporal orientation. For each document, we compute the overall fraction of temporal expressions identified by SUTime and TMV that are assigned to the past. The union measure is the fraction of sentences that are identified as containing at least one temporal expression assigned to the past. We compute the correlation of these measures across documents.

Table 18: The Correlation Structure of Temporal Indicators Across Corpora – Future Indicators

Institution		Categorical	Numerical	Tense	Union
ECB Speeches	Categorical		-0.12	-0.07	0.19
	Numerical	-0.12		0.44	0.47
	Tense	-0.07	0.44		0.89
	Union	0.19	0.47	0.89	
Fed Speeches	Categorical		-0.03	0.15	0.36
	Numerical	-0.03		0.15	0.32
	Tense	0.15	0.15		0.83
	Union	0.36	0.32	0.83	
IMF	Categorical		0.11	0.17	0.34
	Numerical	0.11		0.23	0.31
	Tense	0.17	0.23		0.92
	Union	0.34	0.31	0.92	
European Commission	Categorical		0.16	0.07	0.25
	Numerical	0.16		0.27	0.44
	Tense	0.07	0.27		0.86
	Union	0.25	0.44	0.86	
European Council	Categorical		0.02	0.22	0.43
	Numerical	0.02		0.09	0.35
	Tense	0.22	0.09		0.84
	Union	0.43	0.35	0.84	
European Parliament	Categorical		0.04	0.08	0.24
	Numerical	0.04		0.17	0.35
	Tense	0.08	0.17		0.81
	Union	0.24	0.35	0.81	

Notes: Table displays the correlation structure of the document-level measures of temporal orientation. For each document, we compute the overall fraction of temporal expressions identified by SUTime and TMV that are assigned to the future. The union measure is the fraction of sentences that are identified as containing at least one temporal expression assigned to the future. We compute the correlation of these measures across documents.

Table 19: Probability of SUTime Temporal References Conditional on TMV Temporal References

Corpus	Probability	Past	Present	Future
ECB Speeches	$\Pr(\cdot)$	0.121	0.086	0.084
	$\Pr(\cdot \mid \text{Past Tense})$	0.313	0.061	0.063
	$\Pr(\cdot \mid \text{Present Tense})$	0.090	0.090	0.083
	$\Pr(\cdot \mid \text{Future Tense})$	0.071	0.090	0.116
Fed Speeches	$\Pr(\cdot)$	0.155	0.089	0.082
	$\Pr(\cdot \mid \text{Past Tense})$	0.345	0.059	0.064
	$\Pr(\cdot \mid \text{Present Tense})$	0.109	0.097	0.081
	$\Pr(\cdot \mid \text{Future Tense})$	0.100	0.095	0.116
IMF Speeches	$\Pr(\cdot)$	0.112	0.081	0.071
	$\Pr(\cdot \mid \text{Past Tense})$	0.282	0.059	0.047
	$\Pr(\cdot \mid \text{Present Tense})$	0.086	0.084	0.067
	$\Pr(\cdot \mid \text{Future Tense})$	0.078	0.088	0.117
European Commission	$\Pr(\cdot)$	0.101	0.100	0.083
	$\Pr(\cdot \mid \text{Past Tense})$	0.311	0.095	0.065
	$\Pr(\cdot \mid \text{Present Tense})$	0.073	0.103	0.074
	$\Pr(\cdot \mid \text{Future Tense})$	0.082	0.087	0.143
European Council	$\Pr(\cdot)$	0.122	0.090	0.071
	$\Pr(\cdot \mid \text{Past Tense})$	0.304	0.072	0.057
	$\Pr(\cdot \mid \text{Present Tense})$	0.079	0.095	0.064
	$\Pr(\cdot \mid \text{Future Tense})$	0.094	0.084	0.146
European Parliament	$\Pr(\cdot)$	0.086	0.118	0.053
	$\Pr(\cdot \mid \text{Past Tense})$	0.203	0.096	0.034
	$\Pr(\cdot \mid \text{Present Tense})$	0.065	0.124	0.051
	$\Pr(\cdot \mid \text{Future Tense})$	0.078	0.107	0.088

Notes: Table displays the probability of observing at least one numerical or categorical temporal reference in a sentence (identified by SUTime), conditional on observing a tense temporal reference in the same sentence (identified by TMV). We compute the probabilities conditional on different values of the tense temporal expression, for different values of the numerical/categorical temporal expression.

Table 20: Probability of TMV Temporal References Conditional on SUTime Temporal References

Corpus	Probability	Past	Present	Future
ECB Speeches	Pr(·)	0.334	0.706	0.210
	Pr(· | Past Num. or Cat.)	0.556	0.606	0.127
	Pr(· | Present Num. or Cat.)	0.158	0.803	0.239
	Pr(· | Future Num. or Cat.)	0.162	0.767	0.311
Fed Speeches	Pr(·)	0.408	0.664	0.194
	Pr(· | Past Num. or Cat.)	0.597	0.554	0.129
	Pr(· | Present Num. or Cat.)	0.195	0.801	0.224
	Pr(· | Future Num. or Cat.)	0.219	0.761	0.305
IMF Speeches	Pr(·)	0.281	0.699	0.227
	Pr(· | Past Num. or Cat.)	0.467	0.615	0.151
	Pr(· | Present Num. or Cat.)	0.141	0.787	0.235
	Pr(· | Future Num. or Cat.)	0.124	0.743	0.346
European Commission	Pr(·)	0.268	0.714	0.249
	Pr(· | Past Num. or Cat.)	0.492	0.618	0.182
	Pr(· | Present Num. or Cat.)	0.160	0.814	0.198
	Pr(· | Future Num. or Cat.)	0.128	0.717	0.380
European Council	Pr(·)	0.361	0.651	0.185
	Pr(· | Past Num. or Cat.)	0.596	0.537	0.131
	Pr(· | Present Num. or Cat.)	0.193	0.815	0.155
	Pr(· | Future Num. or Cat.)	0.178	0.647	0.303
European Parliament	Pr(·)	0.327	0.805	0.282
	Pr(· | Past Num. or Cat.)	0.559	0.678	0.238
	Pr(· | Present Num. or Cat.)	0.217	0.894	0.251
	Pr(· | Future Num. or Cat.)	0.171	0.830	0.417

Notes: Table displays the probability of observing at least one tense temporal reference in a sentence (identified by TMV), conditional on observing at least one numerical or categorical temporal reference in the same sentence (identified by SUTime). We compute the probabilities conditional on different values of the numerical/categorical temporal expression, for different values of the tense temporal expression.

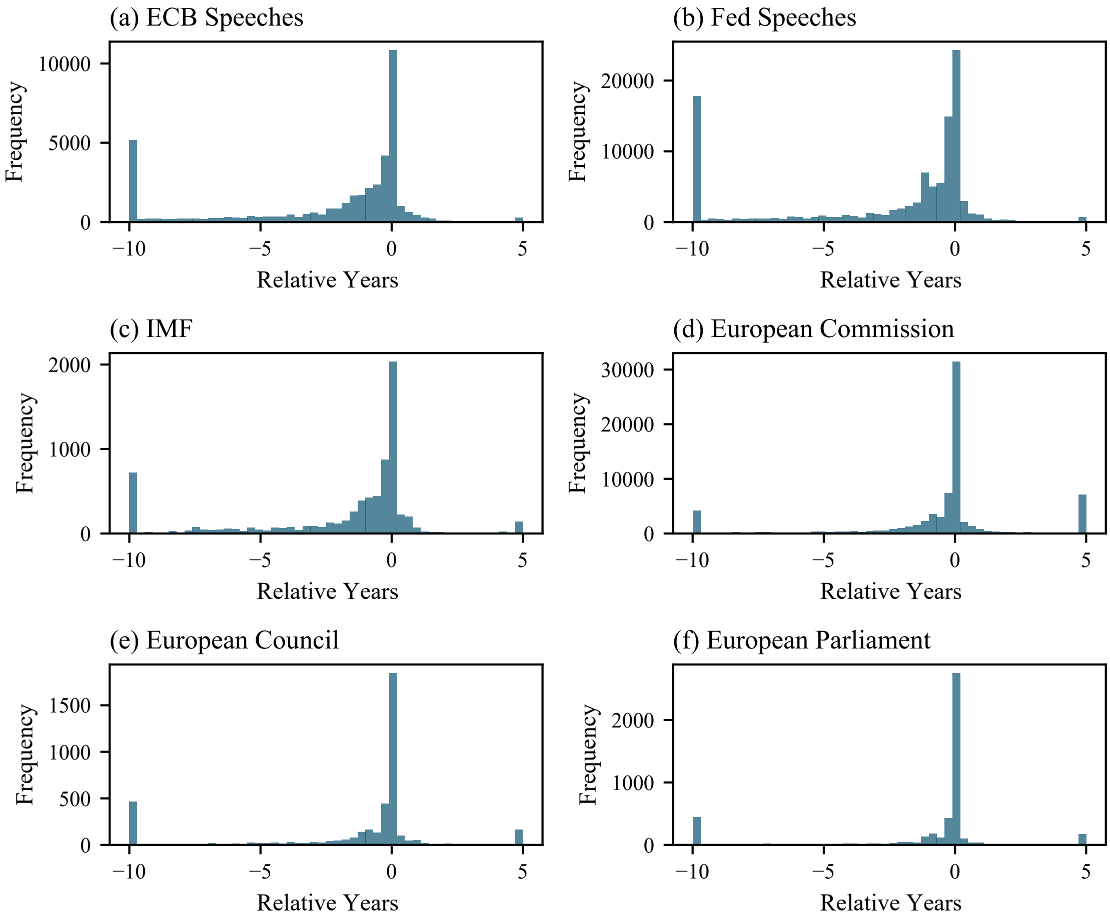
Table 21: Table Displaying Characteristics of the Distribution of Temporal Measures Across Documents Across Corpora

	Parser	Categorical				Numerical				Tense				Union			
	Statistic	Median	Std.	Skew	Kurt.	Median	Std.	Skew	Kurt.	Median	Std.	Skew	Kurt.	Median	Std.	Skew	Kurt.
Orientation	Dataset																
Past	ECB Speeches	0.14	0.14	1.76	8.95	0.69	0.21	-1.01	3.78	0.13	0.09	1.32	5.46	0.24	0.12	0.80	3.75
	Fed Speeches	0.16	0.17	1.59	7.34	0.75	0.22	-1.23	4.45	0.18	0.15	2.22	10.49	0.32	0.17	1.22	5.14
	IMF	0.14	0.17	1.42	6.26	0.67	0.20	-0.48	3.27	0.12	0.09	2.20	12.51	0.22	0.11	1.85	11.56
	European Commission	0.10	0.18	1.85	7.71	0.47	0.22	-0.05	2.56	0.10	0.07	1.58	7.52	0.19	0.11	1.08	5.05
	European Council	0.18	0.20	1.53	6.47	0.53	0.19	-0.67	3.24	0.17	0.09	1.23	6.02	0.26	0.11	0.75	4.13
	European Parliament	0.00	0.31	1.58	4.35	0.50	0.40	0.05	1.49	0.12	0.14	1.67	7.27	0.25	0.21	1.07	4.53
Future	ECB Speeches	0.40	0.21	0.06	2.65	0.08	0.13	1.48	5.41	0.14	0.07	1.27	6.17	0.28	0.10	0.73	4.33
	Fed Speeches	0.36	0.23	0.24	2.71	0.06	0.12	2.35	11.31	0.12	0.07	1.85	21.01	0.25	0.12	0.57	4.57
	IMF	0.38	0.22	0.16	2.67	0.12	0.14	1.73	9.03	0.15	0.07	0.93	4.55	0.27	0.11	0.58	3.31
	European Commission	0.33	0.27	0.38	2.51	0.20	0.19	0.69	2.99	0.15	0.07	0.89	4.96	0.27	0.11	0.74	4.32
	European Council	0.36	0.23	0.28	3.01	0.20	0.17	1.33	6.31	0.11	0.05	1.24	7.02	0.21	0.08	1.20	7.47
	European Parliament	0.00	0.31	1.50	4.09	0.00	0.32	1.49	3.95	0.12	0.12	1.59	7.88	0.27	0.21	0.99	4.48

Notes: Table shows the distribution of the overall measures of temporal orientation per document. For the union, categorical (SUTime), numerical (SUTime), and tense (TMV) measures of future and past orientation, we report the median and upper and lower quantiles of the distribution across documents.

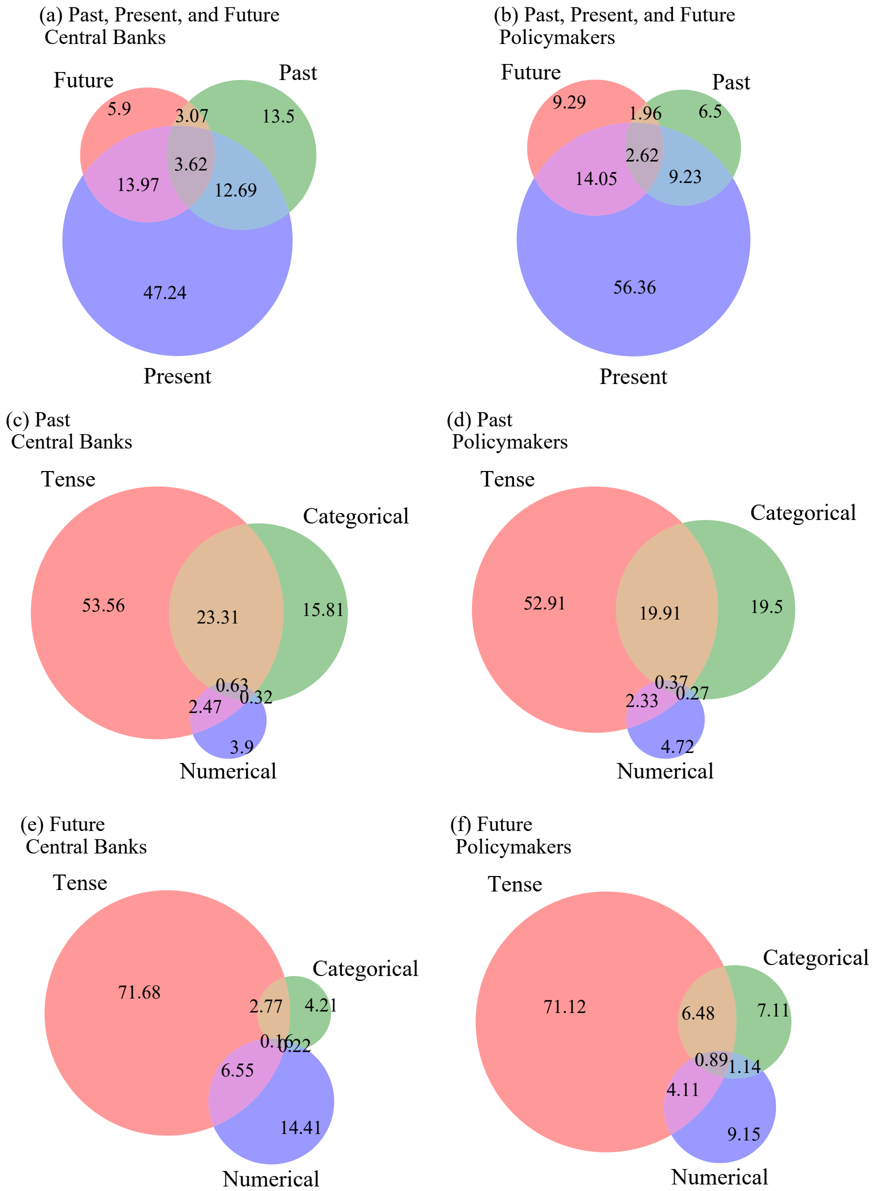
9 Figures

Figure 1: Histograms of Numerical Temporal References Across Corpora



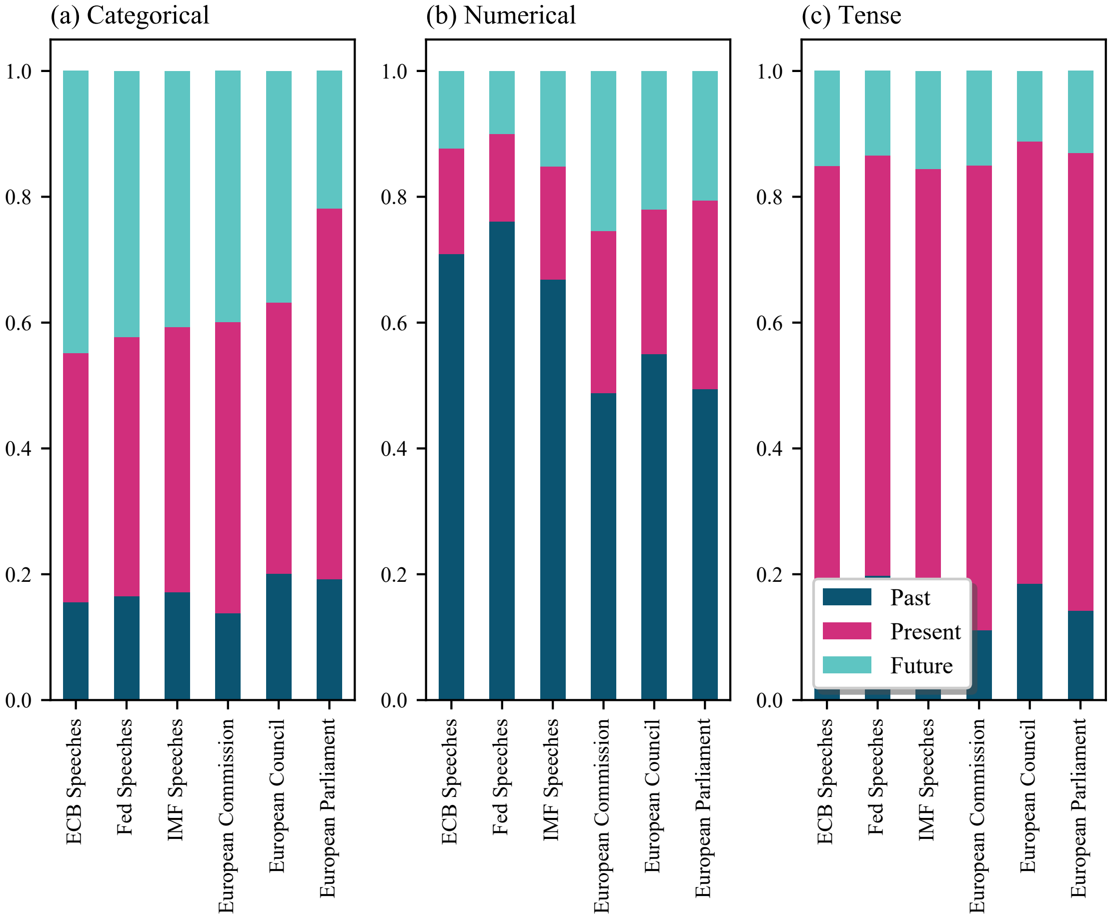
Notes: Figure displays histograms for numerical temporal references across corpora, for 50 bins. Note that we have gathered all references with relative dates below -10 years into a single bin. We have performed an equivalent operation for all references with relative dates above 5 years. This is to aid visualisation.

Figure 2: Representation of The Structure of Temporal References Across Sentences



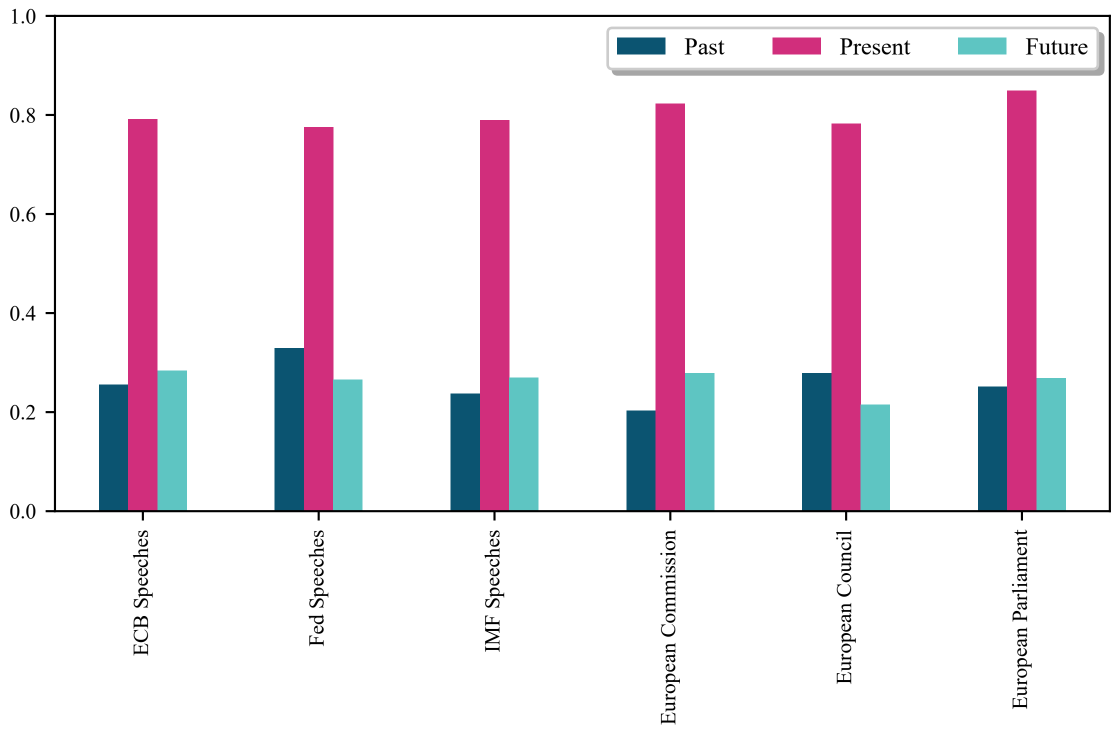
Notes: Figure presents Venn diagrams representing the assignment of sentences to past, present, and future. Column 1 displays diagrams based on data from central banks (ECB and Fed), column 2 displays diagrams based on data from other policymakers (European Commission, European Council, European Parliament, IMF). Panel (a) and (b) show the allocation of sentences to past, present, and future according to the Union measure studied in the main text. Sentences can be assigned more than one temporal orientation according to this measure. Panel (c) and (d) shows the relation between sentences assigned to the future, and indicates whether sentences were assigned to future according to the TMV (tense), or SUTime (categorical and numerical). Sentences can be assigned to the future by more than one parser. Panel (e) and (f) display the same exercise for past temporal expressions.

Figure 3: Past, Present and Future Orientation Across Corpora



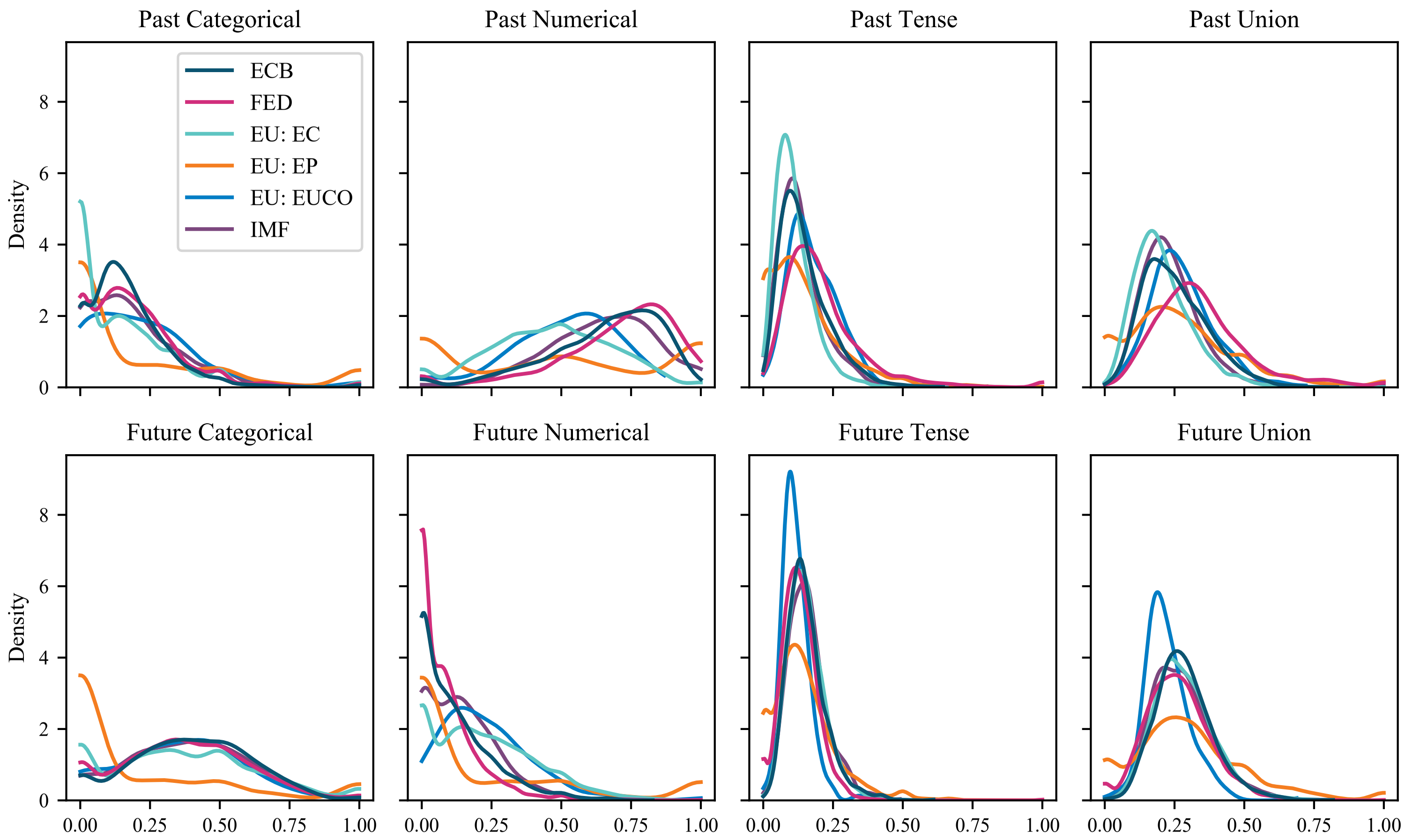
Notes: Figure displays the breakdown of temporal references across the broad categories of past, present and future, sub-divided by corpora.

Figure 4: Union Temporal Orientation Across Corpora



Notes: Figure displays the breakdown of temporal references across the broad categories of past, present and future, sub-divided by corpora.

Figure 5: Kernel Density Plots by Corpora



Notes: Each panel displays a representation of the distribution of the per-document fraction of sentences containing at least one temporal reference, across the different parsers.