Whatever it takes to understand a central banker

Embedding their words using neural networks.

Central bankers go data driven
De Nederlandsche Bank

Johannes Zahner

Philipps-Universität Marburg

May 2022

Table of Contents

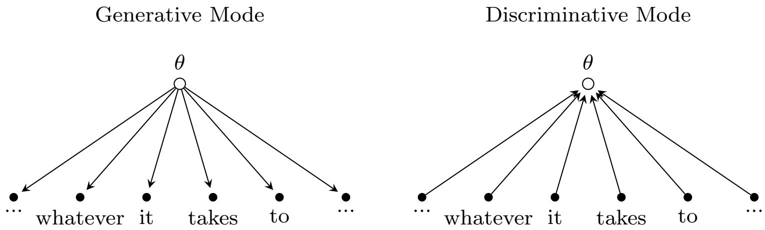
Motivation I

"Within our mandate, the ECB is ready to do whatever it takes to preserve the euro. And believe me, it will be enough."

–Mario Draghi (2012)

Philipps Universität Marburg

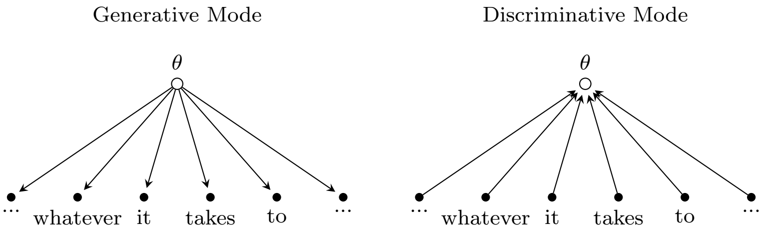
Motivation II



Notes: The illustration is an adaption from Lowe (2021, p.10).

→ Communication discrimination is most effective when the language used is *stable*, *homogeneous*, and *represented in its richness*.

Motivation II



Notes: The illustration is an adaption from Lowe (2021, p.10).

→ Communication discrimination is most effective when the language used is *stable*, *homogeneous*, and *represented in its richness*.

Motivation III

Current approaches of text quantification:

1. Dictionary approaches: *"language is not merely a bag of words but a tool with particular properties which have been fashioned in the course of its use"* Harris (1954) → *strong prior* of the informativeness of the vast majority of communicated terms.
2. Public language model: Not adequate in quantifying *technical monetary policy language*.

Motivation III

Current approaches of text quantification:

1. Dictionary approaches: *"language is not merely a bag of words but a tool with particular properties which have been fashioned in the course of its use"* Harris (1954) → *strong prior* of the informativeness of the vast majority of communicated terms.
2. Public language model: Not adequate in quantifying *technical monetary policy language*.

Motivation III

Current approaches of text quantification:

1. Dictionary approaches: *"language is not merely a bag of words but a tool with particular properties which have been fashioned in the course of its use"* Harris (1954) → *strong prior* of the informativeness of the vast majority of communicated terms.
2. Public language model: Not adequate in quantifying *technical monetary policy language*.

Motivation III

Current approaches of text quantification:

1. Dictionary approaches: *"language is not merely a bag of words but a tool with particular properties which have been fashioned in the course of its use"* Harris (1954) $\rightarrow$ *strong prior* of the informativeness of the vast majority of communicated terms.
2. Public language model: Not adequate in quantifying *technical monetary policy language*.

Paper Overview I

1. We collect a *novel text-corpus* unparalleled in size and diversity.
   → 23.000 speeches by 130 central banks.
2. Introduction of a novel *machine learning* algorithms for text quantifying + comparison of a multitude of different algorithms according to objective criteria.
3. Develop a *language model* previously unseen in monetary policy (and likely economics)
4. Demonstrations of the language model in various applications, such as comparing central bank objectives, measuring the effect of central bank communication in times heightened uncertainty, and creating a gender bias index.
5. The language model is *publicly available*[1], making this new string of research accessible to other researchers.

[1] sites.google.com/view/whatever-it-takes-2021

Paper Overview I

1. We collect a *novel text-corpus* unparalleled in size and diversity.
   → 23.000 speeches by 130 central banks.
2. Introduction of a novel *machine learning* algorithms for text quantifying + comparison of a multitude of different algorithms according to objective criteria.
3. Develop a *language model* previously unseen in monetary policy (and likely economics)
4. Demonstrations of the language model in various applications, such as comparing central bank objectives, measuring the effect of central bank communication in times heightened uncertainty, and creating a gender bias index.
5. The language model is *publicly available*[1], making this new string of research accessible to other researchers.

[1] sites.google.com/view/whatever-it-takes-2021

Paper Overview I

1. We collect a *novel text-corpus* unparalleled in size and diversity.
   → 23.000 speeches by 130 central banks.
2. Introduction of a novel *machine learning* algorithms for text quantifying + comparison of a multitude of different algorithms according to objective criteria.
3. Develop a *language model* previously unseen in monetary policy (and likely economics)
4. Demonstrations of the language model in various applications, such as comparing central bank objectives, measuring the effect of central bank communication in times heightened uncertainty, and creating a gender bias index.
5. The language model is *publicly available*[1], making this new string of research accessible to other researchers.

[1] sites.google.com/view/whatever-it-takes-2021

Paper Overview I

1. We collect a *novel text-corpus* unparalleled in size and diversity.
   → 23.000 speeches by 130 central banks.
2. Introduction of a novel *machine learning* algorithms for text quantifying + comparison of a multitude of different algorithms according to objective criteria.
3. Develop a *language model* previously unseen in monetary policy (and likely economics)
4. Demonstrations of the language model in various applications, such as comparing central bank objectives, measuring the effect of central bank communication in times heightened uncertainty, and creating a gender bias index.
5. The language model is *publicly available*[1], making this new string of research accessible to other researchers.

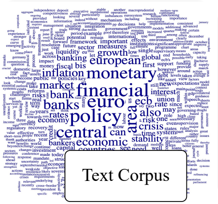
[1] sites.google.com/view/whatever-it-takes-2021

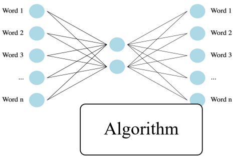
Paper Overview I

1. We collect a *novel text-corpus* unparalleled in size and diversity.
   $\rightarrow$ 23.000 speeches by 130 central banks.
2. Introduction of a novel *machine learning* algorithms for text quantifying + comparison of a multitude of different algorithms according to objective criteria.
3. Develop a *language model* previously unseen in monetary policy (and likely economics)
4. Demonstrations of the language model in various applications, such as comparing central bank objectives, measuring the effect of central bank communication in times heightened uncertainty, and creating a gender bias index.
5. The language model is *publicly available*[1], making this new string of research accessible to other researchers.

[1] sites.google.com/view/whatever-it-takes-bz2021

Paper Overview II





abolishes	0.0420	0.1062	0.0651	0.0660	0.0175	0.0156	...
abolishing	-0.0282	-0.0061	0.0503	0.0913	0.0521	0.0282	
abolishment	0.0252	0.0096	0.0209	0.0287	0.0342	0.0023	
abolition	-0.0458	-0.0191	0.0084	0.0602	0.0539	0.0260	
abominable	-0.0651	0.0610	0.0199	-0.0042	0.0174	0.0624	
aboriginal	0.0072	-0.0641	0.0084	-0.1039	0.0805	0.0443	
abort	0.0672	-0.0163	0.1202	-0.0392	-0.0187	-0.0019	
aborted	0.0419	-0.0322	0.0298	0.0468	-0.0424	0.0270	
abortion	0.0079	-0.0181	0.1283	0.0257	-0.0356	0.0852	
...							

Language Model

Corpus

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

$\rightarrow$ Quantification of central bank communication using machine learning (representation of W in $\theta \rightarrow W$)

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

→ Quantification of central bank communication using machine learning (representation of W in $\theta \rightarrow W$)

Philipps Universität Marburg

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

$\rightarrow$ Quantification of central bank communication using machine learning (representation of W in $\theta \rightarrow W$)

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

→ Quantification of central bank communication using machine learning (representation of W in $\theta \to W$)

Philipps Universität Marburg

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

$\rightarrow$ Quantification of central bank communication using machine learning
(representation of W in $\theta \rightarrow W$)

Literature Overview I

- Relevance of *central bank communication*: Gürkaynak, Sack, Swanson (2005); Blinder, Ehrmann, et. al (2008)
- *Text* as data: Gentzkow, Kelly, Taddy (2019); Baker, Bloom, Davis (2016); Hansen, McMahon (2016); Ehrmann, Talmi (2017); Kalamara, Turrell, et. al (2020)
- Economics/central banking and *machine learning*: Chakraborty, Joseph (2017); Athey (2017); Genberg, Karagedikli (2021)

$\rightarrow$ Quantification of central bank communication using machine learning (representation of W in $\theta \rightarrow W$)

Table of Contents

Language model

- A language model maps a text corpus into an n-dimensional space.
- The model itself can be arbitrarily simple.
- We propose (word) embeddings as a novel language model for quantifying central bank communication.

> *"Approaches [...] which use embeddings as the basis for mathematical analyses of text, can play a role in the next generation of text-as-data applications in social science."*
>
> – *Gentzkow et al. (2019)*

Philipps Universität Marburg

Language model

- A language model maps a text corpus into an n-dimensional space.
- The model itself can be arbitrarily simple.
- We propose (word) embeddings as a novel language model for quantifying central bank communication.

"Approaches [...] which use embeddings as the basis for mathematical analyses of text, can play a role in the next generation of text-as-data applications in social science."

– Gentzkow et al. (2019)

Distinction between algorithms[2]

Count based

- LDA (topic modelling)
- GloVe example

Prediction based

- Word2Vec
- *Doc2Vec*

Pre-trained ***(general language models)***

- GloVe6B
- Word2Vec

Evaluation

[2] Mikolov et al. (2013), Jeffrey Pennington et al. (2014), Blei et al. (2003), Moody (2017)

Distinction between algorithms[2]

Count based

- LDA (topic modelling)
- GloVe example

Prediction based

- Word2Vec
- *Doc2Vec*

Pre-trained *(general language models)*

- GloVe6B
- Word2Vec

Evaluation

[2] Mikolov et al. (2013), Jeffrey Pennington et al. (2014), Blei et al. (2003), Moody (2017)

Distinction between algorithms[2]

Count based

- LDA (topic modelling)
- GloVe example

Prediction based

- Word2Vec
- *Doc2Vec*

Pre-trained *(general language models)*

- GloVe6B
- Word2Vec

Evaluation

[2] Mikolov et al. (2013), Jeffrey Pennington et al. (2014), Blei et al. (2003), Moody (2017)

Word2Vec I

"You shall know a word by the company it keeps"

– *Firth (1957)*

Distributional hypothesis (Harris, 1954): a *word's meaning is based on the context* in which it appears. $\rightarrow$ two words appearing in the same context mean the same thing.

Word2Vec I

"You shall know a word by the company it keeps"

– *Firth (1957)*

Distributional hypothesis (Harris, 1954): a *word's meaning is based on the context* in which it appears. → two words appearing in the same context mean the same thing.

Word2Vec I

"You shall know a word by the company it keeps"

– *Firth (1957)*

Distributional hypothesis (Harris, 1954): a *word's meaning is based on the context* in which it appears. → two words appearing in the same context mean the same thing.

Word2Vec II

What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *rise* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec II

What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}\ \textit{rise}\ \underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$$

Word2Vec II

What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *increase* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Philipps Universität Marburg

Word2Vec II

What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *fall* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

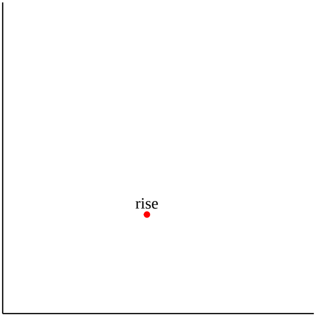
Word2Vec II

What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *GDP* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec II

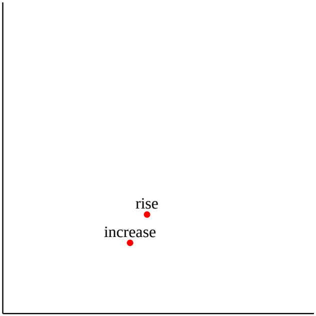


What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *rise* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec II

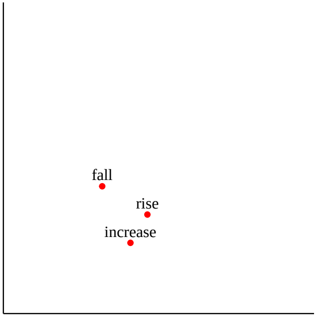


What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *increase* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec II

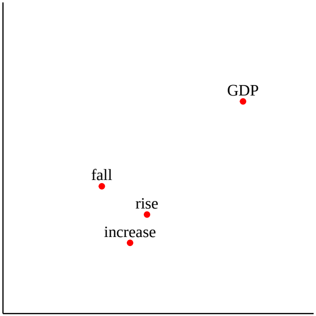


What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ *fall* $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec II



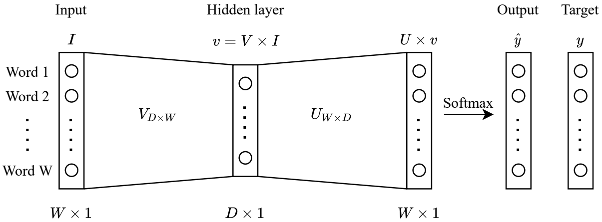
What is context?

Ignazio Visco (2022): *"inflation would be above 3% on average in 2022...*

$\underbrace{\textit{"reflecting the sharp}}_{\textit{3 word context}}$ GDP $\underbrace{\textit{in energy prices"}}_{\textit{3 word context}}$

Word2Vec IV

Ignazio Visco (2022):
... inflation would be above three percent on average in 2022 reflecting the sharp rise in energy prices.



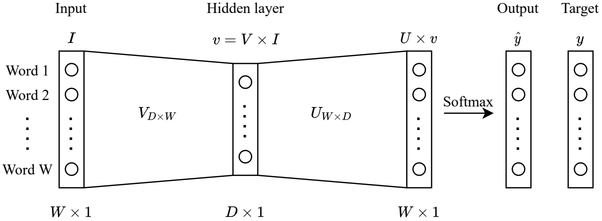
Notes: Model architecture of a feed-forward neural network with three layers. The word-embedding matrix is the projection of the input layer into the hidden layer.

Word2Vec formal

Word2Vec IV

Ignazio Visco (2022):

... inflation would be above three percent on average in 2022 reflecting the sharp rise in energy prices.

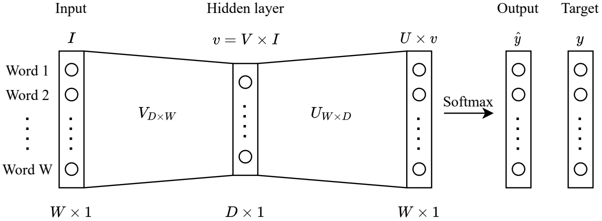


Notes: Model architecture of a feed-forward neural network with three layers. The word-embedding matrix is the projection of the input layer into the hidden layer.

Word2Vec formal

Word2Vec IV

Ignazio Visco (2022):
...inflation would be above three percent on average in 2022 reflecting the sharp rise in energy prices.

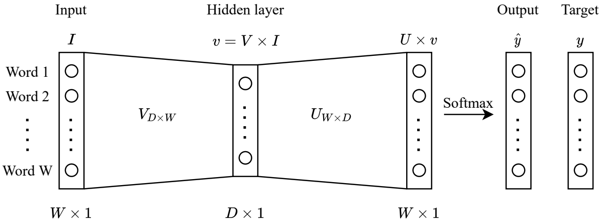


Notes: Model architecture of a feed-forward neural network with three layers. The word-embedding matrix is the projection of the input layer into the hidden layer.

Word2Vec formal

Word2Vec IV

Ignazio Visco (2022):
...inflation would be above three percent on average in 2022 reflecting the sharp rise in energy prices.



Notes: Model architecture of a feed-forward neural network with three layers. The word-embedding matrix is the projection of the input layer into the hidden layer.

Word2Vec formal

Doc2Vec

We use an extension of Word2Vec, called Doc2Vec to create word and document embeddings. Results:

	dim_1	dim_2	dim_3	...
word_1	0.41	-0.32	-0.05	
word_2	0.21	-0.22	0.01	
word_3	0.01	-0.03	-0.31	
...				

→ vector representation for *each word* in the corpus

	dim_1	dim_2	dim_3	...
doc_1	0.22	0.12	-0.26	
doc_2	-0.4	0.01	-0.08	
doc_3	0.51	0.01	-0.04	
...				

→ vector representation for *each document* in the corpus

Evaluation of word embeddings

What are the 'most similar' words to...?

inflation	unemployment	output
core_inflation	unemployment_rate	nonfarm_business
inflation_expectations	natural_rate	sector
economic_slack	joblessness	per_hour
underlying_inflation	jobless	output_growth
inflation_outlook	labor_force	producers
price_inflation	unemployed	manufacturing_output
...	...	...

Note: The underscore is used to highlight collocations. 'Similarity" is measured using cosine similarity.

→ It is evident that our language model is capable of grouping words with semantically similar meanings.

Evaluation of word embeddings

What are the 'most similar' words to...?

inflation	unemployment	output
core_inflation	unemployment_rate	nonfarm_business
inflation_expectations	natural_rate	sector
economic_slack	joblessness	per_hour
underlying_inflation	jobless	output_growth
inflation_outlook	labor_force	producers
price_inflation	unemployed	manufacturing_output
...	...	...

Note: The underscore is used to highlight collocations. 'Similarity" is measured using cosine similarity.

→ It is evident that our language model is capable of grouping words with semantically similar meanings.

Table of Contents

Applications

There is a myriad of potential applications. In the paper we explore the following:

1. Comparing central banks according their *objectives*
2. Indicator of the *ECB's commitment* to act as a lender of last resort
3. Evaluate *gender bias* in the technical language of central bankers
4. (Forecasting central bank policy surprises)

Applications

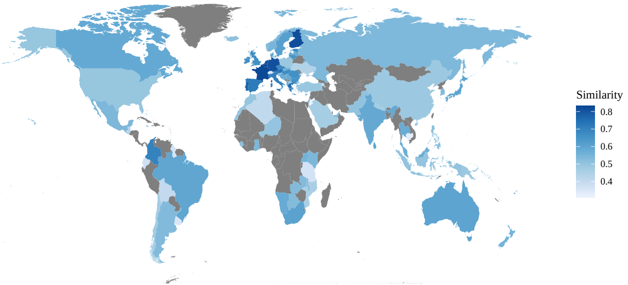
There is a myriad of potential applications. In the paper we explore the following:

1. Comparing central banks according their *objectives*
2. Indicator of the *ECB's commitment* to act as a lender of last resort
3. Evaluate *gender bias* in the technical language of central bankers
4. (Forecasting central bank policy surprises)

Application 1: Monetary policy framework classification I

We investigate factors that influence central bank similarity, using the central bank document similarity (towards the ECB, FED, Reserve Bank of New Zealand (RBNZ)) index as a dependent variable:

Figure: Central banks' similarity to the ECB.



Application 1: Monetary policy framework classification II

Independent variable:

- Monetary policy framework classification
  - Cobham's (2021) classification of *de jure* monetary policy frameworks on annual basis
  - Ten target variables (inflation target, money supply target, ...)
  - 32 mutually distinct categories (loosely structured discretionary targets to fully converging inflation targets)
- Macroeconomic conditions
- Euro area dummy

Application 1: Monetary policy framework classification III

Table: Regression results: Monetary Policy Framework classification

	Similarity					
	RBNZ		FOMC		ECB	
	(2)	(3)	(5)	(6)	(8)	(9)
ITs	*0.11**** (0.03)		*0.09**** (0.02)		*0.16**** (0.03)	
- FIT		0.17*** (0.03)		0.12*** (0.02)		0.12*** (0.03)
- LIT		0.11*** (0.03)		0.10*** (0.02)		0.18*** (0.02)
- FCIT		0.05 (0.04)		0.10*** (0.04)		0.14*** (0.04)
- LCIT		0.06* (0.03)		0.02 (0.03)		0.09*** (0.03)
[...]						
Macro. Controls	yes	yes	yes	yes	yes	yes
Observations	821	821	825	825	821	821
R^2	0.18	0.24	0.19	0.22	0.31	0.37

Note: The notations is adapted from Cobham (2021): ITs = inflation targets; LIT = loose inflation targeting; LCIT = loose converging inflation targeting; FIT = full inflation targeting; FCIT = full converging inflation targeting

Application 1: Monetary policy framework classification III

Table: Regression results: Monetary Policy Framework classification

	Similarity					
	RBNZ		FOMC		ECB	
	(2)	(3)	(5)	(6)	(8)	(9)
ITs	0.11***		0.09***		0.16***	
	(0.03)		(0.02)		(0.03)	
- *FIT*		*0.17****		0.12***		0.12***
		(0.03)		(0.02)		(0.03)
- *LIT*		*0.11****		0.10***		0.18***
		(0.03)		(0.02)		(0.02)
- FCIT		0.05		0.10***		0.14***
		(0.04)		(0.04)		(0.04)
- LCIT		0.06*		0.02		0.09***
		(0.03)		(0.03)		(0.03)
[...]						
Macro. Controls	yes	yes	yes	yes	yes	yes
Observations	821	821	825	825	821	821
R^2	0.18	0.24	0.19	0.22	0.31	0.37

Note: The notations is adapted from Cobham (2021): ITs = inflation targets; LIT = loose inflation targeting; LCIT = loose converging inflation targeting; FIT = full inflation targeting; FCIT = full converging inflation targeting

Application 1: Monetary policy framework classification III

Table: Regression results: Monetary Policy Framework classification

	Similarity					
	RBNZ		*FOMC*		ECB	
	(2)	(3)	(5)	(6)	(8)	(9)
ITs	0.11***		0.09***		0.16***	
	(0.03)		(0.02)		(0.03)	
- *FIT*		*0.17****		*0.12****		0.12***
		(0.03)		(0.02)		(0.03)
- *LIT*		*0.11****		*0.10****		0.18***
		(0.03)		(0.02)		(0.02)
- FCIT		0.05		0.10***		0.14***
		(0.04)		(0.04)		(0.04)
- LCIT		0.06*		0.02		0.09***
		(0.03)		(0.03)		(0.03)
[...]						
Macro. Controls	yes	yes	yes	yes	yes	yes
Observations	821	821	825	825	821	821
R^2	0.18	0.24	0.19	0.22	0.31	0.37

Note: The notations is adapted from Cobham (2021): ITs = inflation targets; LIT = loose inflation targeting; LCIT = loose converging inflation targeting; FIT = full inflation targeting; FCIT = full converging inflation targeting

Philipps Universität Marburg

Application 1: Monetary policy framework classification III

Table: Regression results: Monetary Policy Framework classification

	Similarity					
	RBNZ		*FOMC*		*ECB*	
	(2)	(3)	(5)	(6)	(8)	(9)
ITs	0.11***		0.09***		0.16***	
	(0.03)		(0.02)		(0.03)	
- *FIT*		*0.17****		*0.12****		*0.12****
		(0.03)		(0.02)		(0.03)
- *LIT*		*0.11****		*0.10****		*0.18****
		(0.03)		(0.02)		(0.02)
- FCIT		0.05		0.10***		0.14***
		(0.04)		(0.04)		(0.04)
- LCIT		0.06*		0.02		0.09***
		(0.03)		(0.03)		(0.03)
[...]						
Macro. Controls	yes	yes	yes	yes	yes	yes
Observations	821	821	825	825	821	821
R^2	0.18	0.24	0.19	0.22	0.31	0.37

Note: The notations is adapted from Cobham (2021): ITs = inflation targets; LIT = loose inflation targeting; LCIT = loose converging inflation targeting; FIT = full inflation targeting; FCIT = full converging inflation targeting

Philipps Universität Marburg

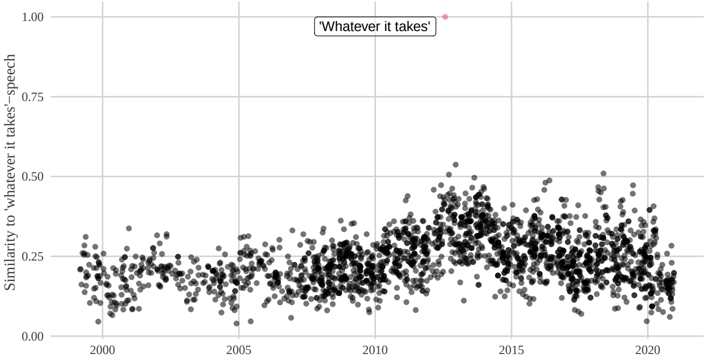
Application 1: Monetary policy framework classification IV

- Adopting an inflation target appears to significantly increase the similarity to an "inflation targeting central banks", i.e. *monetary policy framework → public communication*
- The effect remains positive and statistically significant when controlling for membership in the ea or macroeconomic effects.
- We find different relative weights on the specific inflation target depending on the institution's historical background.
- The adoption of an inflation target remains a highly significant variable when we use the word embeddings.
- One of the factors driving central bank similarity appears to be the adoption of a mutual target.

Application 2: Whatever it takes I

- Focus on the effect of central bank communication in times of heightened uncertainty.
- Focal point is *whatever it takes* speech.
- The speech is widely interpreted as the ECB signaling its willingness to act as a lender of last resort if necessary.
- We calculate the cosine distance between the ECB's remaining speeches to the 'whatever it takes'-event, thereby creating a time-series of an lender of last resort index.

Application 2: Whatever it takes II



Application 2: Whatever it takes III

- To investigate whether the similarity to that speech can calm financial markets in times of heightened uncertainty, we run the following regression:

$$\Delta spread_{10y,t} = wit_{simil,t} + Unc_t + wit_{simil,t} \times Unc_t + X_t + \epsilon_t \qquad (1)$$

- Unc := three different specifications as uncertainty measures
  - $VSTOXX$
  - Decomposition of the VSTOXX into uncertainty (UC) and risk aversion (RA) based Bekaert et. al (2021)
  - ECB's daily CISS index (Hollo et. al, 2021)
- X := a set of control variables[3]

[3] Dummy for the *wit* speech, Moodys agency ratings for Greek bonds, European and U.S. stock prices, monetary policy surprises (Altavilla et. al, 2019), and a dummy for the ECB's different central bank presidents.

Application 2: Whatever it takes IV

Table: Regression results: Whatever it takes

	$\Delta spread_{10y}$		
$Unc_t =$	$VSTOXX_{pd,t}$	$CISS_{pd}$	UC_{pd}
wit_{simil}	1.416^{***}	0.353^{**}	0.485^{***}
	(0.482)	(0.161)	(0.179)
$wit_{simil} \times Unc_t$	-0.070^{***}	-2.911^{**}	-0.020^{***}
	(0.026)	(1.262)	(0.007)
Unc_t	0.016^{***}	0.675^{**}	0.005^{***}
	(0.006)	(0.287)	(0.002)
RA_{pd}			-0.0001
			(0.001)
wit_{dummy}	-1.303^{***}	-1.140^{***}	-1.424^{***}
	(0.317)	(0.406)	(0.278)
L($\Delta spread_{10y}$, 1)	0.248^{**}	0.249^{**}	0.249^{**}
	(0.115)	(0.115)	(0.115)
Constant	-0.318	-0.125	-0.123
	(0.283)	(0.235)	(0.267)
Controls	Yes	Yes	Yes
Observations	2,028	2,028	2,028
R^2	0.116	0.113	0.116

Note: The test statistics are calculated with heteroscedasticity and autocorrelation robust (HAC) standard errors.

Philipps Universität Marburg

Application 2: Whatever it takes V

- All uncertainty measures show a positive correlation with bond spreads.
- $\text{wit}_{\text{simil}}$ is positive and highly significant
- At low uncertainty ($VSTOXX < 19$), the coefficient is positive and becomes negative with increasing uncertainty.
- A possible explanation for the initial positive effect would be that a *whatever it takes* speech has exactly the opposite effect at low uncertainty, i.e. when financial markets are calm, such a speech could be interpreted as a signal of impending troubles.
- Both the sign of the variable of interest remains the same using other uncertainty indicies.

→ Speeches can lower the spread between government bonds when tensions are high and may thus be part of a targeted forward guidance strategy.

Table of Contents

Conclusion

- Quantifying central bank communication has developed to be a substantial entity in monetary policy, with dictionary approaches at the forefront of current techniques.
- We expand the literature on three fronts:
  1. A *text-corpus that is unparalleled in size* and diversity.
  2. Introduction of embeddings, a novel approach from computational linguistics to quantifying texts
  3. Provision of *high quality text-representations* for central bank communication → *open source:* sites.google.com/view/whatever-it-takes-bz2021
- We highlighted the broad applicability illustrating three/four examples in the fields of measuring objectives, financial uncertainty, and gender bias.

Conclusion

- Quantifying central bank communication has developed to be a substantial entity in monetary policy, with dictionary approaches at the forefront of current techniques.
- We expand the literature on three fronts:
  1. A *text-corpus that is unparalleled in size* and diversity.
  2. Introduction of embeddings, a novel approach from computational linguistics to quantifying texts
  3. Provision of *high quality text-representations* for central bank communication → *open source:* sites.google.com/view/whatever-it-takes-bz2021
- We highlighted the broad applicability illustrating three/four examples in the fields of measuring objectives, financial uncertainty, and gender bias.

Conclusion

- Quantifying central bank communication has developed to be a substantial entity in monetary policy, with dictionary approaches at the forefront of current techniques.
- We expand the literature on three fronts:
  1. A *text-corpus that is unparalleled in size* and diversity.
  2. Introduction of embeddings, a novel approach from computational linguistics to quantifying texts
  3. Provision of *high quality text-representations* for central bank communication → *open source:* sites.google.com/view/whatever-it-takes-bz2021
- We highlighted the broad applicability illustrating three/four examples in the fields of measuring objectives, financial uncertainty, and gender bias.

Philipps Universität Marburg

Conclusion

- Quantifying central bank communication has developed to be a substantial entity in monetary policy, with dictionary approaches at the forefront of current techniques.
- We expand the literature on three fronts:
  1. A *text-corpus that is unparalleled in size* and diversity.
  2. Introduction of embeddings, a novel approach from computational linguistics to quantifying texts
  3. Provision of *high quality text-representations* for central bank communication → *open source:* sites.google.com/view/whatever-it-takes-bz2021
- We highlighted the broad applicability illustrating three/four examples in the fields of measuring objectives, financial uncertainty, and gender bias.

Thank you very much for your attention!

Literature Overview II

- In the last years, embeddings have entered the realm of monetary policy:
  - Measure similarity in Twitter tweets (Masciandaro et al., 2020),
  - Development of a real-time economic sentiment index (Aguilar et al., 2021)
  - Improvement of the Euro Area uncertainty index (Azqueta-Gavaldon et al., 2019)
  - Decomposition of central bank vague talk (Hu and Sun, 2021),
  - Measure central banker disagreement (Apel, Grimaldi, and Hull, 2019)
- Generally, economic research relies on general language models trained on a general text corpus such as Wikipedia.
- To the best of our knowledge, we are the first to train embeddings on a specific text corpus and apply the language model to a variety of applications.

Text Source

Our text-corpus:

- Collection of speeches published by the BIS[4]
- Complementary meta-information (speaker, title, ...)
- Minimum application of text pre-processing (Mikolov, Yih, et al., 2013)
- Identification of collocations, f.e. *quantitative_easing* (Blaheta and Johnson, 2001)

→ 23.000 documents, more than 100 million individual word-tokens, more than 130 central banks worldwide and over 1,000 individual speakers.

[4] We enrich the corpus with documents gathered from central bank websites (e.g. minutes, press conferences, transcripts, ...).

Textual stability

Is the communication *rhetoric stable*?

- The central bank's jargon –relative word frequency– is stable *across institutions*
- The central bank's jargon is stable *across time*

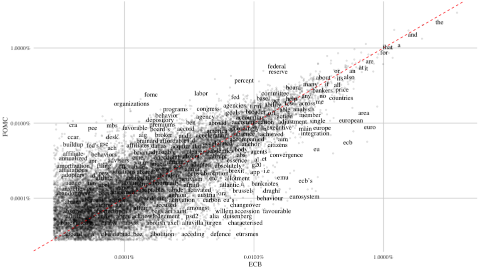


Illustration of jargon between FOMC and ECB

Word2Vec III

Intuitively, we use an algorithm that does the following:

```
Place all words randomly in space
 for each pair of words:
  if in the same context:
     move together
  else:
     move apart.
```

$\rightarrow$ Word2Vec (Mikolov et. al, 2013)

Word2Vec formally

The objective of the network is to maximize the log-likelihood:

$$L = \frac{1}{T} \sum_{t=1}^{T} \log p(w_t|w_c). \tag{2}$$

with

$$p(w_t|w_c) = \frac{exp(v_{w_t}^T v_{w_c})}{\sum_{w=1}^{W} exp(v_w^T v_{w_c})} \tag{3}$$

w_t: target word
w_c: context words

- Iterative optimization algorithm, stochastic gradient descent (Chakraborty and Joseph, 2017)
- window size m is five (Word2Vec) and eight (Doc2Vec)

back

Application 3: Gender bias I

If human *discriminations* and *biases* that are prevalent in the corpus, they persist in the vector data of the embeddings matrix. Example: The most masculine and feminine academic professions according to our embeddings:

Table: Top academic profession for each gender

Female pronouns	Male pronouns
childcare	fashion
wildlife	physics
nursing	architecture
pre-school	mechanics
welfare	computer
education	automation
...	...

Application 3: Gender bias I

"We should mirror the society we serve"

–Christine Lagarde (2020)

"despite all the progress, certain traditional roles are entrenched in people's heads. And this applies not only to men but also to us women. We too are receptive to stereotypes."

–Sabine Lautenschläger (2017)

Application 3: Gender bias I

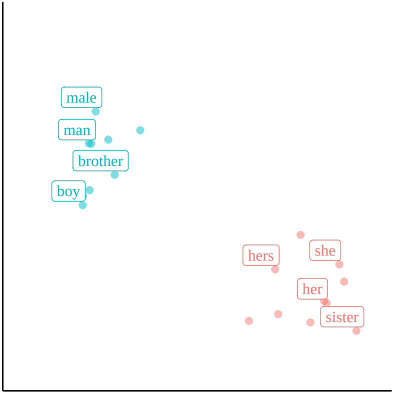
"We should mirror the society we serve"

–Christine Lagarde (2020)

"despite all the progress, certain traditional roles are entrenched in people's heads. And this applies not only to men but also to us women. We too are receptive to stereotypes."

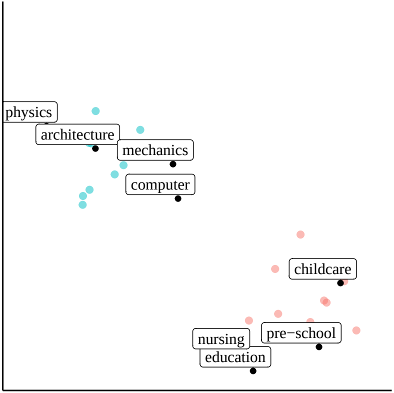
–Sabine Lautenschläger (2017)

Application 3: Gender bias II



- The relative norm distance (RND) measures a group's association with a (potentially) neutral word.
- The *latent bias* of either group can be estimated by their distance towards the neutral term.

Application 3: Gender bias II



- The relative norm distance (RND) measures a group's association with a (potentially) neutral word.
- The *latent bias* of either group can be estimated by their distance towards the neutral term.

Application 3: Gender bias III

To test for underlying biases in our embedding, we collect study programs and their respective gender ratios in Bachelor programs across Europe:

Table: Regression results - gender bias

	Dependent variable:
	Relative norm distance
Fraction of female students	*0.039****
	(0.013)
Constant	-0.030^{***}
	(0.008)
Observations	67
R^2	0.113

Philipps Universität Marburg

Gender bias IV

- Central banks have recognized the problem of gender diversity, f.e. ECB launched its diversity program in 2010
- Has there been a *change in language over time?*
- Focal point: Sabine Lautenschläger: *"100 years of women's suffrage - equality, freedom and democracy"*

Gender bias IV

- Central banks have recognized the problem of gender diversity, f.e. ECB launched its diversity program in 2010
- Has there been a *change in language over time?*
- Focal point: Sabine Lautenschläger: *"100 years of women's suffrage - equality, freedom and democracy"*

Philipps Universität Marburg

Gender bias IV

- Central banks have recognized the problem of gender diversity, f.e. ECB launched its diversity program in 2010
- Has there been a *change in language over time?*
- Focal point: Sabine Lautenschläger: *"100 years of women's suffrage - equality, freedom and democracy"*

Gender bias V

Figure: Similarity of all European Central Bank speeches to Sabine Lautenschläger's speech.

Notes: This graph illustrates the cosine distance between a speech and Sabine Lautenschläger's speech. The solid lines illustrate the mean effect in the corresponding time window.

Gender bias VI

Figure: Similarity of central bank embeddings to Sabine Lautenschläger's speech.

Notes: This graph illustrates the cosine distance between a speech and Sabine Lautenschläger's speech.

Gender bias VII

Results:

- We find significant and economically relevant *positive trend over time*.
- We find *heterogeneity* across central banks in the adoption of gender as a topic.
- Economic recessions do not seem to affect the concentration on gender equality. However, macroeconomic factors have a substantial and significant impact on our index.

→ Evidence of the presence of gender bias in central bankers' language, as well as its lack of persistence.

GloVe

GloVe uses the ratio of word co-occurance probabilities in a corpus to extract word embeddings. The researchers find substantial improvements over Word2Vec. Example:[5]

doc1: *"[...] many euro area countries entered the financial [...]"*
doc2: *"[...] the current macroeconomic policy framework in the euro area [...]"*

	many	euro	area	countries	entered	the	financial	...
many	0	1	0	0	0	0	0	
euro	1	0	2	0	0	1	0	
area	0	2	0	1	0	0	0	
countries	0	0	1	0	1	0	0	
entered	0	0	0	1	0	1	0	
the	0	1	0	0	1	0	1	
financial	0	0	0	0	0	1	0	
...								

back

[5] The context window for estimating this particular co-occurence matrix is 1, i.e. the left and the right word. Jeffrey Pennington et al. (2014) use context windows between 2 and 10.

External evaluation I: Word prediction.

Algorithm	Accuracy	(standard deviation)
Word Embeddings		
Doc2Vec Bow	0.846	0.007
Doc2Vec Bow Pre	0.844	0.009
GloVe	0.831	0.008
Doc2Vec PVDM	0.803	0.009
Doc2Vec PVDM Pre	0.800	0.017
Word2Vec Skipgram	0.678	0.007
GloVe 6B	0.646	0.008
Word2Vec GoogleNews	0.546	0.016
Word2Vec Bow	0.502	0.009
LDA	0.064	0.014

Note: The table shows the evaluation results across the different algorithms introduced in the previous section. The accuracy was evaluated as the Number of correct predictions / Total number of predictions. With regards to the specifications: Bow = (Distributed) Bag Of Words; PVDM = Paragraph Vector Distributed Memory; Pre = pretrained embeddings were used as more efficient starting points.

back

External evaluation II: Prediction of short term interest rates.

Algorithm	3-month Euribor	3-month FFR
Document Embeddings		
Doc2Vec Bow Pre	0.74	0.61
Doc2Vec Bow	0.75	0.59
Doc2Vec PVDM	0.70	0.48
Doc2Vec PVDM Pre	0.67	0.52
LDA	0.55	0.42
Word Embeddings		
Doc2Vec PVDM Pre	0.41	0.35
Doc2Vec Bow Pre	0.40	0.28
Doc2Vec PVDM	0.44	0.22
GloVe	0.38	0.22
Word2Vec GoogleNews	0.36	0.31
GloVe 6B	0.34	0.19
LDA	0.25	0.22
Doc2Vec Bow	0.21	0.25
Word2Vec Bow	0.20	0.21
Word2Vec Skipgram	0.19	0.21

Note: The accuracy was evaluated on a classification task with five categories + one outside option if the model was unsure. Therefore the expected performance would be $1/6 \approx 0.17$. With regards to the specifications: Bow = (Distributed) Bag Of Words; PVDM = Paragraph Vector Distributed Memory; Pre = pretrained embeddings were used as more efficient starting points.

back

Philipps Universität Marburg

Intrinsic evaluation I: Similarity in word embeddings.

inflation	unemployment	output
core_inflation	unemployment_rate	nonfarm_business
inflation_expectations	natural_rate	sector
economic_slack	joblessness	per_hour
underlying_inflation	jobless	output_growth
inflation_outlook	labor_force	producers
price_inflation	unemployed	manufacturing_output
actual_inflation	labor_market	factory
disinflationary	economic_slack	hourly_compensation
inflation_rate	unemployment_rates	business_equipment
disinflation	participation_rate	labor_costs

Note: The table shows the most similar terms to the words *inflation*, *unemployment* and *output* according to the cosine distance. The underscore is used to highlight collocations.

back

Intrinsic evaluation IV: Homonym evaluation.

Table: Intrinsic Evaluation: Homonym across language models.

Doc2Vec	GloVe6B	Word2Vec GoogleNews
basel_committee	zurich	abbr
basle	basle	Tst
capital_accord	zürich	iva
basel_accord	bern	tHe
bcbs	switzerland	Neurol
basle_committee	stuttgart	BASLE
basel_ii	hamburg	PARAGRAPH
basel_iii	cologne	tellus
consultative	lausanne	Def.
minimum_capital	schaffhausen	Complementarity

Note: The table shows for the Doc2Vec and the two genereal corpus models the ten most similar words to the word *basel* according to the cosine distance of the underlying word embeddings. The underscore is used to highlight collocations.

back

Philipps Universität Marburg

Extention of Word2Vec model by adding a document unique feature vector:

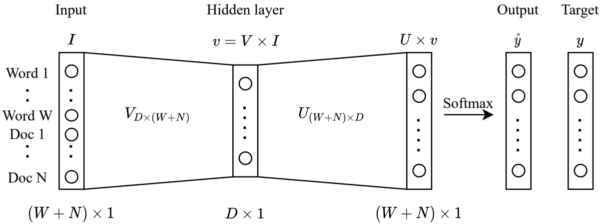


Illustration of Mikolov et al. 2014's Doc2Vec model.

Notes: Illustration of the Doc2Vec model architecture (Mikolov et al. 2014). The only difference to Word2Vec is the additional document ID being fed into the neural network.

→ Each *word* has a vector representation and each *document* has a vector representation.

Doc2Vec

Extention of Word2Vec model by adding a document unique feature vector:

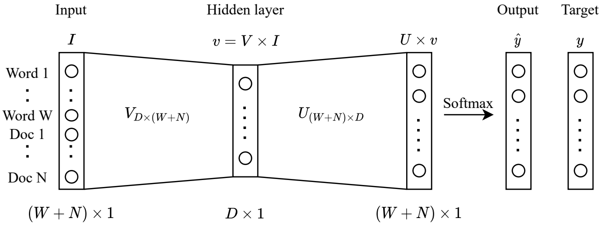


Illustration of Mikolov et al. 2014's Doc2Vec model.

Notes: Illustration of the Doc2Vec model architecture (Mikolov et al. 2014). The only difference to Word2Vec is the additional document ID being fed into the neural network.

→ Each *word* has a vector representation and each *document* has a vector representation.

Evaluation II

Table: Intrinsic Evaluation II: Similarity to the term *basel*

Doc2Vec	GloVe6B	Word2Vec GoogleNews
basel_committee	zurich	abbr
basle	basle	Tst
capital_accord	zurich	iva
basel_accord	bern	tHe
bcbs[6]	switzerland	Neurol
basle_committee	stuttgart	BASLE
basel_ii	hamburg	PARAGRAPH
basel_iii	cologne	tellus
consultative	lausanne	Def.
minimum_capital	schaffhausen	Complementarity

Note: The underscore is used to highlight collocations.

→ Since our language model is *context-specific*, the issue with certain homonyms seems less prevalent than in language models trained on a more general context.

[6] Basel Committee on Banking Supervision (BCBS)

Philipps Universität Marburg