Titre: Title:	European Central Banks monetary policy decisions: A dataset of two decades of press conferences
Auteurs: Authors:	Thierry Warin, & William Sanger
Date:	2018
Type:	Article de revue / Article
Référence: Citation:	Warin, T., & Sanger, W. (2018). European Central Banks monetary policy decisions: A dataset of two decades of press conferences. Data in Brief, 20, 794-798. https://doi.org/10.1016/j.dib.2018.08.061

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/5169/
Version:	Version officielle de l'éditeur / Published version Révisé par les pairs / Refereed
Conditions d'utilisation: Terms of Use:	CC BY

Document publié chez l'éditeur officiel
Document issued by the official publisher

Titre de la revue: Journal Title:	Data in Brief (vol. 20)
Maison d'édition: Publisher:	Elsevier
URL officiel: Official URL:	https://doi.org/10.1016/j.dib.2018.08.061
Mention légale: Legal notice:	© 2018 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Data Article

European Central Bank's monetary policy decisions: A dataset of two decades of press conferences

Thierry Warin [a,*], William Sanger [b,c]

[a] *HEC Montréal, Canada*
[b] *Polytechnique Montréal, Canada*
[c] *CIRANO, Canada*

A R T I C L E   I N F O

A B S T R A C T

The dataset gathers all press conferences in text form made by the first three Presidents of the European Central Bank since its inception in 1998. Press conferences are composed of two elements: (1) introductory statements and (2) questions and answers (Q&A) from journalists. They serve as the main communication vehicle about the monetary policy decision of the ECB's Governing Council. From 1998 to 2016, a total of 205 press conferences have been delivered. The dataset is structured into two main sections: (1) 205 introductory statement of the Presidents of the ECB explaining the monetary policy of the Institution and (2) 205 answers provided by the Presidents to the journalists, hence a total of 410 statements (914,499 words or 3050 pages) about the monetary policy and its context, in text form.

Specifications Table

Subject area	*Economics*
More specific subject area	*Monetary Policy, Institutional Communication*
Type of data	*Text corpus, Figures, Excel-file*

* Corresponding author.
  E-mail address: thierry.warin@hec.ca (T. Warin).

How data was acquired	*Data extracted from the Institution's portal*
Data format	*Raw, filtered*
Experimental factors	*For the second part of the dataset (Q&As), the data concern the answers provided by the ECB*
Experimental features	*Texts were structured using the R programming language. This database could be used to understand the perception of institution through its communication.*
Data accessibility	*The Data is made available in the supplementary material coming with this article.*

Value of the data

- The dataset offers a unique access to almost 20 years of institutional communication.
- The dataset can be used at once to analyze the speeches of the introductory statements of the ECB as well as the answers provided by the Presidents to the journalists during the Q&A segment.
- The dataset can be used to analyze structured as well as unstructured data. The dataset has been precoded: on top of a corpus of text machine-readable (CSV file), the number of words, and questions asked by intervention have been recorded.

1. Data

The dataset provided with this article is a CSV file gathering all Introductory Speeches of the European Central Bank. Throughout the database, a total of 205 statements are provided in raw format. Moreover, after detailing the monetary policy of his institution, the President of the European Central Bank traditionally reply to the journalists covering the Introductory Speeches. The answers to those questions are provided in the CSV file. The next table describes the different variables of the data file shared with this article.

Description of the variables of the dataset:

Name of the variable	Type	Description
date	date	date of the Speech
president	{1; 2; 3}	name of the President (1: W. Duisenberg; 2: J.-C. Trichet; 3: M. Draghi)
firstPart	text	Introductory Speech
firstPartVp	{1; 0}	1: the Vice-President intervened during the first part; 0 otherwise
secondPart	text	Answers to the journalists
secondPartExternalIntervention	{1; 0}	1: someone else than the President intervened during the second part; 0 otherwise
vicePresident	{1; 0}	1: the Vice-President intervened during the second part; 0 otherwise
numberOfAnswers	number	number of answers provided by the President during the Q&A section
link	text	link to the original document (redirecting on the ECB's online portal)

To the best of our knowledge, this is the first dataset allowing for a comprehensive analysis of the relevant communications for the ECB monetary policy [1]. This paper has also some academic and
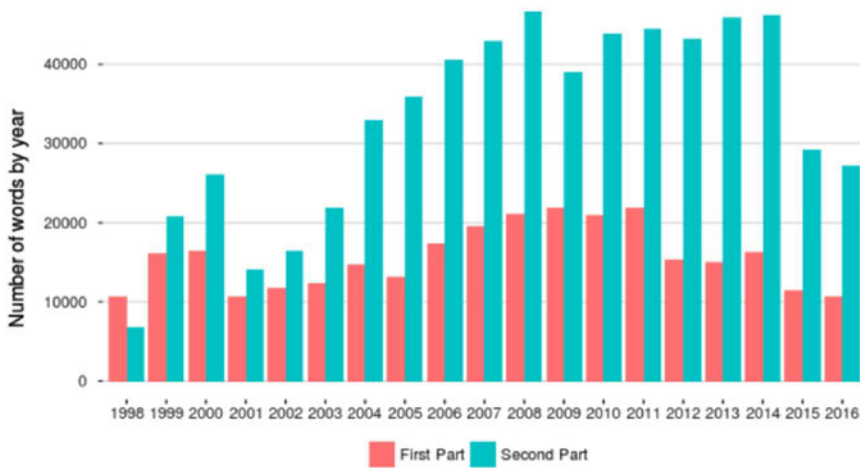
Fig. 1. Number of words by year for each part of the Introductory Statements of the ECB (1998–2016).

managerial implications [2,3]. As such, through the use of unstructured data (Presidents' speeches), researchers can provide a quantitative perspective on the communications generated by the ECB regarding the Eurozone. In term of managerial implications, the same methodologies could be applied to CEO communications or firms annual reports and their implications in terms of perceptions from the financial markets.

This dataset gathers a corpus of introductory statements from October, 13th of 1998 to December, 8th of 2016. In total, 205 press conferences are detailed. Moreover, the answers provided by the three Presidents complete this dataset. In addition to the text, a first layer of coding has been provided for both parts of the statements(the number of words). For the second part of the statements, we manually counted the number of answers provided to the journalists.

2. Experimental design, materials and methods

Data was manually acquired through the Institution's online portal.[1] The dataset is presented as tidy data [4], i.e. each line is an observation, each column is a variable and each element of the dataset is a value (either text, dummy variable, date…).

The number of words could be assessed using the following R code of the stringr library in R [5]:

$$speeches\$nbOfWordsFirstPart < -str\_count(speeches\$firstPart, "//S+") \tag{1}$$

In Fig. 1, we provide the total number of words by year for each part of the dataset.

Fig. 2 breaks down the number of words for each President during the first part of the Introductory Speeches.

Finally, Fig. 3 provides an overview of the number of words for the Q&A part of the press conferences by each President. From this database, quantitative methodologies of text processing could be applied. As such, this database could be extended and complement research on Central Bank communications, especially in the Eurozone, as well as financial markets' perceptions.

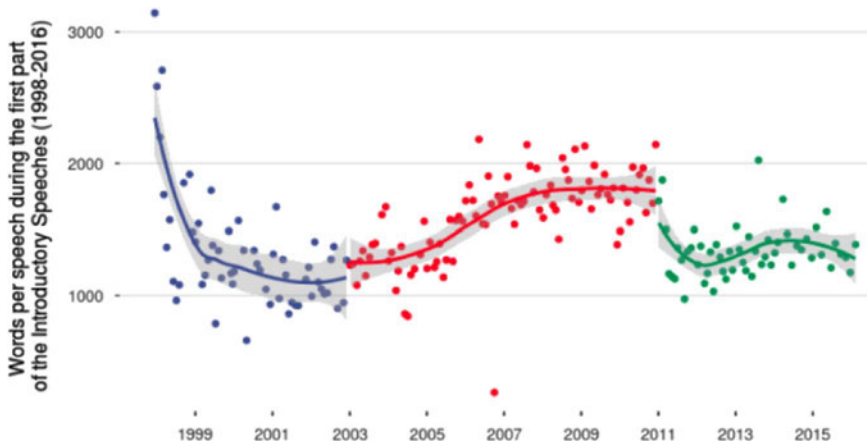[1] https://www.ecb.europa.eu/press/pressconf/.

Fig. 2. Number of words by each President of the ECB during the first part of the press conferences (blue: Wim Duisenberg; red: Jean-Claude Trichet; green: Mario Draghi).
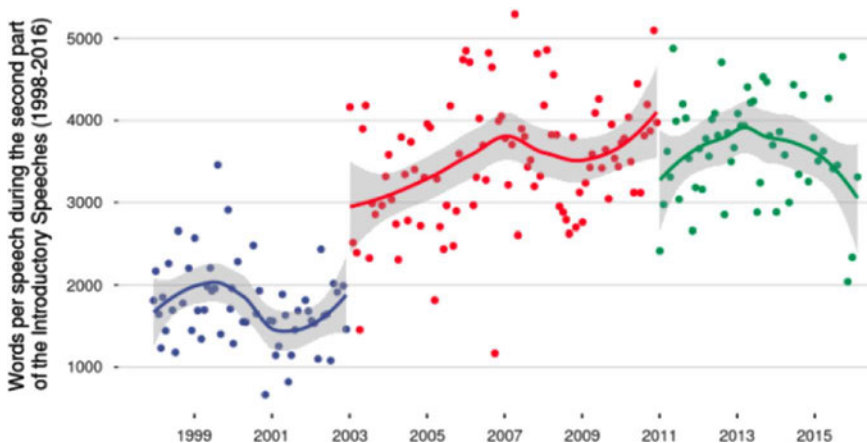

Fig. 3. Number of words by each President of the ECB during the second part of the press conferences (blue: Wim Duisenberg; red: Jean-Claude Trichet; green: Mario Draghi).

Acknowledgements

Transparency document. Supplementary material

Transparency data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2018.08.061.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.dib.2018.08.061.

References

[1] A. Fourçans, R. Vranceanu, The ECB monetary policy: choices and challenges, J. Policy Model. 29 (2) (2007) 181–194. https://doi.org/10.1016/j.jpolmod.2006.12.004.
[2] C. Rosa, (2008), *Talking less and moving the market more: is this the recipe for monetary policy effectiveness?: evidence from the ECB and the Fed.* CEPDP, 855. Centre for Economic Performance, London School of Economics and Political Science, London, UK. ISBN 9780853282501.
[3] D. Amaya, J.-Y. Filbien, The similarity of ECB's communication (SSRN Scholarly Paper ID 2615131), Social Science Research Network, Rochester, NY, 2015 ⟨https://papers.ssrn.com/abstract=2615131⟩.
[4] H. Wickham, Tidy data, J. Stat. Softw. 59 (10) (2014).
[5] H. Wickham, stringr: modern, consistent string processing, R J. 2 (2) (2010).