Collapsing Financial Markets: Unsupervised Modelling of Coronavirus and Trade War News

Carlos Moreno-Pérez [1] and **Marco Minozzo** [2]

[1]University of Verona, Bank of Spain

[2]University of Verona

Modelling with Big Data & Machine Learning:
Measuring Economic Instability
The Bank of England, the Federal Reserve Board and King's College

5 November 2020

Introduction

- We investigate the US financial markets' reactions to news from January 2019 to the first of May 2020.
- As **text data** we use the headlines and the snippets of the articles of the New York Times.
- To understand the **content** and the **sentiment** of the news, we construct text variables using **unsupervised machine learning** algorithms.

Latent Dirichlet Allocation

- To identify the content or theme of the news, we use **Latent Dirichlet Allocation** or **LDA** (Blei, Ng and Jordan, 2003).

Topic	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6
3. Economy / Fed	**economi**	**econom**	**bank**	**cut**	**rate**	**feder**
	0,068	0,062	0,05	0,043	0,037	0,029
29. Coronavirus	**coronaviru**	**test**	**pandem**	**viru**	**spread**	**outbreak**
	0,217	0,057	0,053	0,051	0,037	0,037
33. Brexit	**minist**	**prime**	**brexit**	**may**	**britain**	**european**
	0,077	0,065	0,051	0,05	0,042	0,039
51. Trade war	**china**	**trade**	**deal**	**war**	**chines**	**talk**
	0,17	0,085	0,066	0,058	0,052	0,034
54. Climate change	**chang**	**climat**	**fire**	**california**	**australia**	**water**
	0,135	0,08	0,076	0,054	0,031	0,017

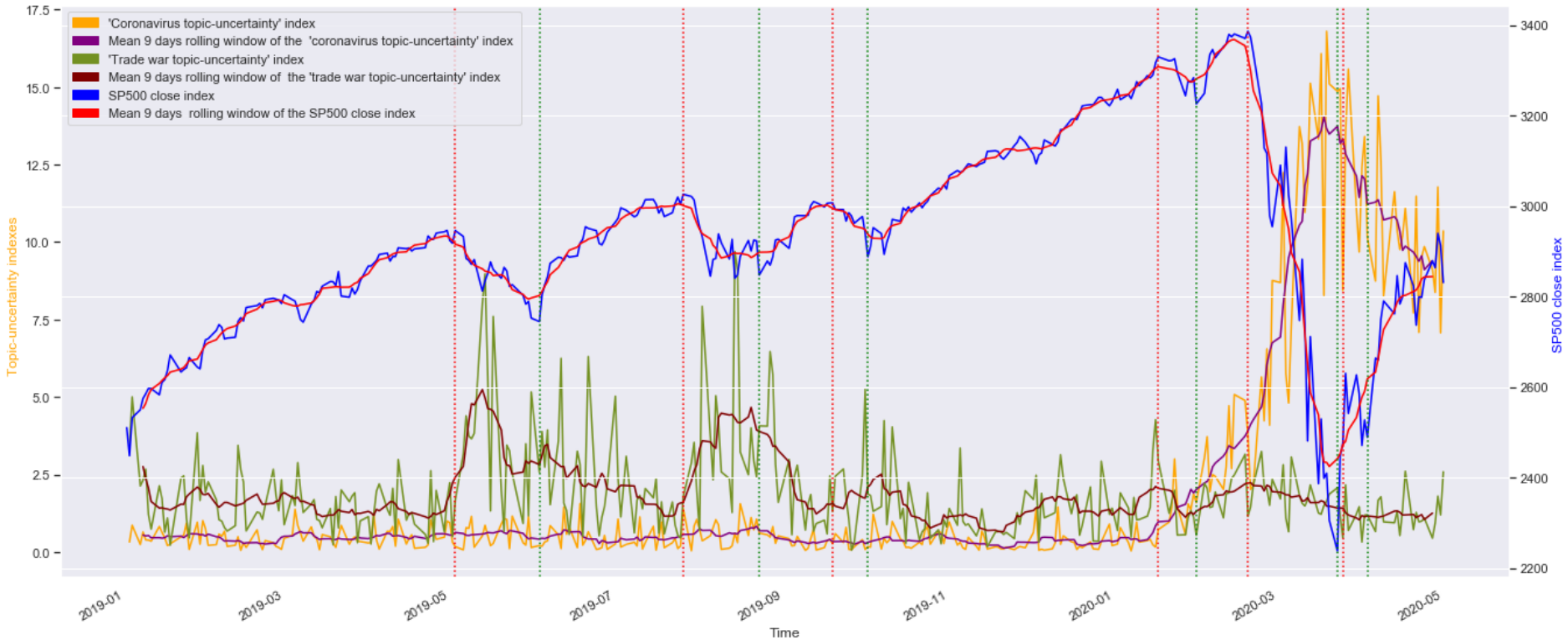
Skip-Gram and K-Means

- To understand the sentiment of the news, we use the **Skip-Gram** model and **K-Means** to build a list of words related to uncertainty.
- We build a daily **uncertainty index** by counting the frequency of the words of this uncertainty list (or uncertainty dictionary).
- To construct this **uncertainty dictionary**, we combine all the words that are in the same clusters of the words 'fear', 'fears', 'worries', 'uncertain' and 'uncertainty' since they share similar semantic meaning.

List of words of the cluster containing the word 'fears'.

analysts, bond_yields, central_banks, climb, damage, drop, exports, factories, fears, fell, financial_markets, fueled, gas, grew, growing, higher, highest, increase, increasing, oil, oil_prices, plunge, policymakers, prices, producers, rate, rattled, rise, rising, slide, slowdown, slowing, slows, slump, spike, supply, tourism, tumbled, worsening.

‘Coronavirus’ and ‘trade war’ uncertainty indexes



The **‘coronavirus’ uncertainty index** is represented by the yellow line; the purple line is the moving average with a 9 days rolling window. The **‘trade war’ uncertainty index** is represented by the green line; the brown line is the moving average with a 9 days rolling window. The blue line is the SP500 close price index.

EGARCH Results

Output for the EGARCH model. Each column tag indicates the financial index (returns) used as a dependent variable in each EGARCH regression. The terms b_2 and b_3 are the EGARCH mean equation's coefficients of the 'coronavirus' and 'trade war' uncertainty indexes, respectively.

	SP500	Nasdaq	Dow Jones	VIX	Treasury yields
b_0	0.96286^{***}	0.09649^{***}	-0.04374^{***}	3.64329^{***}	-0.59790^{***}
	(0.00031)	(0.00008)	(0.00001)	(0.44556)	(0.11712)
b_1	-0.59205^{***}	-0.51618^{***}	-0.55645^{***}	-0.43120^{***}	-0.89927^{***}
	(0.00019)	(0.00031)	(0.00013)	(0.06531)	(0.04520)
b_2	-0.01657^{***}	-0.08537^{***}	-0.11033^{***}	1.04079^{***}	-0.32359^{***}
	(0.00001)	(0.00011)	(0.00004)	(0.02388)	(0.04565)
b_3	-0.75570^{***}	-0.14842^{***}	-0.10868^{***}	0.96973^{***}	-0.18454^{***}
	(0.00024)	(0.00013)	(0.00003)	(0.23096)	(0.04242)
b_4	0.00549^{***}	0.02249^{***}	0.01525^{***}	-0.36312^{***}	0.06286^{***}
	(0.00000)	(0.00001)	(0.00001)	(0.01899)	(0.00851)
θ	0.14934^{***}	0.14304^{***}	0.23023^{***}	0.29421^{***}	0.86025^{***}
	(0.00005)	(0.00005)	(0.00008)	(0.06468)	(0.05467)

Thank you for your attention!