Making text count for macroeconomics: What newspaper text can tell us about sentiment and uncertainty

Eleni Kalamara
King's College London

Arthur Turrell
Bank of England

George Kapetanios
King's College London

Sujit Kapadia
European Central Bank

Chris Redl
Bank of England

The views expressed do not reflect those of the Bank of England, European Central Bank or their policy committees.

Motivation

- Sentiment and uncertainty driving force of business cycles and the decisions of agents (Keynes, 1936)

Motivation

- Sentiment and uncertainty driving force of business cycles and the decisions of agents (Keynes, 1936)
- Measures of sentiment and uncertainty extracted from news text shown to be highly correlated to macro variables (Baker, Bloom and Davis, 2016)

Motivation

- Sentiment and uncertainty driving force of business cycles and the decisions of agents (Keynes, 1936)
- Measures of sentiment and uncertainty extracted from news text shown to be highly correlated to macro variables (Baker, Bloom and Davis, 2016)
- Text-based metrics have advantages of cost, timeliness and scope - real time window on what influences millions of households' views.

Motivation

- Sentiment and uncertainty driving force of business cycles and the decisions of agents (Keynes, 1936)
- Measures of sentiment and uncertainty extracted from news text shown to be highly correlated to macro variables (Baker, Bloom and Davis, 2016)
- Text-based metrics have advantages of cost, timeliness and scope - real time window on what influences millions of households' views.
- Could function like soft data (e.g. surveys) - indicators for policymakers, and inputs into forecasts.

The question

Which text metrics, and text sources, are best as indicators and inputs into forecasts

Run a horse race with many text metrics and several text sources

Data

Text data – UK daily newspapers

Newspaper	Unique articles	% of total articles	Circulation (thousands per day)	First article	Last article
The Guardian	284,293	54.9	138	11/01/1990	11/06/2018
Daily Mirror	139,027	26.8	563	01/03/1995	11/06/2018
Daily Mail	94,747	18.3	1,265	06/01/1990	12/06/2018
Total	518,067	100	1,966	-	-

Descriptive statistics of articles from selected UK newspapers.

Turning Text into time series

Algorithmic-based metrics

Text metric types

Positive and negative dictionary	**Boolean**	**Computer science-based**
Financial stability (Correa et al., 2017)	Economic Uncertainty (Alexopoulos, Cohen et al., 2009)	VADER sentiment (Gilbert, 2014)
Finance oriented (Loughran and McDonald, 2013)	Monetary policy uncertainty (Husted, Rogers and Sun, 2017)	'Opinion' sentiment (Hu et al., 2017; Hu and Liu, 2004)
Afinn sentiment (Nielsen, 2011)	Economic Policy Uncertainty(Baker, Bloom and Davis, 2016)	sentence_sentiment
Harvard IV (used in Tetlock (2007))	Single word counts of "uncertain", "econom", and "sustainab"	punctuation_sentiment
Anxiety-excitement (Nyman et al., 2018)		
term frequency-inverse document frequency (tf-idf) on "uncertain" and "econom"		

The three broad categories of deterministic text metrics used.

From raw text to time series at the article level

Text	alexopoulos_09	stability_sentiment	tf_idf_uncert	vader_sentiment	sentence_sentiment_policy
Global GDP growth picked up during 2016 and has been strong over the past year (Section 1.1). Weighted by countries' shares of UK exports, global growth is estimated to have remained at 0.8% in 2017 Q4. That pace of growth is expected to persist in the near term, above expectations in November. Survey indicators of output (Chart 1.1) and new orders remain robust, particularly in the euro area and United States. Measures of business and consumer confidence are also healthy... (continues)	0	0.04	0.00	0.99	0.30
The economy has struggled and is in a bad state with disappointing performance, unhappy consumers, low confidence with high uncertainty. The policy of the central bank, the Bank of England, will be constrained to be lower for longer. The Bank of England has said that the economy faces a number of risks, particular with respect to monetary policy. Pundits are increasingly concerned about the high risk of a hard Brexit.	1	-0.11	0.05	-0.90	-0.08
The current direction of policy is very bad.	0	-0.25	0.00	-0.54	-0.33
The current direction ofpolicyis very good.	0	0.25	0.00	0.44	0.33
Although currentpolicy is bad for savers, the weather is very good.	0	0.00	0.00	-0.15	0.00
... the current policy is bad for savers. However, the weather is very good.	0	0.00	0.00	-0.15	-0.33

Selected text metrics applied to example text. ▸ Stationarity?

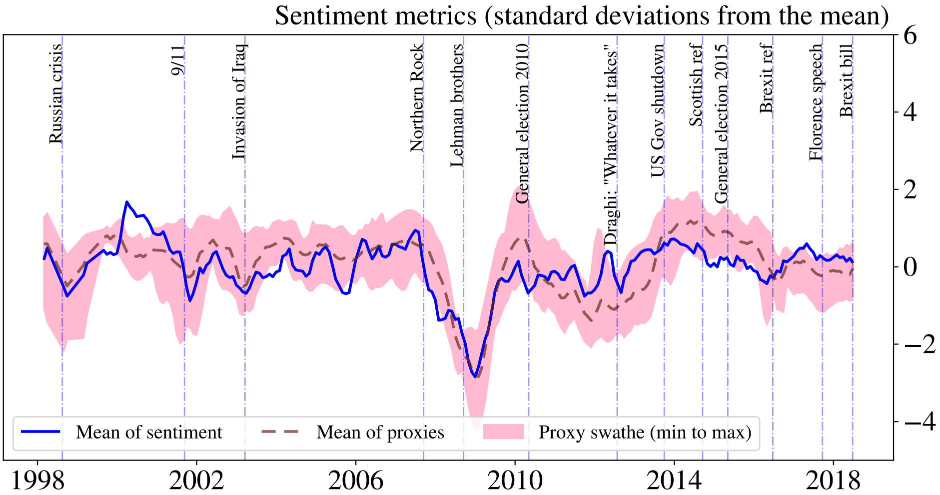
How do text metrics compare with proxies for sentiment and uncertainty?

How do text metrics compare with proxies?

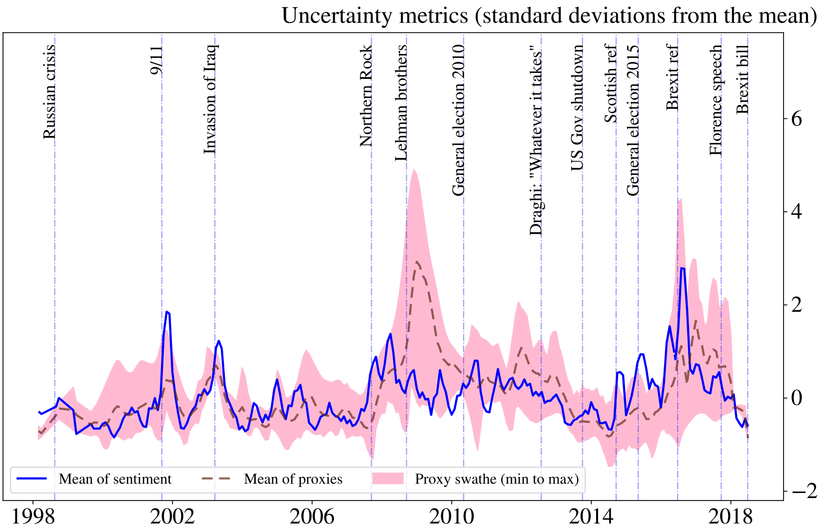
- Proxies are large set of soft data (e.g surveys) and hard (e.g VIX) sentiment and uncertainty see Proxies
- Granger causality texts: mixed results but most metrics Granger cause at least one proxy at 3 month horizon. GC test: sentiment GC test: uncertainty
- Results across newspapers are combined with a weight proportional to their reach suggested by Kennedy and Prat (2017). See weighting details

Visual evidence of correlation with proxies

Mean sentiment using *The Daily Mail*

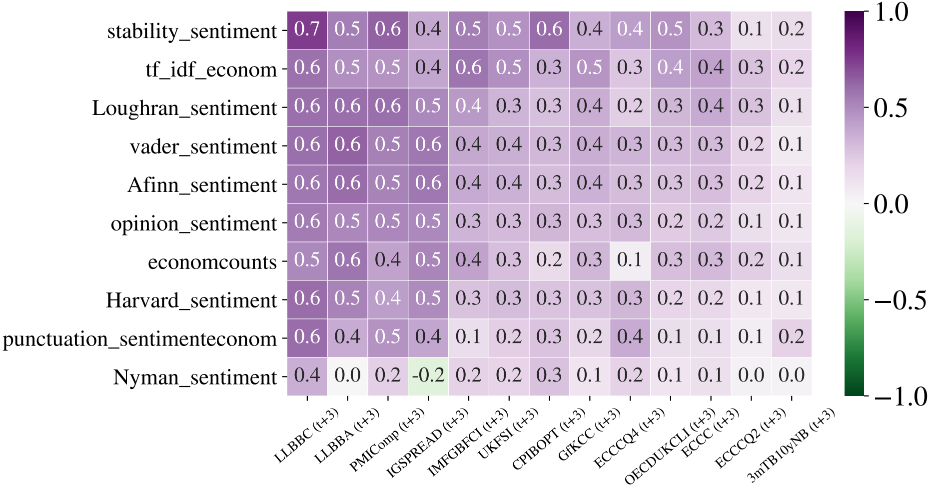


Mean uncertainty using *The Daily Mail*

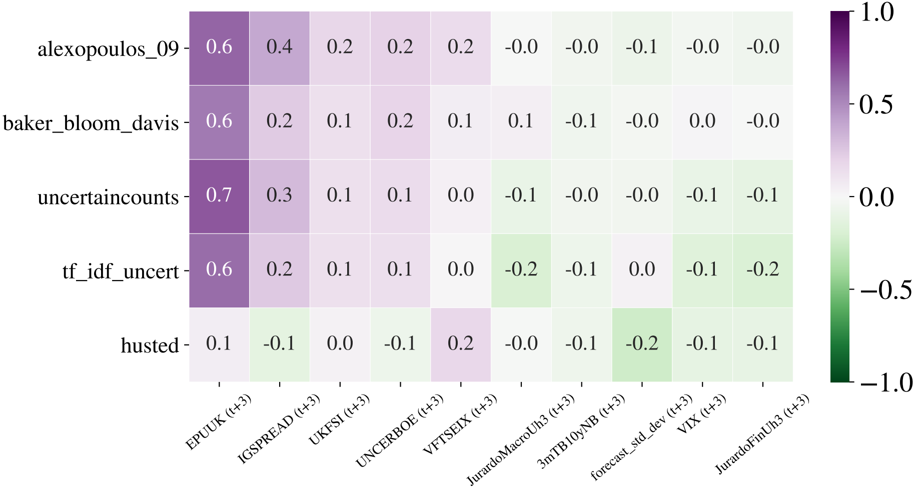


Quantitative evidence of correlation with proxies

Quantitative evidence: correlations at three month horizon (sentiment)



Quantitative evidence: correlations at three month horizon (uncertainty)



Can newspaper text help to predict the future?

Forecast environment

Forecast Environment

- In-sample and out-of sample forecast exercises (rolling window)
- In-sample period of 36M, direct forecast at 6M horizon.
- Metric of success; Looking at RMSE ratios

$$Ratio = \frac{\text{RMSE}_{\text{Text}}}{\text{RMSE}_{\text{Benchmark}}}$$

where $\text{RMSE}_{\text{Benchmark}}$ and $\text{RMSE}_{\text{Text}}$ are out-of-sample root mean squared errors.

  - > 1 : negative performance.
  - < 1 : positive performance.

▸ Targets used.

Forecast exercise 1

Base model is an AR(1): simple but hard to beat.

Text metric-based model is an AR(1) with addition of a single text-indicator:

$$y_{t+h} = \alpha + \beta y_{t+h-1} + \gamma x_t + \epsilon_{t+h}$$

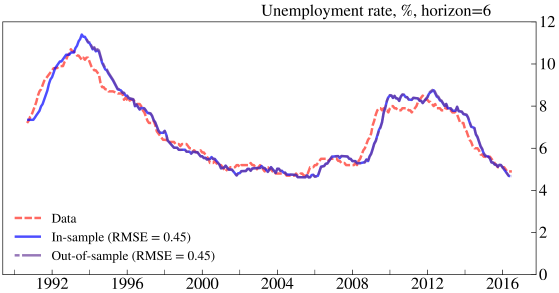
where *y* = target variable, *x* = text indicator, *h*= horizon

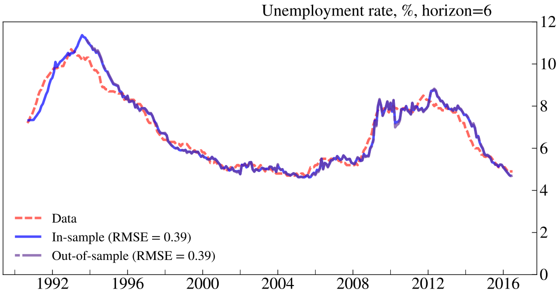
For quarterly targets: use all three months of text metric data from the previous quarter:

$$y_{t+h} = \alpha + \beta y_{t+h-1} + \sum_{i=0}^{2} \gamma x_{t+i/3} + \epsilon_{t+h}$$

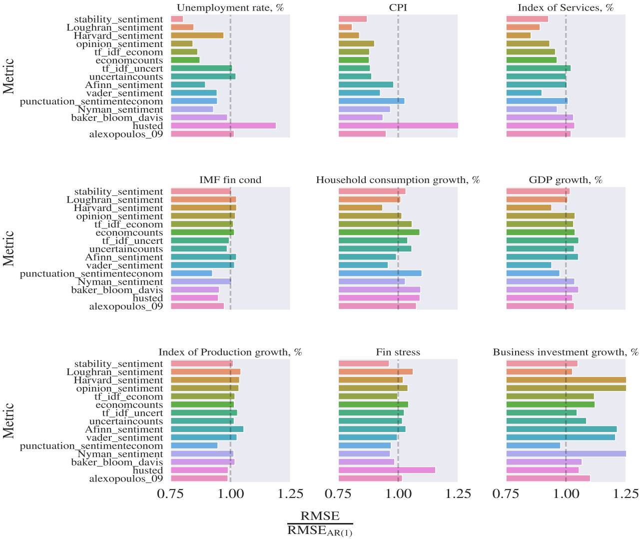
Forecast results

Example: Forecasting unemployment rate. AR(1) with and without ‘tf_idf_econom’ text metric at 6 month horizon





Which text metric does the best?



Nested Diebold-Mariano test for statistical significance

	Target	CPI	Household consumption growth, %	Index of Services, %	Unemployment rate, %
Newspaper	Metric				
Daily Mail	Harvard_sentiment	*	**	*	
	Loughran_sentiment	**			
	Nyman_sentiment				**
	alexopoulos_09	*			
	economcounts	***			
	husted				*
	opinion_sentiment	**			
	stability_sentiment	*			
	tf_idf_econom	*			
	uncertaincounts	*			
Daily Mirror	Harvard_sentiment	*			
	Loughran_sentiment	*			
	economcounts	*			
	stability_sentiment	**			
Guardian	Loughran_sentiment	*			
	alexopoulos_09	*			
	economcounts	**			
	opinion_sentiment				*
	stability_sentiment	**			*
	tf_idf_econom	***			**
	tf_idf_uncert	**			
	uncertaincounts	**			

Statistically significant differences in RMSE at a horizon of $h = 6$ are shown. *, **, *** denote rejection of the null, of no difference in RMSE relative to an AR(1), at the 10%, 5%, 1% levels respectively. Only those targets for which at least one metric-newspaper pair had a p-value of less than 10% are included*.

Forecast exercise 2

Base model is a factor model + AR(1)

Text metric-based model is the base model with addition of a single text-indicator:

$$y_{t+h} = \alpha + \beta \cdot y_{t+h-1} + \sum_j \gamma_j \cdot F_{jt} + \eta \cdot x_t + \epsilon_{t+h}$$

where y = target variable, x = text indicator

For quartely targets:

$$y_{t+h} = \alpha + \beta \cdot y_{t+h-1} + \sum_{i=0}^{2} \left(\sum_j \gamma_{ji} \cdot F_{j,t-1+i/3} + \eta_i \cdot x_{t-1+i/3} \right) + \epsilon_{t+h}$$

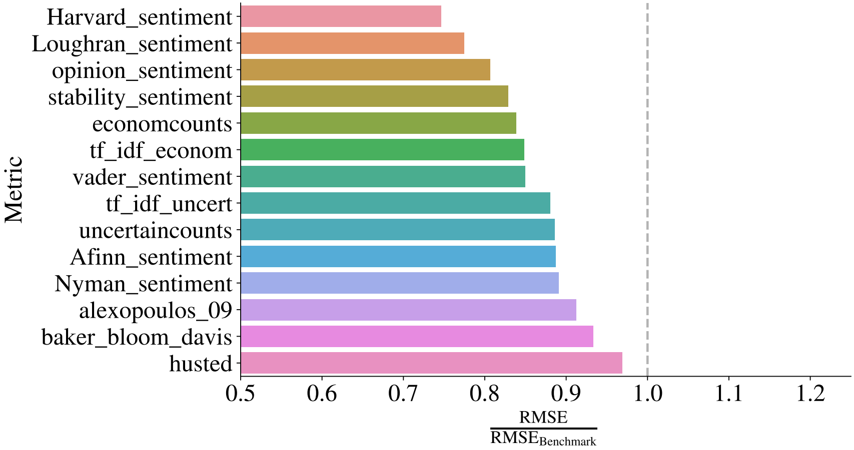
Forecast results

Nested Diebold-Mariano test for statistical significance

Newspaper	Target / Metric	CPI	Household consumption growth, %	GDP growth, %	Index of Services, %	Unemployment rate, %
Daily Mail	Afinn_sentiment					**
	Harvard_sentiment	**	*	**		
	Loughran_sentiment	*				*
	Nyman_sentiment	*				*
	alexopoulos_09	*				
	baker_bloom_davis	*				
	economcounts	*				
	opinion_sentiment	*				**
	stability_sentiment	*				
	tf_idf_econom	*				
	tf_idf_uncert	*				
	vader_sentiment	*		**	*	
Daily Mirror	Harvard_sentiment	*				
	stability_sentiment					*
	vader_sentiment	*				
Guardian	Harvard_sentiment	***		***		
	Loughran_sentiment	***		**		**
	Nyman_sentiment	***				**
	economcounts	**				*
	opinion_sentiment	***				***
	stability_sentiment	***				*
	tf_idf_econom	**		***		*
	tf_idf_uncert					*
	vader_sentiment	***				

Statistically significant differences in RMSE with a horizon of $h = 6$ are shown. *, **, *** denote rejection of the null, at the 10%, 5%, 1% levels respectively. Only those targets for which at least one metric-newspaper pair had a p-value of less than 10% are included*.

RMSEs of the text based models relative to the factor benchmark for GDP growth.



RMSEs with text relative to the factor model. The text metrics used as inputs to the forecast exercise are weighted averages across the three newspapers.

Forecasting using a high dimensional feature space

The Idea

- Text inherently high dimensional.
- Algorithmic metrics offer a pre-defined way to turn text into series.
- Extract as much useful information as possible from the text .

The Idea

- Text inherently high dimensional.
- Algorithmic metrics offer a pre-defined way to turn text into series.
- Extract as much useful information as possible from the text .
- **Model based metrics: provide a model with a plethora of features from the text and let it learn what weight to attach to each feature.**

Steps to follow

- Create term-frequency matrices using the union of the dictionaries and terms up to 3-grams. ▸ see example

- Use Machine Learning methods to forecast the target variables.

- Models include: Lasso (Tibshirani, 1996), Partial least squares (PLS) (Chin, 1998), Ridge regression (Hoerl and Kennard, 1970), Support vector Machines (Drucker et al., 1997), Elastic net (Zou and Hastie, 2005) and Random forest (Breiman, 2001).

Forecasting set up

- **Approach 1:**
  - Benchmark model: AR(1)
  - feature-based model is the benchmark model with addition of the high dimensional feature space, z_t:

$$y_{t+h} = \alpha + \beta \cdot y_{t-1} + \eta \cdot z_t + \epsilon_{t+h}$$

- **Approach 2:**

- Benchmark model: factor model

$$y_{t+h} = \alpha + \beta \cdot y_{t+h-1} + \sum_j \gamma_j \cdot F_{j,t} + e_{t+h}$$

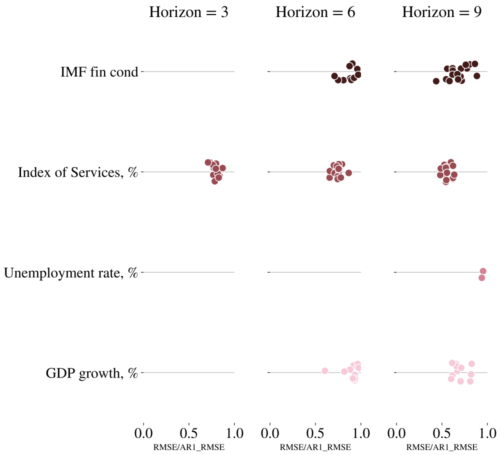
- Estimate residuals:

$$e_{t+h} = \eta \cdot z_t + v_{t+h}$$

- Use estimated residuals , $\hat{e}_{t+h}$, in the forecast of y using OLS.

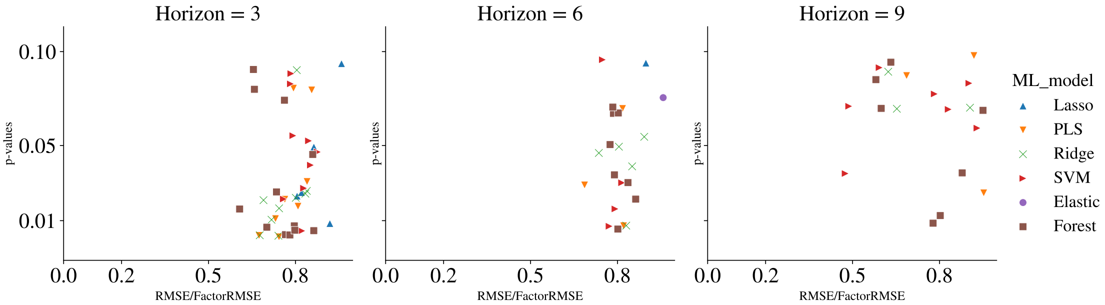
Forecast Results

Approach 1: Best targets across horizons



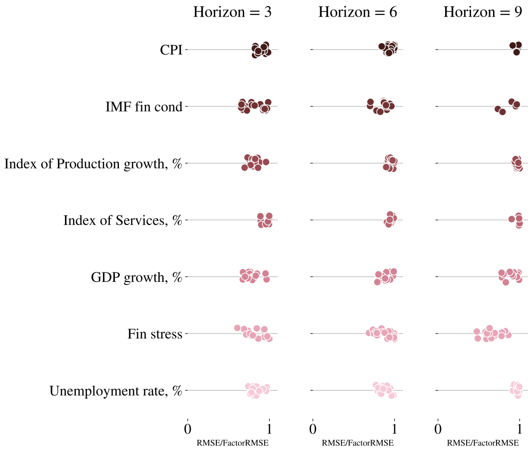
Each point denotes the RMSE using a different ML method and newspaper relative to the AR1. ▸ See D-M test

Approach 1: Best ML models across horizons



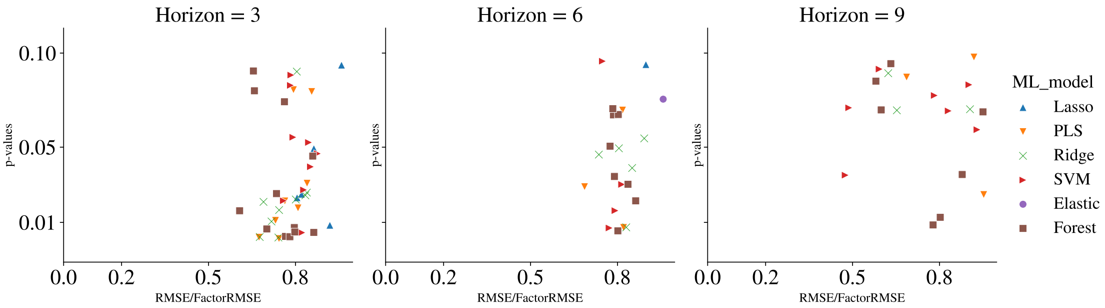
y-axis: Diebold and Mariano test p-values, x-axis: Out-of-sample RMSE relative to the AR(1).

Approach 2: Best targets across horizons



Each point denotes the RMSE using a different ML model and newspaper relative to the factor model. See D-M test

Approach 2: Best ML models across horizons



y-axis: Diebold and Mariano test p-values, x-axis: Out-of-sample RMSE relative to the factor model.

Discussion and Summary

Summary

News is not just noise!

- Text based time series strongly correlated to proxies 6 months ahead.
  - Worse for financial data, worst for uncertainty

Summary

News is not just noise!

- Text based time series strongly correlated to proxies 6 months ahead.
  - Worse for financial data, worst for uncertainty
- Relative forecast improvements at h = 3,6,9 steps ahead.

Summary

News is not just noise!

- Text based time series strongly correlated to proxies 6 months ahead.
    - Worse for financial data, worst for uncertainty
- Relative forecast improvements at h = 3,6,9 steps ahead.
- Best performance for real economy variables.

Summary

- Algorithmic best metric:
  - No single metric dominates all others. Single measures based on counting a single word (e.g tf_idf_econom) perform well too.

Summary

- Algorithmic best metric:
  - No single metric dominates all others. Single measures based on counting a single word (e.g tf_idf_econom) perform well too.

- Machine learning methods can extract more information from text, giving relative forecast improvements for a wider range of variables and for longer horizons.

Thank you.

Appendix

Augmented Dickey-Fuller tests

	The Guardian	No. obs.	The Daily Mirror	No. obs.	The Daily Mail	No. obs.
Afinn_sentiment	-03.03**	313	-04.28***	247	-03.05**	261
Harvard_sentiment	-04.17***	312	-07.83***	248	-02.94**	257
Loughran_sentiment	-02.20	313	-04.77***	248	-05.96***	265
Nyman_sentiment	-03.92***	303	-05.86***	248	-04.66***	264
alexopoulos_09	-03.97***	314	-05.64***	248	-04.30***	260
baker_bloom_davis	-04.14***	314	-14.69***	250	-13.09***	266
economcounts	-03.23**	313	-01.77	242	-03.47***	263
husted	-07.38***	315	-16.02***	250	-14.34***	266
opinion_sentiment	-03.26**	313	-06.34***	249	-04.13***	262
punctuation_sentimenteconom	-04.14***	307	-05.72***	247	-05.20***	262
stability_sentiment	-04.64***	314	-04.22***	248	-04.50***	265
sustainabcounts	-03.19**	308	-03.24**	245	-02.69*	255
tf_idf_econom	-03.57***	304	-03.01**	248	-02.48	257
tf_idf_uncert	-05.00***	312	-09.34***	249	-09.53***	265
uncertaincounts	-02.13	301	-02.55	241	-04.08***	261
vader_sentiment	-03.02**	313	-04.08***	246	-04.29***	264

Results of the Augmented Dickey-Fuller test on all text metrics. No. observations differ as no. of lags chosen using AIC criterion. Asterisks denote p-values; 1%: ***, 5%: **, 10%: *. At 1%, can reject null for all metrics for at least 1 of the 3 papers except for 'sustainabcounts'. Go back

Granger causality tests – sentiment

	LLBBA	CPIBOPT	PMIComp	UKFSI	3mTB10yNB	LLBBC	IGSPREAD	GfKCC	OECDUKCLI	ECCC	ECCCQ4	ECCCQ2	IMFGBFCI
economcounts	12.45***	8.57***	4.91***	8.98***	3.95***	1.97	1.82	2.91*	4.87***	2.69*	1.64	2.15	0.17
stability_sentiment	6.66***	4.78***	15.09***	5.18***	6.51***	2.98*	5.08***	1.03	1.87	0.89	0.26	0.60	3.65**
punctuation_sentimenteconom	2.42	1.51	2.09	13.27***	4.76***	2.75*	0.65	3.81**	4.41***	3.70**	4.80***	3.00*	2.28
tf_idf_econom	11.65***	13.10***	4.17***	4.84***	4.26***	2.91*	1.49	1.80	1.11	0.97	1.09	1.08	0.43
Loughran_sentiment	2.99*	6.50***	3.50**	1.83	2.83*	1.55	1.63	2.25	0.84	2.10	2.15	1.10	0.28
Harvard_sentiment	6.70***	0.51	3.14*	1.81	3.27*	2.79*	3.23*	2.05	1.53	1.34	1.13	0.21	0.76
Nyman_sentiment	3.89**	2.17	2.94*	1.57	2.33	3.28*	1.96	1.69	0.68	1.30	1.41	0.45	0.49
opinion_sentiment	3.07*	1.50	2.92*	1.95	3.63**	1.54	2.88*	1.22	0.77	1.04	0.65	0.44	0.54
vader_sentiment	1.93	5.14***	1.98	2.68*	2.07	0.91	0.62	1.18	1.07	1.00	1.02	1.21	0.40
Afinn_sentiment	1.85	1.31	2.36	0.80	3.21*	2.03	1.13	0.77	1.01	0.96	0.49	0.15	0.32
sustainabcounts	0.19	0.31	1.88	0.65	0.97	0.63	2.51	0.64	0.39	0.83	1.01	1.37	0.57

Results of a test looking at whether the sentiment text metrics Granger cause the proxies for sentiment, at a three month horizon. The text metrics are averaged using the weighted reach of the three newspapers' text. Asterisks denote p-values; 1%: ***, 2%: **, 5%: *. ▸ Go back

Granger causality tests – Uncertainty

	UKFSI	VFTSEIX	forecast_std_dev	EPUUK	IGSPREAD	JurardoFinUh3	3mTB10yNB	JurardoMacroUh3	UNCERBOE	VIX
tf_idf_uncert	6.09***	2.40	2.31	3.73**	1.52	1.22	1.37	0.96	0.60	0.88
alexopoulos_09	1.82	2.20	2.35	0.98	2.11	3.05*	2.04	1.49	1.94	0.83
uncertaincounts	1.54	2.10	1.85	0.87	1.92	2.53	1.92	1.44	1.45	0.52
baker_bloom_davis	2.77*	3.05*	0.50	1.94	1.74	0.62	0.51	0.65	0.75	1.70
husted	1.88	1.11	1.96	1.29	1.09	0.08	1.06	1.53	0.44	0.41

Results of a test looking at whether the uncertainty text metrics Granger cause the proxies for uncertainty, at a three month horizon. The text metrics are averaged using the weighted reach of the three newspapers' text. Asterisks denote p-values; 1%: ***, 2%: **, 5%: *. ▸ Go back

Targets used

Target	Description	Type	Frequency
LFSURATE	LFS unemployment rate	Real	Monthly
CPIall	CPI	Real	Monthly
IOS	Index of Services	Real	Monthly
IOP	Index of Production	Real	Monthly
ABJR.Q	Household Consumption	Real	Quarterly
gan8.q	Business Investment	Real	Quarterly
UKFSI	Chatterjee et al. (2017) Financial Stress Index	Financial	Monthly
IMFGBFCI	IMF UK Financial Condition Index	Financial	Monthly

▸ Go back

Proxies used

Name	Description	Type
LLBBC	Lloyds Business Barometer – confidence	Sentiment
LLBBA	Lloyds Business Barometer – activity over next 12 months	Sentiment
ECCCQ2	Gfk/EC financial situation of household over next 12 months	Sentiment
OECDUKCLI	OECD UK composite leading indicator	Sentiment
ECCCQ4	Gfk/EC general economic situation over the next 12 months	Sentiment
CPIBOPT	CBI Business Optimism	Sentiment
PMIComp	Composite measure of PMI	Sentiment
GfKCC	GfK Consumer Confidence	Sentiment
ECCC	European Commission Consumer Confidence	Sentiment
3mTB10yNB	Nominal 10 year yield less Treasury bill 3 month yield	Uncertainty, sentiment
IGSPREAD	Investment Grade Corporate Spread	Uncertainty, sentiment
JuradoFinUh3	Jurado, Ludvigson and Ng (2015) financial uncertainty 3 months ahead	Uncertainty
JuradoMacroUh3	Jurado, Ludvigson and Ng (2015) macroeconomic uncertainty 3 months ahead	Uncertainty
UKFSI	Chatterjee et al. (2017) financial stress index	Uncertainty
VIX	CBOE volatility index	Uncertainty
UNCERBOE	Bank of England uncertainty measure	Uncertainty
VFTSEIX	FTSE volatility	Uncertainty
EPUUK	Baker, Bloom and Davis (2016) economic policy uncertainty index for UK	Uncertainty
forecast_std_dev	UK Treasury collected standard deviation of professional forecasts of GDP, 3 months ahead	Uncertainty
IMFGBFCI	IMF UK Financial Condition Index	Uncertainty

Descriptions of the proxy time series and what they are used for.

Go back

Augmented Dickey-Fuller tests

	The Guardian	No. obs.	The Daily Mirror	No. obs.	The Daily Mail	No. obs.
Afinn_sentiment	-03.03**	313	-04.28***	247	-03.05**	261
Harvard_sentiment	-04.17***	312	-07.83***	248	-02.94**	257
Loughran_sentiment	-02.20	313	-04.77***	248	-05.96***	265
Nyman_sentiment	-03.92***	303	-05.86***	248	-04.66***	264
alexopoulos_09	-03.97***	314	-05.64***	248	-04.30***	260
baker_bloom_davis	-04.14***	314	-14.69***	250	-13.09***	266
economcounts	-03.23**	313	-01.77	242	-03.47***	263
husted	-07.38***	315	-16.02***	250	-14.34***	266
opinion_sentiment	-03.26**	313	-06.34***	249	-04.13***	262
punctuation_sentimenteconom	-04.14***	307	-05.72***	247	-05.20***	262
stability_sentiment	-04.64***	314	-04.22***	248	-04.50***	265
sustainabcounts	-03.19**	308	-03.24**	245	-02.69*	255
tf_idf_econom	-03.57***	304	-03.01**	248	-02.48	257
tf_idf_uncert	-05.00***	312	-09.34***	249	-09.53***	265
uncertaincounts	-02.13	301	-02.55	241	-04.08***	261
vader_sentiment	-03.02**	313	-04.08***	246	-04.29***	264

Results of the Augmented Dickey-Fuller test on all text metrics. No. observations differ as no. of lags chosen using AIC criterion. Asterisks denote p-values; 1%: ***, 5%: **, 10%: *. At 1%, can reject null for all metrics for at least 1 of the 3 papers except for ‘sustainabcounts’. Go back

Granger causality tests – sentiment

	punctuation_sentimenteconom	stability_sentiment	tf_idf_econom	economcounts	Loughran_sentiment	vader_sentiment	Harvard_sentiment	Afinn_sentiment	opinion_sentiment	Nyman_sentiment	sustainabcounts
LLBBC	14.21***	8.85***	3.06*	2.41	3.88**	4.47***	7.08***	2.96*	2.10	2.22	1.17
IGSPREAD	6.67***	3.58**	4.85***	8.29***	4.20***	3.22*	3.83**	4.43***	2.49	1.84	2.24
PMIComp	7.11***	2.02	1.94	1.36	7.58***	4.64***	3.22*	7.11***	5.26***	1.58	0.56
IMFGBFCI	1.87	3.83**	11.71***	6.84***	3.53**	2.85*	1.89	1.71	0.78	0.41	0.29
CPIBOPT	4.93***	8.65***	4.41***	3.43**	3.10*	2.53	1.76	1.81	2.59	1.30	1.19
LLBBA	6.42***	3.11*	0.47	1.25	4.25***	6.43***	5.73***	3.49**	2.65	1.26	0.59
UKFSI	8.78***	6.56***	5.95***	2.82*	1.77	1.67	2.50	1.54	1.03	0.20	0.65
ECCCQ4	11.95***	3.44**	1.94	1.72	1.69	2.94*	2.36	1.35	0.63	0.27	0.52
OECDUKCLI	4.41***	8.03***	3.76**	2.06	1.59	2.22	1.38	1.67	1.42	1.08	0.39
GfKCC	3.96***	2.21	5.17***	4.49***	3.01*	2.73*	2.06	1.14	0.66	0.24	0.17
ECCC	2.54	2.61	3.95***	3.17*	2.23	3.12*	0.84	0.96	0.23	0.17	0.34
ECCCQ2	0.66	0.74	2.10	1.76	2.40	1.52	0.49	0.56	0.77	0.65	0.12
3mTB10yNB	1.28	0.17	0.52	1.14	0.51	1.31	0.48	0.32	0.36	0.08	0.45

Results of a test looking at whether proxies of sentiment Granger cause the text metrics, at a three month horizon. The text metrics are averaged using the weighted reach of the three newspapers' text. Asterisks denote p-values; 1%: ***, 2%: **, 5%: *. Go back

Granger causality tests – uncertainty

	alexopoulos_09	uncertaincounts	tf_idf_uncert	baker_bloom_davis	husted
EPUUK	58.94***	66.22***	36.26***	20.93***	11.27***
VFTSEIX	13.99***	7.81***	3.05*	3.11*	6.76***
UNCERBOE	7.13***	4.55***	1.71	3.31*	1.44
IGSPREAD	7.19***	1.94	0.89	1.60	1.69
JurardoMacroUh3	2.80*	2.46	1.69	0.18	3.08*
UKFSI	1.42	0.94	1.12	1.74	0.33
3mTB10yNB	0.45	0.22	1.54	0.83	1.57
VIX	0.74	0.50	0.61	1.33	0.84
JurardoFinUh3	0.51	0.58	0.95	0.02	1.40
forecast_std_dev	0.70	0.43	0.28	0.06	1.29

Results of a test looking at whether proxies of uncertainty Granger cause the text metrics, at a three month horizon. The text metrics are averaged using the weighted reach of the three newspapers' text. Asterisks denote p-values; 1%: ***, 2%: **, 5%: *. ▸ Go back

Weighting applied when pooling newspapers

Newpaper	Reach	Weighted Reach
The Guardian	0.144	0.228
Daily Mirror	0.148	0.234
Daily Mail	0.338	0.536
Total	0.63	1

Reach is the share of the population that reports getting news from each source (Kennedy and Prat, 2017).

▸ Go back

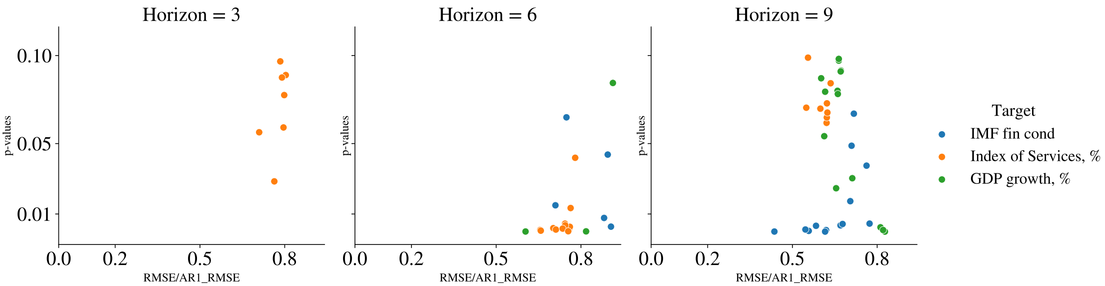
Example of a $N \times P$ term frequency matrix

Date	policy	risk	economy	bank	limit
2017-01-01	3	0	0	1	0
2017-02-01	2	1	2	3	0
2017-03-01	1	0	0	0	0
2017-04-01	1	0	0	0	0
2017-05-01	1	0	0	0	0
2017-06-01	1	0	0	0	0
2017-07-01	1	0	0	0	0
2017-08-01	1	0	0	0	0
2017-09-01	0	0	0	0	0
2017-10-01	1	0	0	1	0

$N \times P$ dimensional term frequency matrix aggregated to monthly frequency; Dimensions $N = 325$(=total number of months), $P = 5656$ =(Total number of features), $P >> N$.

▸ Go back

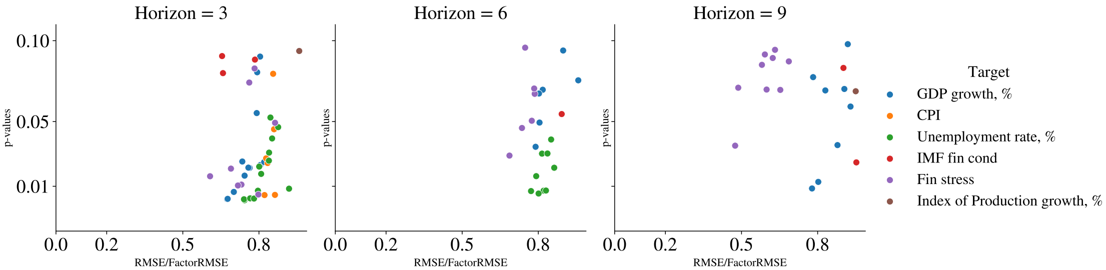
D-M test: Best targets across horizons



y-axis: Diebold and Mariano test p-values, x-axis: Out-of-sample RMSE relative to the AR(1).

Go Back

D-M test: Best targets across horizons



y-axis: Diebold and Mariano test p-values, x-axis: Out-of-sample RMSE relative to the factor model. Go Back

References

Alexopoulos, Michelle, Jon Cohen, et al. 2009. "Uncertain times, uncertain measures." *University of Toronto Department of Economics Working Paper*, 352.

Baker, Scott R, Nicholas Bloom, and Steven J Davis. 2016. "Measuring economic policy uncertainty." *The Quarterly Journal of Economics*, 131(4): 1593–1636.

Breiman, Leo. 2001. "Random forests." *Machine learning*, 45(1): 5–32.

Chatterjee, Somnath, Ching-Wai (Jeremy) Chiu, Sinem Hacioglu-Hoke, and Thibaut Duprey. 2017. "A financial stress index for the United Kingdom." Bank of England Staff Working Paper 697.

Chin, Wynne W. 1998. “The partial least squares approach to structural equation modeling.” *Modern methods for business research*, 295(2): 295–336.

Correa, Ricardo, Keshav Garud, Juan M Londono, Nathan Mislang, et al. 2017. “Constructing a Dictionary for Financial Stability.” Board of Governors of the Federal Reserve System (US).

Drucker, Harris, Christopher JC Burges, Linda Kaufman, Alex J Smola, and Vladimir Vapnik. 1997. “Support vector regression machines.” 155–161.

Gilbert, CJ Hutto Eric. 2014. “Vader: A parsimonious rule-based model for sentiment analysis of social media text.”

Hoerl, Arthur E, and Robert W Kennard. 1970. “Ridge regression: Biased estimation for nonorthogonal problems.” *Technometrics*, 12(1): 55–67.

Hu, Guoning, Preeti Bhargava, Saul Fuhrmann, Sarah Ellinger, and Nemanja Spasojevic. 2017. "Analyzing Users' Sentiment Towards Popular Consumer Industries and Brands on Twitter." 381–388, IEEE.

Hu, Minqing, and Bing Liu. 2004. "Mining and summarizing customer reviews." 168–177, ACM.

Husted, Lucas F., John Rogers, and Bo Sun. 2017. "Monetary Policy Uncertainty." Board of Governors of the Federal Reserve System (U.S.) International Finance Discussion Papers 1215.

Jurado, Kyle, Sydney C Ludvigson, and Serena Ng. 2015. "Measuring uncertainty." *The American Economic Review*, 105(3): 1177–1216.

Kennedy, Patrick, and Andrea Prat. 2017. "Where Do People Get Their News?" Columbia Business School Research Papers 17-65.

Keynes, John Maynard. 1936. *The general theory of employment, interest, and money.* Springer.

Loughran, Tim, and Bill McDonald. 2013. "IPO first-day returns, offer price revisions, volatility, and form S-1 language." *Journal of Financial Economics*, 109(2): 307–326.

Nielsen, Finn Årup. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs." *arXiv preprint arXiv:1103.2903.*

Nyman, Rickard, Sujit Kapadia, David Tuckett, David Gregory, Paul Ormerod, and Robert Smith. 2018. "News and narratives in financial systems: exploiting big data for systemic risk assessment." *Bank of England Staff Working Papers*, 704.

Tetlock, Paul C. 2007. "Giving content to investor sentiment: The role of media in the stock market." *The Journal of finance*, 62(3): 1139–1168.

References v

Tibshirani, Robert. 1996. “Regression shrinkage and selection via the lasso.” *Journal of the Royal Statistical Society. Series B (Methodological)*, 267–288.

Zou, Hui, and Trevor Hastie. 2005. “Regularization and variable selection via the elastic net.” *Journal of the royal statistical society: series B (statistical methodology)*, 67(2): 301–320.