How digital media markets amplify news sentiment

Lara Marie Berger

University of Cologne

28th of August 2024
EEA ESEM, Rotterdam

Introduction

(How) does the internet change media content?

Introduction

(How) does the internet change media content?

- The media is known to make the world seem more sensational than it actually is (Ryu, 1982; Soroka, 2016; Soroka and Krupnikov, 2021, e.g.).
  $\rightarrow$ Nobody reports about a plane that arrives on time.
- Sensational headlines resonate stronger with readers (Kuiken et al., 2017).

Introduction

(How) does the internet change media content?

- The media is known to make the world seem more sensational than it actually is (Ryu, 1982; Soroka, 2016; Soroka and Krupnikov, 2021, e.g.).
  $\rightarrow$ Nobody reports about a plane that arrives on time.
- Sensational headlines resonate stronger with readers (Kuiken et al., 2017).
- The internet comes with incentives that might enhance this bias:
  - Each headline is in competition with thousands of others.
  - Ad-revenues often depend on click-rates.
  - Journalists receive immediate feedback.

Introduction

(How) does the internet change media content?

- The media is known to make the world seem more sensational than it actually is (Ryu, 1982; Soroka, 2016; Soroka and Krupnikov, 2021, e.g.).
  $\rightarrow$ Nobody reports about a plane that arrives on time.
- Sensational headlines resonate stronger with readers (Kuiken et al., 2017).
- The internet comes with incentives that might enhance this bias:
  - Each headline is in competition with thousands of others.
  - Ad-revenues often depend on click-rates.
  - Journalists receive immediate feedback.

$\rightarrow$ *Does the digitization of media markets enhance sensationalist bias?*

$\rightarrow$ *If so, does this change something for readers?*

This paper

I analyze a huge dataset of news articles online and offline and conduct experiments with professional journalists and readers.

This paper

I analyze a huge dataset of news articles online and offline and conduct experiments with professional journalists and readers.

Preview of results:

1. Online headlines are more often emotional than offline headlines.

This paper

I analyze a huge dataset of news articles online and offline and conduct experiments with professional journalists and readers.

Preview of results:

1. Online headlines are more often emotional than offline headlines.
2. A driver of this difference is the incentive for journalists to capture readers' attention.

This paper

I analyze a huge dataset of news articles online and offline and conduct experiments with professional journalists and readers.

Preview of results:

1. Online headlines are more often emotional than offline headlines.
2. A driver of this difference is the incentive for journalists to capture readers' attention.
3. The emotionality of headlines can translate into alterations in readers' emotions and belief updating.

Are headlines online more emotional than offline?

Data: 339,865 news articles about the economy from major German news outlets from 2003-2022 obtained through *LexisNexis*

Are headlines online more emotional than offline?

Data: 339,865 news articles about the economy from major German news outlets from 2003-2022 obtained through *LexisNexis*

Data for main analyses: Only economic news, includes article content

news outlet	time-frame	N online	N offline	N total
BILD	01/01/2017 - 01/06/2022	4,680	3,092	7,772
Der Spiegel	01/02/2003 - 01/06/2022	79,192	14,296	93,488
Die Welt	05/07/2009 - 01/06/2022	65,226	24,535	89,761
Die Zeit	10/01/2009 - 01/06/2022	50,643	64,260	114,903
Rheinische Post	10/09/2020 - 01/06/2022	11,348	22,593	33,941
all from above	all available points in time	211,089	128,776	339,865

▸ Robustness Datasets

Are headlines online more emotional than offline?

Empirical Strategy

- find differences in headline tonality in the data ▸ Regression equations
- note: no causal claims possible

Are headlines online more emotional than offline?

Empirical Strategy

- find differences in headline tonality in the data ▸ Regression equations
- note: no causal claims possible

Classification

- fine-tune pre-trained machine learning model to classify headlines into three classes (roBERTa) ▸ Choice of Classifier

Are headlines online more emotional than offline?

Empirical Strategy

- find differences in headline tonality in the data ▸ Regression equations
- note: no causal claims possible

Classification

- fine-tune pre-trained machine learning model to classify headlines into three classes (roBERTa) ▸ Choice of Classifier

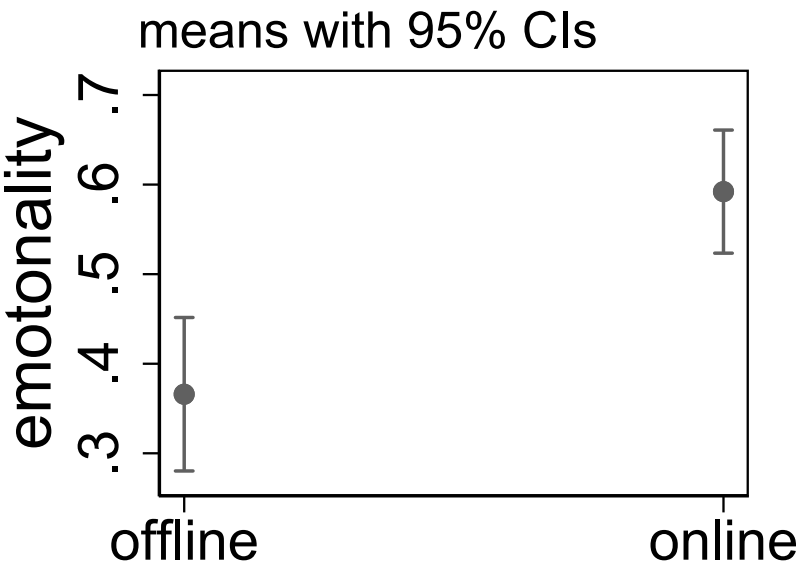
Main Outcomes

- classes: positive (1) , neutral (0), negative (-1)
- emotionality: positive/negative (1), neutral (0)

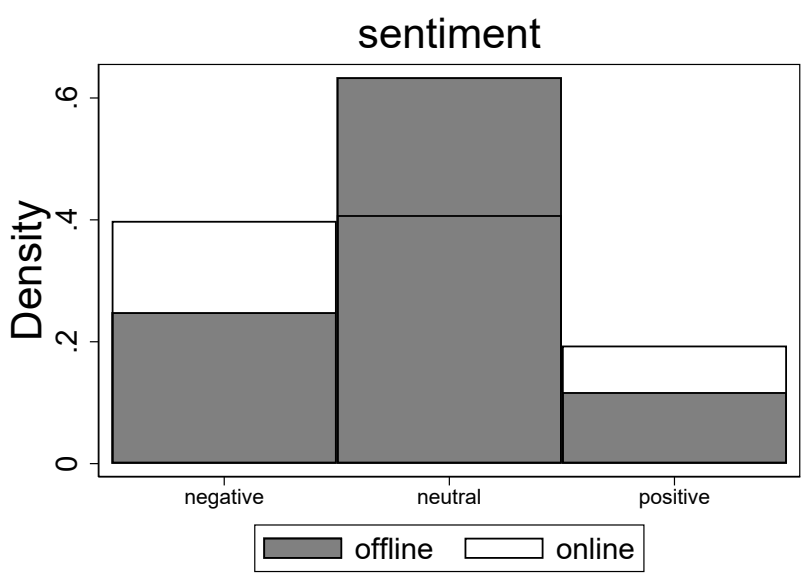
Are headlines online more emotional than offline?

Results:

(a) Means and confidence intervals



(b) Distribution



Are headlines online more emotional than offline?

Robustness:

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles
- non-economic headlines ▸ Data ▸ Figures

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles
- non-economic headlines ▸ Data ▸ Figures
- non-German news outlet ▸ Data ▸ Figures

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles
- non-economic headlines ▸ Data ▸ Figures
- non-German news outlet ▸ Data ▸ Figures
- other popular sentiment classifiers ▸ Data ▸ Figures

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles
- non-economic headlines ▸ Data ▸ Figures
- non-German news outlet ▸ Data ▸ Figures
- other popular sentiment classifiers ▸ Data ▸ Figures
- different points in time ▸ Figures

Are headlines online more emotional than offline?

Robustness:

- subset of matched articles
- non-economic headlines ▸ Data ▸ Figures
- non-German news outlet ▸ Data ▸ Figures
- other popular sentiment classifiers ▸ Data ▸ Figures
- different points in time ▸ Figures

Result 1: *Headlines of news outlets are written more often emotionally for online audiences than for offline audiences.*

What causes the difference?

Experiment with Professional Journalists

What causes the difference?

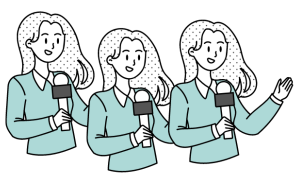
Experiment with Professional Journalists

- **Sample:** 201 full-time journalists recruited via email by biggest German journalist association. Age: 18-80, work experience: 2-42 years, work for outlets from the entire political spectrum.

What causes the difference?

Experiment with Professional Journalists

- **Sample:** 201 full-time journalists recruited via email by biggest German journalist association. Age: 18-80, work experience: 2-42 years, work for outlets from the entire political spectrum.
- **Experimental Design:** Journalists have to choose an either positive, neutral or negative headline for a given article about the German economy. *Randomization* occurs on the journalist's incentives.



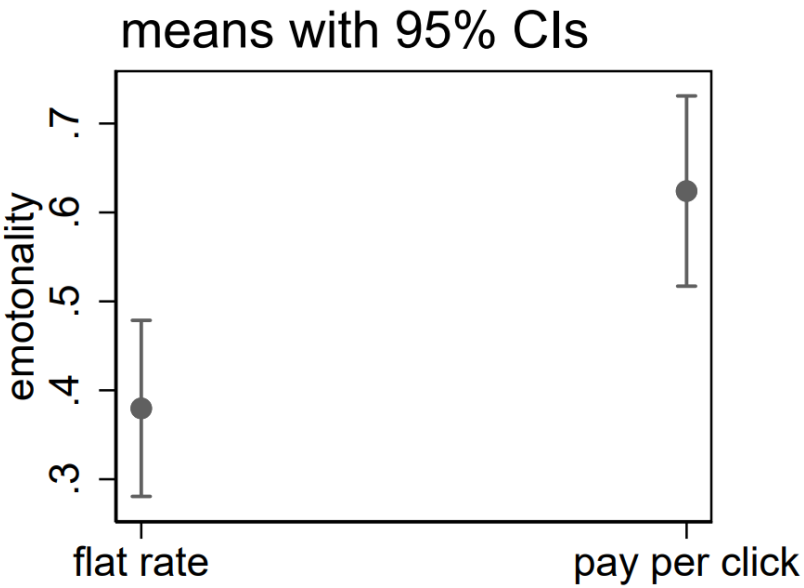
control-group
receives flat payment

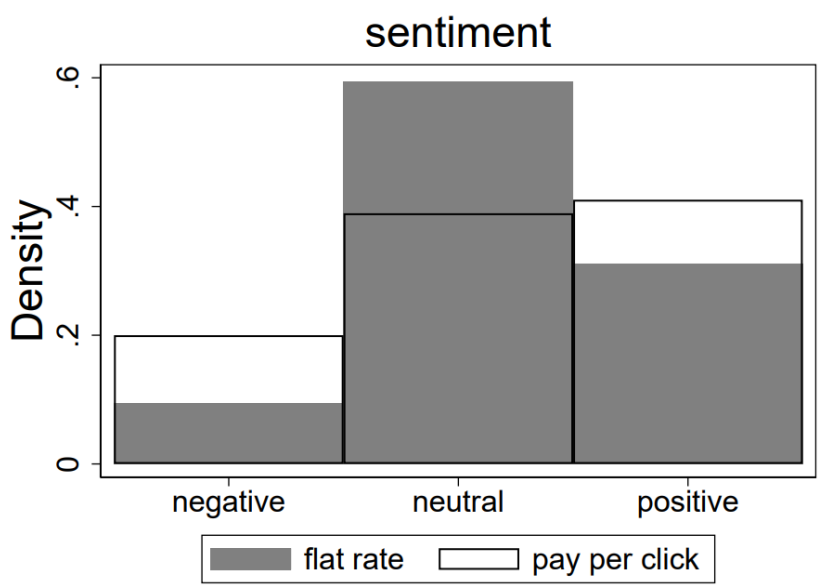
treatment-group
is paid per click on headline

▸ Overview ▸ Article ▸ Headlines

What causes the difference?

Experiment with Professional Journalists





What causes the difference?

Result 2: *Incentives to generate attention can induce journalists to select emotional headlines.*

What causes the difference?

Assumption: Incentives to capture attention are higher in digital markets than in analog markets.

What causes the difference?

Assumption: Incentives to capture attention are higher in digital markets than in analog markets.

- Plausible reason: different degrees of competition

What causes the difference?

Assumption: Incentives to capture attention are higher in digital markets than in analog markets.

- Plausible reason: different degrees of competition
- Work in progress: Experiment with journalists that varies the degree of headline competition.

What causes the difference?

The role of competition

1. **Evidence from headline competition in offline markets**
   - If the degree of competition is really related to headline tonality, frontpage headlines should be more emotional than print headlines on other pages.

What causes the difference?

The role of competition

1. **Evidence from headline competition in offline markets**
   - If the degree of competition is really related to headline tonality, frontpage headlines should be more emotional than print headlines on other pages.
   - Empirical evidence with robustness datasets: Frontpage headlines are indeed on average more emotional and more negative. ▸ Figures

What causes the difference?

The role of competition

1. **Evidence from headline competition in offline markets**
   - If the degree of competition is really related to headline tonality, frontpage headlines should be more emotional than print headlines on other pages.
   - Empirical evidence with robustness datasets: Frontpage headlines are indeed on average more emotional and more negative. ▸ Figures
2. **Evidence from headline competition in online markets**
   - Inspired by Meyer et al. (2022) I exploit an exogenous variation in the competition of headlines in online markets induced by a legal dispute in 2013 in Germany.

What causes the difference?

The role of competition

1. **Evidence from headline competition in offline markets**
   - If the degree of competition is really related to headline tonality, frontpage headlines should be more emotional than print headlines on other pages.
   - Empirical evidence with robustness datasets: Frontpage headlines are indeed on average more emotional and more negative. ▸ Figures
2. **Evidence from headline competition in online markets**
   - Inspired by Meyer et al. (2022) I exploit an exogenous variation in the competition of headlines in online markets induced by a legal dispute in 2013 in Germany.
   - DiD framework: Lower competition reduced the average emotionality of the online headlines of the affected outlet. ▸ Figures

How does this influence readers?

Experiment with Readers

- 299 students are randomly exposed to one of the headlines from the journalist experiment

How does this influence readers?

Experiment with Readers

- 299 students are randomly exposed to one of the headlines from the journalist experiment
- elicitation of short-term emotions

How does this influence readers?

Experiment with Readers

- 299 students are randomly exposed to one of the headlines from the journalist experiment
- elicitation of short-term emotions
- elicitation of beliefs regarding content of article (incentivized)

How does this influence readers?

Experiment with Readers

- 299 students are randomly exposed to one of the headlines from the journalist experiment
- elicitation of short-term emotions
- elicitation of beliefs regarding content of article (incentivized)
- short, partially related investment choice (incentivized)

How does this influence readers?

Experiment with Readers

Results:

How does this influence readers?

Experiment with Readers

Results:

1. Those who read an emotional headline have beliefs that are more distorted from the content in the article. ▸ Figure

How does this influence readers?

Experiment with Readers

Results:

1. Those who read an emotional headline have beliefs that are more distorted from the content in the article. ▸ Figure
2. Those who read a positive headline feel better. ▸ Figure

How does this influence readers?

Experiment with Readers

Results:

1. Those who read an emotional headline have beliefs that are more distorted from the content in the article. ▸ Figure
2. Those who read a positive headline feel better. ▸ Figure
3. No statistically significant difference between the investment decisions.

How does this influence readers?

Experiment with Readers

Results:

1. Those who read an emotional headline have beliefs that are more distorted from the content in the article. ▸ Figure
2. Those who read a positive headline feel better. ▸ Figure
3. No statistically significant difference between the investment decisions.

Result 3: *The emotionality of headlines can translate into alterations in readers' emotions and belief updating.*

Open questions

Limitations

- Readers experiment only looks at *one side of the coin*:

Open questions

Limitations

- Readers experiment only looks at *one side of the coin*:
  - More emotional headlines might distort readers' beliefs and thus could be bad for overall knowledge.

Open questions

Limitations

- Readers experiment only looks at *one side of the coin*:
  - More emotional headlines might distort readers' beliefs and thus could be bad for overall knowledge.
  - More emotional headlines might however also induce readers to read more and could thus be good for overall knowledge.

Open questions

Limitations

- Readers experiment only looks at *one side of the coin*:
  - More emotional headlines might distort readers' beliefs and thus could be bad for overall knowledge.
  - More emotional headlines might however also induce readers to read more and could thus be good for overall knowledge.
- Work in progress: New readers' experiment that analyses this trade-off in detail.

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

I find that:

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

I find that:

1. Headlines of news outlets are written more often emotionally for online audiences than for offline audiences.

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

I find that:

1. Headlines of news outlets are written more often emotionally for online audiences than for offline audiences.
2. Incentives to generate attention induce journalists to select emotional headlines.

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

I find that:

1. Headlines of news outlets are written more often emotionally for online audiences than for offline audiences.
2. Incentives to generate attention induce journalists to select emotional headlines.
3. The emotionality of headlines can translate into alterations in readers' emotions and belief updating.

Conclusion

I explore how digital markets influence the sensationalist bias of news headlines.

I find that:

1. Headlines of news outlets are written more often emotionally for online audiences than for offline audiences.
2. Incentives to generate attention induce journalists to select emotional headlines.
3. The emotionality of headlines can translate into alterations in readers' emotions and belief updating.

Main Contribution: First paper (to the best of my knowledge) to show how the internet can amplify the sensationalist bias in the news media.

References I

- Kuiken, Jeffrey, Anne Schuth, Martijn Spitters, and Maarten Marx (2017). "Effective headlines of newspaper articles in a digital environment". In: *Digital journalism* 5.10, pp. 1300–1314.
- Meyer, Tim, Anna Kerkhof, Carmelo Cennamo, and Tobias Kretschmer (2022). "Competition for attention on information platforms: the case of local news outlets". In: *Working paper*.
- Ryu, Jung S. (1982). "Public affairs and sensationalism in local tv news programs". In: *Journalism mass communication quarterly* 59.1, pp. 74–137.
- Shapiro, Adam Hale, Moritz Sudhof, and Daniel J. Wilson (2022). "Measuring news sentiment". In: *Journal of econometrics* 228.2, pp. 221–243.

References II


Soroka, Stuart (2016). "Gatekeeping and the negativity bias". In: *Oxford research encyclopedia of politics*.


Soroka, Stuart and Yanna Krupnikov (2021). *The increasing viability of good news*. Elements in Politics and Communication. Cambridge University Press.

Descriptive Datasets (incl. Robustness)

Data for main analyses: Only economic news, includes article content

news outlet	time-frame	N online	N offline	N total
BILD	01/01/2017 - 01/06/2022	4,680	3,092	7,772
Der Spiegel	01/02/2003 - 01/06/2022	79,192	14,296	93,488
Die Welt	05/07/2009 - 01/06/2022	65,226	24,535	89,761
Die Zeit	10/01/2009 - 01/06/2022	50,643	64,260	114,903
Rheinische Post	10/09/2020 - 01/06/2022	11,348	22,593	33,941
all from above	all available points in time	211,089	128,776	339,865

Data for robustness checks: Includes all topics, but headlines only

news outlet	time-frame	N online	N offline	N total
BILD	01/01/2017 - 17/05/2022	66,720	97,576	164,296
Der Spiegel	01/01/2021 - 31/12/2021	19,122	4,896	24,018
The New York Times	01/01/2021 - 09/11/2021	25,230	29,216	54,446

▸ Main Dataset ▸ Robustness Results

Choice of Classifier

Classification of Main Outcomes:

- many different sentiment classification approaches available
- choice of the classifier can play a crucial role for outcomes and different text types require different classifiers (Shapiro, Sudhof, and Wilson, 2022)
- Approach: Let RAs classify random subset of headlines, rate performance of different popular classifiers by the overlap with these classifications

Figure: Comparison of Sentiment Classifiers

	Sentiment		*Emotionality*	
algorithm	**accuracy**	**macro F1**	**accuracy**	**F1**
SentiWS	0.4773	0.3729	0.5413	0.3174
LM	0.6106	0.5257	0.6373	0.5436
VADER	0.5920	0.5554	0.6800	0.6428
pre-trained roBERTa	0.6853	0.6638	0.7200	0.7301
fine-tuned roBERTa	0.7253	0.7057	0.7413	0.7581

Are headlines online more emotional than offline?

Comparison:

$$tonality_i = \beta_0 + \beta_1 * online_i + \epsilon_i \quad (1)$$

Are headlines online more emotional than offline?

Comparison:

$$tonality_i = \beta_0 + \beta_1 * online_i + \epsilon_i \quad (1)$$

$$tonality_{ijkt} = \beta_0 + \beta_1 * online_i + \beta_2 X_{ijkt} + \epsilon_{ijkt} \quad (2)$$

- *control variables:* tonality of content, article length, agency content dummy, outlet fixed effects, topic fixed effects, time fixed effects
- standard errors are bootstrapped with 50 replications based on 5 clusters at the level of the news outlet

▸ back

Are headlines online more emotional than offline?

Results:

Table 2: OLS Estimates - Emotionality of Headlines in Standard Deviations

	(1)	(2)	(3)	(4)	(5)
online	0.4524***	0.3107***	0.3685***	0.3923***	0.2544***
	(0.0730)	(0.0440)	(0.0604)	(0.0595)	(0.0312)
content emotionality		0.7206***			0.6570***
		(0.0353)			(0.0330)
article length					-0.0047
					(0.0171)
agency content					0.0726*
					(0.0376)
topic FE	no	no	yes	no	yes
time FE	no	no	no	yes	yes
Constant	-0.2809***	-0.6064***	-0.2580***	0.000	-0.0145
	(0.0874)	(0.0266)	(0.0919)	(0.0516)	(0.0168)
R^2	0.0482	0.1705	0.0889	0.0262	0.1667
Observations	339,865	339,865	339,865	339,865	339,865

Are headlines online more emotional than offline?

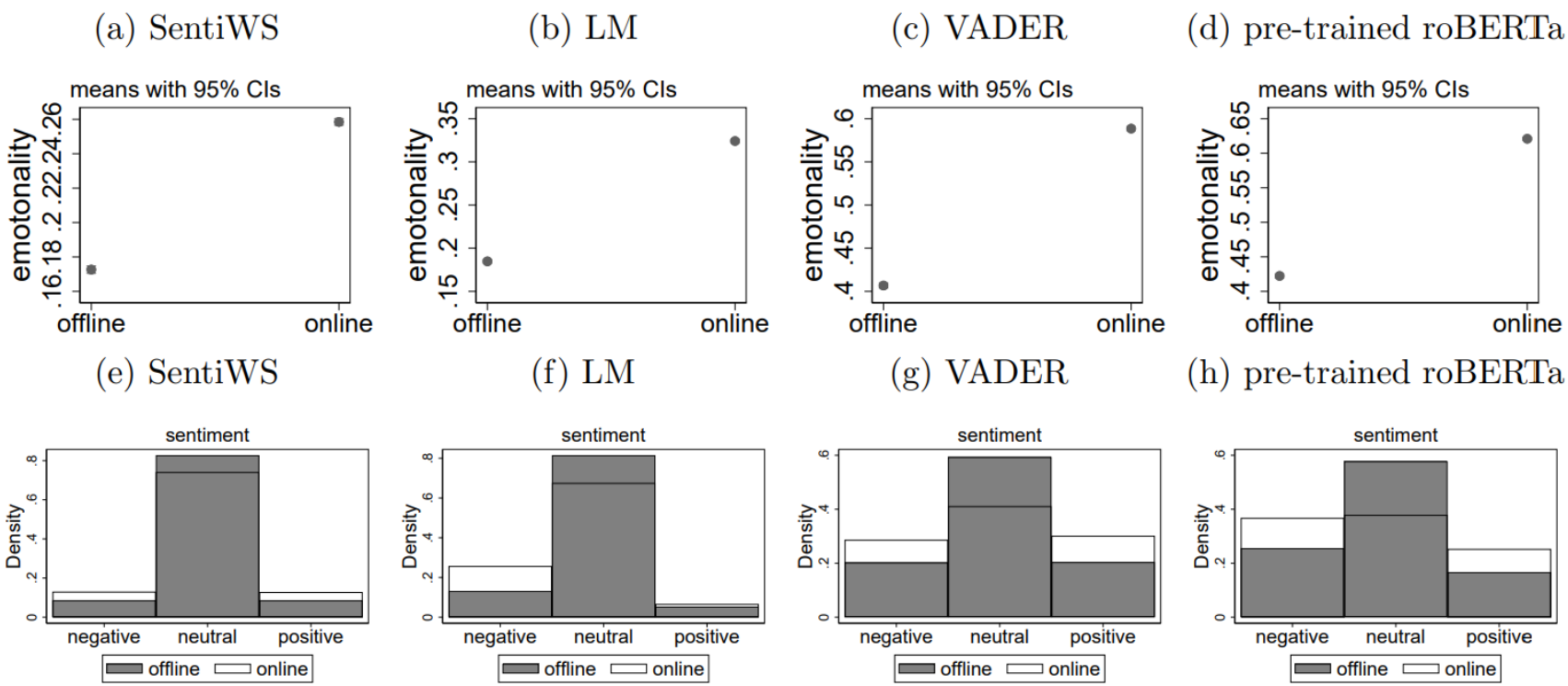
Results:

Table 3: OLS Estimates - Sentiment of Headlines in Standard Deviations

	(1)	(2)	(3)	(4)	(5)
online	-0.1065***	-0.0599***	-0.1078***	-0.0878***	-0.0622***
	(0.0373)	(0.0139)	(0.0287)	(0.0245)	(0.0176)
content sentiment		0.7488***			0.7371***
		(0.0435)			(0.0409)
article length					-0.0362**
					(0.0177)
agency content					-0.0043
					(0.0222)
topic FE	no	no	yes	no	yes
time FE	no	no	no	yes	yes
Constant	0.0661	0.2540***	0.0692	0.0000	0.0143
	(0.0406)	(0.0241)	(0.0606)	(0.0108)	(0.0128)
R^2	0.0027	0.2768	0.0188	0.0013	0.2765
Observations	339,865	339,865	339,865	339,865	339,865

Are headlines online more emotional than offline?

Figure 3: Robustness Check: Classifications with Different Algorithms

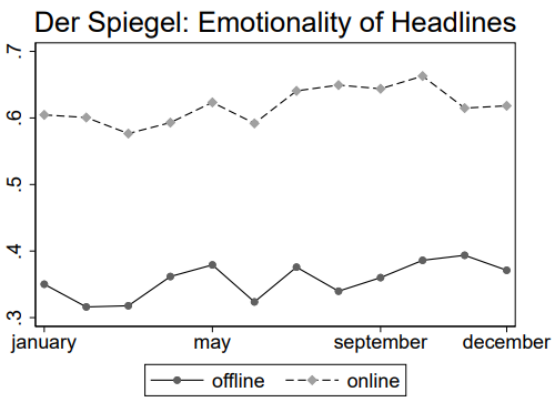


▸ back

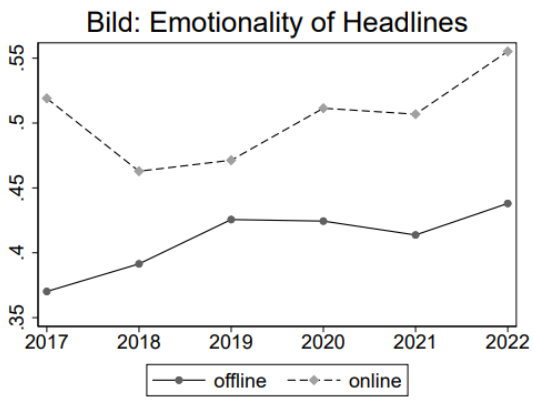
Are headlines online more emotional than offline?

Figure 4: Robustness Check: Classifications with Different Datasets

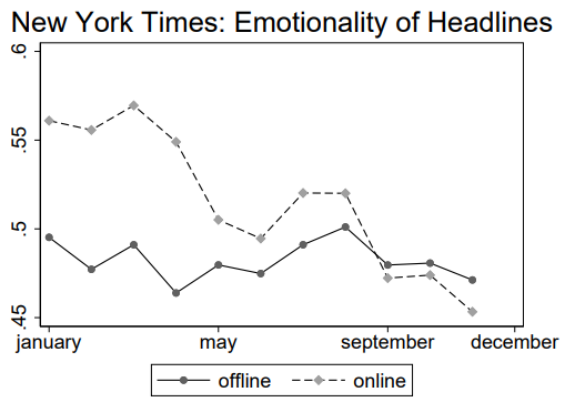
(a) Der Spiegel



(b) BILD



(c) The New York Times

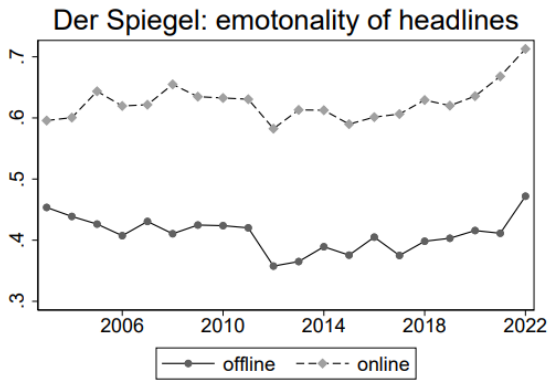


back

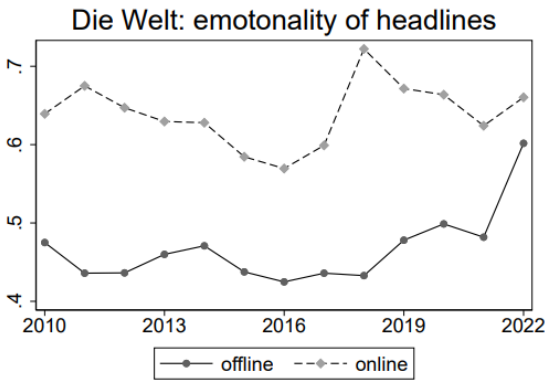
Are headlines online more emotional than offline?

Figure 2: Emotionality and Sentiment of Headlines over Time

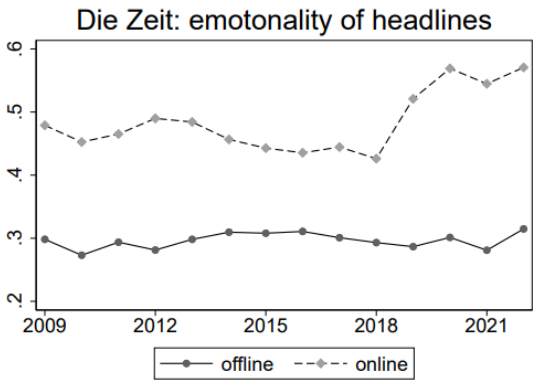
(a) Der Spiegel



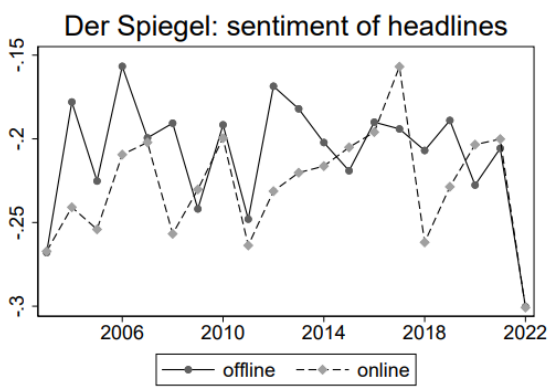
(b) Die Welt



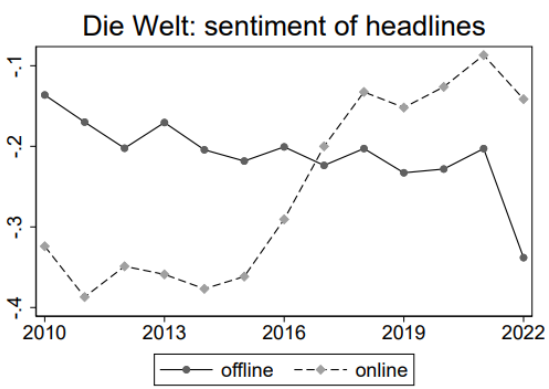
(c) Die Zeit



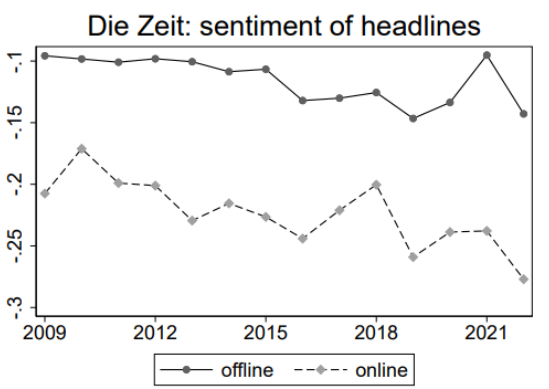
(d) Der Spiegel



(e) Die Welt



(f) Die Zeit



▸ back

Are headlines online more emotional than offline?

Topic heterogeneity

Are there topics for which online headlines are especially often emotional (or negative) while offline headlines are not?

- estimate interaction terms for different topic dummies with the online dummy
- headlines for topics that tend to polarize such as "environment" or "immigration" are formulated especially emotional in the online sphere
- no systematic difference detectable for sentiment

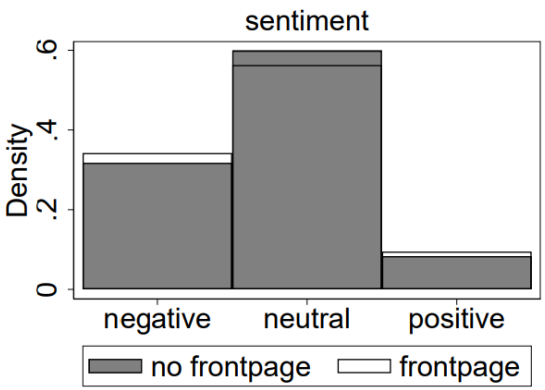
Keep in mind that all topics are "subtopics" of articles on economic issues.

back

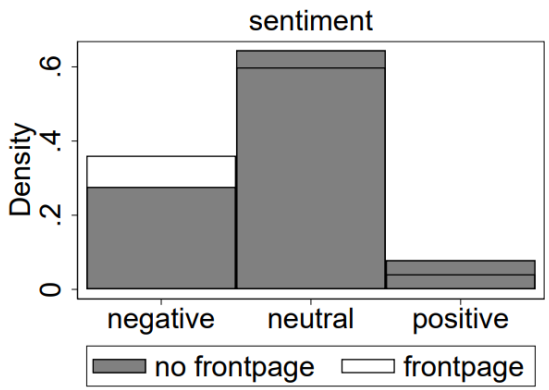
What causes the difference?

Figure 7: Sentiment of Frontpage- and Non-Frontpage Headlines

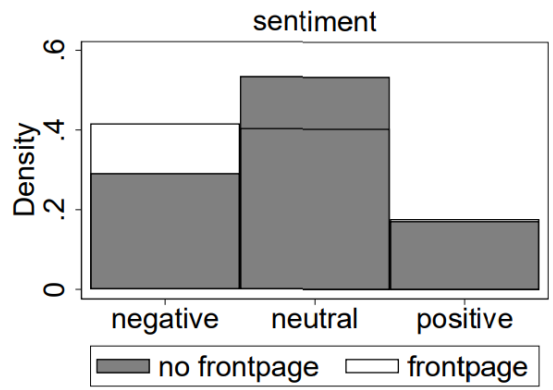
(a) BILD



(b) Der Spiegel



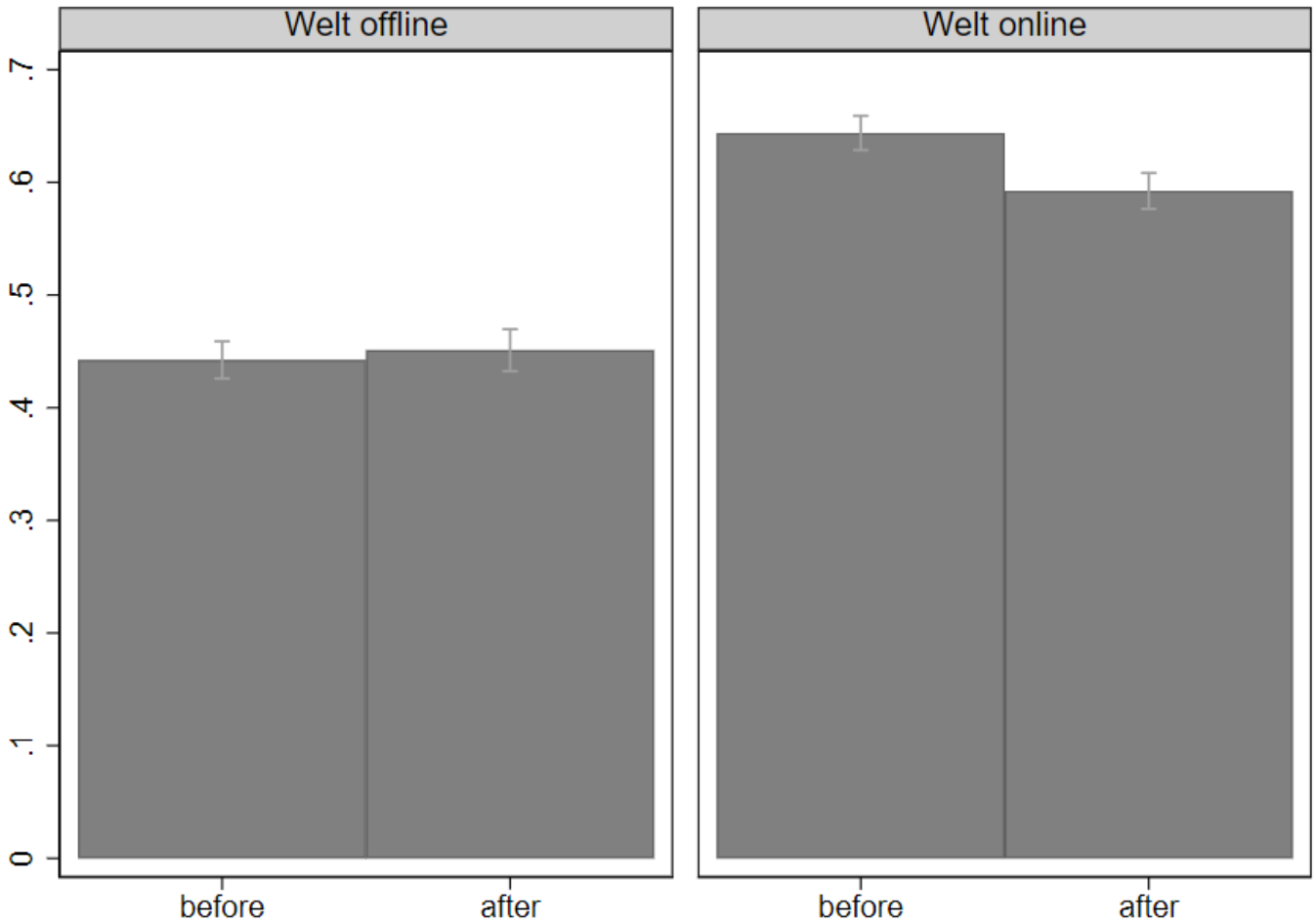
(c) The New York Times



▸ back

What causes the difference?

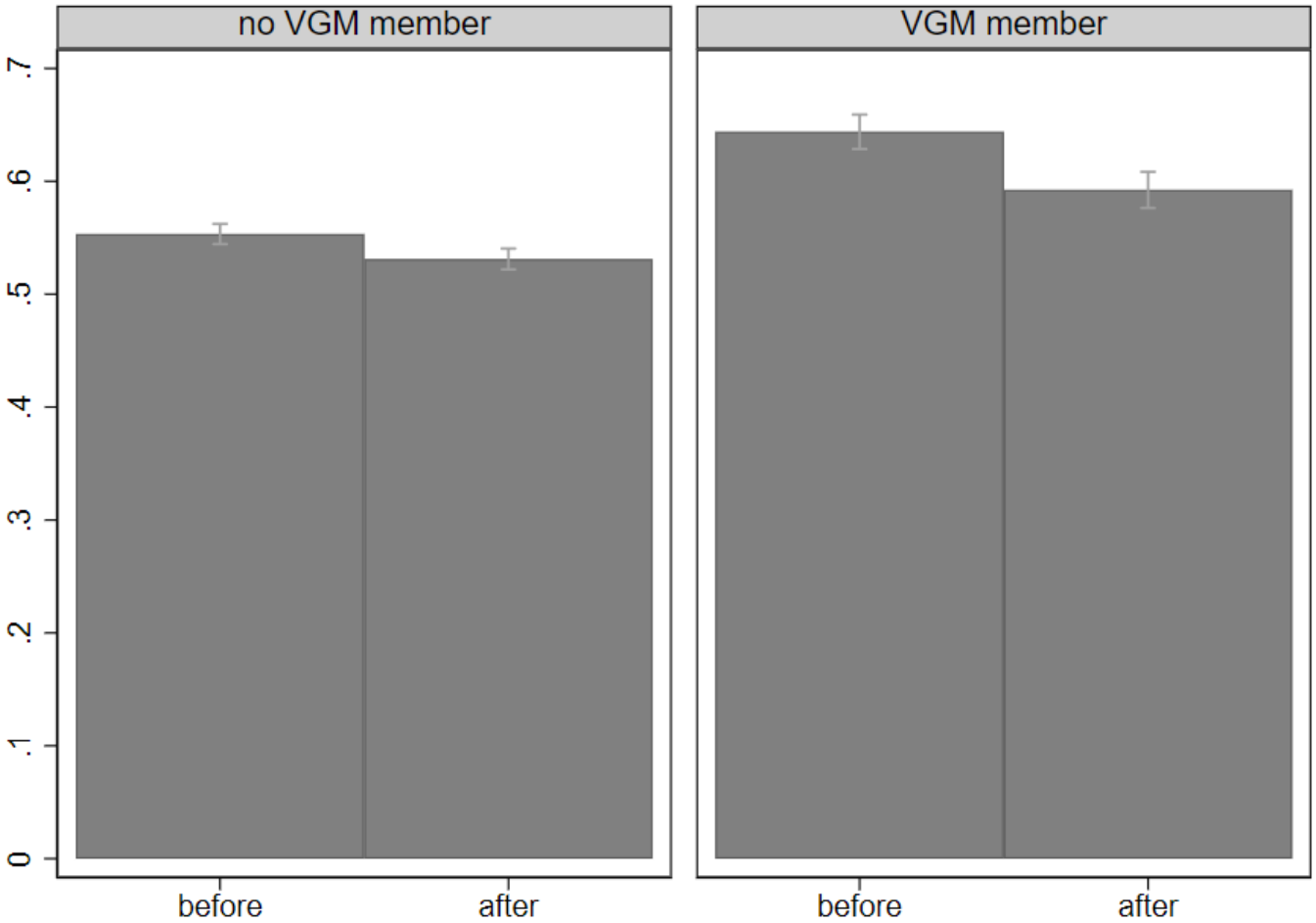
(b) Emotionality of the online and offline headlines of the removed outlet



back

What causes the difference?

(a) Emotionality of the online headlines of the removed and non-removed outlets



back

What causes the difference?

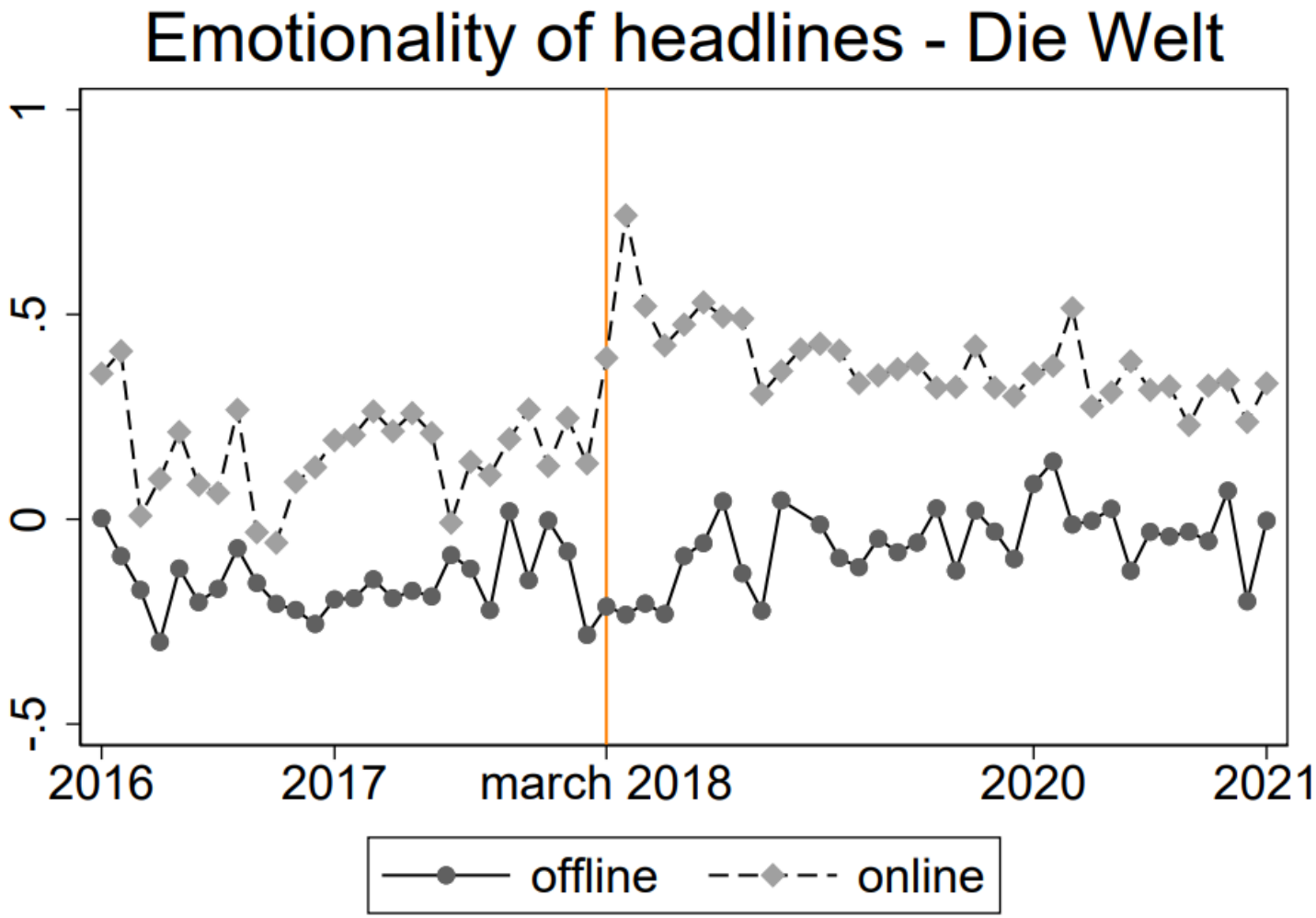
Suggestive evidence from descriptive data

1. **Automated Finance News at Welt Online**
   - Lack of space limit in the online sphere allows the publication many articles even if they only target a very niche readership.
   - *Welt Online* implemented a rather primitive "robo-writer" for finance news in March 2018
     ("Stock X performed better than index Y today")
   - Online and offline emotionality develops parallel before introduction (placebo-tests) and anticipation highly implausible ▸ Event Graph
   - Difference-in-difference reveals that introduction of robo-news increased the share of positive and negative headlines

▸ back

What causes the difference?

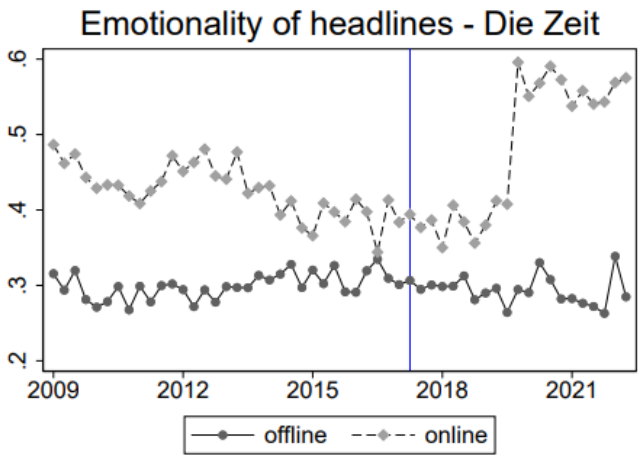
Figure 5: Emotionality of Welt Headlines over time (Agency only)



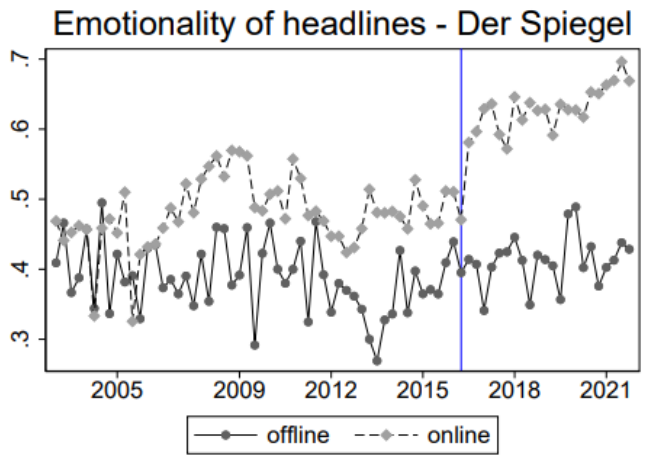
What causes the difference?

Figure 6: Emotionality of Headlines over time (No-Agency only)

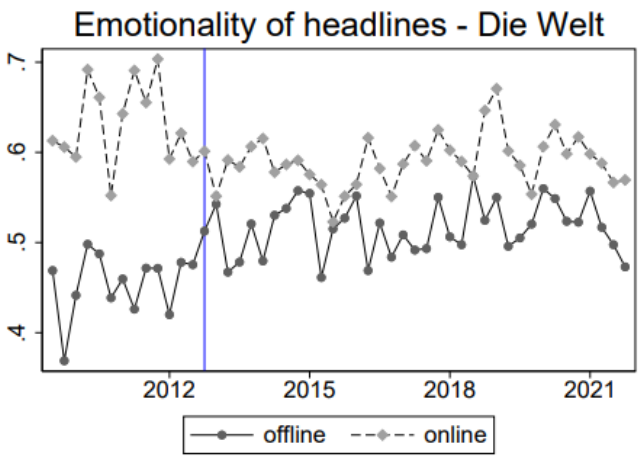
(a) Die Zeit



(b) Der Spiegel

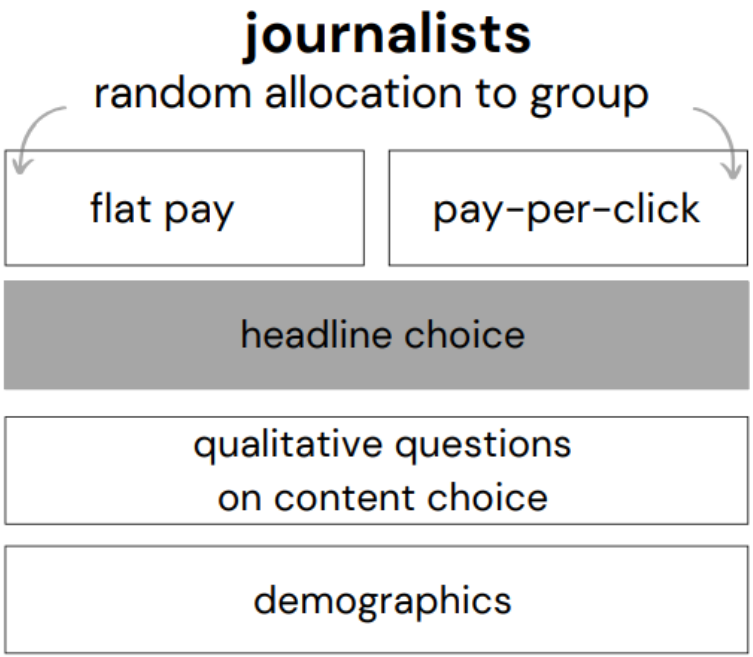


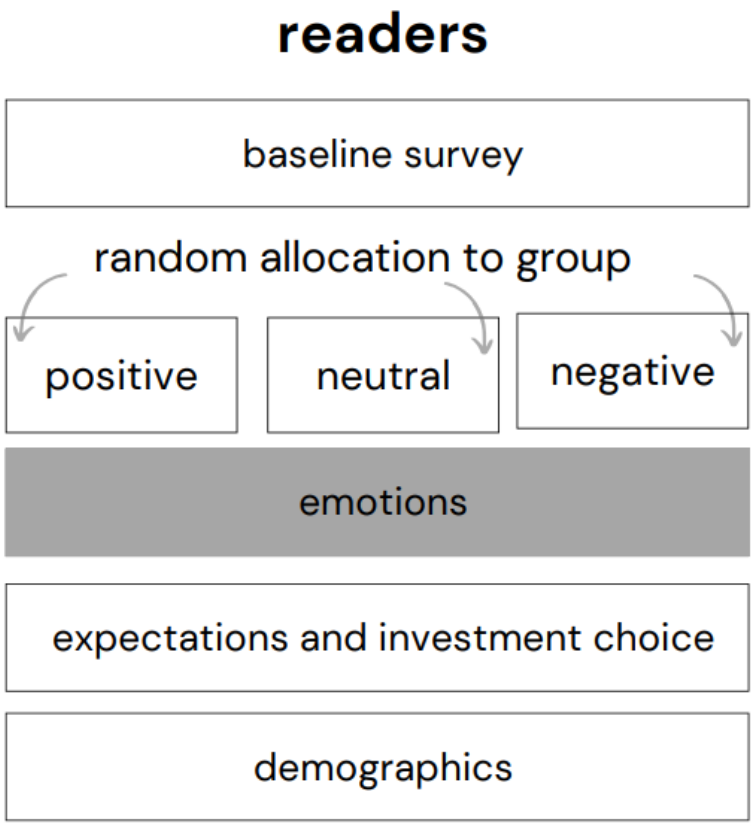
(c) Die Welt



back

What causes the difference?





back

What causes the difference?

"In ihrem im Oktober veröffentlichten Herbstgutachten gehen führende Wirtschaftsforschungsinstitute von einem Wachstum des Bruttoinlandsprodukts um 2,4 Prozent in 2021 aus. Im Frühjahr hatten sie noch damit gerechnet, dass in diesem Jahr ein Anstieg um 3,7 Prozent zu erwarten sei.

Die wirtschaftliche Lage in Deutschland sei nach wie vor von der Coronapandemie gekennzeichnet, hieß es. Im Verlauf des Jahres 2022 dürfte die deutsche Wirtschaft aber wieder die Normalauslastung erreichen. Laut Prognose der Institute steigt das Bruttoinlandsprodukt im Jahr 2022 um 4,8 Prozent. In ihrer Frühjahrsprognose gingen die Institute nur von einem Plus um 3,9 Prozent für das nächste Jahr aus."

Quelle: dpa

▸ headline choices ▸ back

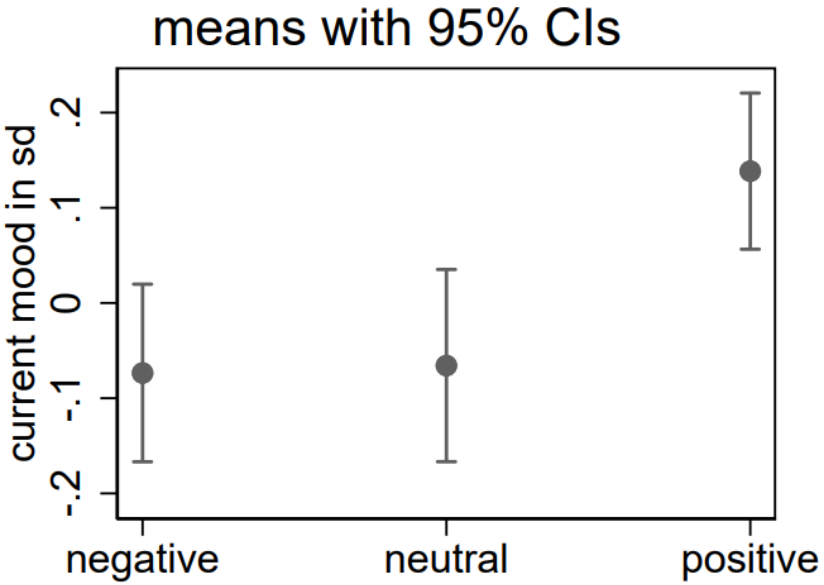
What causes the difference?

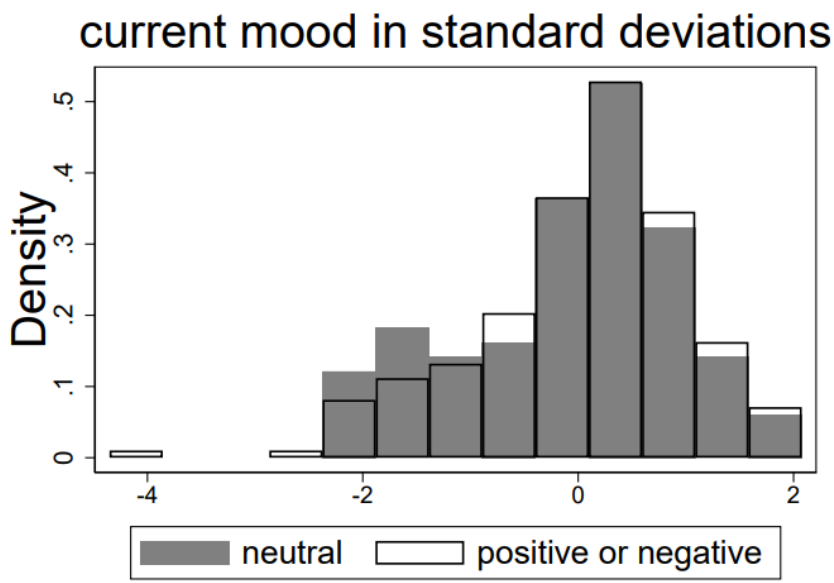
Welche dieser drei Überschriften würden Sie am ehesten über die untenstehende Meldung setzen?

- **Prognose macht Mut: 2022 soll die deutsche Wirtschaft wieder stark wachsen** ○
- **Prognose: So wird sich die deutsche Wirtschaft in nächster Zeit entwickeln** ○
- **Prognose macht Angst: 2021 läuft für deutsche Wirtschaft schlechter als erwartet** ○

▸ article ▸ back

How does this influence readers?



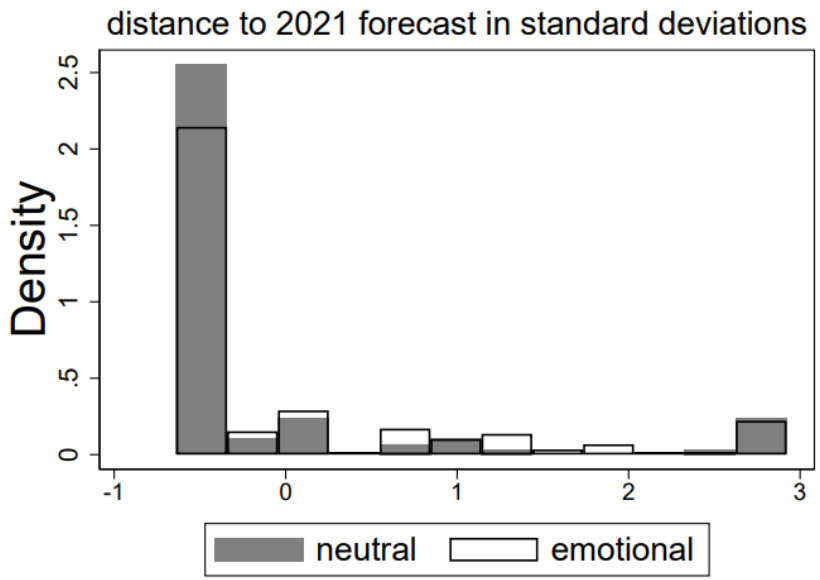


back

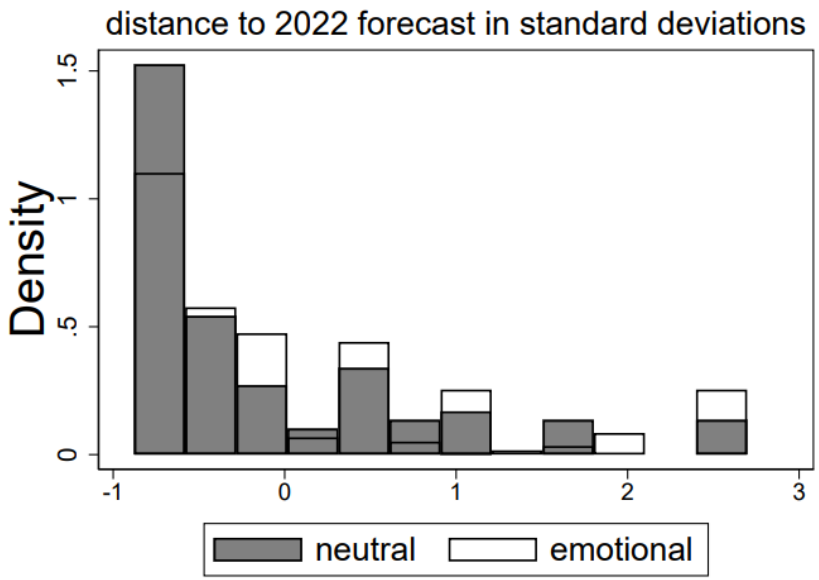
How does this influence readers?

Figure 11: Belief Updating by Readers

(a) 2021



(b) 2022



back